NAB Broadcast Engineering

CONFERENCE PROCEEDINGS 2008

Welcome to the place where creativity originates, inspiration comes alive and dreams become reality.

NABSHOW

Where Content Comes to Life™

A collection of papers
from the 62nd NAB Broadcast
Engineering Conference

Conferences: April 11-17, 2008
Exhibits: April 14-17, 2008
Las Vegas Convention Center • Las Vegas
www.nabshow.com

2007-2008 NAB BROADCAST ENGINEERING CONFERENCE COMMITTEE

Mr. Steve Fluker, (Chairman)
Director of Engineering
Cox Radio, Orlando
4192 N. John Young Parkway
Orlando, FL 32825

Mr. Talmage Ball
VP, Engineering
Bonneville International
55 North 300 West
Salt Lake City, UT 84180

Mr. Dave Converse
VP & Director of Engineering
ABC/Disney Television Stations Group
500 S. Buena Vista Street
Burbank, CA 91521

Mr. Al Grossniklaus
Director of Engineering & Operations
WTHR NBC 13
1000 North Moridian Street
Indianapolis, IND 46204

Mr. Gary Kline
Vice President of Engineering & IT
Cumulus Media
3535 Piedmont Road
Building 14 - Suite 1400
Atlanta, GA 30339

Mr. Wayne Kube
Technology Manager, Broadcast Media
Belo Corporation
400 S. Record Street
Dallas, TX 75202

Mr. John Merrill
Director of Engineering
CBS 5 KPHO-TV
4016 N Black Canyon Highway
Phoenix, AZ 85017

Mr. Paul Shulins
Director of Technical Operations
Greater Media, Boston
55 Morrissey Blvd.
Boston, MA 02125

SBE Representatives:

Mr. John L. Poray
Executive Director, SBE
9102 North Meridian Street, Suite 150
Indianapolis, IN 46260

Mr. Frederick Baumgartner, CPBE
Dir Broadcast Engineering
Qualcomm MediaFLO
E-148 D, 5775 Morehouse Dr
San Diego, Ca 92121-1714

TABLE OF CONTENTS

BROADCAST ENGINEERING CONFERENCE OPENING SESSION
Sunday, April 13, 2008
9:00 AM – 9:30 AM 1

DIGITAL OPPORTUNITIES FOR RADIO
Sunday, April 13, 2008
9:30 AM – 12:00 PM 3

Conditional Access: The Next Stage in HD Radio™ Evolution
Tom Rucktenwald, NDS 5

Managing Radio Metadata for Multiplatform Digital Distribution
Daniel Mansergh, KQED Public Radio 15

The Future of Radio in a Changing World
Dave Wilson, Consumer Electronics Association.. 21

Digital Opportunities for Radio
Melinda Driscoll, American Public Media | Minnesota Public Radio, Laura Jensen, NPR, Nick Kereakos, American Public Media | Minnesota Public Radio 31

DTV BROADCASTING FOR MOBILE AND HANDHELD
Sunday, April 13, 2008
9:30 AM – 5:30 PM 45

Managing the Unmanageable: Transforming Media for Multi-chanel, Multi-platform Delivery
John Pallett, Telestream, Inc. 47

Workflow in the New Media Landscape: Production, Post and Final Delivery
Michael Castro, National Geographic.............. 51

Local and Ultra Local Content in Broadcast Mobile TV
Richard Lhermitte, ENENSYS Technologies.. 55

The Role of File-based Workflows and Metadata in Repurposing Content for the Web and Mobile TV
Mark Bishop and Alex Holtz, Thomson Grass Valley 62

Leveraging FEC Advances to Optimize DVB-H Networks
Marshall Porter, Jerry Miller and Thomas Stockhammer, Digital Fountain...................... 67

New Techniques for Mobile TV Broadcasting Based on ISDB-T
Masahiro Okano, Yoshikazu Narikiyo, Makoto Taguchi, Kenichi Tsuchida, Masayuki Takada, Hiroyuki Hamazumi, Yasuhiro Ito and Kazuhiko Shibuya, NHK Science & Technical Research Laboratories 73

Preserving SFN in a Broadcast Network Using IP Distribution
Nicolas Fannechere, ENENSYS Technologies.. 79

Improving Mobile TV Streaming: Lessons Learned for Successful Deployments
Boris Felts and Ian Locke, Envivio Inc. 83

RADIO TECHNOLOGY ADVANCEMENTS
Sunday, April 13, 2008
1:00 PM – 5:30 PM 91

Radio Broadcasters: Building File-Based Networks
Eric Wiler, Jones Radio Networks and Gary Pelkey, WEGENER.. 93

HD Audio Quality and Netcasting
Greg Ogonowski, Orban/CRL Systems, Inc. 99

From ITM to ITWOM: Correcting, Completing, and Updating the Longley-Rice Irregular Terrain Model
Sid Shumate, Givens & Bell, Inc. 105

Can the Public Internet Be Used for Broadcast?
Simon Daniels, Audio Processing Technology 111

A New Approach to Peak-to-Average-Power Reduction for FM+IBOC Transmission
Philipp Schmid, Nautel, Inc. 118

Field Tests for Service Area and Handover Service in T-DMB
Sang-Hun Kim, Guy-Young Kim and Man-Sik Kim, Korean Broadcasting System 129

An Improved Coverage Prediction Method for HD Radio
John Kean, NPR Labs - National Public Radio 137

Advances in Digital Measurement Techniques for FM Broadcast
Frédéric Allard, Vincent Philippe and Tony Peterle, Audemat 146

COMMUNICATING WITH MANAGEMENT
Monday, April 14, 2008
10:30 AM – 12:00 PM 157

DTV RECEPTION ISSUES
Monday, April 14, 2008
10:30 AM – 12:00 PM 159

New Neighbors: Can Wireless Microphones and Consumer Devices Coexist In the White Spaces?
Christopher Lyons, Shure Incorporated 161

ALTERNATIVE STL TECHNOLOGIES
Monday, April 14, 2008
1:00 PM – 2:30 PM 167

The HD Radio™ STL: Issues, Options and Technologies
Bob Band, Harris Corporation 169

Best of Synchronous with the Best of IP
Guy Gampell, APT 176

Robust HD Radio Exporter to Exgine Architecture
Timothy B. Anderson, Harris Corporation 182

TV NEWS AND LIVE PRODUCTION
Monday, April 14, 2008
1:00 PM – 5:30 PM 191

Advanced Video Image Technologies for Sports TV Productions
Kimihiro Tomiyama, Kensuke Hisatomi, Miwa Katayama, Yuichi Iwadate, Tetsuya Hayashida and Hirotaka Maruyama, NHK Science & Technical Research Laboratories; T. Goji Etch, Kinki University and Tetsuo Yoshida, Hitachi Kokusai Electric Co. Ltd 193

Managing Multiformat Images for the Broadcast News Environment
Karl Paulsen, AZCAR Technologies............ 199

Best Practices: Using IP in Broadcast TV
Joel Wilhite, Harmonic Inc........................... 206

Live Integrated Production Systems Streamline Live Workflow
Ken Swanton, Broadcast Pix 217

High Definition Electronic News Gathering (HD-ENG) Field Test Report
Walter Sidas, Gregory Coppa and Robert P. Seidel, CBS.. 226

NBC Universal's New IPTV Distribution System
Robert Goldfarb and Shadrach Kisten, NBC-Universal.. 253

DIGITAL RADIO SUMMIT
Monday, April 14, 2008
3:00 PM – 5:00 PM 257

Bandwidth & Frequency Allocation Issues in International Digital Radio AM & FM Broadcasting
Charles W. Kelly, Nautel, Ltd. 259

New Standards and Codecs for European Digital Radio Broadcasting
Olaf Korte, Fraunhofer IIS 266

Mobile Coverage Optimization by Polarization Diversity in VHF and UHF Propagation
Myron D. Fanton, PE, Electronics Research, Inc. 271

Does Your Yotta Byte?
Andrew P. Janitschek, Radio Free Asia 274

NEW TECHNOLOGIES FOR RADIO LISTENING
Tuesday, April 15, 2008
9:00 AM – 11:30 AM 283

RF Simulcasting over IP Networks
Junius Kim, Harris Corporation 285

Practical Considerations of Radio Broadcast Operations in an Arbitron PPM Market
Larry Paulausky, Greater Media, Inc. 293

Consumer Ratings of Impaired Audio at Various Signal/Noise Ratios
Ellyn G. Sheffield, Towson University, John Kean, and Dan Schwab, NPR Labs – National Public Radio 299

VIDEO CONTENT CREATION & MANIPULATION
Tuesday, April 15, 2008
9:00 AM – 5:00 PM 303

File Formats in Television Archiving and Content Exchange
Peter Thomas, Blue Order Solutions AG 305

Non-Real-Time Services
Richard Chernock, Triveni Digital 311

BXF - The Promise of Reduced Costs and Increased Revenues
Chris Lennon, Harris Corporation 315

Seam Carving for Video
Mike Knee, Snell & Wilcox Ltd 320

An Integrated, File-based Production Workflow for HD Television: Expected Impact and Challenges
Luk Overmeire, Matthias De Geyter, Tom Beckers, Steven Van Assche, Luc Andries, VRT Medialab 330

From MXF to SOA
Ernesto Santos, MOG Solutions 337

Migration to All-IP Infrastructures for Distribution of Broadcast Services
Tom Lattie, Harmonic Inc 344

Forensic Marking for HD VoD and Broadcast Services
Pascal Marie, Thomson 348

From Camera to the Home -- Managing Aspect Ratio through the Production and Distribution Process
Larry Thaler, NBC-Universal 356

The State of Broadcast Automation
Sid Guel, Broadcast Automation Consulting 361

AUDIO OVER IP
Tuesday, April 15, 2008
1:00 PM – 2:30 PM 369

Rapid Radio Deployment Pack Emergency Edition
Pierre Robidoux, Canadian Broadcasting Corporation / Société Radio-Canada 371

Advanced Tech for IP Remotes
Steve Church, Telos Systems 378

FUTURE BROADCAST TECHNOLOGIES – A WORLDWIDE PERSPECTIVE
Tuesday, April 15, 2008
2:30 PM – 4:00 PM 387

NEXT GENERATION PUBLIC ALERTING
Tuesday, April 15, 2008
4:00 PM – 5:00 PM 389

MONITORING AND MEASUREMENTS IN THE BROADCAST PLANT - RADIO
Wednesday, April 16, 2008
9:00 AM – 12:00 PM 391

VSWR Measurements in Broadcast Transmission Systems
Tim Holt, Director, Bird Technologies Group 393

The Role of the Detector in Spectrum Analyzer Measurement of Hybrid Digital Signals
David Maxson, Broadcast Signal Lab, LLP 397

RF Measurement Techniques for Broadcast Engineers
Myron D. Fanton, P.E., Electronics Research, Inc. 405

RF Signal Performance Measurements of Consumer FM Receivers and Coverage Effects
John Kean, NPR Labs - National Public Radio 411

MONITORING AND MEASUREMENTS IN THE BROADCAST PLANT – TV
Wednesday, April 16, 2008
9:00 AM – 11:00 AM 419

Optimizing the QC Process in a File-based Workflow Facility
Rob Zwiebel, Harris Corporation.................. 421

System Wide Video Quality Assurance
Ralph Bachofen and Rich Chernock, Triveni Digital 425

Controlling and Measuring Loudness for Digital Television Broadcast
Michael Babbitt, Dolby Laboratories............ 430

TELEVISION RF & TRANSMISSION SYSTEMS
Wednesday, April 16, 2008
11:00 AM – 4:30 PM 441

Antennas for Distributed Transmission and Single Frequency Networks
Myron D. Fanton, P.E., Electronics Research, Inc. 443

Evaluation of Buildings Penetration Loss for 100 Buildings in Belgium
David Plets, Wout Joseph, Leen Verloock, Emmeric Tanghe, Luc Martens, Ghent University, Department of Information Technology; and Etienne Deventer and Hugo Gauderis, Flemish Radio and Television Network 446

RF Coverage and Tower Motion
Myron D. Fanton, Electronics Research, Inc. 451

RADIO RF & TRANSMISSION SYSTEMS
Wednesday, April 16, 2008
2:00 PM – 5:30 PM 455

AM Co-location: Money on the Table?
Lawrence Behr, LBA Group, Inc. 457

Radio Transmitter Maintenance "Back to Basics"
Paul Shulins, Greater Media 464

Implications of IBOC Injection Levels above -20dB
Gary Liebisch, Nautel 469

FM IBOC Building Penetration Tests at Elevated Digital Subcarrier Levels
E. Glynn Walden, CBS Radio 474

Free Software Tools for Design of AM Antennas
Van Richards-Smith, Radio TAB Network 502

TECHNOLOGY INNOVATIONS
Thursday, April 17, 2008
9:00 AM – 11:30 AM 513

Super Hi-Vision Transmission Experiment in the 21GHz Band
Hisashi Sujikai, Yoichi Suzuki, Shoji Tanaka and Kazuyoshi Shogen, NHK Science & Technical Research Labs 515

HDMI as Television Application Platform for Interactive and More
Rainer Zwing, and Gerold Lübcke, Thomson; Scott Francis, Greg Herlein and Dan Kyte, Thomson/PRN... 522

11.88 GB/S SDI Continuing the Evolution of SDI
Gareth M. Heywood and Ryan Latchman, Connectivity Technologies, Gennum Corporation 528

Audio Mixing Requirements in Next Generation Broadcast Receivers for Audio Description and Other Enhanced Features
Roland Vlaicu, Dolby Laboratories 546

HDTV System Onboard the Lunar Explorer Kaguya (SELENE)
Seiji Mitsuhashi, Junichi Yamazaki, Junichi Tachino, Masahito Yamauchi, NHK (Japan Broadcasting Corporation); Kazuo Tanimoto, Meisei Electric Co., Ltd.; Rie Honda, Kochi University; Yuichi Iijima, Hironori Maejima, Hisashi Otake and Shinichi Sobue, JAXA 557

CODECS, COMPRESSION SYSTEMS & SCALING FOR VIDEO
Thursday, April 17, 2008
9:00 AM – 12:00 PM 565

10 Bit High Quality MPEG-4 AVC Video Compression
Matthew Compton, TANDBERG Television, Part of the Ericsson Group 567

Practical Applications of Compression Standards
Todd Roth, NEXIO Server Systems, Broadcast Communications Division, Harris Corporation 573

Providing Spatial Scalability Using Scalable Video Coding to Mobile Broadcasting
Sangjin Hahm, Changseob Park, Byungsun Kim, Injoon Cho, Keunsoo Park, Korean Broadcasting System; MunChurl Kim, ICU........................ 592

Scalable Video Coding (SVC) and Broadcast Delivery of 1080P High Definition
Elie Sader, Harmonic Inc.............................. 597

Understanding and Implementing JPEG 2000 Compression for Long-form EFP Acquisition
John Naylor and David Bancroft, Thomson 604

Bridging the Gap with HD Transcoding
Tim Simerly, Texas Instruments.................... 612

Author and Presenter Index 621

Company Index......................................623

Broadcast Engineering Conference Opening Session

Sunday, April 13, 2008
9:00 AM – 9:30 AM

Chairperson: Lynn Claudy, Senior Vice President
NAB, Washington, DC

*We Can Work Together: Advice to DTV and HD Radio Engineers from the Consumer Electronics Retail Community

Diane Warren, Executive Vice President, HD Digital Radio Alliance, San Antonio, TX
Robert Schwartz, Counsel, Consumer Electronics Retailers Coalition, Washington, DC

*Paper not available at the time of publication

Digital Opportunities for Radio

Sunday, April 13, 2008
9:30 AM – 12:00 PM

Chairperson: Paul Shulins
Greater Media, Boston, MA

Conditional Access: The Next Stage in HD Radio™ Evolution
Tom Rucktenwald, NDS, Costa Mesa, CA

Managing Radio Metadata for Multiplatform Digital Distribution
Daniel Mansergh, KQED Public Radio, San Francisco, CA

The Future of Radio in a Changing World
Dave Wilson, Consumer Electronics Association, Arlington, VA

*Seeding the Internet -- Automating Podcasting with Open Source Tools
Frederick Gleason Jr., Paravel Systems LLC, Warrenton, VA

Digital Opportunities for Radio
Melinda Driscoll, American Public Media | Minnesota Public Radio, St. Paul, MN
Laura Jensen, NPR, Washington, DC
Nick Kereakos, American Public Media | Minnesota Public Radio, Saint Paul, MN

*Paper not available at the time of publication

CONDITIONAL ACCESS: THE NEXT STAGE IN HD RADIO™ EVOLUTION

Thomas Rucktenwald
NDS
Costa Mesa, California

ABSTRACT

No longer a theory, HD Radio[1] Conditional Access is here. Developed specifically for radio broadcast managers and engineers, this paper examines the operational aspects of CA as well new and unique abilities enabled by Conditional Access - including advertising substitution.

In addition, this work includes recent examples of real HD Radio deployments, highlighting the challenges and opportunities they have created.

NDS RADIOGUARD™

NDS RadioGuard™, the HD Radio Conditional Access solution, is tested and ready. From its NAB 2007 announcement and booth concept demonstrations, the product is currently available from the major broadcast equipment manufacturers and sellers. The technology, created by NDS, is a complete CA solution that can be installed in a single station or as part of a station group.

RadioGuard[2] technology is included in the next generation HD Radio decoder integrated circuits that begin production in May 2008. This means that Conditional Access can be a standard part of future HD Radios. The new chips are lower-cost, lower power, and have significant new features. The embedded CA technology requires no special deployment effort by the radio manufacturer.

The IC has the decryption methodology, holds entitlements so that the radio knows the programming it should receive, and is individually addressable so that it may receive and execute specific instruction from the broadcaster. Consumers can sign-up using phone or internet.

Broadcast implementation requires a small amount of additional equipment and a possible HD Radio equipment upgrade. The HD Radio Importer must be version 3.0X or higher. This Importer includes the scrambler that encrypts audio and data. The Exporter/Exciter may also require some software upgrade to match the Importer.

The Importer receives its Conditional Access information from the RadioGuard Protector. An Entitlement Control Message Generator provides codes that change the operation of the Scrambler. The Protector also includes a carousel, constantly telling the radio population about the radio entitlements.

The Protector mates to the RadioGuard Initiator. The Initiator accepts radio receiver information, generates entitlements, and provides the system setup and user interface.

The Initiator connects to the National Resource Manager, a device that contains the database of all radios possible within the system, located in the NDS California facility. The NRM also coordinates services among all broadcasters. It provides global IDs for programs that are common across the nation and it provides individual local IDs so that local station programming is correctly identified and received by qualified radios.

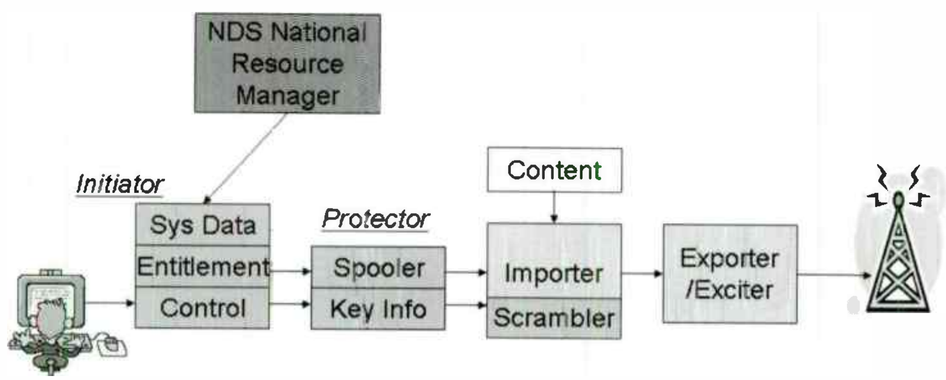


Figure 1 - Basic Station Installation

Connection to the National Resource Manager requires a secure Virtual Private Network connection. This can be done on a LAN-to-LAN or Client-to-LAN basis. It would be wise to install a firewall for the Initiator or put the Initiator behind the existing station/group firewall. Creating the VPN connection to the NDS NRM requires a couple hours of setup and verification between IT departments. Once this is set, it should last for the lifetime of the installation.

Security is built into the architecture. The radio security is internal to the decoder IC and is very difficult to circumvent. The encryption is a moving target, constantly changing with each crypto period. Entitlements and targeting include encryption technology. Equipment connectivity requires secure communications.

Reliability is built into the architecture. If the Importer fails, all the additional channels are down. If the Importer is operating, the programming exists and will be encrypted as directed. If the Protector fails, the programming still transmits and the encryption still exists, although the encryption will not be changing and no radios will receive entitlements. If the Importer and Protector are valid but the Initiator fails, the program will include agile encryption and radios will receive entitlements, but no new radios will be entitled until the Initiator is operational.

STATION INSTALLATION

The Protector should be on the same Ethernet network as the Importer. It may be wise to co-locate the two pieces and use the same Ethernet switch or router. Co-location should eliminate communication problems, an important factor in a successful installation.

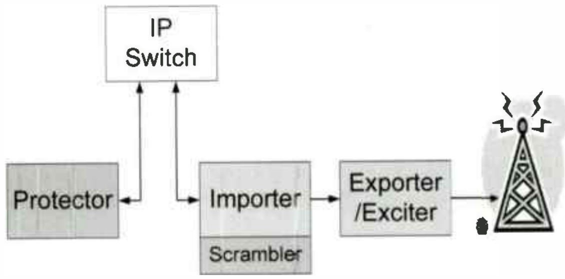


Figure 2 - Protector/Importer Co-location

The Initiator can be anywhere in the facility, as long as it can talk on Ethernet to the Protector. Whereas the Protector and Importer should be co-located, the Initiator could be anywhere convenient for an operator or administrator. In addition, internal station/group communication with the Initiator is via a web browser from any machine capable of a secure connectivity.

The Initiator could also be at a Network Operations Center, not even in the same city as the station or the Protector. The Initiator can control 1000 Protectors; the Initiator can control an entire station group. The Initiator will route radio entitlements to the correct radio station Protectors based upon the setup.

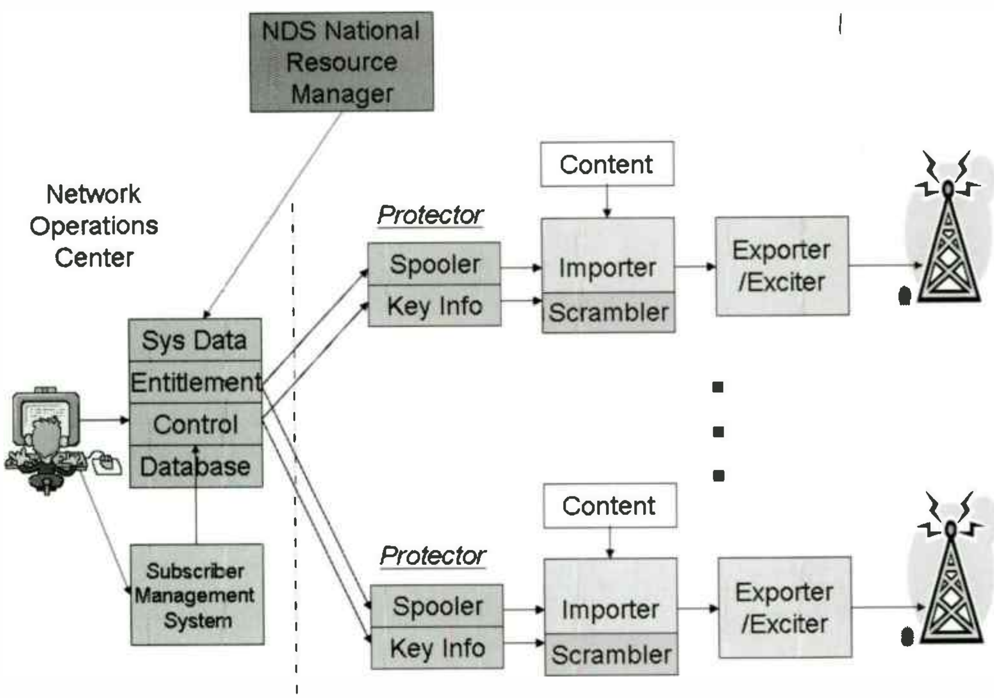


Figure 3 - Initiator Controls 1000 Station Protectors

Most of this setup involves setting the Ethernet address for the Initiator and the Protector. Both devices run on Linux; both have a desktop environment. After setting the Protector, a mouse and keyboard are no longer required and all access can originate from the Initiator.

REAL WORLD CIRCUMSTANCES

As in any development, first tests and pilots uncovered stability issues, bugs, and simple memory leaks for all of the participating equipment manufacturers. These were subsequently resolved.

We have also learned some important installation considerations.

We recommend that the Protector and Importer be co-located and run off the same IP switch. Co-location was one of the original architectural requirements. However, we find that running the Protector and Importer on the same network may be sufficient if network traffic does not have packet collisions. Communication between these two devices is critical for perfect performance.

Analog station over-modulation can cause interference with digital radio reception, particularly in extended hybrid mode. This really has nothing to do with conditional access, but is something that stations will discover as they turn on their HD3. Please follow best HD Radio practices and FCC regulations.
Integration with the National Resource Manager (NRM) is required and this integration requires a virtual private network (VPN) connection. Immediately, this involves IT personnel and with this sometimes comes Information Technology politics. Politics are more difficult to cure than technology issues. LAN-to-LAN VPNs always seem to work. Client-to-LAN VPNs have occasional discrete problems, particularly if the station system closes their client when there is little activity. The client must maintain or re-establish its VPN to accommodate NRM connectivity for this system.

Despite our best efforts, Conditional Access will not overcome "Acts of God." If the HD transmission goes down because the station transmitter site suffers from snow, ice, earthquakes, high wind damage, or flood, CA cannot cure it.

PROVISIONING SERVICES

Successful system use involves operations. The Initiator is easy to control and easy to understand. An operator can control this system using a web browser. Operation is via click boxes, radio buttons, and text boxes.

When setting the broadcast, the operator activates the conditional access offering to the consumer with something euphemistically called a "Sellable Unit."

The timeframe for an ongoing channel or a special event is accommodated by setting an expiration.

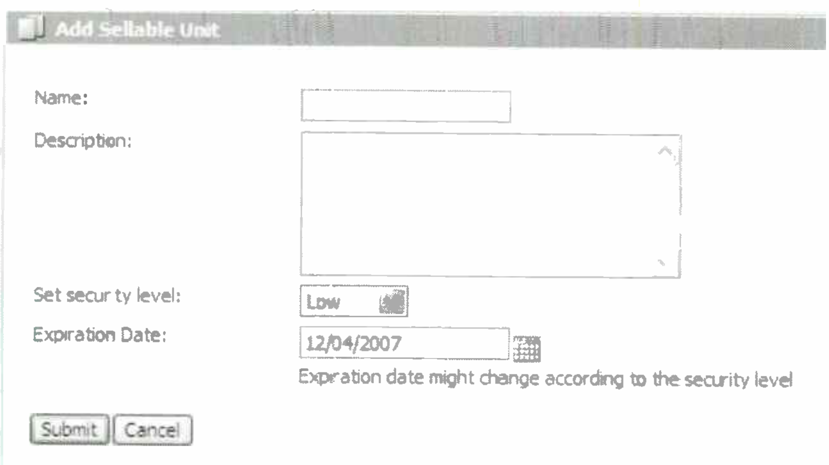


Figure 4 - Add Sellable Unit Screen

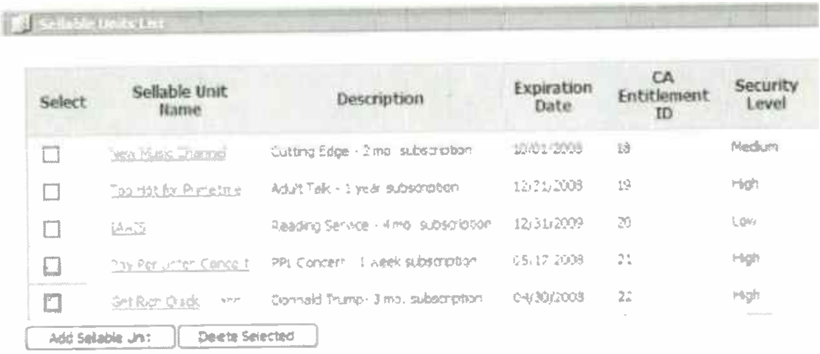


Figure 5 - New Sellable Unit Added to the List

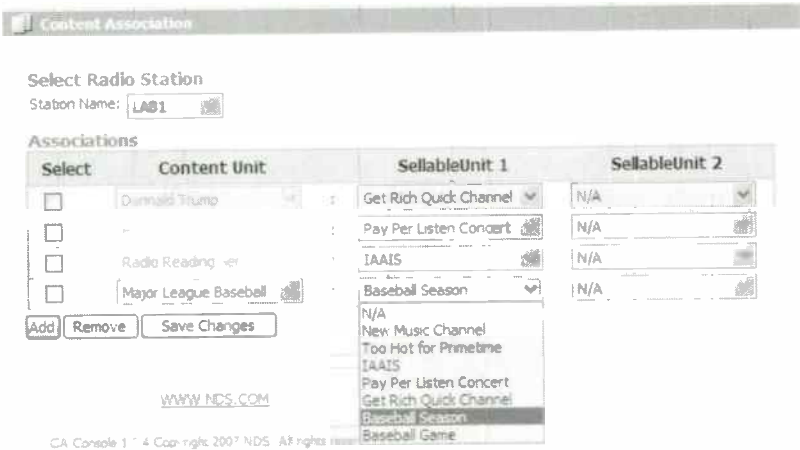


Figure 6 - Add Content Association, Sellable Unit List

Content that will play through the Importer is associated with the "Sellable Unit," logically tying the actual content to the consumer offering.

Entitling a radio using the NDS RadioGuard interface requires the radio ID number. This is an internal electronic serialization of the unit within the HD Radio decoder IC. A CA-equipped radio should automatically display this number when it finds a CA-protected channel where the radio does not have the entitlement.

Alternately, the Radio ID can be displayed by a button-push in the receiver menu.

The consumer provides the radio ID and the operator enters it. The operator chooses the programming the consumer requests and sets the expiration period, which may be unlimited. When the "Submit" button is pushed, the system will activate the radio for the requested service.

Figure 7 - CA-Equipped Radio With Serial Number Display[3]

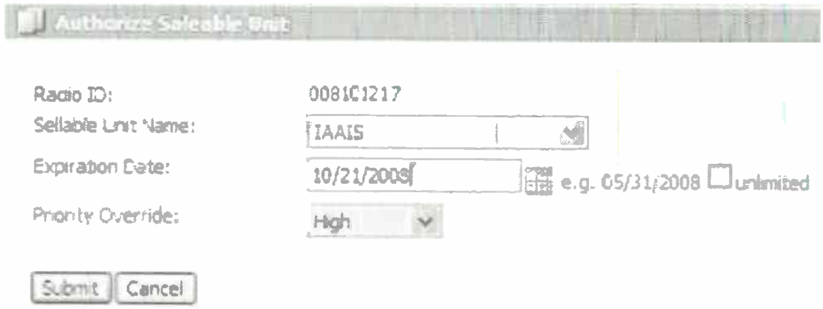


Figure 8 - Individual Radio, Adding Sellable Unit

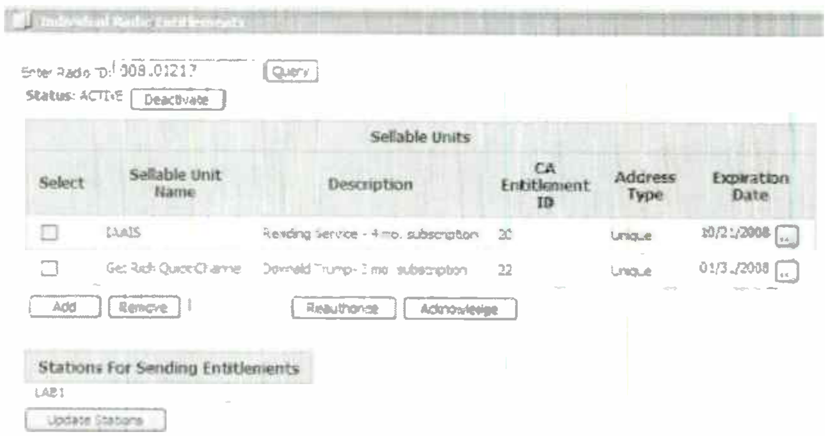


Figure 9 - Authorized Sellable Unit for the Radio

This NDS interface is the simplest activation. Numerous other methodologies exist. Some will be operator controlled and others will be consumer activated via a website or portal.

OTHER ENTITLEMENT METHODOLOGY

NDS has a known API for connectivity. RadioGuard provides a published SOAP interface for this known API. Systems capable of connecting to and controlling RadioGuard entitlements include fundraising software, existing back office installations for the web, professional subscriber management systems, and the Consumer Portal.

Fundraising software, often used by public radio and public good broadcasters, can be easily modified to accommodate RadioGuard entitlement. The simple addition of the SOAP API and a couple of database columns creates the upgraded product.

Some groups already have a back office installation for internet offerings. Extending this to control RadioGuard is a simple process. Again, it is implementing the SOAP API and a couple of database columns.

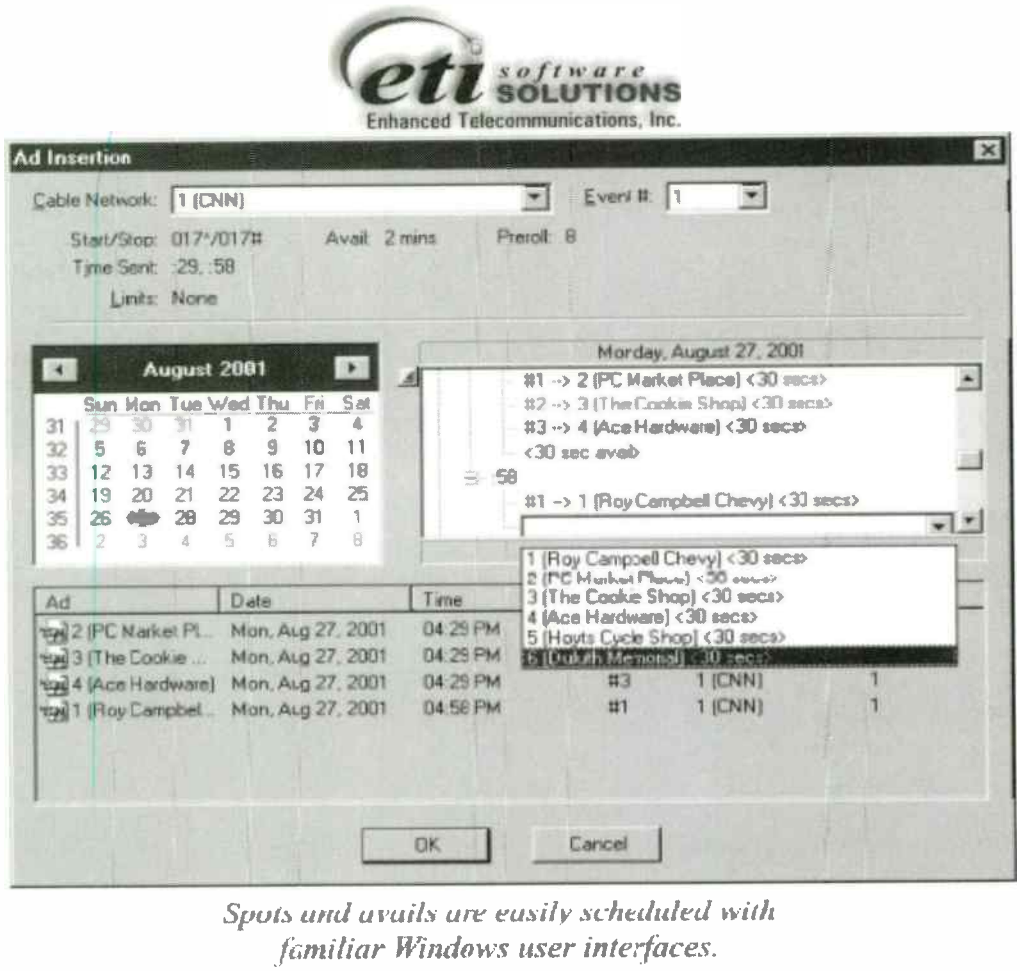


Figure 10 - Subscriber Management System[4]

For those that wish a professional interface, a subscriber management system will work directly with NDS RadioGuard. The figure above shows ETI Software. A purchased SMS can run the system or alternatively, the broadcaster can "rent" a portion of the system and have access to only their subscribers and opt-ins. This latter alternative architecture is deployed for the SES Americom/NRTC IPTV system, allowing second and third tier Telcos to manage their own subscribers. This subscriber management system can be operator controlled or have a web interface that allows consumer self-registration.

Another exciting implementation is the Consumer Portal. Created by Acxiom, this web portal allows the consumer to self-register via the internet.

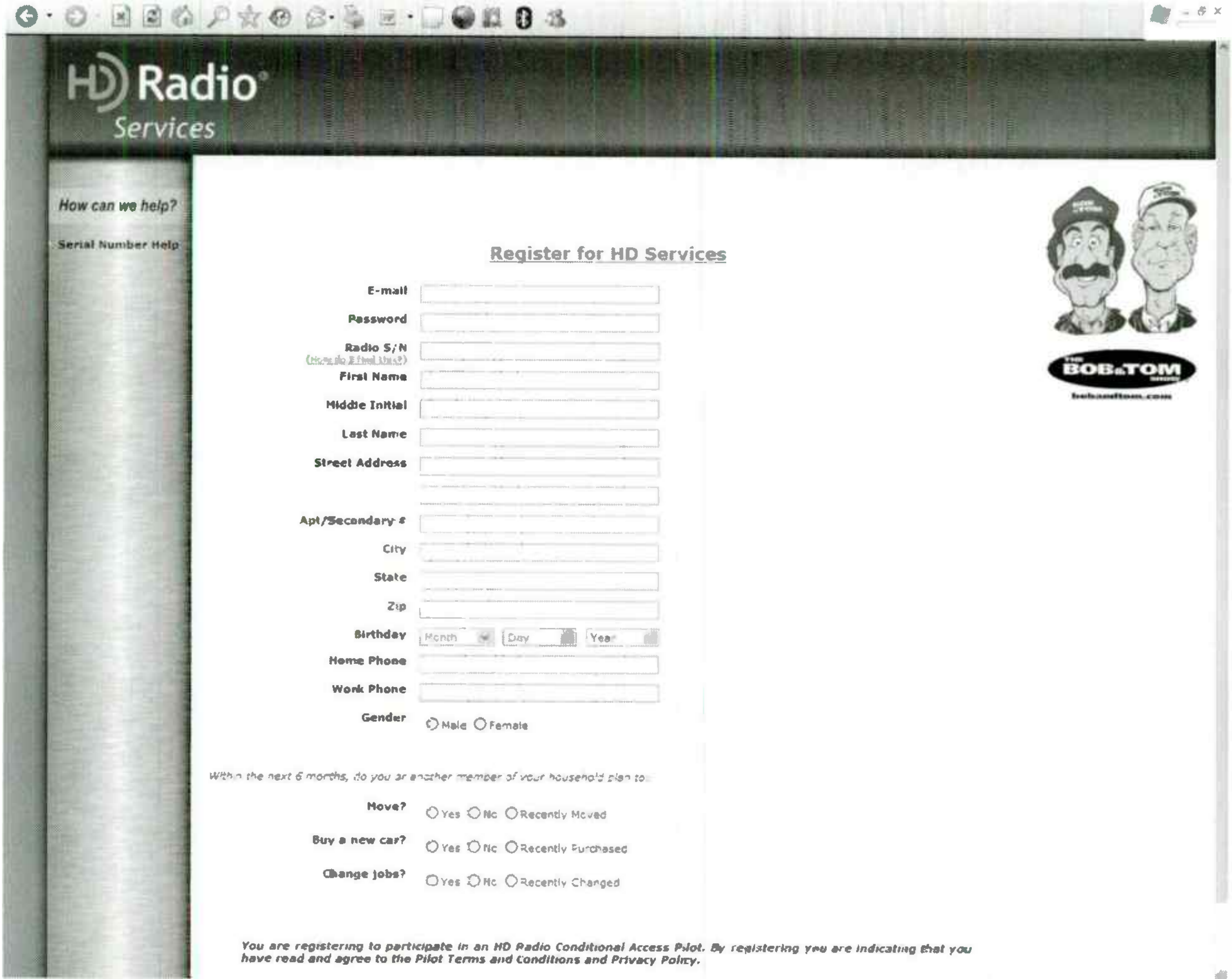


Figure 11 - The Consumer Portal[5]

When the consumer signs up using the internet, the information provided may appear to be something like this:
John Doe, 1234 Main Street, Anywhere, CA 90210, The Tom Show, 04/15/08, $0.00 per month, 12 months

However, this is much more than the radio activation. With a valid name and address, Acxiom is capable of providing extensive information and statistics about the consumer and consumer groups that sign up for the CA-protected services.

The real information looks like:
John Doe, 1234 Main Street, Anywhere, CA 90210, The Tom Show, 04/15/08, $0.00 per month, 12 months, income $50,000-$75,000, less than 1 year at address, $5,000-$9,000 car value, Age of 1st individual at address 26-35. Single, Home Owner, Truck Owner, Has Bank Credit Card, Sweepstakes/Contests, Power Boating, Sports on TV, own a dog, Lifestage Cluster # 58 “Young Workboots”[5]....

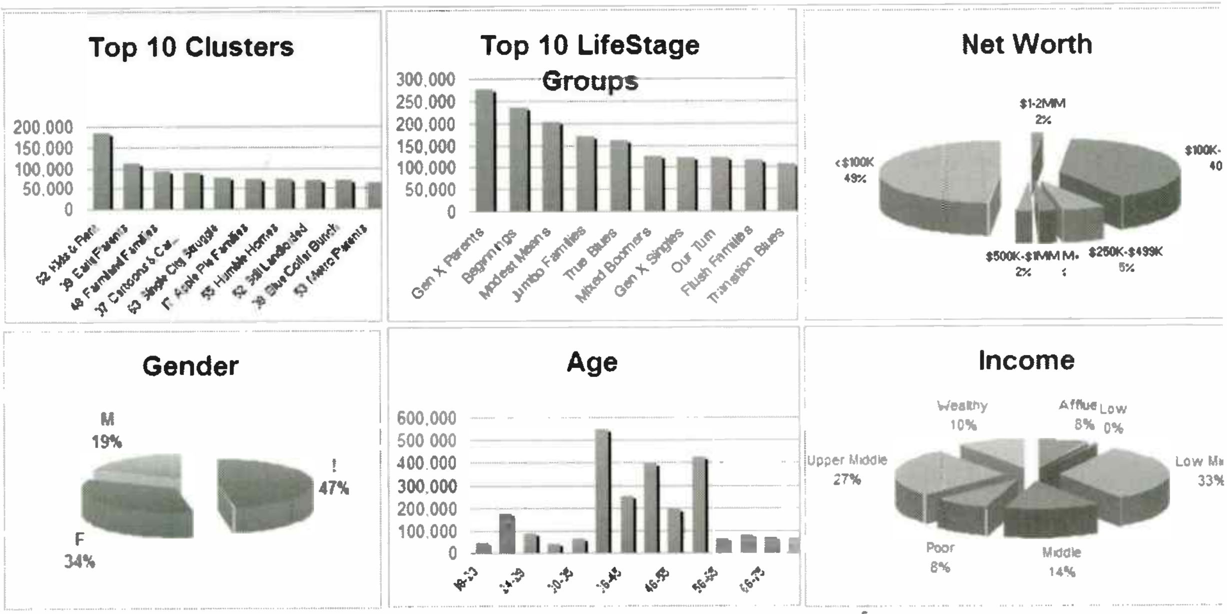


Figure 12 - Consumer Portal Statistical Information[5]

TARGETED ADVERTISING

Information, like that provided by the Consumer Portal, combined with Conditional Access entitlement, fuels new and unique abilities. This includes targeted advertising, sometimes called advertising substitution.

In targeted advertising, a participating consumer may hear a different ad, based upon their preferences and lifestyle, than the transmitted commercial heard by the public. That participating consumer is much more important to the advertiser because they are already a qualified potential customer. The ad substitution mechanism is important to the broadcaster because it means that they can sell the same commercial space multiple times. The consumer finds the substitute ad far more acceptable because it provides information about something that they want or need.

Since bandwidth limitations tend to bias a targeted advertising or ad substitution system away from a parallel stream approach, a recorded file delivery with a playback trigger might be the best methodology. Recording technology patented by iBiquity is part of the licensed patent portfolio. iBiquity's radio stream recording is already successful.

A targeted ad delivery system will send a recorded file. The file will be delivered to and stored on the radio as per RadioGuard targeting and individual radio addressability. As a commercial that allows substitution begins to play in the broadcast stream, the radio triggers a substitute stored file replacing the broadcast ad.

Targeted ads transmit via data carousel at less-than-real-time speed. The entire file must be received without error before it can be used. Partially received files will be stored for future completion, picking up or filling in data not previously received.

Consumers are matched with their preferences and demographics, and their radios receive entitlements for different advertisements based upon these. When an ad plays in the carousel that matches their entitlements, the radio receives and records the file, storing it for future playback.

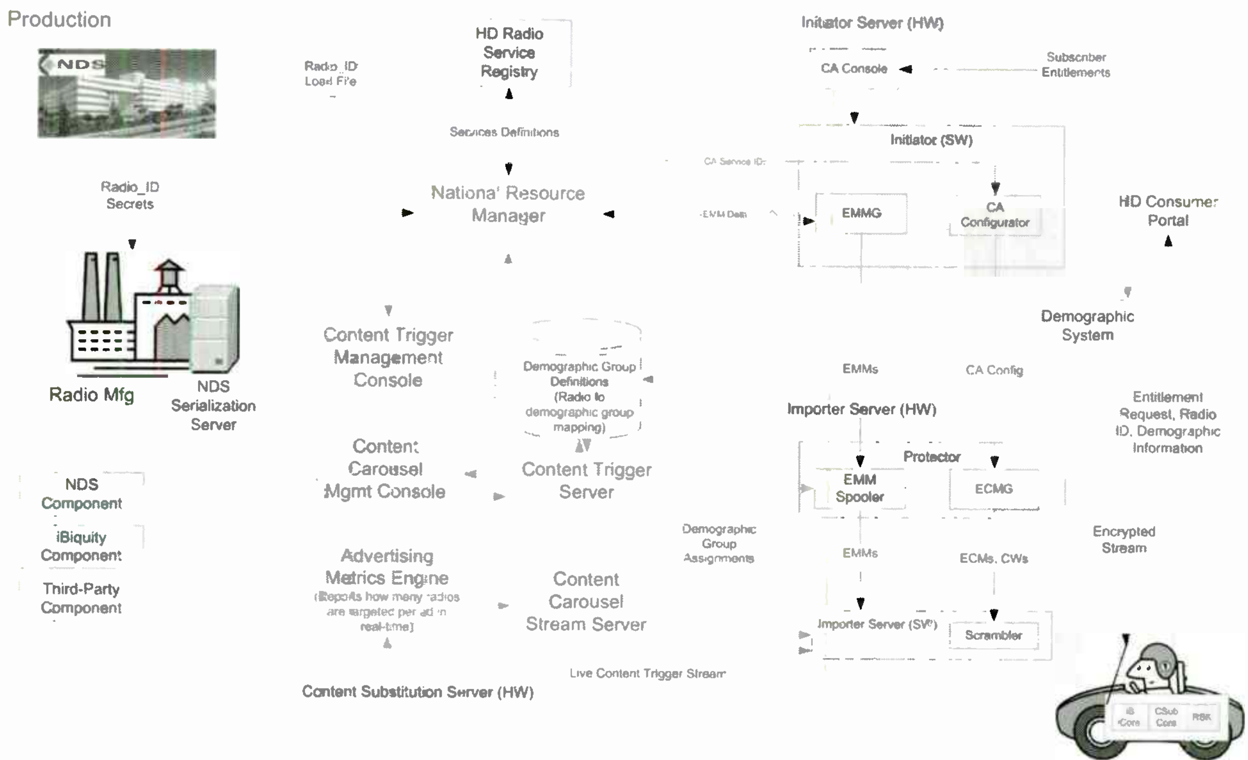


Figure 13 - Targeted Ad Substitution Ecosystem

The components required for this success include:

1. HD Radio transmission system
2. Targeting methodology
3. Broadcast content carousel for sending advertisements
4. Organizing application matching consumer preference and radio IDs with ads
5. Information used for matching and mapping consumer preferences
6. Advertisements
7. Metrics, that is audience/results measurement

All of the above elements exist. The HD Radio transmission system is now well known. Targeting is an integral part of NDS RadioGuard Conditional Access. NDS already deployed targeted advertising substitution for television and will have a concept demonstration for radio at their NAB booth. Consumer information is an integral component of the Consumer Portal. The radio industry has well known advertising brokers. Audience measurement needs only refinement so that it can record and report the granularity of a substitute ad.

TOMORROW'S RADIO PROGRAMMING

Today's technology already provides for tomorrow's programming. If you can dream it, you can make it happen.

Conditional access provides choices. Yes, it does provide possibilities for pay radio, which could be a subscription or it could be a one-time high-value event. A trial for pay subscription services is already under test as a possible application scenario. Who could not imagine a successful pay-per-listen concert for a very popular artist?

However, these are options. Pay radio is just one of the possibilities. Most of the successful programming with this new technology will be free to the public.

Radio Reading Services, for the blind and sight-impaired, must maintain copyright integrity for the publications that they read. IAAIS will be deploying HD radio channels, replacing the SCA service they presently maintain. HD Radio provides sight-impaired listeners with outstanding fidelity, something they cannot get with an SCA receiver, and allows the station, the service, and the consumer to use an off-the-shelf radio.

Public Radio can offer pledge-drive-free programming to their members, the people who have already contributed to the station. They can also offer member channels to those that align themselves with the station. Contribution level need not be the deciding factor for member programming; an opt-in request for this programming may provide the entitlement, making conditionally accessed programming available to all no matter their financial circumstances.

Data applications, like the already successfully tested Navteq navigation pilot, use conditional access for data integrity and targeting. Since RadioGuard is part of the standard HD Radio IC, these types of data services can be a part of any receiver. No additional specialty integrated circuits, which add expense and complexity, are required.

Any station can align themselves with local performance groups or locations. The local jazz club, symphony, theater group, bar, band, or Fair can provide content or be a source. Localism drives terrestrial broadcasting, particularly with new HD Radio channels.

What will drive the consumer to HD Radio?

Creativity is the answer. Look at YouTube. People flock to that internet site because of creativity. A creative person has opportunity. There is no boundary based upon race or gender or any other factor. Could there be a parallel concept for radio? The technology to drive this or any other application either exists or can be configured.

Do you need conditional access? The station should encourage listeners to identify themselves. This builds community, confirms listenership, and provides added revenue. Targeted advertising, where the station can sell the same ad space repeated times, demonstrates the advantage in knowing your listener.

REFERENCES AND ACKNOWLEDGEMENTS

1. HD Radio is a registered trademark of iBiquity Digital Corporation
2. NDS RadioGuard is a registered trademark of NDS Ltd.
3. Thanks to Dice Electronics, Lakewood, CA
4. Thanks to Enhanced Telecommunications, Inc., Norcross, GA
5. Thanks to Acxiom, Little Rock, AR

MANAGING RADIO METADATA FOR MULTIPLATFORM DIGITAL DISTRIBUTION

Daniel Mansergh
KQED Public Radio
San Francisco, California

ABSTRACT

Listeners have high expectations for rich descriptive metadata integrated with audio programming, cultivated by exposure to satellite and cable audio services, portable media players, computerized music libraries and podcasting. As broadcasters and audio program producers leverage the opportunities of multicasting and enhanced PAD support afforded by new digital radio technologies, developing an efficient, high-quality metadata management workflow for multiple distribution channels alongside traditional audio-only production and distribution workflows has become a significant challenge. This paper reviews the current status of metadata standards, tools, and best practices for developing efficient programmatic data management systems, with a focus on the practical implications of deploying a consolidated data/audio workflow in a large radio production and broadcasting organization. An integrative approach to data systems architecture is proposed, drawing from the fields of library and information sciences, database management, internet application development and broadcast engineering.

BACKGROUND

Led by the rollout of digital radio systems with integrated PAD capability and a resurgent interest in RDS as a way to bring comparable features to analog FM broadcasts, broadcasters are beginning to see metadata as an essential component of their over-the-air services. When complementary distribution technologies such as live streams, on-demand streaming archives, and podcasts are factored into the mix, grappling with the disparate metadata requirements for each of these channels can present quite a challenge for organizations without established data workflows.

Current approaches to managing metadata involve regular manual data entry late in the distribution chain, with producers often entering the same or similar information about identical programs into several different tools: an automation system, PAD management software, a website database, an ID3 tag editor, and an online distribution portal, for example.

This is clearly inefficient, but more significantly, the manual approach doesn't scale well. Demands on program producers to supply distinct multiple sets of programmatic metadata continue to increase as new digital services emerge, each with their own data requirements. Content owners wishing to extract value from the "long tail" of library materials in new on-demand services often face a daunting challenge in manually preparing metadata for each distributor, knowing full well that they will have to go through the same process from scratch when a new service arrives.

Ideally, broadcasters should be able to manage their audio assets and programmatic metadata in a consolidated system that minimizes repetitive manual entry by automatically taking in data from many sources and redistributing appropriately formatted metadata sets for each distribution platform. Since no such system exists, KQED Radio decided to build one.

A Unique Case Study

KQED Public Radio is one of the premier public broadcasters in the US, regularly ranking as the most-listened-to public radio station in the country. It produces approximately 20 hours of news and information programming weekly, including a daily news service for other California public radio stations, with 60 staff members at five locations around the state. It also operates a successful website, with live streaming, online archives, and podcasts offered for all locally-produced programs.

Although the extent of the metadata requirements at KQED may be somewhat unique for a single station, it is useful as a case study to understand the complexities of the issues involved in crafting a comprehensive and flexible data management system, no matter what the metadata needs of a particular station, group, or network may be.

AN INTEGRATIVE APPROACH

Since the technologies and expertise required to implement a consolidated metadata system span multiple disciplines not usually associated with broadcast engineering, an integrative methodology drawing from the best practices of these various fields is warranted. This approach serves to leverage the strengths of each discipline while checking their inherent problem-solving biases, opening up new avenues of research and implementation that otherwise would not have been considered.

Discipline 1: Information Science

In any data management project, the fundamental questions of information architecture, including taxonomy, data relationships, and organizational structure must be addressed. Fortunately, experts within the fields of library and information sciences have been actively grappling with these issues as the focus of media archivists has begun to shift away from straight cataloging of physical media to conservation of the contents of those media in permanent digital storage systems. While motivated by the urgent need to preserve large collections of recordings stored on magnetic tape before continuously degrading media makes recovery of the recordings impossible, the related effort to define consistent metadata standards for the cataloging of these digital collections is extremely helpful in informing the development of data management systems for "born digital" assets.

The product of this work is an alphabet soup of metadata standards that all attempt to define an optimal set of data to catalog media assets, but with very different assumptions about the comprehensiveness, searchability, and specificity required of the data and expectations for the skills and objectives of data users.

Thankfully, the Corporation for Public Broadcasting in 2002 began a project to sift through all of these resources and develop the Public Broadcasting Metadata Dictionary (PBCore), a common language for defining and describing metadata elements for media archives and exchanges. With the release of PBCore version 1.1 in 2007, the dictionary gained a clearly defined relational structure and a complete XML Schema Definition (XSD), making it an extremely useful starting point for a metadata management system. Alhough it was developed for public broadcasters, PBCore is free for anyone to use under a Creative Commons license.

For KQED, PBCore was an obvious choice to serve as the underlying data model for a consolidated metadata system. First, the intended use of PBCore and the scope of the KQED project are closely aligned. Second, since PBCore contains all the elements required to describe video as well as audio assets, expanding the metadata management system to serve the requirements of co-licensee KQED Television can easily be accommodated without structural changes. Finally, it provides an open standard platform to facilitate data interchange with other broadcasters and program producers.

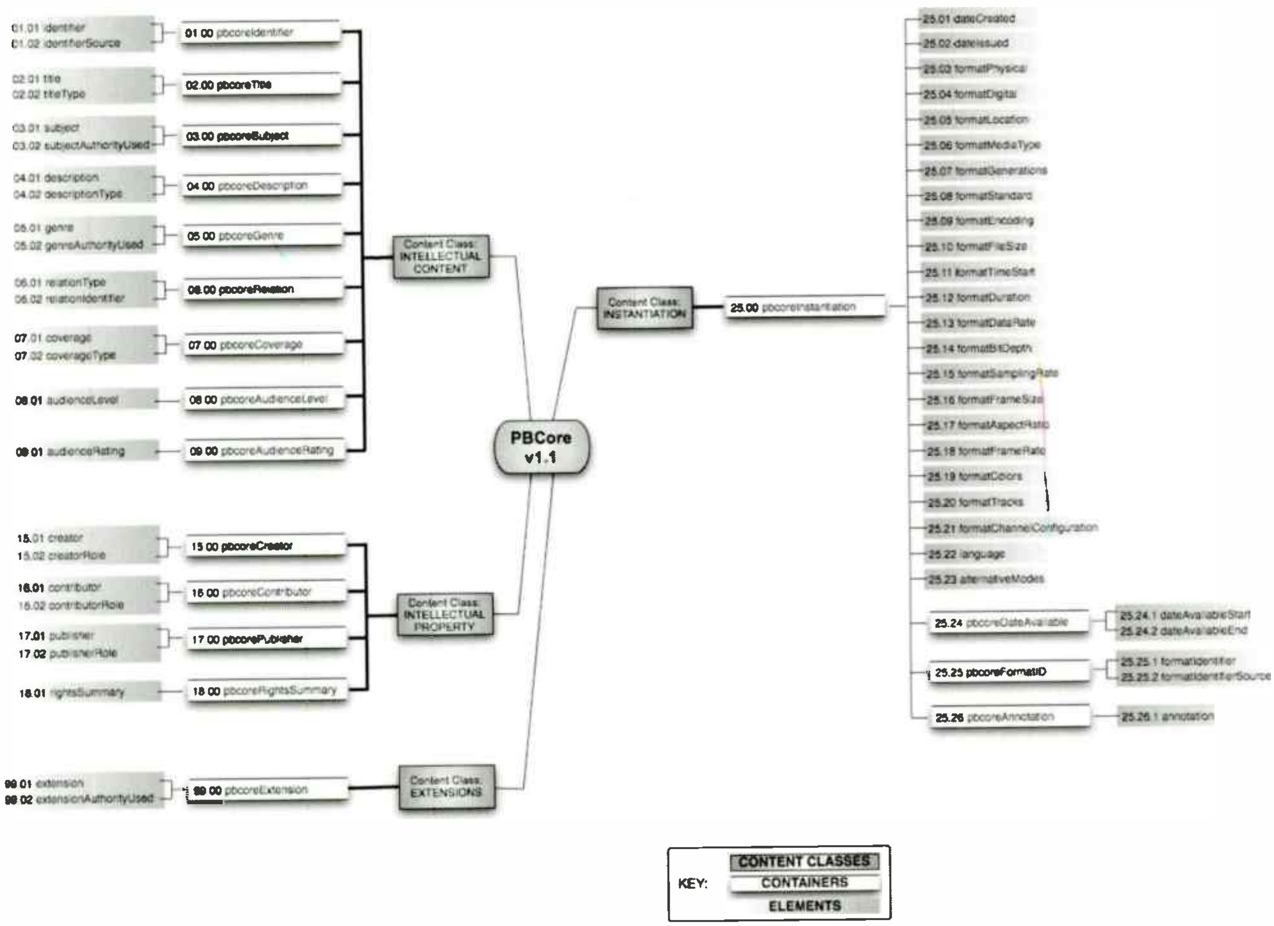


Figure 1. The PBCore 1.1 metadata model. 53 descriptive metadata elements are arranged in 15 containers and 3 sub-containers, all organized under 4 content classes.

Just as the value of a metadata system is judged by the quality and consistency of information it contains, the strength of a data model is only as good as a specific system built upon it. For this reason, a thorough understanding of the principles and assumptions that led to the development of a particular model are essential when determining how it will be implemented. In the case of PBCore, the development history of the dictionary is well-documented on the pbcore.org website, and numerous training and support resources for new development projects are available from the community of existing PBCore users.

Discipline 2: Database Management

With the PBCore XSD as a starting point, defining the structural framework of a SQL-based metadata database would appear to be a fairly straightforward exercise. However, other considerations must be addressed before a functional system can be deployed, beginning with a thorough understanding of where and how data management tasks will be performed within existing production and operational workflows. This requires a primary focus on the end-user experience throughout the database development process, since the users will ultimately determine how successfully the system functions. Ideally, data collection features should be embedded directly into newsroom, production, and automation systems to allow users to easily funnel information into the metadata system at the point of creation.

Once the operational and user requirements are well understood and documented, these should be translated into a set of design specifications for the database system as a whole. Incorporating the expertise of a well-qualified database developer or administrator to provide guidance can be invaluable at this stage, since structural problems in the database can affect performance and scalability of the system as a whole and are virtually impossible to fix later.

At KQED, the developers and information architects on the in-house Interactive Department staff proved to be a key resource as we conducted a user needs assessment and evaluated possible approaches to database structure. Beyond their professional expertise in the mechanics of data systems development and management, their years of experience developing the kqed.org website provided a practical context to the process and contributed insights about human-system interactions that otherwise would have been missed.

One finding that proved to be dramatic in its implications for the course of the metadata system development project was the recognition of the kqed.org Content Management System as the largest customer and de facto archive of schedule information, produced content and associated metadata from KQED Radio. As such, it already contained the majority of the information needed to supply the requirements of PAD displays, podcast RSS feeds, online streaming schedules, and on-demand program archives. However, since it was developed solely as a website solution, the existing data contribution and management tools were inadequate for the expanded requirements of a comprehensive system and the underlying data model wasnÕt structured to allow such additional functions to be added easily.

A full redesign of the kqed.org CMS to accommodate the new requirements was clearly beyond the project scope and would have needlessly complicated the operation of the website. On the other hand, the cost and effort required to develop a monolithic new asset management system that would duplicate much of the existing CMS functionality would be a waste of resources. In the end, we opted for a modular design that distributes the functions of the data system into a federated network of task-optimized applications, with XML and the PBCore XSD enabling data interchange between them.

This approach to data system development yields distinct advantages over traditional all-in-one database applications when scalability and flexibility are more important than uniformity and centralized control. Individual components are optimized for specific functions, and are typically small, light, and easily deployed over time. Existing applications can be integrated into the system, as long as they provide support for data interchange using standard protocols, and new applications can be added to accommodate new requirements without replacing the existing system. Finally, since this methodology is ubiquitous in modern Internet application development, freely available open-source tools and access to a large pool of skilled programmers with experience in this field combine to keep the development costs of a custom system relatively low.

Discipline 3: Application Development

Once the fundamental decision about database architecture has been made, the desired features of the overall system must be translated into detailed design specifications for individual system components. With typical standalone database systems this is a fairly easy task, since the capabilities of user-facing client applications and backend data services are well differentiated. In a distributed system the process is somewhat more complicated, since individual applications may be designed to accomplish a combination of tasks traditionally associated with both clients and servers.

The most practical way to address this complexity is to create a functional map of the proposed system, with each operational step in the proposed data workflow broken down into its constituent parts. By including all data collection, display, transfer, validation, storage, and distribution tasks required of the system in one design document, the specifications of individual applications can be much more easily determined.

The first few applications of a distributed metadata system are the easiest to develop, since they're probably already in place. Many production, automation, and newsroom systems have basic built-in metadata support and some even allow additional user-specified data fields to be added, making them useful for data capture tasks. Website contribution applications, production databases, and PAD management tools are other likely candidates for integration. If required, custom development of data interchange functionality may be justified since customization or enhancements to an existing product can be more cost-effective and fit a particular workflow better than developing a separate data application from scratch.

Once the role of existing applications within the metadata system has been clarified, the design specifications for new applications can be determined. Since much of the custom development required to implement a metadata system is typically related to data interchange and transformation, it is useful to develop a complete listing of the specific fields defined in metadata sets used by any external data sources or destinations to guide the development process.

Real-time data feeds, including RDS, HD Radio MPSD and SPSD, and online streaming playlists are a good place to start, since the data requirements for current services are relatively modest and well understood. Online archives of program material will have similar basic requirements, although they will generally require additional schedule information and long-form alternative versions of program descriptive information optimized to the website's style requirements. Podcast metadata is divided among ID3 tags, basic RSS XML, and any additional schemas required to support specific services, such as iTunes and Media RSS.

Schedule information from traffic and automation systems, as well as RSS feeds from network program providers and online content databases such as Gracenote and freedb are likely to be the easiest data sources to integrate since this metadata is available in easily machine-readable (tagged) formats. Other data sources, such as automated email and program websites, may be parsed by appropriately intelligent data ingest applications if development is justified by the projected operational efficiencies.

In addition to existing requirements, some thought should be given to probable new metadata needs, including images, geotags, granular (segment-level) topic information, and related content locators. Although it is impossible to predict all future metadata requirements, planning to capture information that is likely to be of interest based on current trends will make the contents of a media archive that much more valuable when new media services arrive.

Once the master list of required fields has been compiled, it should be reviewed to identify common information shared by multiple data services. This is actually much more complicated than it would at first appear, since variations in usage and formatting constraints among current data services (such as different character limits for specific fields on display devices from different manufacturers or varied interpretations of certain field names) must be weighed against the desire to minimize duplicative data entry. In some cases, predefined concatenation, formatting, and other text manipulation procedures built into data translation applications or the display layer for a particular data channel may allow common data fields to be reused. In the end, some information may need to be stored in multiple formats to optimize the end-user data experience on different services.

For the KQED project, this process took the better part of six months, aided by the Interactive team and an outside consultant with experience building media archive systems. The end product was a framework of mapping relationships (or "crosswalks," as they are commonly known) for all of the fields that the proposed metadata system would need to handle through the transitions between external data sources, in-house contribution systems, and export destination formats.

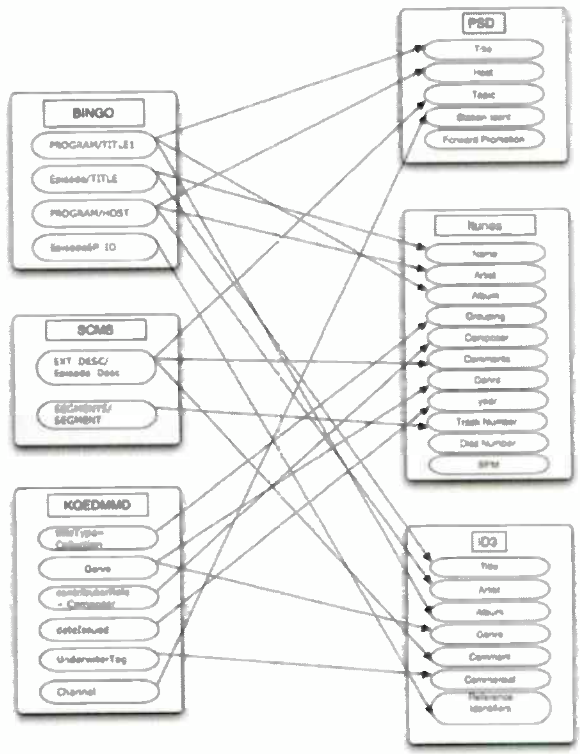


Figure 2. Metadata crosswalks for three common formats. Bingo, SCMS, and KQEDMMD represent contributing data systems; PSD, iTunes, and ID3 are distribution metadata sets.

A thorough exploration of the issues surrounding specific mapping and crosswalk relationships between metadata sets is beyond the scope of this paper; suffice it to say that a complete understanding of the requirements for each data translation is essential to ensure consistent data interchange between applications and external systems.

With the design specifications for the data applications fully documented, deployment can begin. Depending on the requirements and complexity of a specific system, this may take weeks, months, or even years. Existing data source or destination systems will need to be modified, configured, and tested to ensure that they can perform their required data interchange tasks. New applications will need to be acquired or developed in-house, customized, and introduced into the station's data workflow. Again, the modular approach to data application development is a huge advantage in this process, since the deployment can be broken into easily manageable small projects that provide system users with continual improvements in their workflow as new capabilities are introduced.

Discipline 4: Broadcast Engineering

When deploying the system, the traditional strengths in system integration and project management central to the practice of broadcast engineering complement the theoretical and organizational design insights gained from the three preceding data-centered disciplines. Obviously, successfully interfacing a centralized data management system with existing automation, traffic, production, and newsroom systems, station IT networks, servers and workstations, audio format conversion systems, media libraries, and network programming distribution systems, as well as PAD data transmission equipment such as transmitter data links, RDS encoders, and HD Radio Importers, Exporters, and Exciters depends upon the expertise of skilled engineers who know the capabilities of these systems well.

But broadcast engineers bring more than just hands-on technical skills to a metadata system development project. A systemic approach to project design, which is standard practice in broadcast facilities, is central to the development of a distributed metadata system. The essential process of capturing, documenting, and streamlining data workflows is aided by the application of traditional signal flow conventions to data management systems, even though the underlying mechanics of data systems do not always follow the same principles. Core system design considerations, such as usability, accessibility, flexibility, scalability and redundancy are routinely assessed and balanced in broadcast engineering projects.

Finally, broadcast engineers understand the overarching importance of the human element in the deployment of a workflow system. No other single factor will dictate the success of a metadata system more than this one, since the biggest variable in implementing large-scale change in any broadcast or media production organization is in the people involved. Many existing operational patterns and workflows have roots that stretch back decades, and old habits are impossible to change by fiat. By adopting an engineer's perspective and focusing on the weak links in existing processes and systems that cause frustration to users, individual incentives to adopt the new system can be identified and incorporated alongside new components of the metadata system. These incentives will work best if there is a clear connection between new data management procedures and direct, obvious improvements in an individual's daily work.

In the KQED project, this approach prompted the development of additional features to foster user acceptance in several key areas. For people preparing materials for digital distribution, the demands of expanded metadata entry requirements are tempered by consolidating data collection and approval into one system and automating audio conversion processes on a background server. In the newsroom, data gathered by reporters, producers, and editors is available for real-time status updates on production progress in a scheduling and assignment system, so the management of ongoing coverage becomes much more transparent.

In addition to these design insights, having broadcast engineers involved in the development of data workflows and the human-system interface of a metadata system will help as the system is deployed and engineers are called upon for user support, training, ongoing system maintenance, and planning for enhancements to the system as requirements change.

PUTTING IT ALL TOGETHER

The data system development approach proposed here is by no means a linear process. Most of the development will be conducted in parallel with all four disciplines informing each other, suggesting novel approaches to problems that would not otherwise be considered. Even the decisions that must be made before full-scale development work begins, such as which underlying data model to choose and what the database architecture should be, benefit from a thorough exploration from each quarter.

As we experienced at KQED, this may lead the project in surprising directions which ultimately result in a stronger, better-structured and more flexible system. However, an integrative design process does have its disadvantages. It can be time-consuming and expensive if outside consultants need to be retained to participate in the process. Balancing the different objectives of several disciplines can be frustrating, so clear consensus principles need to be agreed upon early in the process. And the depth of expertise required to sort through some of the more difficult issues may not be available locally.

Ultimately, though, the benefits to an organization of a comprehensive study of their metadata needs and development of systems to better manage those requirements will pay dividends as the use of metadata-dependent digital distribution platforms continues to expand. Done well, this process will lead to a tight integration of metadata and media production within an organization that will be essential to its long-term success in the competitive digital media environment.

RESOURCES

AES Technical Committee on Archiving, Restoration and Digital Libraries: www.aes.org/technical/ardl

Dublin Core (ISO 15836): dublincore.org

KQED Interactive: Tim Olson, Brian Underwood, Ken Murphy, Eric Westby, Tobin Mori

NPR Labs: www.nprlabs.org/research/pad.php

PBCore: www.pbcore.org

PRI PSD: psd.publicbroadcasting.net/standards.html

The Future of Radio in a Changing World

DAVE WILSON
Consumer Electronics Association
Arlington, Virginia, USA

ABSTRACT

This paper will discuss modern consumer technology, changing consumer expectations for electronic media, and how they are affecting free local radio. It will consider possible changes to free local radio's technical infrastructure and to its frequency allocation rules. Changes like these could make free local radio more competitive in a changing world.

WHAT CONSUMERS WANT

Back in free local radio's "good old days" when there was basically radio, TV and print advertising, the total "pie" of listeners and advertising dollars was smaller than it is today, but radio's share was considerably larger. Over time as cable TV, satellite TV, satellite radio and the internet have blossomed the total pie of audience members and the total pie of advertising dollars have increased, but free local radio's relative share of the total pie has decreased. Free local radio needs to work on growing its piece of the pie, not by clinging to its old ways while fighting back advances in technology, but rather by embracing technological advances and using them to expand the reach and relevance of free local radio.

The fact is, technological advances have made free local radio less important overall, not just to its listeners, but to its advertisers and consumer electronics manufacturers, as well. Today it is obvious to even casual observers that a lot of time consumers formerly spent listening to the radio is now spent talking on cell phones, listening to portable media players, or listening to satellite radio. Furthermore, some of the things for which consumers used to rely on radio are now being provided by competitive services in a non-audio manner, such as traffic and weather information provided via the internet, or via GPS devices. With consumer demand for these competing services continuing to rise the consumer electronics industry is working to satisfy this demand by providing a steady stream of new and innovative products. There was a time in the last century when meeting consumer demand for audio entertainment systems meant offering radio receivers and record, cassette and/or CD players. Those days are long gone, however, and the percentage of shelf space dedicated to radio receivers in a typical electronics store has diminished over the years.

The consumer electronics industry is like any industry, it seeks to identify consumers' needs and serve them well. To this end the Consumer Electronics Association (CEA) is constantly conducting consumer research to help its members identify consumers' needs. In one such study, from late 2007, CEA attempted to predict what products would be most in demand during the important holiday shopping season. People were asked what consumer electronics product they would most like to receive as a gift.[1] They were not prompted with a list of pre-selected choices. The number one response? Portable MP3/digital audio player. Among the other responses were in-dash satellite radio and satellite radio generally in no specific form. Smart phone, cell phone and iPhone® were also among the responses.[2] What percentage of people said they wanted an AM/FM tuner? Zero. An AM/FM receiver? Zero. An AM/FM radio for their car? Zero. A clock or table radio? Zero. A portable AM/FM radio? Zero.

Now one could rationalize this survey result in several ways. For example, perhaps one might argue that most people already have multiple free local radio receivers. However, it is probably smarter to pay attention to the fact that the consumer electronics product most frequently at the top of peoples' wish lists last year was a portable MP3 player, and that the MP3 player that has generated the most revenue in the CE industry is the radio tuner-less Apple iPod®.[3] People not only want portable media players, they appear to specifically <u>not</u> want AM/FM radios, for when given the choice between portable media players with radios and portable media players without radios they are consistently and overwhelmingly selecting portable media players without radios.

Oh, and another thing that free local radio broadcasters should know. Despite the aggressive efforts of the HD Radio Alliance to market HD Radio in 2007, and the frequency with which HD Radio was promoted on free local radio stations during this period, the term "HD Radio" was not mentioned by a single one of the 1003 people who were asked what consumer electronics product they would most like to receive as a gift.[4] Not once.

At the International Consumer Electronics Show in January, 2008 TWICE magazine published NPD Group market share data for various consumer electronics product categories.[5] These rankings were based on dollar sales at retail. Overall, taking all CE product categories into account, the top five companies in terms of dollar sales at retail were Hewlett Packard, Apple, Sony, Samsung and Canon. Only one of these companies (Sony) is selling HD Radio™ receivers.

Table 1: Jan-Oct 2007 retail dollar sales ranking – all categories

1. Hewlett Packard
2. Apple
3. Sony
4. Samsung
5. Canon

According to NPD data, sales of mobile electronics products made up just three percent of all CE product sales, and the top five companies selling mobile electronics were Garmin, Pioneer, TomTom, Alpine and Kenwood. While Alpine and Kenwood are selling HD Radio receivers, it is worth noting that the number one and number three companies in the mobile electronics category primarily make GPS equipment, and this equipment gives consumers a means of getting traffic information from a source other than free local radio.

Table 2: Jan-Oct 2007 retail dollar sales ranking – mobile electronics

1. Garmin
2. Pioneer
3. TomTom
4. Alpine
5. Kenwood

Mobile electronics = 3% of all CE sales

Sales of home audio products made up just four percent of all CE product sales. The top five companies selling home audio products were Sony, Monster Cable, Bose, Yamaha and Panasonic. Sony and Yamaha are making HD Radio products. However, the fact that Monster Cable ranked number two in revenue generated from home audio equipment is a good indication of how important home theater systems are to this category. In fact, home theater audio systems, audio amplifiers, and speakers account for nearly two thirds of the revenue generated in the home audio category. While the importance of free local radio to the bottom line of a manufacturer varies widely from company to company, in general the consumer electronics industry is generating a very small portion of its revenue from free local radio.

Table 3: Jan-Oct 2007 retail dollar sales ranking – home audio

1. Sony
2. Monster Cable
3. Bose
4. Yamaha
5. Panasonic

Home audio = 4% of all CE sales

Sales in the portable audio category demonstrate what is now obvious to anyone who follows radio or consumer electronics. MP3 players and similar devices have revolutionized the way people buy and listen to music. The top five companies selling portable audio devices are Apple, Sony, SanDisk, Griffin Technology and Bose. Apple, of course, makes its very successful iPod. Griffin Technology makes accessories for the iPod. Sony and SanDisk make MP3 players and accessories. Many of the Sony and SanDisk players, as well as players from other companies, include FM tuners. Despite this, consumers have voted with their wallets and made Apple's FM tuner-less iPod the leader in portable audio devices.

Table 4: Jan-Oct 2007 retail dollar sales ranking – portable audio

1. Apple
2. Sony
3. SanDisk
4. Griffin Technology
5. Bose

Portable audio = 7% of all CE sales

Some in the radio industry have argued that it is important for radio broadcasters to campaign for inclusion of radio tuners in CE products like MP3 players and cell phones. Unfortunately, this thinking misses the point. The radio industry's goal should be to increase consumption of its product. Simply putting radio tuners that provide the same old radio service in other products does not increase consumption of radio. Sure, on the margins it may help get a few people to tune in who otherwise might not have for lack of a tuner, but the fact is people are buying MP3 players, cell phones and other devices to use their non-radio features, and radio would do well to study how and why consumers use the products they are purchasing, and modify the service it is providing to better satisfy its customers. Getting more receivers into the hands of consumers is not going to increase consumers' consumption of radio to any measurable degree.

In 2007 CEA researchers asked consumers why they listened to radio.[6] A clear majority (84 percent) said that the content itself attracted them to radio. Music (59 percent) was the most popular form of content, followed by news (36 percent), weather (14 percent), entertainment (13 percent) and traffic (11 percent). Consumers crave good content, and if free local radio provides it, its customers will listen. If they can find better content elsewhere they will not listen to radio, whether they have a tuner or not.

Consumer surveys and the voting consumers do every day with their wallets confirm that consumers crave good content, and that those who are leaving free local radio for other services are likely doing so for better content. There are plenty of portable media players on the market that include radio tuners, yet the player that most consumers have chosen by far, the Apple iPod, does not include a tuner. What does the fact that people are overwhelmingly choosing a radio tuner-less portable media player over other players with tuners indicate? It indicates that these consumers believe that the content on free local radio cannot measure up to the content on their portable media players, and thus that they believe there is no point in having a radio tuner on their portable media player.

The bottom line is that the number of radio tuners a consumer owns is irrelevant if the consumer does not listen to the radio. Broadcasters need to focus on making the latter happen.

Instead of providing a service that consumers are less and less interested in, the radio industry would do well to study why consumers want the products they are purchasing, and modify the service it is providing to better satisfy modern consumer needs. If consumers want to listen, then they will buy radios. What consumers want is more choices, and more personalized content.

WHAT RADIO BROADCASTERS PROVIDE

While there have been some evolutionary changes over time, the radio broadcasting industry has been providing essentially the same service for nearly a century. Radio has expanded from AM to FM, gone from mono to stereo, added auxiliary data, and is now working on migrating from analog to digital transmission. However, through all of this evolution free local radio's principal product has remained a real time audio stream that consumers tune into and consume in real time. The problem with this system is that it severely limits consumer choice. Depending on how many stations are within range, a consumer may have only five, ten or fifteen things to choose from when listening to free local radio. This was terrific 50 years ago, but it is pathetic today. A satellite radio subscriber has over 100 things to choose from all of the time, and the owner of a portable audio player can have over 1,000 things to choose from. Technology has simply advanced to the point where consumers have far better options when it comes to obtaining audio programming.

Many broadcasters are excited about the concept of adding additional audio streams to their broadcasts using HD Radio technology. While this is exciting, radio broadcasters need to carefully think about what this means for consumers. The person who has been able to receive ten free local radio stations versus 100 satellite programs or 1,000 portable audio player songs might now have 20 versus 100, or 20 versus 1,000 choices. Is doubling the number of free local radio choices really that significant to the consumer?

WHAT RADIO NEEDS TO DO

Radio needs to seriously think about all of the things it can do with the digital transmission system it is in the process of implementing. It needs to think way beyond the traditional service that it provides and consider a future where its offerings might be completely different. This will not be easy, for some in the industry will undoubtedly resist such change, and without a strong majority of stakeholders supporting a path forward it is likely that regulators will do nothing to enable improvements in radio broadcasting. However, everything has to start somewhere, so radio should start working toward an ultimate goal of a much better service right now.

The good news is that a digital system is being deployed, and this technology enables the types of services that 21st century consumers demand. What radio needs to do is thoroughly consider all of the possible ways that it could serve the public and grow its business with this new technology. It also needs to carefully study the needs of modern consumers. With an objective understanding of the needs of modern consumers, as well as an open-minded view of the services that digital radio could provide, free local radio could be well on its way to increasing its relevance in the communications landscape of the future.

Today's consumers want personalized content. They are accustomed to a world where an internet browser keeps track of their geographic location, their demographic information, and their likes and dislikes. Their favorite websites are able to provide the latest news about their communities, their favorite sports teams, their favorite stocks, and their local weather. The information one person receives is likely to be very different from the information received by a person in the next room, even though they might both be connected to the same website. Today's consumers can congregate in communities that are defined by interests, not by geography. They are accustomed to a world where a navigation device keeps track of their location on the road and provides them with locally relevant information. And they have cell phones or smart phones that let them retrieve their customized information virtually anywhere.

A real-time audio stream simply cannot satisfy the modern consumer. It could if spectrum and radio station resources were unlimited, for in such a fantasy world there could be a specific real-time stream for every consumer. But we live in the real world where resources are limited, and in the real world free local radio's main product is becoming less and less effective at satisfying consumers' needs.

A real-time audio stream from a free local radio station serving a typical community will have traffic reports that require everyone living in the eastern suburbs to listen to the traffic conditions in the western, northern and southern suburbs, too. For that person in the eastern suburbs the irrelevant traffic reports may as well be commercials or dead air.

A typical city may have four, five or more major colleges or universities. A typical audio stream from a free local radio station for this city requires listeners from all of these schools to listen to reports of how all of the other schools' teams did, reports that many of them may not be interested in.

Free local radio would be much better off if it could offer the type of personalized service provided by the internet and portable media players. If radio could combine the timeliness and personal relevance of internet news and information with the mobility and personalized selection of entertainment of portable media players, then it might really be on to something. Furthermore, if free local radio could expand its reach beyond local communities, letting people "take the community with them" wherever they go, then it would really be serving its customers well. This is not to say that free local radio should not remain a local service, rather that consumers should be able to stay in touch with their local service even when they move beyond the boundaries of what today is a typical radio station's coverage area. Radio can do these things, if it is willing to think outside the box.

To achieve the above goals radio needs to do two things. It needs to stop trying to provide a real-time audio stream and start providing small, cacheable audio packages instead. It also needs to rethink its method of allocating frequencies to communities, and re-farm its spectrum to dramatically increase its capacity to serve its customers.

Think about a radio station that is providing 48 minutes of music an hour to its listeners. with three minutes of traffic and weather, and nine minutes of commercials. Whether the three minutes of traffic and weather are in three 60 second segments each 20 minutes apart, or six 30 second segments each 10 minutes apart, the time available to cover all of the traffic and weather information for the community is severely limited. If it is a community of any significant size, like a metropolitan area, there are going to be many major roads and traffic conditions for which information is never reported, simply due to lack of time.

On this same station the 48 minutes of music per hour consist of 16 three minute songs that have been selected to appeal to as wide a segment as possible of the station's target audience. However, no matter what music the station selects every song cannot possibly appeal to every person the station is trying to reach. In the past, when consumers' choices were more limited while driving in the car, people were more inclined to continue listening to songs that were not that appealing to them because their other choices were limited. There may have been a few other stations that they would jump to looking for more appealing content, or a limited selection of their own music which they might have had in the form of compact discs or cassettes. Today, however, things are completely different. There are about 17 million satellite radio subscribers in the United States.[7] In April 2007 Apple announced that cumulative sales of its iPod had passed 100 million units worldwide.[8] The satellite radio services and the Apple iPod are the major symbols of the plethora of choices that consumers now have when it comes to keeping themselves informed and entertained in the car. Consumers have a number of other options as well.

The handful of local radio stations available to the typical consumer simply cannot offer service as customized to personal taste as that provided by satellite radio, portable media players and other modern consumer devices. So what can free local radio do to retain and grow its user base heading into the future? It can dramatically alter its service and leverage the type of technology that has made personal portable media players so popular. It can stop being a service that provides a real time audio stream and start being a service that provides targeted audio files that allow consumers to customize their portable media players with timely information targeted to their personal needs.

In the example above it was posited that a typical radio station does not have enough airtime available for traffic reports to cover all of the major roads and their traffic conditions to the extent that consumers would like. (Traffic is being used here as an example. Similar arguments can be made with respect to news, weather, sports and other information.) It was also posited that a real time sequence of songs broadcast by the radio station can never match the individual tastes of each of its listeners as well as the listeners themselves can program to their own tastes using their portable media players. Thus the best path forward for free local radio seems obvious. The time allocated to the transmission of music should be reallocated to the transmission of news, weather, sports, traffic and other information, and each of the various items broadcast by a station should be packaged in its own file with accompanying data that targets the file to a specific demographic and/or geographic group.

The small audio packages would be targeted to very specific people, and would include header information that allows the targeted people to receive them while simultaneously allowing the people who are not interested in these packages to ignore them. For example, a typical audio package might be a 30 second traffic report for the northbound lanes of the main highway heading out of town. The commuters who are traveling home in that direction could then program their receivers to cache these audio packages and play them back either on demand or at scheduled times specified by the commuter. The commuters who live south of town would not be interested in these traffic reports and would therefore not have their receivers programmed to cache them. Instead, their receivers would be programmed to store traffic reports for the highway heading south.

The following example illustrates how this system would work. A typical, basic portable media player might allow the user to select among music and photos stored on the device (see Figure 1).

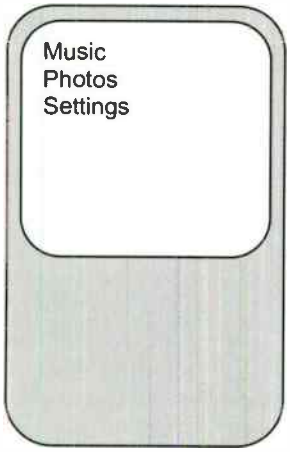


Figure 1: Basic portable media player

Now imagine a portable media player where the main menu has an "information" option in addition to music, photos and settings, as illustrated in Figure 2.

Figure 2: Portable media player with "information" option

When the user selects the "information" option a more detailed list of selections appears based on the free local radio signals that the device is receiving. An example of the options that might be available is shown in Figure 3.

Figure 3: Possible types of information options

Even more detailed options could be located under each of the general information options, as illustrated in Figure 4. In this example under the "traffic" option the user can select the route that he or she takes to work and have the portable media player store and replay traffic updates only for that route. There could be an option to specify timeframes during which specific updates are played so, for example, the southbound updates are played during morning drive and the northbound updates are played during afternoon drive. Or, perhaps there could be preset buttons that allow the user to select pre-programmed traffic information on-the-fly.

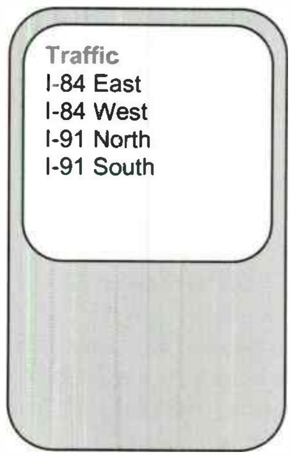


Figure 4: More detailed information options

Not only should consumers be able to customize down to the routes they take while commuting – they should also be able to customize the frequency with which they receive traffic (and other) updates. Figure 5 illustrates one way that this might be done. The same principles would apply to weather updates, news updates, and everything else.

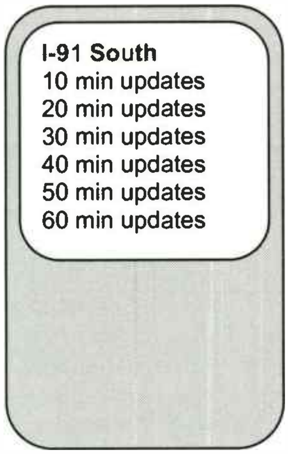


Figure 5: User-selectable frequency of information playback

One way the consumer could determine which radio station's content is received by the portable device would be to select the station under the "settings" heading on the device's main menu, as illustrated in Figure 6. Consumers would select a radio "service provider" just like they select a service provider for their mobile phone service. But of course they could change the radio service provider on their own at any time, and at no cost.

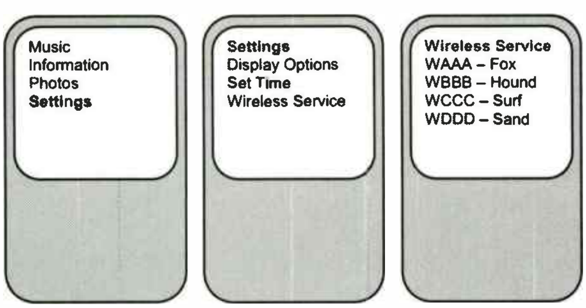


Figure 6: User selection of radio station

The bottom line is that free local radio needs to stop thinking of the service it provides in terms of a live audio stream, and start thinking of it in terms of small customized packets of information that are mass mailed to all portable media players. As with any mass mailings, some people would choose to open them and some would not. In this manner consumers would be able to personalize their radio listening experiences, just like they can personalize their web browsing experiences and their portable media player experiences. The ability to personalize their content is extremely important to consumers. To remain competitive in the future radio must enable its customers to customize their experiences. The technology to do this exists today, the pieces just need to be put together.

It is worth noting that in a system like the one proposed here each of the packets of information distributed by the broadcaster could include an audio and/or graphic sponsor message, so free local radio would remain an ad-supported business.

While the local focus of free local radio is very important, and a critical element of free local radio's success, it is nevertheless very important to fill in the gaps between cities and improve the geographic coverage of radio so that it will be on par with competing services.

COLLOCATED TRANSMITTERS

Providing consumers with virtually uninterrupted local service as they travel across the country is critical to making free local radio competitive with other services. To achieve this, major changes are needed to the way channels are allocated. One change that is needed is the required use of collocated transmitters.

A critical element of a new, more competitive free local radio infrastructure is collocated transmitters. Not some collocated transmitters haphazardly scattered across the country, but a well-organized distribution of transmitters that results in the same group(s) of frequencies being collocated all over the country.

Traditionally radio broadcasters have worked diligently to optimize their coverage areas by carefully selecting specific transmitter sites. The location and authorized power of a station can give it a tremendous advantage (or disadvantage) with respect to other free local radio stations. This system worked well when free local radio was essentially the only source of live audio entertainment available to motorists and others. However, times have changed dramatically in the past decade. Satellite radio, cell phones and the internet offer other options for live audio, and each of them offers far more options than a free local radio receiver. When recorded audio is considered, too, portable media players widen consumers' choices even further.

At any given time the consumer looking for audio content has thousands and thousands of audio programs to choose from. If tuning into one of these options were considered "purchasing" it, and the device used to do the tuning were considered the "shopping mall" where the purchase was made, the choices available to the consumer in the free local radio "mall" (the AM/FM receiver) are dwarfed by the choices available in any of the other malls (satellite radio receivers, internet devices, portable media players, etc.)

There is a concept in retail sales which suggests that there are cases where direct competition between a few competitors actually results in them all being better off than if they were not competing. Imagine, for example, that there are nine shoe stores spread out all over a particular community. One of these shoe stores happens to be in a mall in particular part of town, and is the only shoe store in this mall. In time, several of the other shoe stores relocate to this mall until there are five different shoe stores in this one mall. One might think that the shoe store that was first in the mall might see a drop in business as all of this new competition moves in at the same location. However, cases like this can actually be good, not only for the first shoe store into the mall, but for all of the other shoe stores that moved there, too. This is because the mall becomes known as a great place to shop for shoes, so people are more likely to travel a greater distance to come visit these stores than they are to travel the same distance to visit the other shoe stores in the market that are scattered around in other places. Free local radio can learn from this principle.

Traditionally, the free local radio stations with the best signals in a given market have fought to prevent encroachment into their territory by other free local radio competitors. However, with all of the new competition that free local radio faces from other audio services, the time may have arrived when those stations with the strongest and best located signals in their markets would actually be better off if all of the other free local radio stations in their market were covering the market as well as they are.

Consumers who want to listen to music they like while driving home from work are much more likely to find it on a portable media player with 1,000 choices than on a free local radio receiver with, say, 15 choices. They are also more likely to find it on a satellite radio receiver with over 100 choices than on a free local radio receiver. Consumers shopping for audio programming can now choose to shop at several other "malls" with more stores than the free local radio mall.

For many years free local radio broadcasters have fought station allocation rule changes that would have permitted new stations to come on the air. In the past this made sense because there was only one "shopping mall" for live audio programming (the AM/FM receiver) and new competitors generally took business away from existing stations instead of growing the overall business in "the mall." However, now several new malls have been built and they are drawing consumers away from the free local radio mall. There is the Sirius mall, the XM mall, the internet mall and the portable media player mall, just to name a few. In order to draw consumers back to the free local radio mall, free local radio needs to increase the number of choices it offers consumers. The time has come for free local radio to stop instinctively opposing new allocations and start promoting them.

SINGLE FREQUENCY NETWORKS

Collocation of transmitters would be the first step toward making free local radio's infrastructure more competitive in the modern world. The next step would be to provide seamless nationwide coverage through the use of single frequency networks (SFNs).

The best way for free local radio stations to ensure that all radio receivers can receive service on a particular frequency throughout the entire country is through the use of SFNs. The rollout of digital transmission technology offers an opportunity to develop such networks because one of the features of digital broadcast technology is its ability to overcome multipath interference, which is what the two co-channel signals from an SFN will appear to be to a receiver.

An SFN is simply a network of transmitters with slightly overlapping coverage contours all operating on the same frequency. All of the transmissions from the various transmitters must be synchronized so that in the areas where the signals overlap the two signals either add together constructively or are close enough in time to appear to the receiver to be a main signal and a multipath reflection that are within the receiver's ability to decode.

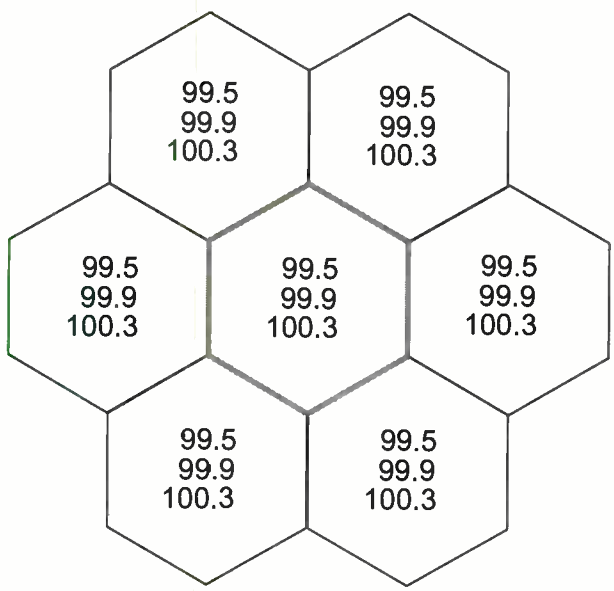


Figure 7: A three frequency, seven transmitter site SFN network

Not only must the radio frequency signals be synchronized, but the content that they are carrying must be synchronized, too. This presents a challenge because in order for free local radio to remain local there must be cases where the programming being carried by one transmitter does not match the programming being carried by a geographically adjacent transmitter. One way that this might be addressed could be to classify certain transmitters in the SFN as dominant signals, and in areas where their signals overlap with those of their neighbors receivers would know to ignore any digital information they receive that does not include the code for the dominant signal. After some defined amount of time, say five or ten seconds, without the receipt of information from a dominant signal, the receiver could be authorized to decode the non-dominant signal. Going back to the dominant signal from the non-dominant signal might require a similar time delay to help ensure that the receiver is actually experiencing good reception of the dominant station. While this would result in some moments of silence for listeners in areas where SFN signals overlap, this would only have to be the case when the programming on the two signals is different. Whenever the programming on the two signals is the same there would be no need to identify one of the signals as dominant, and the receiver would be less likely to mute in the interference areas. Furthermore, if the information being broadcast by the stations is not real time audio, but rather short audio files like those described earlier in this paper, then the consumer would never hear any muting at all because audio playback would only occur after an entire file has been received.

If the SFN concept were to be combined with the transmitter collocation concept it would be possible to have full nationwide coverage on multiple frequencies using one network of towers. In a situation like this in the FM band it would be necessary to eliminate every other channel (*e.g.*, use only 88.3, 88.7, 89.1, etc.) because there is no way to collocate first adjacent channel signals without causing interference. A system like this would result in 50 FM signals in every market. With all of these signals being HD Radio signals, each capable of carrying multiple streams of audio, the result for the consumer would be a selection of programming choices that rivals satellite radio. If the content broadcast over these signals was packaged in the type of small audio files directed at specific demographics that were described earlier, then consumers' ability to personalize their free local radio experiences would be dramatically improved.

The HD Radio system is capable of broadcasting two 48 kb/s streams of compressed audio. There is no reason to think that 48 kb/s would not be sufficient to send 60 one minute audio files, each with headers that include associated target demographic information, in one hour. So, for example, a single FM frequency could be capable of broadcasting 120 one minute audio files per hour, each targeted to a specific demographic. This number is conservative for it does not take into account the fact that bit rates lower than 48 kb/s have been found to produce acceptable audio, nor does it take into account the fact that the overall bit rate available on an FM channel will increase when FM stations stop broadcasting in hybrid mode and start broadcasting in all-digital mode.

Sticking with the 120 one minute audio files per hour per channel example, with 50 channels available at any location this would mean that 6,000 one minute audio clips per hour per location would be available to consumers of FM radio. That works out to be 144,000 one minute audio clips per day. This would offer a tremendous opportunity to provide a service that each consumer could personalize to his or her own taste. For example, if there are ten major roadways over which people commute each day, and traffic reports for each of these roadways, in each direction, were provided every ten minutes, this would require only 120 audio clips per hour, leaving 5,880 audio clips per hour available to provide other content. Maybe one station would try to be an all-traffic station and provide all of these traffic reports itself, or maybe specific stations would focus on specific roadways and work on getting sponsorships from the gas stations, coffee shops, dry cleaners and other stores all along that roadway. And perhaps a station that is focusing on a specific traffic artery would also focus on collecting and distributing localized news and information for the people who live along this commuter route, localized information that they could receive and listen to anywhere in their market.

And just imagine this. Each portable media player that receives free local radio broadcasts could be designed to store cookies that identify which audio clips have been stored and listened to. When the consumer connects the player to the internet this information could be automatically forwarded to the broadcaster and the advertiser, giving the advertiser an immediate, independent report about the reach of the advertising that was purchased. This sort of feedback to the advertiser would make it much easier for the broadcaster to quantify the value being provided, enabling radio advertising to compete more effectively with internet advertising and its ability to quantify clicks.

The potential for revenue growth with a system like this would be tremendous. Keeping with the same example of 120 one minute audio files per hour per station, each station would have 2,880 one minute audio files to send per day. If each of these files contains a 15 or 30 second sponsorship message that would be 2,880 sponsorship messages per day. Compare this with a heavily loaded station that might be broadcasting 24 commercials per hour today. Twenty four commercials per hour works out to an inventory of 576 spots per day. Migrating to the new system proposed here would mean a 400 percent increase in the amount of inventory available to sell. When the inventory of a product increases significantly without a corresponding increase in demand, the price of the product has to drop. However, one of the great benefits of this plan is that there would be a corresponding increase in demand because a station's entire inventory of spots would become drive time inventory. The audio files transmitted by radio stations would be stored on consumers' devices and played back when consumers want to hear them. Thus, the difference between the value of an ad broadcast at 2:00 am and an ad broadcast at 7:00 am would narrow significantly because the ad broadcast at 2:00 am would no longer be less likely to be heard. In fact, the commercial broadcast at 2:00 am might even become more valuable than the commercial broadcast at 7:00 am because the commercial broadcast at 7:00 am might be part of a traffic or weather report that becomes obsolete very quickly, while the commercial broadcast at 2:00 am might be part of an audio file that is listened to repeatedly throughout the day.

CELLULAR NETWORKS

While a nationwide system of 50 all-digital SFNs in the FM band appears to be the best way to improve the free local radio service, this paper has not explored every technical aspect of such a plan and it is possible that there could be some technical challenges that prove insurmountable.

If the SFN concept is too difficult to implement then perhaps a cellular based allocation plan should be considered. It would not be as efficient as the SFN scheme at increasing consumer choice in the FM broadcast band, but it would still significantly improve consumers' choices over the existing allocation scheme, and it would also provide nationwide coverage.

A cellular based system in the FM band could have 12 signals co-located at each transmitter site, meaning that there could be 12 signals everywhere throughout the country. The difference between the cellular system and the SFN system is that, in the cellular system, geographically adjacent signals would be on different frequencies. It is because of the need to use different frequencies for geographically adjacent signals that fewer signals would be available at each transmitter site.

Assuming the signals in the cellular system were HD Radio signals then each one would be capable of broadcasting multiple audio signals, and the service provided to consumers would at least be more competitive with other audio delivery services than free local radio is today. And, as with the SFN proposal, migration to a system that broadcasts short audio files instead of a real time stream would further improve the competitiveness of the service.

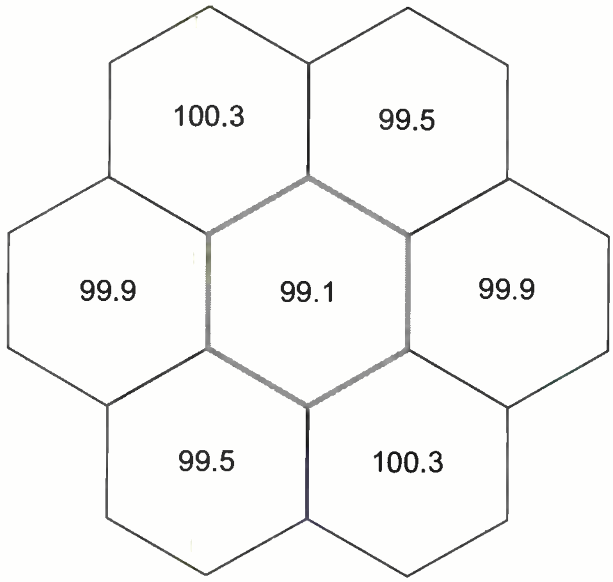


Figure 8: Cellular-based coverage of an area using four frequencies

OTHER THOUGHTS

The purpose of this paper is to urge those free local radio broadcasters who are not already thinking outside the box to start doing so, and to urge those who are already thinking outside the box to start thinking farther outside the box. Absent such thinking, and absent action based on such thinking, free local radio may be in for a steady decline over the next decade or two.

With this in mind, here are some other thoughts that could be developed further, but that are dependent on other things happening first. Thus, it seems a bit premature to develop them too much at this point.

Thought #1: If the capacity of the FM band were increased, either using one of the principles described in this paper or in some other way, then perhaps it would be appropriate to reorganize the AM band by moving most (or all) existing stations to the FM band. If some stations remain in the AM band perhaps they could be on reallocated channels that are spaced far enough apart to allow every station in the AM band to be a clear channel station. Or, if no stations remain in the AM band perhaps this band could be reallocated for unlicensed use, at power levels higher than the current Part 15 limits,[9] essentially allowing anyone to start up a low power broadcast station without prior approval from the Federal Communications Commission. In this manner the demand for very localized community-based broadcast stations would be addressed.

Thought #2: It is possible that the future of free local radio could be one where there are companies that operate the technical infrastructure of collocated transmitters, and separate companies that provide the content.

Thought #3: The television industry is currently working to define a method for broadcasting Advanced Television Systems Committee (ATSC) digital TV signals to mobile and handheld devices.[10] Two mobile/handheld digital TV systems have been developed and demonstrated as part of this effort.[11] Also, television broadcasters have already established guidelines for implementing SFNs.[12] If consumer use of the FM radio band declines significantly in future years there might come a time when reallocating all or part of the FM band for deployment of ATSC mobile/handheld SFNs would make sense. Such an action would not necessarily mean that the FM band would become a TV band, for the ATSC standard enables the broadcast of non-TV content, too.

CONCLUSION

Free local radio is now where local TV was a quarter century ago when cable television was beginning to blossom. A growing selection of alternative sources of content is making free local radio a smaller and smaller part of the overall picture. In the case of TV the past 25 years have seen steady market share erosion for ABC, CBS and NBC affiliates. Figure 9 shows Cable Television Advertising Bureau data for the past two decades.[13]

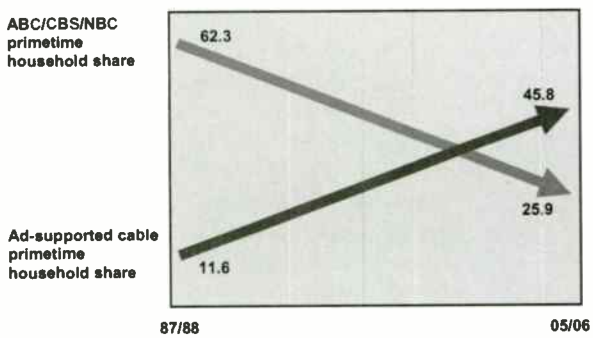


Figure 9: TV viewing trend

Free local radio will undoubtedly experience a trend similar to TV as satellite and internet services continue to grow. If it wants to slow, or perhaps even reverse this trend it needs to dramatically change its technical infrastructure.

People who think that putting a free local radio receiver in all portable electronics will make radio more competitive need to understand that consumers are not turning away from free local radio because they do not have access to receivers, they are turning away because they prefer to listen to other things. Broadcasters must work to provide a service that competes more effectively with the internet, satellite radio and portable media players. Consumers will buy receivers that enable them to acquire content that they desire. If they do not desire the content then they will not listen, period. Putting more receivers in their hands will not change this.

There is no question that the changes proposed in this paper would not be easy. There is a lot of entrenched thinking, and there are a lot of people with parochial interests who will resist changes like those proposed here. Ultimately, it is not even clear that there would be enough support to make changes like this happen. However, it is clear that without significant change free local radio's position in the world will continue to erode over time. The major changes needed to save it have to start somewhere. Why not here? Why not now?

REFERENCES

1 *14th Annual CE Holiday Purchase Patterns Study*, Consumer Electronics Association research report, November 2007.
2 iPhone is a registered trademark of Apple, Inc.
3 iPod is a registered trademark of Apple, Inc.
4 HD Radio is a registered trademark of iBiquity Digital Corporation.
5 "2007 Market Share Reports by Category," *TWICE*, January 7, 2008, pp. 74-80.
6 *5 Technology Trends to Watch – 2008*, Consumer Electronics Association.
7 "SIRIUS Exceeds 8.3 Million Subscribers," Sirius Satellite Radio press release, January 3, 2008. "New Genesis is First Hyundai to Offer XM NavTraffic Real-Time Traffic Service," XM Satellite Radio press release, January 13, 2008.
8 "100 Million iPods Sold," Apple, Inc. press release, April 9, 2007.
9 Title 47 of the Code of Federal Regulations, Part 15.
10 "ATSC Receives Proposals for Mobile and Handheld Standard," Advanced Television Systems Committee press release, June 22, 2007.
11 "Mobile DTV Systems Square Off At CES," *TWICE*, January 21, 2008 (http://www.twice.com/article/CA6524156.html).
12 "ATSC Approves New Recommended Practice, Design of Synchronized Multiple Transmitter Networks," Advanced Television Systems Committee press release, September 25, 2004.
13 "Long Term Total TV Household Share Trends," 2007 TV Facts, Cable Television Advertising Bureau, http://www.onetvworld.org, January 27, 2008.

Digital Opportunities for Radio

Melinda Driscoll
American Public Media | Minnesota Public Radio
Saint Paul, MN
Laura Jensen
National Public Radio
Washington, DC
Nick Kereakos
American Public Media | Minnesota Public Radio
Saint Paul, MN

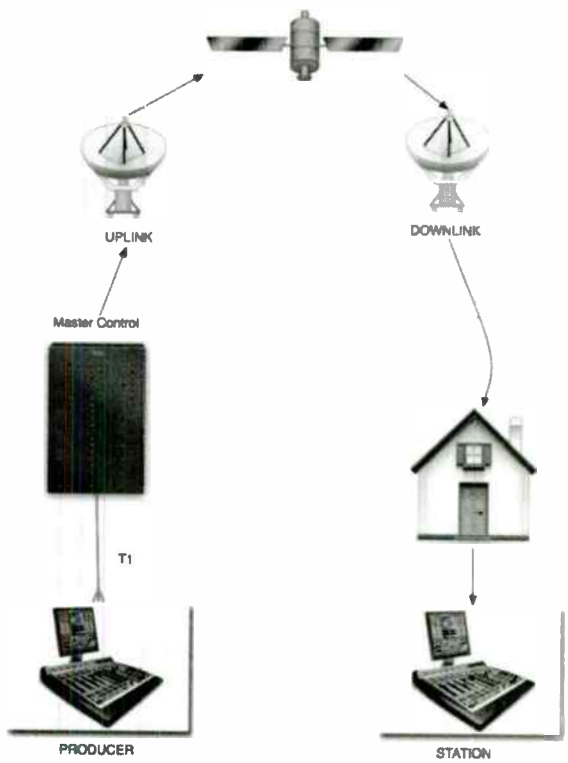
ABSTRACT

Making your content live in multiple platforms is key to your success. With an understanding of converging technologies and thoughtful production methods you can pinpoint whom you want to reach, how you want to reach them and be smart about it. From HD Radio and podcasts to mobile devices and PRSS ContentDepot®, you'll learn how to turn technical challenges into opportunities across a converging media landscape. Learn to work smart, not hard, while positioning your content for multiple platforms.

TRADITIONAL RADIO PRODUCTION WORKFLOWS

In traditional radio content production models, produced content is targeted at single channel terrestrial broadcast radio. Regardless of whether the radio content is local, distributed, or even syndicated, it is still a single channel medium limited within the constraints of the aural listening experience. In recent years this reality has kept the production process and audience digestion straightforward and linear. Respectively, what are now recognized as "traditional" radio production models were created.

TRADITIONAL LINEAR MODEL



These models are characterized by building a singular listening experience and the only written data to accompany them is limited to internal radio network messaging to inform program directors of content for each episode. Forward promotion within the "media space" of the program was limited to promotion opportunities within the content itself.

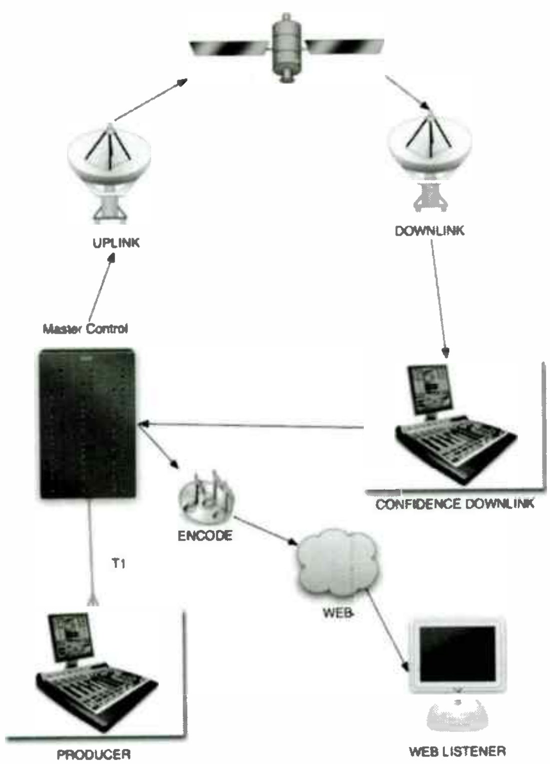
SHIFTING LANDSCAPE

During the infancy of the new media world radio content generally made its way to consumers as a repurposed entity. This meant that the same linear context existed on a new channel, but wasn't necessarily created with multiple mediums in mind.

Linear content combined with early low quality and lossy encoders meant that listening to radio content on the web had limited market and audience penetration and may have been considered more of an experiment and a learning tool rather than a legitimate distribution channel. This is evidenced by an early common use of the web as a preview medium, keeping the data size low enough to remain cost effective while giving the consumer an idea of what the "real" content sounded like. This exploration of new media channels helped content producers and distributors gather audience feedback, suggestions, and perspectives from listeners. Additionally, it provided an experimentation period in which content providers could update and fine tune methodologies without a large audience base a stake.

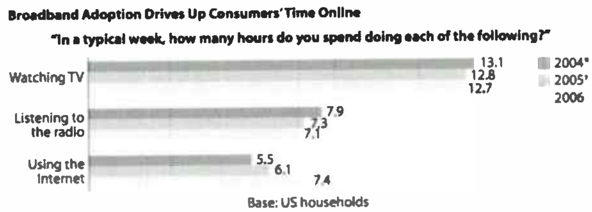
Early new media adoptions of mainstream broadcast media were often traditionally produced content that had been finalized for its original medium respectively. Extreme examples include broadcasters recapturing their own content on the consumer end to repurpose it for the new media platform. The only metadata added to these files and streams were rudimentary and simple such as name, title, and copyright. However, respective to the relative data and audience listening metrics of new media at the time many argued that this minimal investment was appropriate in terms of scope.

CAPTURING A REAL TIME DOWNLINK



ELEMENTS OF CHANGE

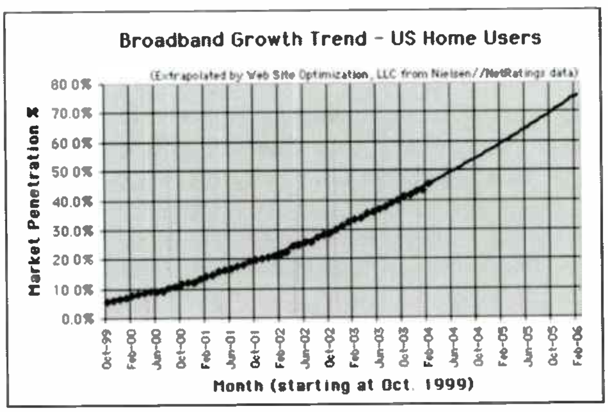
Consumer audiences are larger and now more data-connected; data connections are faster in homes and businesses, and iPods and MP3 players are commonplace devices. Storage media is cheaper, portable devices are smaller, rechargeable battery life is longer, and time shifting of consuming media is an everyday activity. These elements have supported significant changes in the way audiences are consuming radio and broadcast media.



Source: Forrester's NACTUS 2006 Benchmark Survey
*Source: Forrester's Consumer Technographics® 2004 North American Benchmark Study
+Source: Forrester's Consumer Technographics® 2005 North American Benchmark Study

BROADBAND GROWTH

According to Nielsen/NetRatings in a study released in December 2006, more than 78% of residential web users access the Internet are using a high-speed broadband connection. The research company also indicated that broadband users spend 33% more time online than dial users equaling nearly 35 hours per month; they also viewed twice as many web pages.



Source: Extrapolated by Web Site Optimization, LLC from Nielsen//NetRatings data

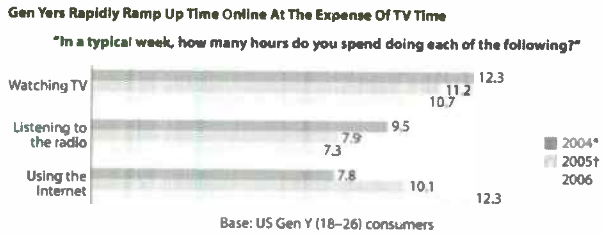
CONSUMER DELIVERABLES

Increased data connections, lower consumer price point entries for devices, and storage have made a stronger and more compelling case for broadcasters to place more energy in the new media space. As a result the industry has experienced an increase in quantity, and quality is on an upswing.

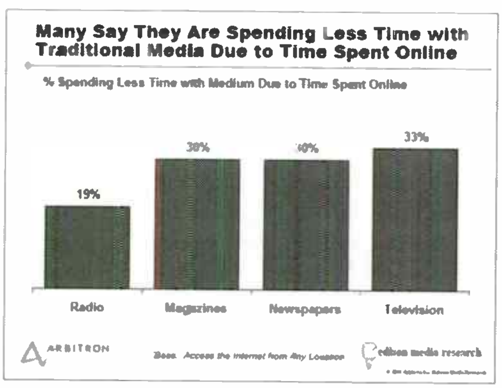
NEW PLATFORMS AND CHANNELS

The steady growth of the new media space, financial accessibility, and steady integration into consumer lifestyle choices has shifted and changed the equation for broadcasters. Today the landscape is converging into a rich, multi-channeled, on-demand, and interactive medium. User generated content lives in and on the same workspace as media companies, stimulating content growth and real time "up to the minute" commentary from both the audience and the content creators' perspectives.

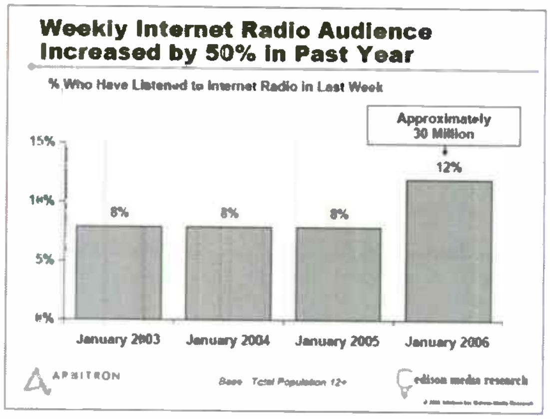
For a media organization to remain viable and continue a growth trajectory inside this space, they must have a compelling web presence, make content portable, be perpetually available, and able to "plug in" to the multiple data aggregators available. Spending time on Google Video or Apple's iTunes demonstrates this reality upfront and instantly. While appointment listening may be the mainstream in more traditional demographic groups, this does not hold true for younger consumers and growing audiences.



Source: Forrester's NACTUS 2006 Benchmark Survey
*Source: Forrester's Consumer Technographics® 2004 North American Benchmark Study
+Source: Forrester's Consumer Technographics® 2005 North American Benchmark Study



Source: Arbitron/Edison Media Research Internet and Multimedia 2006: On-Demand Media Explodes



Source: Arbitron/Edison Media Research Internet and Multimedia 2006: On-Demand Media Explodes

"57% of online adults have used the internet to watch or download video, and 19% do so on a typical day."
Source : PEW INTERNET & AMERICAN LIFE PROJECT

EDUCATION

This experience holds true in educational circles where the new media space is being extended as an interactive educational tool. American Public Media's education collection on Apple's iTunes U platform is a perfect example of a novel cross-pollinating of both media and education. By integrating and aggregating comprehensive metadata and both producer and audience perspectives, this new approach has promise to evolve new ways of listening, learning, and experiencing highly crafted content for longer periods of time and potentially enhancing the impact and shelf life of the product. Additionally the listening and consumption of the content transforms to a non-linear and perpetual experience.

American RadioWorks®

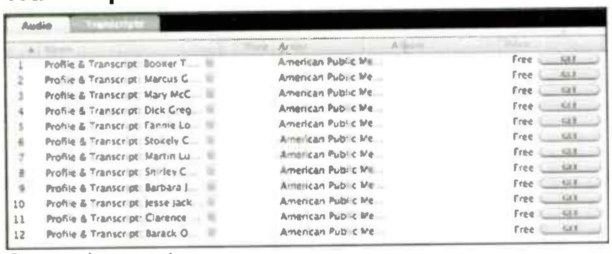
An American Public Media course on iTunes U

Audio

1	Say It Plain: A Century of Great Af...	50:32	Michele Norris, Kat...	AmericanRadioWorks	Free	GET
2	1895: Speech to Atlanta Cotton S...	3:20	Booker T. Washingt...	AmericanRadioWorks	Free	GET
3	1921: Explanation of the Univers...	3:38	Marcus Garvey	AmericanRadioWorks	Free	GET
4	1939: "What Does American Dem...	5:44	Mary McLeod Bethu...	AmericanRadioWorks	Free	GET
5	1963: Civil Rights speech after B...	21:23	Dick Gregory	AmericanRadioWorks	Free	GET
6	1964: Democratic National Conve...	8:11	Fannie Lou Hamer	AmericanRadioWorks	Free	GET
7	1966: "Black Power" speech at UC...	53:05	Stokely Carmichael	AmericanRadioWorks	Free	GET
8	1968: "I've Been to the Mountaint...	43:12	Martin Luther King...	AmericanRadioWorks	Free	GET
9	1974: "The Black Woman in Cont...	27:18	Shirley Chisholm	AmericanRadioWorks	Free	GET
10	1974: Watergate Hearings, House...	13:01	Barbara Jordan	AmericanRadioWorks	Free	GET
11	1988: "Keep Hope Alive" Democr...	50:19	Jesse Jackson	AmericanRadioWorks	Free	GET
12	2001: "Be Not Afraid" American E...	47:42	Clarence Thomas	AmericanRadioWorks	Free	GET
13	2004: Democratic National Conve...	16:08	Barack Obama	AmericanRadioWorks	Free	GET

Course audio

Transcripts



1	Profile & Transcript: Booker T...	American Public Me...	Free	GET
2	Profile & Transcript: Marcus G...	American Public Me...	Free	GET
3	Profile & Transcript: Mary McC...	American Public Me...	Free	GET
4	Profile & Transcript: Dick Greg...	American Public Me...	Free	GET
5	Profile & Transcript: Fannie Lo...	American Public Me...	Free	GET
6	Profile & Transcript: Stokely C...	American Public Me...	Free	GET
7	Profile & Transcript: Martin Lu...	American Public Me...	Free	GET
8	Profile & Transcript: Shirley C...	American Public Me...	Free	GET
9	Profile & Transcript: Barbara J...	American Public Me...	Free	GET
10	Profile & Transcript: Jesse Jack...	American Public Me...	Free	GET
11	Profile & Transcript: Clarence...	American Public Me...	Free	GET
12	Profile & Transcript: Barack O...	American Public Me...	Free	GET

Course documents

WHERE TO PUT THE ENERGY

A large question for all media companies when faced with this reality is how and where to define what space to thrive in. The approaches of both aggregator and curator are equally legitimate, however broadcasters are wrestling about where to focus and why. American Public Media has chosen to curate its content and subsequently "plug into" leading aggregators to connect most effectively and interact with audiences. National Public Radio (NPR) has taken a similar approach choosing to build and host the creation of both original content specific to new media, plus developed new channels for its established content. The NPR Music site launched in November 2007 explores new strategic business models for content aggregation of NPR and station-produced music media.

NPR Music home page

AN OVERVIEW OF THE CURRENT PUBLIC RADIO LANDSCAPE

This is an exciting time for us in public radio, as we wrestle with new business models and platforms for content delivery. As we look to new media, we can't overlook existing formats. Findings in a 2007 study by Arbitron and Edison Media Research indicated that consumers continue to rely on AM/FM radio even while they start to use emerging digital platforms, including podcasting, satellite radio, and HD radio. Faced with these evolving media platforms, public radio producers have been expanding their content formats and delivery mechanisms. The report ("The Infinite Dial 2007: Radio's Digital Platforms") is available at
www.arbitron.com/downloads/digital_radio_study_2007.pdf.

Let's take a quick look at some of the initiatives currently underway in public radio.

On-Air AM/FM Broadcast

Public radio has a wealth of content types produced in different formats for different uses on AM/FM radio. Completed programs included branded series that are produced on a daily or weekly basis for stations to air based on a standardized program clock. Typically these shows are formatted for one to two hour time blocks. Examples include news shows such as NPR's "Morning Edition" and "All Things Considered" and cultural and entertainment programming such as American Public Media's "A Prairie Home Companion" and "Performance Today."

During the past year, we have seen major public radio producers develop significant new content series for traditional broadcast with an eye towards expanding audiences. A multi-platform approach to new programming such as NPR's "Bryant Park Project" includes web components such as a blog and podcasts as well as on-air offerings.

Content series produced for public radio also include discrete modules intended for inclusion in other long form content. These modules can take the form of 60 and 90 second segments on astronomy or other topics that can fill short breaks in programming. StoryCorps is an example of a national initiative to record the voices and life stories of everyday people via mobile StoryBooths. Recorded conversations are preserved at the Library of Congress and are produced for broadcast on public radio and the web.

HD Radio

NPR Labs initiated public radio's Tomorrow Radio Project in 2003. The team researched multi-channel HD uses for public radio and has continued to champion HD radio as well as other methods to help public radio stations get content to their listeners more effectively. Public radio program producers have responded to the opportunities provided by HD outlets by developing new programming content specifically designed for air on secondary channels. Examples include 24 x 7 music services such as "Groove Salad", "XPoNential Radio", and "Folk Alley" distributed by NPR, and Spanish language streams like NPR's "Radio Ahora" and PRI's "BBC Mundo."

To take advantage of HD Radio capabilities to display digital Program Associated Data (PAD) or Program Services Data (PSD), NPR Lab's is currently spearheading activities to further network-distributed PAD for the public radio system's most listened-to programs, and they are working to influence receiver manufacturers to deploy improved PAD designs to better serve the public radio audience.

Podcasting

Podcasting delivers an audio file, typically an MP3 file, available online for downloading. Consumers can then hear the podcast whenever they want from their computers or on portable media players. NPR has teamed with over 50 public radio stations and producers to offer a public radio podcasting service. The podcasting program makes available programming that may otherwise not be offered as part of on-air content. Podcasting provides a format for public radio producers to expand their creative content and distribution for consumers that want to hear diverse programming and real voices.

Like many broadcasters, American Public Media is using podcasting as a way to both better serve existing audiences through this channel as well as reach new audiences and experiment developing not-for-broadcast content. For example, our radio programs like Marketplace®, News From Lake Wobegon®, and The Writer's Almanac® are available as podcasts. Additionally, we're producing a non-broadcast content as podcasts such as a personal finance video podcast called "Money Clip" hosted by Marketplace's Economics Editor Chris Farrell.

Mobile Delivery

In 2007, NPR launched NPR Mobile. Through collaboration with ten partner stations, NPR and local station news and features are available to users of mobile phones and other devices. NPR mobile includes NPR Mobile Voice, a voice service available via a phone call, and NPR Mobile Web that delivers text, audio, and pictures directly to users through a web site developed for mobile phone viewing.

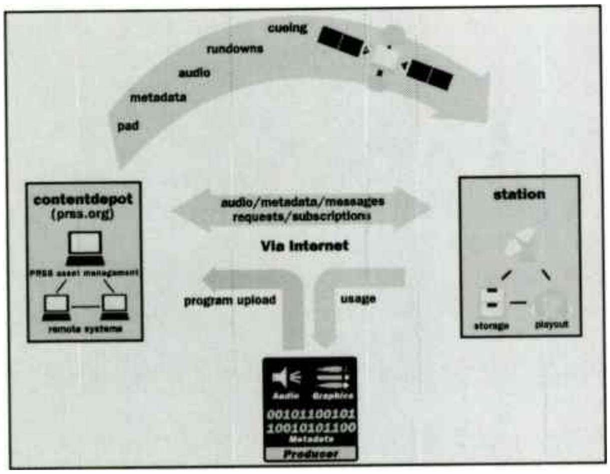
A National Program Distribution Network

Public radio benefits for a centralized distribution network through which the majority of AM/FM-based content is delivered to public radio stations throughout the United States. The Public Radio Satellite System® (PRSS) is managed by NPR Distribution and is a unique, cooperative enterprise. Each Participating Station is a stakeholder in the collective assets of, and services provided by, the satellite system. Stations that are part of the network own their own downlink and uplink equipment. The satellite transponder capacity and the national operating system equipment located in Washington, DC, are owned by the Public Radio Satellite Interconnection System Charitable Trust.

The PRSS serves more than 400 downlinks and over 250 program producers and distributors. Many additional stations also receive programming sent over the satellite through local connections with downlink stations.

In 2006, the PRSS launched a new distribution system. The PRSS® ContentDepot® combines Internet and satellite technologies to offer an integrated suite of services for public radio. All types of content can be delivered through the system, including live programs for airing by stations in real-time, pre-recorded audio files, one time specials, promotional materials, promos, graphics, text, and metadata.

The ContentDepot® uses a standards-based architecture that builds off of public broadcasting metadata standards (PBCore) and audio standards (MP2) to make content consistent and automatable by stations. The ContentDepot® provides a public radio metadata archive of programming distributed through the system.



Creating Automation-Friendly Content

Station automation requires consistency. In order to deliver content in a manner for stations to automate, producers must:

- Adhere to consistent formats, program segmentation, and air windows for file-based content.
- Upload content by the upload deadlines
- Name content appropriately for station searches and playout systems

Online Content

Content producers and distributors have exploited online capabilities to expand distribution options. In 2007, NPR launched NPR Music, a music discovery site that makes available NPR and partner station produced content. Content available on the site includes pieces from on-air broadcasts as well as pieces produced specifically for the site (including video).

Many independent producers have developed their own websites and offer free or revenue-based services for stations to download content for air and for consumers to stream audio and obtain podcasts. The Public Radio Exchange is a nonprofit service for distribution, peer review, and licensing of radio pieces. http://www.prx.org/

As individual content producers and media companies look at the various platforms now available for getting their products to consumers new efficiencies in production workflows are all the more critical.

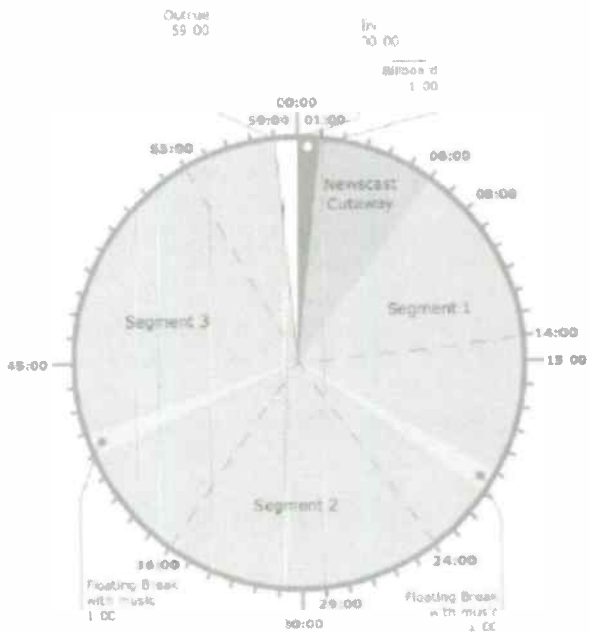
RE-DESIGNING THE PRODUCTION MODEL

Utilizing new channels and platforms requires a new approach to the production cycle, an approach that is smart and coordinated to align program content, metadata, and targets. Prioritizing what data should be compiled and how the content will be used must be decided early in the process to yield maximum flexibility and listener impact. The public radio system began examining best practices and new production methods as part of the release of the ContentDepot®, public radio's new network distribution system.

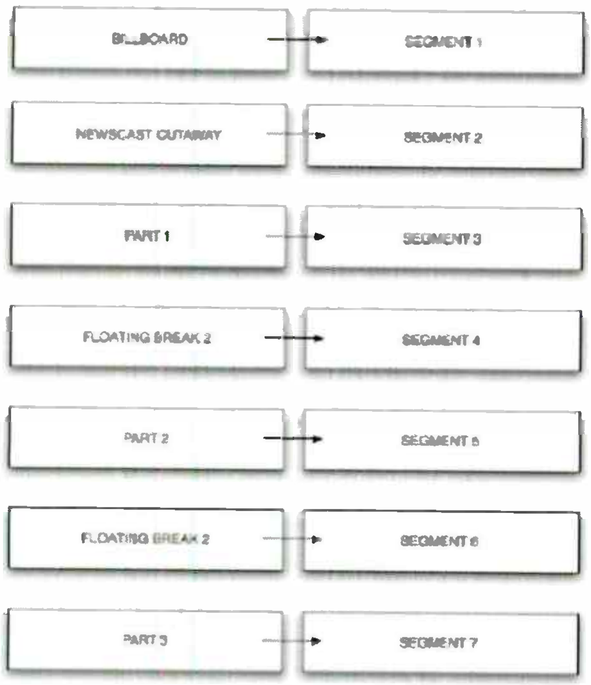
FLEXIBLE MODEL

Programs that are not delivered in real-time via the ContentDepot® are referred to as "store forward" programs. The program segments replace and mirror traditional program identification breaks. All store forward programs are delivered to stations locally via a satellite storage receiver. This model provides maximum flexibility at the station level, underwriting, breaks, and weather may all be inserted in and around the program segments however the local station or network wishes.

Traditional Clock



Segmented Program



METADATA

The ContentDepot® requires very basic metadata entered to facilitate the delivery of programming to stations and to deliver the basics of the show each day to program directors and staff at the local level. These metadata fields address items like start and kill dates/times, title, and artist. This simplicity somewhat parallels what the industry experienced on early new media channels, where the metadata was simplistic and basic enough to provide instructions needed for baseline operation at the station level. This metadata set is clearly not intended for public consumption, but is rather an insider industry tool. The illustration below illustrates how stripped down this metadata is, contained in the cart chunk header of a MPEG1 Layer 2 - 256 kbps stereo .wav file.

Cart Chunk

```
       Version: 0101
         Title: APrairie 080112 SGMT 01 APHC Par
        Artist: Garrison Keillor
         CutID: 61861
      ClientID:
      Category:
Classification:
        OutCue:
     StartDate: 2008/01/12
     StartTime: 18:00:00
       EndDate: 2008/01/13
       EndTime: 23:59:59
     ProdAppId: ContentDepot
ProdAppVersion: 1.0
       UserDef:
      LevelRef: 0
      Reserved:
           URL:
```

MORE DATA

Contemporary audience expectations for the consumption of program associated metadata has forced public broadcasters like American Public Media and National Public Radio to provide regular, up-to-date, high quality audio, visual, and interactive content to listeners. As a result we've embraced these expectations in an effort to keep up with audiences and to provide the experience and interaction they expect. Today a new program could not and would never be launched successfully without a compelling web component, audio archives, on demand listening, and a feedback mechanism or public forum. As examples, examine the public metadata available to any Internet user for the programs A Prairie Home Companion®, and The Splendid Table®.

Example 1: A Prairie Home Companion®: RSS Feed Via Google

APM: A Prairie Home Companion

⊟ A Prairie Home Companion

Frozen, Frosty Fitzgerald: We're back from vacation and broadcasting live from The Fitzgerald Theater. With this week's special guests, award winning English trumpet soloist Alison Balsom, dangerously sweet country singer Suzy Bogguss, and sitting in with The Guy's All-Star Shoe Band, steel guitar wiz Joe Savage, guitarist Dan Neale, and the multi-talented John Niemann. Also with us, The Royal Academy of Radio Acting, Tim Russell, Sue Scott, and Tom Keith, The News From Lake Wobegon, and much more this week on A Prairie Home Companion.

http://prairiehome.publicradio.org

FROZEN, FROSTY, FITZGERALD

T WEEK SHOW

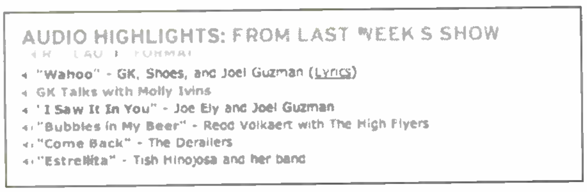
Suzy Bogguss

January 19, 2008

This week on A Prairie Home Companion, we're back from vacation and broadcasting live from The Fitzgerald Theater. With this week's special guests, award winning English trumpet soloist Alison Balsom, dangerously sweet country singer Suzy Bogguss, and sitting in with The Guy's All-Star Shoe Band, steel guitar wiz Joe Savage, guitarist Dan Neale, and the multi-talented John Niemann. Also with us, The Royal Academy of Radio Acting: Tim Russell, Sue Scott, and Tom Keith, The News From Lake Wobegon, and much more this week on A Prairie Home Companion.

LAST WEEK'S SHOW

All About the January 12th show with Molly Ivins, Tish Hinojosa, Joe Ely, Joel Guzman and more >>



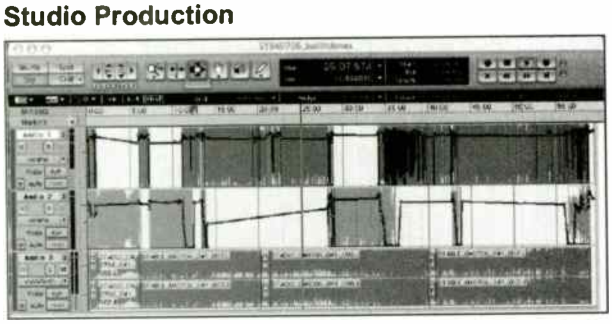
AUDIO HIGHLIGHTS: FROM LAST WEEK'S SHOW

- "Wahoo" - GK, Shoes, and Joel Guzman (Lyrics)
- GK Talks with Molly Ivins
- "I Saw It In You" - Joe Ely and Joel Guzman
- "Bubbles in My Beer" - Reed Volkaert with The High Flyers
- "Come Back" - The Derailers
- "Estrellita" - Tish Hinojosa and her band

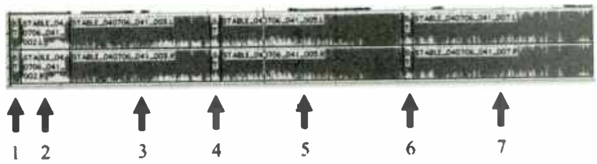
Example 2: The Splendid Table®

The Splendid Table® is a weekly national distributed program produced by American Public Media. The program is produced weekly in seven segments each week, has a weekly podcast, and online web archive, discussion section, regularly updated website, and a dynamic RSS module. Here's how the production cycle plays out:

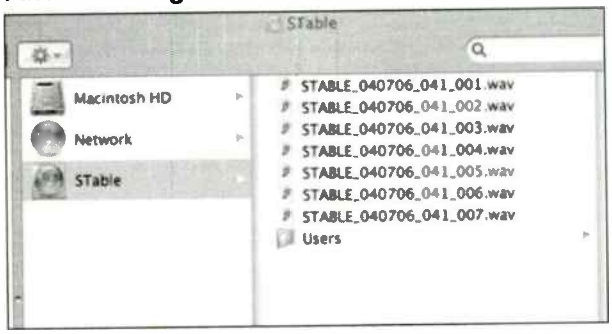
Studio Production



Program Segments



Finalized Program Files



Content Depot

Episode File(s)

Segment 1: Billboard - CUTID60218
File: Splendid_075_SGMT01 0.91 Mb 01/10/2008
Length: 00:01:00

Segment 2: 5 Min News Hole - CUTID60219
File: Splendid_075_SGMT02 4.56 Mb 01/10/2008
Length: 00:05:00

Segment 3: Part 1 - CUTID60220
File: Splendid_075_SGMT03 11.9' Mb 01/10/2008
Length: 00:13:03

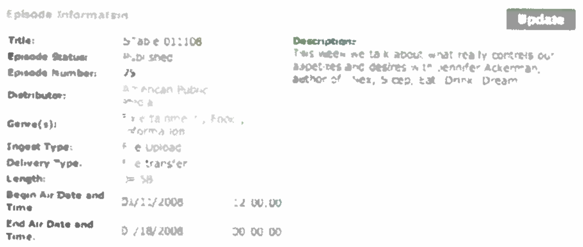
Segment 4: Station Break 1 - CUTID60221
File: Splendid_075_SGMT04 0.89 Mb 01/10/2008
Length: 00:00:59

Segment 5: Part 2 - CUTID60222
File: Splendid_075_SGMT05 18.15 Mb 01/10/2008
Length: 00:19:52

Segment 6: Station Break 2 - CUTID60223
File: Splendid_075_SGMT06 0.89 Mb 01/10/2008
Length: 00:00:59

Segment 7: Part 3 - CUTID60224
File: Splendid_075_SGMT07 - 16.54 Mb 01/10/2008
Length: 00:18:06

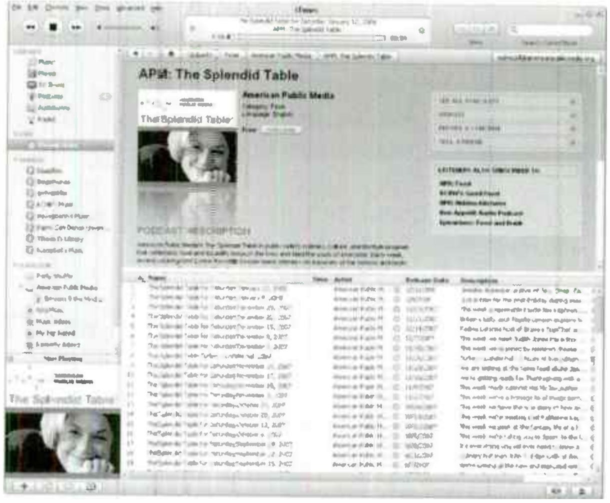
Episode Summary



Web Site

On Demand

Podcast

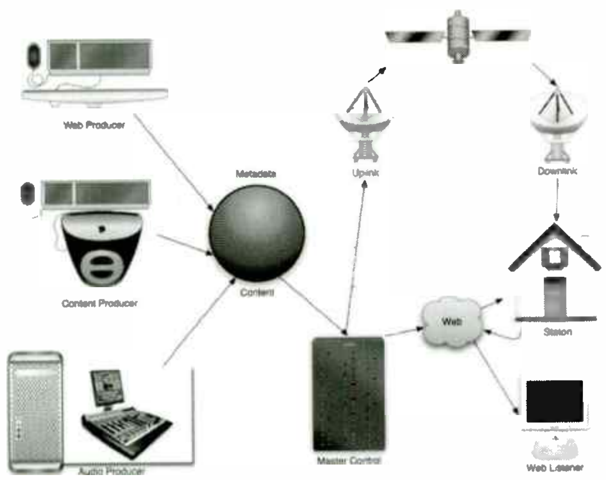


OPPORTUNITIES

The metadata necessary to facilitate the needs of the ContentDepot" are somewhat simple, and focused completely at the station management level to facilitate the broadcast of the program. At the same time the expectations of our audiences continue to grow and they expect more via the web, broadband, mobile devices, WiFi, and HD Radio. In order to keep up with this growth curve we suggest that a purposeful realignment of program elements and associated metadata assembly is required to realize maximum efficiencies, impact, and relevance for both stations and audiences. The distinction between what we as a network are now providing to stations and also direct to consumers is virtually the same. The end products may be adjusted slightly to match targets, however we provide RSS plug in modules, program rundowns, timings and cues, archives, and feedback mechanisms to both stations and additionally directly to listeners.

As a result, we conclude that the production flow can and should be unified upstream to pool the ideas and efforts of producers, technical staff, writers, radio staff, online staff, business development, marketing, and sales. This does not suggest that all these perspectives belong in the control room, but rather strives to fundamentally address that all these parts make up the support structure of the whole, all of which have a stake in the deliverables. While America Public Media and National Public Radio have updated their production methods to utilize the ContentDepot" , we also see a future opportunity to re-tool how metadata, production, and distribution could all be combined into a cohesive, flexible, and scalable model. Furthermore, a revised model can be applied to Ðand streamline Ðcontent produced for non-broadcast channels.

THE NEW MODEL



WHY DOES ALL OF THIS MATTER?

Reflecting back on the earlier point about changes in media consumption and delivery, audiences have new ways and new expectations for consuming the content we produce. If we want to continue to reach our audiences, we need to be where they are. To survive and thrive with these changes, we need to rethink our production processes to enable our content to reach audiences in these new ways.

MARKETING CONSIDERATIONS FOR THE NEW MODEL

As we plan our production workflow changes to adapt to these changes in audience use and delivery technologies, we need to remember that the changes should be driven by our audience needs and wants. In other words, our new production processes shouldn't focus only on enabling channel delivery, they need to factor in the audience's use of these channels. In short, start at the end. Where is your target audience looking for content? How are they finding it? How are they using it? Once we know the answers to these questions, we can design the final output of our content to be optimized for our audience's use.

These are fundamentally marketing issues: our production processes need to ensure that our content is:

- Available – our content is available where our audience is looking
- Findable – our content is presented, indexed, tagged so our audience will find it
- Usable – our content is presented in a way that's accessible and friendly to our audience's behaviors

Example: Make it available

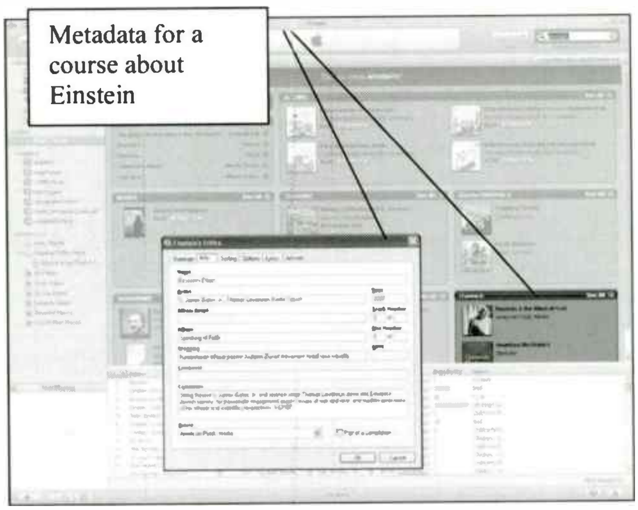
Where are audiences looking for Internet radio? In the last fiscal year Minnesota Public Radio's audience for its live Internet radio streams grew over 100%. The reason? We made our streams available where the audience was looking for Internet radio.

- The Windows Media Radio guide
- Web search engine searches for "internet radio", especially "Classical internet radio"
- Shoutcast, SnapFM, NetVibes, the Open Directory Project, NPR Music, iTunes Radio directory, etc.

Example: Make it findable

Once we know where our audiences are looking for our content, we need to be sure that they can find our content there. Often, that means paying particular attention to the search functions in the applications and directories where our audiences are looking.

For example, in September 2007 American Public Media launched a new site on Apple's iTunes U education platform as part of Apple's "Beyond Campus" launch on iTunes U. Knowing that the iTunes U audience would be looking for content in a much different way than we usually present it to our public radio audience, we endeavored to tag our content using common academic terms. In addition, we familiarized ourselves with the iTunes U search function to determine which metadata elements are most factored into search. We established a workflow to insert keywords and descriptions accordingly.



In the new model alluded to earlier, metadata considerations for the iTunes U outlet would be one of many that is built into the production process.

Example: Make it usable

One of the biggest challenges in preparing content for audiences that are using it in increasingly diverse ways is making sure the content is easy for them to use. Sometimes what seems like the smallest detail in metadata syntax can make a big difference in our content's usability. Podcasts demonstrate a great example of this particular challenge. Depending on the user's MP3 player and how the user sorts her podcasts, our choices in metadata and syntax can make or break a podcast's usability.

As we transition to smarter production processes, focusing on the end product and all the ways our audience will interact with our content can help us make smart choices about metadata, syntax, format, and other elements that affect our ability to serve our audiences.

BRAND CONSIDERATIONS FOR THE NEW MODEL

In addition to making sure our content reaches our audience, we also need to make sure it's appropriately attributed to our organizations. Building brand considerations into the production workflow can be a particularly challenging issue in the world of syndicated media. Video and portable device screens are small, adding a brand name to bare bones content like RSS feeds can be cumbersome to the user, and syndicated content can appear in virtually any form, in any location. So, what can we do? The points below outline the guiding principles we apply when choosing how to brand digital content when we don't always know how it will be used.

1. Inventory the known and most common uses.

   Video example: YouTube, YouTube embedded on other sites, Vimeo, Gather, on your site, video podcast

2. Find examples of effective and ineffective branding approaches in those uses.

National Geographic video on YouTube with well-recognized yellow rectangle

National Geographic video as it appears when YouTube video is embedded on other sites with YouTube logo added

3. Create a branding system that works for your users, your company, and your content.

   The Current is an alternative music station from Minnesota Public Radio. We often produce video when musicians come to the station to perform live. As we have experimented with delivering these videos online via outlets like YouTube and Gather.com as well our own Web site, we considered these outlet in our approach to branding the videos.

The Current video on The Current's web site

The Current video on YouTube

The Current video on Gather

4. Field test, monitor, and adjust your approach as needed.

 After developing an online brand approach, often we discover issues we hadn't anticipated. A field test is a good way to work these discoveries into the plan. Try out your online branding approach with a small audience, in a somewhat controlled environment to get your audience to help you discover the issues you hadn't anticipated.

REVENUE CONSIDERATIONS FOR THE NEW MODEL

While the whole industry is still trying to figure out viable and sustainable revenue models for online content, advertising-supported revenue is the most common model currently in place.

In the public radio world where we operate, underwriting support is the primary revenue source enabling us to distribute our online content. As a result, we benefit from optimizing our content production methods to easily accommodate underwriting messages and inventory. Additionally, when we don't have an underwriter for particular online content we want to use that advertising space for our own promotional messages.

Content production processes need to allow us to quickly and dynamically insert underwriting messages in a variety of formats. The example below illustrates how the audio for a radio show episode is delivered to the radio broadcast distribution system. At the same time, an underwriting message is appended for the podcast audio and an audio prepend and images are dynamically served in front of the online streamed audio. This is just one example of how flexible our production systems need to be to accommodate a variety of advertising formats and inventory.

Broadcast distribution system:
Program audio

Program audio — **Podcast:**
Advertising prepend + Program audio

On-demand stream:
Advertising prepend + Program audio + Advertiser image

Advertising revenue model benefits from new production model

SUMMARY

Podcasts, online video, cell phones, smart phones, ubiquitous wireless, broadband growth, HD radio Đall of these trends present a challenge and an opportunity for us as content producers to reach and serve our audiences in new and exciting ways. In order to adapt to these changes, a thoughtful evaluation of our production workflow is required. Such assessment will include both processes and systems and will consider how changes in the media landscape have influenced audience expectations and response. This ongoing analysis will position us to take most advantage of the opportunities these changes present.

DTV Broadcasting for Mobile and Handheld

Sunday, April 13, 2008
9:30 AM – 5:30 PM

Chairperson: Mark Richer
Advanced Television Systems Committee, Washington, DC
Chairperson: Dave Converse
ABC/Disney Television Stations Group, Burbank, CA

*Panel
Panelists: David Virag, Thomson Silicon Components, Indianapolis, IN
Wayne Luplow, Zenith/LG Electronics Corporation, Lincolnshire, IL
Dave Glenn, Ion Media TV, Saint Petersburg, FL
John Godfrey, Samsung, Washington, DC

Managing the Unmanageable: Transforming Media for Multi-channel, Multi-platform Delivery
John Pallett, Telestream, Inc., Nevada City, CA

*Producing vs. Repurposing for Multiple Platforms
Pete Sockett, WRAL TV, Raleigh, NC

*Make Your Mobile TV Network a Reality
Jean Macher, Thomson Grass Valley, Beaverton, OR

Workflow in the New Media Landscape: Production, Post and Final Delivery
Michael Castro, National Geographic Television, New York, NY

Local and Ultra Local Content in Broadcast Mobile TV
Richard Lhermitte, ENENSYS Technologies, Rennes, France

The Role of File-based Workflows and Metadata in Repurposing Content for the Web and Mobile TV
Mark Bishop, Thomson Grass Valley, Beaverton, OR

Leveraging FEC Advances to Optimize DVB-H Networks
Marshall Porter, Digital Fountain, Fremont, CA

New Techniques for Mobile TV Broadcasting Based on ISDB-T
Masahiro Okano, NHK Science & Technical Research Laboratories, Tokyo, Japan

Preserving SFN in a Broadcast Network Using IP Distribution
Nicolas Fannechere, ENENSYS Technologies, Rennes, France

Improving Mobile TV Streaming: Lessons Learned for Successful Deployments
Boris Felts, Envivio Inc., South San Francisco, CA

---

*Paper not available at the time of publication

MANAGING THE UNMANAGEABLE: TRANSFORMING MEDIA FOR MULTI-CHANNEL, MULTI-PLATFORM DELIVERY

John Pallett

Telestream, Inc. • Nevada City, California

There's a significant and lucrative, new source of revenue in the media market – consumers are increasingly willing to pay money to download and view content on personal computers, mobile phones, and handheld devices.

However, this opportunity comes with a catch. As a content owner, in order to exploit the growth in the mobile, handheld, and Web market, you must determine how to repackage and publish your media for playback on a broad range of different platforms, to a broad range of distribution channels.

It may look easy on paper: connect the video archive box with the iTunes box, and you're making money! The first step – publishing a single file to a single distribution channel – can be deceptively easy. This early success may elicit thoughts such as, "Well, I made one on my laptop – how difficult will a hundred be?"

It probably will be difficult, because a project is not a process. Creating systems that successfully replicate production steps dependably and economically – even in low volumes –to multiple distribution channels for multiple platforms is not a trivial task.

MULTI-PLATFORM PROJECTS POSE SIGNIFICANT CHALLENGES

Different types of publishing projects pose unique challenges. One-off and small projects – a few files to a few channels – do not pose significant hurdles for storage and network traffic; and desktop tools may suffice. These projects are typified by the traditional challenges of working with video files and metadata.

However, when turning a small project into a replicable, reliable daily or weekly process with the infrastructure and workflow management that is required to sustain increased file volume and/or an increasing number of distribution channels and playback platforms, a new breed of challenges arise. The most significant challenges are in the domains of infrastructure, repeatability, personnel, and automation – all thwarting your efforts to maintain high quality, on-time delivery, and production cost efficiencies.

In this paper, we'll review these two sets of challenges – traditional challenges inherent in any video production workflow, and new, workflow management challenges – those that occur when video production workflows increase in scale, repetitiveness, and number of distribution channels and platforms. We'll also discuss where software solutions can – and cannot – help you overcome these challenges.

TRADITIONAL CHALLENGES

Some of the traditional challenges facing small video production projects are obvious. When generating any video file, the output format must be considered, and often, each playback platform's unique and specific requirements.

Requirements likely include unique frame sizes, bit rates, wrappers, and file size limits. Complicating matters, different platforms often have different hardware compatibility requirements (most notably 3GP and IPTV variants). For example, not every 3GP file works on all mobile phones. The underlying video and audio formats must also be considered, as must the method used for wrapping (multiplexing) the audio and video into a single file.

Visual requirements must also be considered: the underlying video may need to be cropped, re-shaped, processed for inverse-telecine from 29.97 or 25fps to 24fps, or otherwise optimized to conform to each platform's requirements.

Last, delivery details requiring consideration: the packaging of the video, adding a preview movie and poster art, plus MD5 hashing and delivery protocol via FTP or proprietary file transfer methods.

These traditional challenges are applicable to virtually every video workflow, big or small, and should be no surprise. In many cases, desktop tools suffice to address them, and are particularly well-suited to discovering a solution through trial and error. However, what happens when a workflow must be repeated weekly, or scaled up to hundreds of assets and channels?

NEW WORKFLOW CHALLENGES

As media volume grows and production cycles increase in frequency (or even become near-real time), and distribution channels and delivery platforms multiply, new workflow management challenges associated with scalability and repeatability come into play. Lets examine some of these workflow management challenges.

Repetitive Edits

Some workflow challenges are brought about by your own internal marketing requirements. You may need to remove black, add bumpers and trailers, and perform branding tasks, consuming valuable time in an edit suite. Manual editing is not particularly expensive for a few clips, but when dealing with dozens to hundreds of clips each day or week, it can become costly, or worse – error-prone.

How much does an edit bay cost to staff and operate full-time? How can you insure that the correct bumper is being applied to the right clip? These questions are usually answered by developing a process around the editing requirements, not marketing or operational requirements. As a result, the process can limit the flexibility of marketing and operations to achieve the broad-reaching, frequently-changing demands of channels and individual consumers.

User-Generated Content

A second, emerging marketing requirement is support for user-generated content as source material: amateur-sourced video from cell phones, DV cameras, and other consumer devices. Web site developers are all too familiar with the challenges of dealing with such content: consumer devices that create non-compliant files, incomplete submissions via the Web, completely invalid files, and files with unacceptable content.

Hardware

Next, consider the workflow challenges inherent in hardware requirements for your workflows. When encoding from tape, dealing with hundreds of files may require 24 by 7 staffing for tape deck operation. If so, a process must be devised, tested, and implemented to manage both the personnel and the video assets. When you're transcoding or editing at this volume, your video material typically bounces between several different servers during multiple processing steps.

Depending upon format, consider also that even a few dozen full-length video episodes may consume a terabyte of storage. How can you be sure that you don't exceed storage requirements on each server or edit bay-halting your workflow process dead in its tracks? How do you optimize network connectivity between systems in each step of the workflow process?

Answers to these questions are often two-fold: first, you need to design a hardware environment that matches the process; second, you need to manually manage the location of files (or use a single, central storage) as they move through the process. What if the process changes?

Delivery

If you're delivering media over the Internet, you may plan your connectivity for one set of providers, but what if their requirements change? Many delivery platforms require that all submitted content matches their internal catalog. How can you be sure that what you're submitting matches what the provider expects? Today, these processes are largely managed – and bottlenecked – by manual effort.

Final Considerations

Finally, some of the platform delivery requirements that were fairly easy to meet as a small project, become process management challenges when scaling up a workflow. Example: how can you be sure that you are matching the right video file with the right metadata, and the right preview image, for the right provider?

Whether you're dealing with traditional video file challenges or the more significant, workflow management challenges, a single point is worth noting: with mobile and Web streaming delivery, more and more of the workflow is file-based.

WORKFLOW MANAGEMENT SOFTWARE

As we move increasingly to file-based workflows, it certainly would be beneficial if workflow-focused software could be utilized to automate production, reduce failures, and report exceptions, among other goals. Task automation and failure reduction has the significant benefit of dramatically reducing dead time between tasks and re-dos, thus, overall shortened production times.

Naturally, there are some things that software cannot do – it can't automate complicated edits which require artistic decisions, for example. However, software should be able to – and increasingly *can* – manage your infrastructure, monitor your workflow, identify bottlenecks, simplify, and even fully automate key tasks.

There are other advantages beyond repeatability and automation. One key aspect of software-based solutions is that they tend to be very flexible.

A unique nature of multi-platform delivery is that it changes rapidly –providers have been known to change their specification every few weeks! For this reason, software-based solutions are almost always better poised to address these changes than pure hardware solutions.

Increasingly, software vendors are responding to market demands by tackling many of these challenges, both from the traditional perspective and increasingly from the workflow management perspective.

What specific problems are being solved today? Where are software solutions heading? To answer that, let's look at the challenges one at a time.

Flexible Support for Wide-Ranging, Ever-Changing File Formats

To meet the challenge of complex and changing file format requirements, the makers of file-based transcoding solutions including FlipFactory, Episode, and Compressor are continually updating their format support to include the latest requirements for new playback platforms.

In addition, many of these solutions can automatically and effectively apply rules for tasks including cropping, re-sizing, and inverse telecine, removing manual labor from the process and reducing human errors, shortening time to delivery.

Devices You Share, Upgrade, Configure

Ever-changing format requirements have also driven hardware designers to integrate highly-flexible and configurable embedded software components directly in the product. In the past, capture cards with a fixed, single file format were dedicated fixtures in an edit system.

Today, flexibility is king. Innovative products such as Pipeline support multiple formats, and offer software updates to ensure always-current format support.

Automating Task Routing

Some products on the market today are dedicated to solving tasks *within* workflows. For example, software now exists which can *sniff* user-generated content, mining each file for file and format characteristics, and metadata. This data is extracted and analyzed by the workflow automation rules configured for this workflow.

The extracted data is used to validate (or reject) media with specific metrics, and can also be used to automatically vector the media through a prescribed path or set of steps among all possible steps, based on specific file/format characteristics, and metadata values present.

Automating Visual Processing

Other technology is being developed which speeds – or even fully automates – complex, visual processing tasks. For graphics and branding workflows, solutions such as GraphicsFactory include operator-driven design tools to create templates. These templates, along with scripts, allow the simple application of logos, bumpers, trailers, and other branding tasks on incoming spots, for example.

The upshot: some branding workflows can be moved from an edit bay to a fully-automated workflow, as a single step.

Beyond graphics, other components can now automatically detect – and remove – black from source material, and automatically perform quality monitoring by identifying characteristics including macro-blocking and interlacing artifacts.

WORKFLOW AUTOMATION IS KEY

As powerful and beneficial as these tools are, these time-saving components can not realize their full potential in the absence of an automated workflow system. The arena of automated workflow is where the truly exciting software work is being conducted, and is key to the next big breakthrough in the video industry. A variety of systems are currently available, and interesting trends are looming on the horizon.

MAM Systems

Metadata requirements have traditionally been handled by MAM systems, which excel at connecting video assets with metadata tags. MAM systems often include plug-ins to permit metadata transformation. However, keeping these plug-ins up-to-date with provider requirements is a challenge unless an open standard is utilized.

MAM system capabilities typically address only some of the traditional single-file challenges. What about workflow management? A developing area in MAM systems – and across the industry in general – is recognition of the requirement for a more robust set of workflow management tools – tools which extend beyond the needs of simple asset management.

MAM systems vendors are developing increasingly-sophisticated workflow tools in their solutions. With roots in traditional asset management – cataloging, storing, finding, and using media – many workflows are a natural extension of these systems' strengths. These systems also excel at solving highly-complicated asset management problems.

Recognizing the real value in video workflows, however, MAM systems are starting to address transforming workflows; not just storing and moving assets, but changing, combining, and creating them.

While this is new ground for many MAM vendors, they are actively pursuing this path. Nonetheless, there is still work to be done by MAM companies to develop the tools specific to video workflows.

In contrast, video transcoding solutions have long offered technology that addresses problems specific to video, but lack the workflow management depth of MAM systems.

Transcoding Systems

Transcoding solutions like FlipFactory typically include a workflow automation layer, enabling full or partial automation of some workflows. These solutions not only automate file-based transcoding, packaging, and delivery, but also manage their internal storage to ensure that they do not fail due to lack of storage space or storage capacity failures.

These vendors are also beginning to implement features to reach further into other aspects of workflows, including managing graphics applications and user-generated-content validation.

Even with these advances, there is much work to be done before transcoding solutions can address the broader challenges of global enterprise storage, network management, and bottleneck identification, to name a few.

MXF: a Good Start

MXF, as a standard, has opened new opportunities for workflow management by allowing rules – and metadata – to be carried within files. In and of itself, MXF does not offer any workflow advantage; vendors typically implement different portions of the MXF standard, often creating files that are not compatible across different systems.

However, when a single MXF standard is developed and used by a customer, it can become a specification for the workflow that the customer wants to automate. This *one document* approach can be very powerful when vendors agree to support it, and interesting workflow opportunities develop.

Still, customers must be wary of relying upon custom solutions that implement today's workflow when considering how to expand in the future.

And, in fact, this typifies today's workflow management software solutions. Whether involving consultant, or vendors, or both – customers still end up with the responsibility and cost of building of their own solutions by integrating many different parts. The value of providing out-of-the-box workflow management solutions is a remarkable opportunity – and challenge – for all vendors.

Will there come a day when file-based workflow management tools can solve video workflow automation problems out-of-the-box? The answer, of course, is not *every* workflow. That said, there's a move toward highly-configurable, out-of-the-box workflow management systems that *will* solve most workflows.

THE FUTURE LOOKS AUTOMATED

Where today's workflow management tools may not excel at video and today's video software may not excel at workflow management, solutions are on the way. A new generation of *video workflow management* tools are being developed that promise to merge both workflow tools and video tools into a system that can take advantage of a variety of components. In effect, these tools provide a model of an automation layer on top of the task layer.

These systems will come in a variety of flavors, but most immediately will likely evolve through tighter integration between MAM, transcoding, and other software solutions. With these systems in place, supporting flexible workflow configurations and the integration of 3rd-party components and systems, the ability to automate file-based video workflows will largely become a question of picking the right software components and hardware.

Of course, we aren't there yet. There are operational problems that need to be solved today. How does one choose which software vendors to work with? Is there a standard that can help future-proof software solutions within the next generation of workflow automation layer solutions?

I think that most software vendors would agree upon one thing: adding future-proof interoperability acronyms such as SOA and WSDL to a marketing datasheet – and product – is easy. This is no secret. It is fairly easy to wrap software solutions in an interoperability layer when the time is right. Because of this, choosing a software vendor based upon a supported standard is almost certainly the wrong decision criteria.

Whether a vendor supports a given standard is less important than whether they can – and will – adapt to your needs. To measure this, it may be worth looking at how willing the vendor is to work with you as your workflow grows.

Given a software vendor, one indication of adaptability is how well integrated with other vendors they are. The more third-party products a system integrates with, the more likely that system will continue to be integrated and inter-operable in the future.

With that said, the future is looking bright for video workflow management solutions. Custom systems today are already being developed that remove some of the most tedious – and fragile – manual aspects of video workflows. Out-of-the-box, video workflow management platforms are not far in the future.

And when they're established, a whole new set of marketing opportunities will emerge, allowing multi-platform distribution, customization, and mass content delivery in ways consumers only dream about today.

WORKFLOW IN THE NEW MEDIA LANDSCAPE: PRODUCTION,POST AND FINAL DELIVERY

MICHAEL CASTRO
National Geographic Television
Washington DC

The National Geographic Society (NGS) was created one hundred and twenty years ago. In 2007, NGS created National Geographic Global Media (NGGM), in service to the Society's mission to inspire people to care about the planet. NGGM is comprised of National Geographic's renowned magazine and book publishing divisions, National Geographic Entertainment (NGE), National Geographic Digital Media (NGDM) and National Geographic Television (NGT).

NGT has been producing documentaries for more than 40 years, has won merely 130 Emmy awards, and is the producer of one of the longest running and most acclaimed series on television, *National Geographic Explorer*. NGT was established to expose the most extraordinary explorers and scientists to the largest audience possible, which in the 1960s was broadcast television. Today the Society distributes content through a variety of media platforms such as NationalGeographic.com and National Geographic Channel and the Society has created a Digital Studio to handle production and distribution of short form programming.

Starting in 2004, NGT made the decision to completely switch to High-Definition (HD). Since the beginning, explorers have always adapted innovative technologies to document their scientific study and geographic findings. This decision was made to future proof National Geographic content for potential distribution platforms even before there was any mention of broadband distribution on the web or programs being sent to your handheld. The plant was retooled from cameras to editing suites and all programming from that point forward was delivered in HD with 5.1 surround even when everyone was still sparring over HD formats. During the change over, NGT decided to separate the HD costs during program budgeting to track the difference between Standard Definition (SD) and HD production. At the time, the costs were tracking at about 30% above SD production. Costs have come down considerably year to year. But even with the additional costs, the decision has paid off and has secured National Geographic's content for the future.

Currently, NGT produces over one hundred hours of HD documentaries a year. We have a high shooting ratio so that 100 hours translates to over six thousand hours of High Definition video shot each year since 2005. Most of it is beautifully composed pictures of wild life, nature, man made wonders, world cultures and most recently conservation issues.

Aside from these assets, there are also thousands of hours of existing film. With all this content available, National Geographic is uniquely positioned to take advantage of all new distribution platforms as they became available. As our TV crews scoured the globe to shoot the stories for our programs, the additional footage not used in the documentaries becomes an asset for all the other divisions. NGT is a primary content supplier for all the other outlets of our Global Media group:

- NG Channel NGC (2/3 owned by FOX)
- NG Channels International (NGCI)
- NGDM
- NGT International (NGTI)
- NG Giant Screen Films
- NG Home Entertainment
- NG Digital Motion (Footage Sales)

We repurpose all of our content.

The challenge is how do you make these thousands of hours available to all the other divisions; make it easily searchable and available? Or how do you feed the beast successfully and still deliver your programs on time and budget?

Building a production strategy

Nowadays a content supplier, whether part of a major organization or an independent production company, is not just delivering a program anymore, they are planning a production with multiple requirements. For example: Domestic and International show masters, associated short form internet and mobile content, bonus material for the home video DVD, possibly up converting some of the material to be used in features and providing material for potential stock footage sales.

Whether it's for our clients or for our internal use, at the end of the day for the front end content supplier, it is all a deliverable. What's most important is to keep in mind all the deliverables up front. So you can build a strategy for each production.

Here are some of the things we keep in mind while planning one of our documentaries:

- How will it be distributed?
  1. Broadcast & Satellite - contracted production.
  2. VOD -often requiring different formatting & scenes.
  3. Short form web distribution - often with a different focus than the program being produced.
  4. Home Video DVD - usually needing bonus material.
  5. HD Footage Sales - generic beauty & "money shots" that other 3rd parties will be interested in purchasing.
  6. International vs. Domestic focus – planning for this opens the opportunity to shoot alternate scenes. For example, interviewing a variety of International experts on a subject in order to appeal to a worldwide audience.

- What does the client require?

  7. Technical Specifications
     a. Video Formats
     b. Audio requirements
  8. Delivery specifications
     a. Domestic delivery
        i. Formatted (for commercial breaks). Usually a shorter running time than international deliveries.
        ii. Imperial measurements
        iii. Audio masters
        iv. CGI masters
     b. International delivery
        i. Clean, Seamless copy
        ii. Usually re-edited for longer running times.
        iii. Metric measurements
  9. Promotional deliverables
     a. Background material.
     b. Behind the scenes material.
     c. High resolution still pictures for print promotion (2k and 4k image cameras will make this much easier).
     d. Producer or scientist blog to promote the program or provide additional insights at broadcast.
  10. Rights and Clearances (we'll get into this later).

- What does your own group need to complete the production?

  11. Shooting back plates for CGI material.
  12. Capturing generic scenes and natural audio for transitional purposes.

All of this is built into our planning process for a production. From all of this planning a budget and production schedule is devised.

But before shooting begins, there are a few more factors to consider:

Choosing a primary format. Starting with the right format will solve numerous problems later. However, this is not always possible so if you need to include another HD or SD format, know the pitfalls of it before going in.

Know your post process BEFORE you shoot a single frame! Know how it's going to be finished, on what platform it will be edited and what your final finishing format will be. In this new landscape of file formats and diversified distribution platforms, relying on post to fix any production problems is not only risky but is tougher than ever.

Planning like this will make the whole process much smoother.

Managing Communication

While this sounds obvious, it's one of the hardest, most difficult things to manage because with communication comes expectation and often times expectations are misunderstood.

One of the ways we keep communication active is to keep our other divisions informed of what were up to on a regular basis. We have monthly meetings that go through every production we're currently working on. We also publish a Pipeline Report which is distributed to all NG areas monthly so they will be aware of what is being worked on.

We also keep the Global Media groups informed of where in the world we are shooting and for how long. This gives the different groups the opportunity to weigh in and ask for additional footage that might be on their shopping list. For example, if we're shooting in India at the Taj Mahal, our Archives group may ask us for a beauty shot of the Taj Mahal at sunset. This gives us the ability to maximize the crew on the ground to everyone's benefit. Obviously we're under strict deadlines and budgets, so this only works when there is time for that extra effort. But it can also be a relief to a tight budget. If the requesting group is willing to pay for the shooting day to acquire the shot, then they get a beautiful shot for 1/3rd the price it would normally cost and we get a little relief on the shooting budget. It works out to everyone's advantage.

Managing technical expectations

This is particularly difficult because of all the differing video formats, resolutions rates, file formats, storage possibilities, codec's and constant technical innovations.

At National Geographic, we strive to sit down ahead of time and go through the producer's expectations, what they hope to accomplish, how they want to screen and edit their material. We often have some pretty heated debates about how to technically accomplish what they expect. Also our producers are constantly looking to incorporate the latest technology in their productions. Using the latest imaging technology is great and exciting but there also has to be a workflow path – how are we going to download the images, edit with the images and eventually archive the material?

There are more potential solutions now than ever. The key is picking the right applications and technical work flow for your particular needs.

Currently our plant is tooled for 720p production. This format was chosen because it is the broadcast standard for The National Geographic Channel, which is a joint venture between Fox and National Geographic.

However, now that HD formats have had a chance to shake out over the past few years, we are planning a transition to 1080. Also the Home Video marketplace is currently dictating 1080i or 1080p. While we could easily up convert the 720p material, we prefer to keep our HD footage in its native format.

But when it comes to shooting in the field, we end up with one primary format (either 720p or 1080i) and a few other formats such as HDV and P2. That's because these cameras tend to be smaller and cheaper. We can't risk putting a $60,000 HD camera in harms way. We lose cameras every year to a variety of factors: the elements (arctic weather or hot humid conditions), accidents (cameras dropped into the ocean or off a cliff) or predation. We lost a half dozen cameras alone last year to hungry wild animals. While our primary shooting format is Varicam or Cine Alta, our second camera, the one we sacrifice, will be smaller and less expensive.

But cameras are getting smaller, lighter, less expensive and delivering better HD pictures every day. It seems there is a new camera every 6 months with greater capabilities. For us, the smaller the better with the highest HD resolution.

Because we shoot such a large amount of footage, we don't have the ability to store our material one to one. But instead digitize it in the Unity at a lower resolution acceptable for editing and screening. We later go back and re-digitize the finished program one to one during the on-line process.

Why don't we just finish our programs in DNX 220 or even Apple's Pro Res HD? Because while these delivery formats are acceptable to some Domestic clients, they currently do not live up to the standard of National Geographic and many of our international clients expect full resolution HD.

Now that storage rates have dropped in price so dramatically and FCP has increased functionality for HD final finishing, we are looking at Final Cut Pro as a potential lower cost solution to editing, on-lining, color correction and audio mixing.

Capturing the footage & metadata

Our footage is always embargoed while the program is in production and up until all the contracted deliverables has been satisfied. After that, the actual program footage is embargoed for the length of the license term. However, the "outs" become available to all our other departments for repurposing.
This is where the metadata becomes extremely important. During the entire process of acquiring the footage and editing the program, the producers have been building the metadata associated with it. From log notes to annotated EDL's to the rights information (location & personnel agreements, contracts, stock footage used) and everything else.

The question is: where does all that information go and how do we get it to all the necessary areas for repurposing?

We have installed front end production software that captures the metadata, a proxy of the video, log notes, rights information and gives the producer the ability to string together scenes creating a rough EDL.

All this metadata eventually transfers to the Digital Archive which becomes available to all other areas of National Geographic through their Digital Asset Management system .

A number of years ago, our Digital Archives group installed a MAM system to store all the video content for the Global Media group and the Society in general. Besides the thousands of hours arriving from National Geographic TV each year, they also have to contend with:

- Short form (or web) file based content. A relatively new addition considering the age of the Archives.
- Large screen content from our IMAX division.

- Content from our footage partners (other footage libraries represented by National Geographic).
- The Society's historical video record, which includes original video from explorations, lectures, events, scientific discoveries, critter cam and others.

Content is prioritized and ingested based on content gap analysis data as well as a review of the program, availability of physical elements, and a basic rights snapshot.

MAM streamlines the process for the research and retrieval of assets for various distribution needs. The Digital Archive group created an NG-specific cataloguing style guide that outlines in detail the "rules" for cataloguing. Basic and broad keywords and categories were standardized and a common language, style and hierarchy are developed for each clip and video asset (outtake, program master, species, location, etc).

This system is available internally through our IT network and on the web for external users.

Rights and Clearances

Unless you are a news organization, getting the proper rights and clearances is one of the most critical aspects of a production and one that is often overlooked. It's great to get that perfect shot, but if the proper clearances haven't been secured, you've just wasted your time and money while increasing your anxiety.

Nowadays, with consumer generated content through web video and broadband distribution, plus the proliferation of reality shows where everyone is a star, most people are pretty savvy about using their image, their location (house, office, land, etc), or video they may own.

It used to be a lot easier, securing rights only for your program and its contents. Now rights contracts and clearance agreements have to be broad and all encompassing. Rights need to be secured in any manner, in all media currently known or subsequently devised, in perpetuity, worldwide. This guarantees that you can use the material in all platforms, for whatever length of time. You are securing the rights not only for broadcast distribution but for internet, broadband, satellite, Home Video, DVD, mobile distribution and future platforms. But securing unlimited broad rights can be costly.

Besides securing the rights to distribute the material, it is important to determine *how* the material is going to be used…for the program itself, for promotion of the program, for promotion of your channel or brand and any other possible uses now and in the future that can be thought of.

All this information needs to be linked back to the original content, so retrieving a clip, shot or scene also contains the necessary rights information.

It's imperative to have all the legal documents in order. Making sure there are releases for everyone who appears in the program, location releases for all locations that appear and contracts for all the non-original video.

While this is labor intensive and takes up a lot of time during the production, it pays off later when other areas want to use a clip or scene from a previous production.

In Conclusion

While everything mentioned here is necessary in this new hyper video landscape; planning your production considering all possible outlets, working through the challenges and technical aspects of production, obtaining the rights and clearances to your material, it's important to remember the primary goal: to produce an interesting, entertaining, honest and sometimes important program that will inform and sometimes inspire people to take action.

At National Geographic, that brings us back to our mission of caring about the planet. But content also needs to be future proofed in preparation for potentially new uses. You never know what will come up next. A good example of this is, we are now looking at ways to incorporate 3D production into some of our shooting. As the head of our Digital Media group has mentioned to me many times, "don't throw anything away" because in the end, it's all going to be usable.

Local and Ultra Local Content in Broadcast Mobile TV

Richard LHERMITTE
ENENSYS Technologies
France

ABSTRACT

Broadcast Mobile TV is meant to be one of the next 'killer application' for Mobile operators and content providers.

After in depth research, consulting groups announce that MobileTV Broadcast will create successful services in the coming years.
Broadcast MobileTV is already covered by different standard and has been launched using them all (DVB-H, T-DMB, FLO).
These deployments and trails are using existing national services (main popular channels). User feedbacks analyze shows that there is a real need for local contents like regional channels. Ultra local content, like specific mall/area live services or stadium dedicated content has been pointed out as a logical development.

The aim of this paper is to identify the different type of Mobile TV contents (National, Local and Ultra local), in order to check whether Local Mobile TV Services is feasible looking at possible business models and technical aspects.

Indeed, Ultra local content means that there are many contents to be handled at once, hence the way they are managed must be different than National and Local one whether be it on the business or on the technical side.

MOBILE TV

Mobile TV already exists. Most of the time, the actual Mobile TV is defined, managed and controlled by Mobile Network Operator. (MNO). These MNO use their cellular network to deliver the video content. To do that, they create their own content bouquet including Live and download (VOD) content. With this actual Mobile TV business model, this is the MNO who is fully in charge of the Mobile TV service: content selection (and creation for some of them), content delivery, associated business model and end user management (billing). Since one unique and centralized actor operates, Mobile TV services are quite easy to put in place and quite fast to deploy (re-use of existing cellular network). However Video service offer uses a network and handsets which have not be design to achieve that kind of performance originally.

It ensures these main following inconvenient:

- limitation of the cellular network in terms of simultaneous supported end users and in terms of supported bandwidth per video channel
- non suitable handsets which do not offer a satisfying experience: poor video quality due to non adapted screen size and resolution, high battery consumption.

BROADCAST MOBILE TV INTRODUCTION

TV Broadcast is not new. It started in the analogue format and switched to digital approximately 12 years ago. Broadcast technology is very well adapted to unlimited number of end users. Associated to digitalization, it offers the capability to deliver many different services per used frequency.
Hence solutions and standards have been defined to offer capability to use this broadcast technology to deliver TV contents to mobile devices.
For now, multiple standards are available and have already been deployed (trial and commercial offers). The most famous ones are:

- DVB-H (Digital Videao Broadcasting – Handle)
- FLO (Forward Link Only)
- T-DMB (Terrestrial Digital Media Broadcasting)
- CMMB (Chinese Mobile Multimedia Broadcasting - under finalization)
- ...

BROADCAST MOBILE TV: EXISTING END USER FEED BACK

Multiple Mobile TV Broadcast trials have been conducted and commercial Broadcast Mobile TV services have been deployed, bringing accurate technical results and end users' feedback.

There was no major technical discovery; Broadcast technology can fit all major Mobile TV requirements, offering a nice service quality and significant user experience. In that way that End Users declared they

were willing to pay a fixed month subscription fee.
The average day watching time appeared to be around 16 minutes.

The surprise came while analyzing the day usage: Mobile TV content had been watched at home, for as much as half of its overall watching time.

The last major feedback is about content in itself: Most of the time, consumers watched best-adapted content like short news, sport and entertainment content. On top, consumers were requesting local TV content. Local and dedicated content were the two main user requests.

MOBILE TV CONTENT CONSTRAINTS

The biggest part of Mobile TV content comes from existing contents which are, most of the time, created by National live content providers for a large audience using usual TV sets (SD resolution).

The first step to take to get dedicated Mobile TV content is to adapt the existing screen size and resolution, as well as events' duration: dedicated Mobile TV content must have clear and easy to see scene (not to many details), and program duration must be reduced.
The second step is to create specific Mobile TV content, content which are made and designed for Mobile TV users only.
Content adaptation and related cost must be handled by one single actor of the Mobile TV ecosystem; it could be the content provider but also a MNO which takes the opportunity to get involved in the video content creation.

Wireless technologies use many frequencies: due to the high number of live video contents, it is hard to get available frequencies for a new service. That is why this is necessary to share one single frequency (ie the bandwidth) between all contents: National and Local. There is even more to it, since it may not be possible to broadcast all services at once, providers need to make a selection among them.
Content selection, content adaptation, content creation are the key factors to get a great Mobile TV service quality.

The business model related to the content access brings also some constraints on the services management side. When content is free to air (free to access), the service is quite easy to deploy. When content is not free (post paid or pre-paid, pay per month/week/days or pay per event), it implies content protection and user management. Content protection means content scrambling; it is generally done at the content compression stage and is linked to user management. User management implies to get a centralized user database with user right definition, attribution and content access policies but also a centralized billing system.
Paid contents require centralize content management.

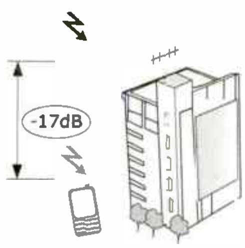
Another feedback from trial and commercial deployment indicates that users consider Mobile TV services as an extension of their usual TV. He/she would like to have the same experience as in front of his/her TV set: an easy way to access content and an easy way to switch from one content to another.
We generally attribute this ability to the handset and more precisely to its user interface, i.e. the handset capacity to give easy access to the content. The handset must also be aware of all available contents/services and make it easily available. That is what the content declaration, which is generally include in the Electronic Service Guide, does. This ESG is broadcasted in parallel of the video content.
Because we could not imagine that users need to make a specific and manual procedure to see all available contents in each region he will visit, content availability must appear automatically when users first enter a new content coverage area.

MOBILE TV DEVICES CONSTRAINTS

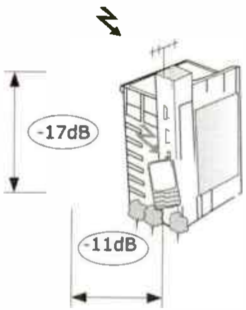
Mobile devices are quite small and light. Associated screen sizes are proportional. It means dedicated screen resolution which, in the case of Mobile TV, must be adapted. QVGA screen resolution is now largely used and could be considered as a standard resolution. In terms of video display, the screen size must not be too small.

However, using a large screen size consumes a lot of energy. Autonomy is indeed another constraint brought by mobile handset. Energy is consumed by the screen, content processing/decoding and content reception (RF and front end part).
Broadcast Mobile TV technologies have taken into account these constraints by adapting content delivery for example. DVB-H Time Slicing is one of these solutions. Time Slicing consists of time division (like TDMA) where the handset receives all content at once, during only 10% of the content total duration meaning receiving the content burst by burst. Such a solution is well adapted when all contents are managed simultaneously from one central point.

The last Mobile handset constraints is related to signal reception. Wireless broadcast technologies have been initially designed for outdoor reception using dedicated antennas (satellite dish, RF antenna). For classical UHF frequencies, the RF antenna is placed on a high point with a specific orientation to guarantee satisfying signal quality and reception. This is not the case for Mobile TV handsets which do not have any external antenna and which reception occurs at the 'ground' level. Studies demonstrated that there is 17 looses dB for the RF reception for Mobile TV handsets.



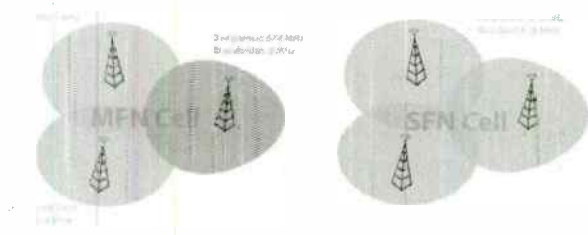
We discovered previously that half of the time consumers watch Mobile TV services at home, meaning indoor reception. And the other half of the time, the same feedback indicates that most of the time, reception was made inside a building. We can conclude that good indoor reception is a key factor for service adoption. However, reception within a building is 11 dB less that outside the same building.



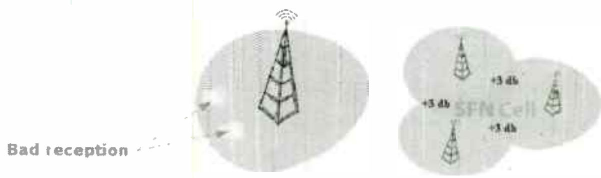
Globally, Mobile TV handsets have 28dB less in the quality signal reception when compared to a classical RF broadcast reception.
We can easily conclude that signal transmission has to be efficient with deep density to comply with these constraints.

MOBILE TV COVERAGE CONSTRAINTS

According to the previous conclusion, a good coverage of Mobile TV services is a key to service success.
An existing solution to increase broadcast signal quality and signal density is to deploy multiple Single Frequency Network (SFN). Each area/region could be covered using one frequency which is broadcasted by multiple transmitters.



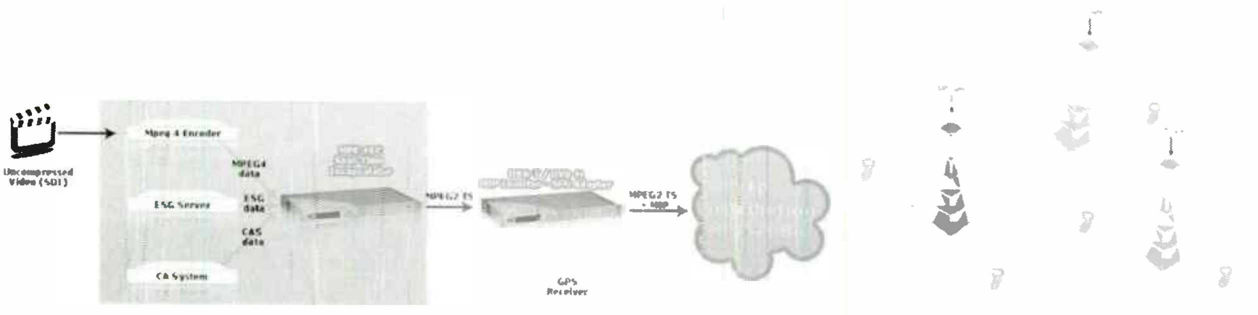
SFN also allow to guarantee reception in specific areas that one transmitter is not able to cover, like shadow area at the entrance of a building.



Transmitters belonging to the same SFN cell radiate the same RF symbols, over the same frequency, at the very same time, using the very same RF modulation. Any receiver located in the area covered by the SFN cell will receive signal from one transmitter, or another, without any difference. And moving from an area covered by one transmitter to another within the same SFN cell is transparent at the reception level.

Perfect synchronization between all transmitters within the SFN cell is feasible only by the use of a SFN adapter at the head end side.
SFN adapter is centrally located at Head End side to build a unique source that all transmitters will have to broadcast in a synchronized way.
Since SFN adapter and transmitters are not located at the same place, the only common timing reference available in the world is GPS satellites. This is why in SFN networks, SFN adapters and transmitters are connected to GPS receivers providing them with 1 PPS and 10MHz signals.

Needed to be inserted at the central head end or before delivering the stream to the transmitters, SFN synchronization implies a major constraint to content delivery and associated architectures.



MOBILE TV CONTENT DEFINITION

According to the previous section, we could easily distinguish two major kinds of contents: National content and Local content. Going more in detail and regarding specific areas where video content could be interesting to catch, another kind of content could be imagined: this is what we call "Ultra Local content".

Mobile TV National content definition

National content are associated to the most popular content which are broadcasted to the entire dedicated population. The associated coverage must be the largest. The content is based on:

- Generalist subjects,
- Dedicated popular thematic like sport, music, weathers forecast, …

Well know channels using National content are CNN, EuroSport, …

In Mobile TV context, due to limited number of channels, such channels are selected to avoid service overlapping or information redundancy.

User feedback and audience estimation provides this well known ratio: 80% of the population is watching 20% of the global content focusing on the most popular channels. That is why even if Mobile TV users were requested dedicated content, these content are the base of Mobile TV service.

Mobile TV Local content definition

The main characteristic of National content, which is broadcasting generalist programs, is also a drawback. Users request more "proximity" information.

These contents are focusing on a specific region, on a specific town or part of a state, generally associated to regional channels.

Most of the time, this content is created close to the place where they are broadcasted, i.e. inside these associated regions.

Some example are 4NBC for New York dedicated channel and TvBreiz for dedicated French Brittany region.

Due to the high number of specific areas, the number of such kind of content could be high but is also limited by the fact that the public is not big enough to enable concurrence.

According to the associated business model, this kind of contents and channels could be managed as National one.

Mobile TV Ultra Local content definition

More and more people are looking for "proximity" information related to where they are. These kinds of information, focusing on a specific area, are only available inside this specific area and could be of no importance outside. Most typical example is stadium dedicated content; this could be content related to the stadium event and events outside the stadium but of the same interest, and content you are able to receive only inside the stadium. Imagine that you are going to see a football match, immediately inside the stadium you have the possibility to watch video content that summaries the previous results of the teams, that displays the last goals, and that offers you the possibility to watch the other games which are played simultaneously in another stadium. This is Ultra Local content. And you could find such kind of content in many different areas like malls (shopping information), museums, public areas, theme parks, …

The number of such kind on contents and channels created for and inside specific areas, could be very high. This content has also the particularity to be available to everybody, independent from the National Mobile TV subscribing, whatever the receiver device is.

According to these facts, high number of content and independent content provider, such kind of content could not be managed as National or Local one and must have their own delivery mechanism. But this Ultra Local content must be easily accessible from end user point of view; the Mobile TV device must be able to display automatically the available Ultra Local content and services without specific user action (like channel scanning). Ultra local content referencing in National broadcast delivery is one of a key problematic.

MOBILE TV – NATIONAL AND LOCAL CONTENT MANAGEMENT

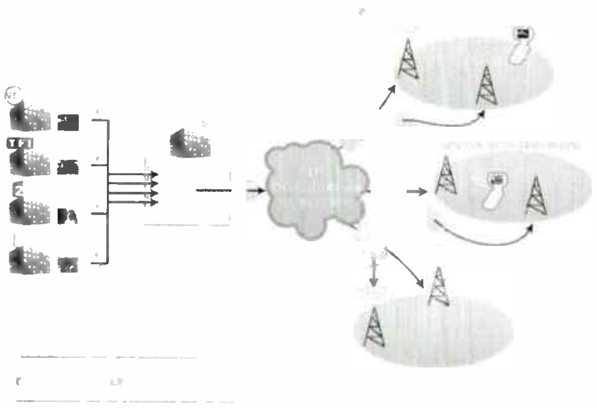
Mobile TV management could be summarized in:

- Making aggregation of all contents
- Making content repurposing
- Delivering content to all transmitter sites
- Complying with Mobile TV constraints

National and Local services management (and delivery) can be done using centralized or distributed architecture. National services are generally managed from a central head end (Centralized architecture) when Local services could be managed from regional points (Distributed architecture) or from a central point (Centralized architecture)

According to broadcast network characteristics (distribution network, SFN / MFN, …) each of these two content management has benefits and limitations.

National services management: centralized architecture

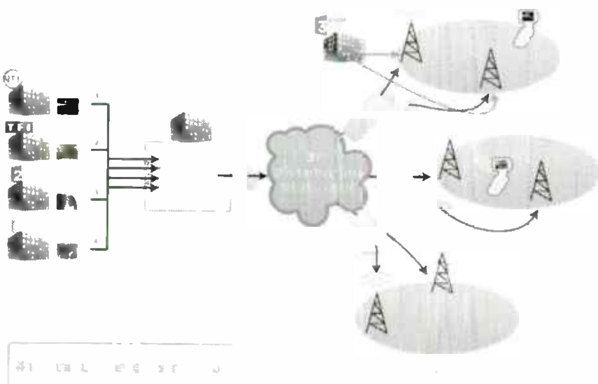


Centralized architecture implies that all contents are managed from a central point: receiving all content feeds and managing all of them simultaneously.
Main advantage is that it is the easiest way to manage Electronic Service Guide (ensuring the declaration and referencing of all contents) and one of the best way to control Users content access and billing, achieving a secure CAS deployment. This architecture also offers the easiest and more powerful capability to optimize the bandwidth allocation using Statistical coding for example. And off course, this is the most efficient way to guarantee the full compliance with Mobile TV constraints.

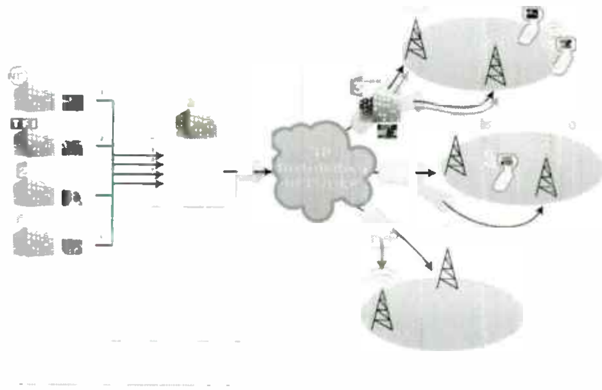
Local services management: distributed architecture

There are two ways to make Local content distributed architecture.

Distributed architecture #1 is the most distributed one; Local content, which is created in one dedicated local area, is inserted at the last point of the distribution network: the transmitter site.



Distributed architecture #2 is characterized by the fact that local content in inserted at one point, just before the local area global content distribution.

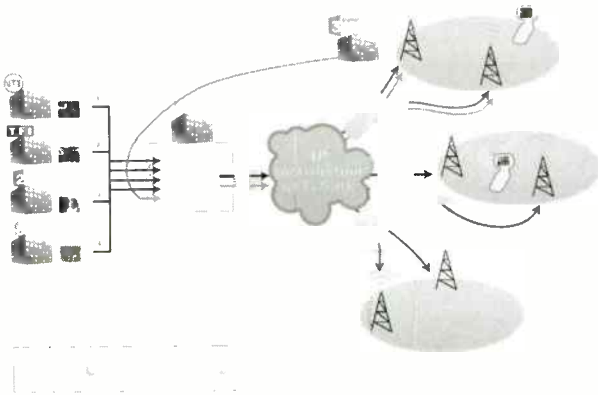


Distributed architecture #1 is not quite valid due the fact that is almost impossible to preserve SFN. Trying to make such architecture could impact the overall area transmission making lots of perturbation. Also, considering the needs of one "specialized and complex box" which makes local content reception and insertion at each transmitter site, the cost of deployment for such architecture is very high.
Architecture #2 complies with SFN constraint and do not implies any specific box at any transmitter site. But it implies to have a dedicated distributed network for each region. For example, for money reason, it is not reasonable to have one satellite distribution per local area. Cable network is quite mandatory between each transmitter site considering this architecture.

These two architectures enable each region to be quite independent and each regional content provider to manage directly its equipments, its content, its advertising. its program grid, ...
However, one main issue is that user management and CAS deployment are made difficult since there are generally centralized.

Sharing the same frequency (mux) for National and Local content, both architectures imply a modification of already created National multiplex, the insertion of Local content and the modification of the ESG. These cases are not the best ones to guarantee the best quality of services, and would require a more complex supervising and monitoring system.

Local services management: centralized architecture



This latest Local content management architecture enables the management of Local content like National one, including full compliance with Mobile TV constraints, a nice possibility of bandwidth allocation optimization (Stat Mux) and the simplest way to manage the ESG and CAS. Centralized ESG also guarantees a full description and referencing of all contents. And this architecture allows an easy way to handle multiple Mobile Network Operators for same content and service.
On the other hand, this architecture has the disadvantage to consume more distribution network bandwidth as all content (National + Local) are broadcasted for the central point to all regional points and transmitters. As for Distributed architecture #1, this architecture requires a dedicated product in front of all transmitters. Product which has to make the Local content selection according to the area where it is located.
A way to solve these technical issues is to optimize the centralized head end by putting the maximum of 'intelligence' at this point. For example, there is no need for extra Forward Error Correction on the national distribution network; also, SFN synchronization is be done for all regions simultaneously, ESG is computed simultaneously for all regions as well as all require 'references tables'. This permits to guarantee all Mobile TV constraints from a central point, to verify it easily, as well as deploying Transmitter front end products which are less 'intelligent' (require less processing power) having an as lower as possible price for this equipment.

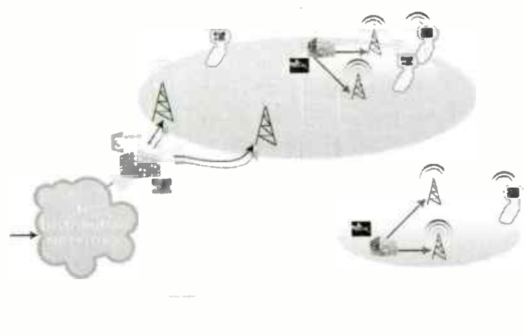
In term of Business model, this architecture does not allow Local content provider to be independent. Local services could be monopolized by one Operator meaning that Local service will not have the possibility to be received by all users.

MOBILE TV – ULTRA LOCAL CONTENT MANAGEMENT

Ultra Local content delivery brings some other issues. Due to its potential high number, global management could be difficult and specific distribution could be complex. As explained before, Ultra Local content must be independent from the Mobil Network Operator. That is why this content could not:

- share the 'main' frequency/network/mux which is linked to an operator.
- be transported by a Mobile Operator network.

Consequence is that Ultra Local content and associated provider must have the possibility to deliver/broadcast their content independently from existing contents/services. This could be done only by using a specific frequency and/or using another technology. It could be feasible that National and Local contents are broadcasted using technology A and Ultra Local content using technology A or technology B.



When technologies are different, the problematic is at the receiver level: having the capability to receive content from multiple ways using multiple technologies / standards on potential different frequencies bands.
For global content delivery, the problematic is at the referencing level. We could not imagine end users having to manually 'scan and search' for new content each time he will enter on a specific area; even if there is panels which indicate that dedicated content is available in this area. We have to find a way to reference content which is not available everywhere and is different on each area. It's another point that forces to not have a centralized management of the Ultra Local content

Such architecture implies that Mobile TV service providers will not be able to validate all the delivered content to its end users and must accept end user terminal to display other services than its selected offer.
In the same philosophy, Ultra Local content could be protected by Conditional Access and ultra local user management system must be available to provide the

users' rights to access to it. For example, entering in a stadium will only allow stadium spectators to watch the dedicated content; in this case, scratch card could be used to provide access code; the CAS must support it and Mobile handset too.

Now, Ultra Local content services could be envisaged. Technical aspects must be clarified and must be found to allow a nice user experience:

- content broadcasting on separate frequency and / or using a different technology
- content referencing in a global way for an easy access
- content protection with 'universal' CAS mechanism (no proprietary solution)
- end user management for content access located to specific area.

CONCLUSION

Mobile TV users request Local content. However, Broadcast Mobile TV technologies bring constraints to services deployment that could enter in conflict with this local content delivery. National and Local services deployment and management can be achieved with special features to comply with Broadcast Mobile TV constraints using centralized architecture or 'special' distributed architecture.

Ultra local contents must be independent from other contents/services; it implies that Ultra local contents must have their own deployment architecture. And independent architecture goes with the problematic to have easy access to this content and content protection.

When FM radios were deployed, it permits to have lot of independent and ultra local audio services. An ideal situation could be to have the same feasibility for video content. Nevertheless, now, Broadcast technologies do not offer the same flexibility; and business model for video are more complex. Adding some advance features to existing one could achieve, in the future, this goal.

THE ROLE OF FILE-BASED WORKFLOWS AND METADATA IN REPURPOSING CONTENT FOR THE WEB AND MOBILE TV

Mark Bishop and Alex Holtz
Integrated Production Solutions, Thomson Grass Valley
Jacksonville, Florida

ABSTRACT

With many organizations recognizing that the majority of their future revenue growth will come from alternative ways to deliver their valuable content, file-based workflows can provide the ability to customize content for multiple delivery streams through a Create Once Publish Everywhere (COPE) framework. Leveraging static and temporal metadata with an integrated system from preproduction through to publication, media outlets will transform themselves from traditional broadcasters to multi-distribution content publishers — a key to securing their economic future.

MORE THAN 31 FLAVORS

Growing up I loved visiting Baskin Robbins and choosing from what seemed like an endless array of flavors. Once I made my choice of ice cream, cone and sprinkles, I walked out with a big smile because I got exactly the treat I wanted. When Burt Baskin and Irv Robbins first created their franchise model around a flavor for every day of the month, they probably never anticipated introducing more than 1,000 unique and delicious ice cream flavors.[1]

The growing challenge facing many content providers today is in many ways similar to the demanding kid in an ice cream store: each client wants content their own way and for good reason. The diversity of devices accessing content is growing every day from desktop, laptop, set-top box, mobile phones, iPods, PDAs and other Internet appliances. The rendering and display requirements vary significantly from device to device including composition requisites for varying advertising formats and auxiliary data that may or may not accompany the video content. Finally the network services for target markets can vary greatly from 56Kbps on some mobile devices to more than 20Mbps in the latest triple play IPTV networks.

These demands require content providers to tailor their assets with specific encoding formats, resolutions and bitrates needed by the market. However developing special content for every combination of device, network and application is next to impossible. Not only is this approach expensive and very time consuming, but when combined with a manual process, it can often lead to inconsistent results in the final product.

CREATE ONCE, PUBLISH EVERYWHERE

A properly designed multimedia content repurposing system solves the problem by taking the master content and automatically conforming it to suit the needs of the target platform. With the Create Once Publish Everywhere (COPE) framework, only a single copy of the content in its original form is maintained and the system can repurpose it to fit the desired scenario in an automated fashion.

Using COPE, the traditional and well-understood workflow that creates the primary media for television distribution doesn't need to change. The workflow is simply supplemented through the insertion of metadata that represents the specific areas of content that are going to be repurposed in one or more ways. Users create associations not only with the audio and video essence, but also with descriptive metadata — the who, what, when, and where.

Once proper metadata tags are inserted within a piece of content, it is the publication and distribution medium itself which determines how those tags are used. Directives can include re-editing and re-framing instructions to extract exactly the desired piece of material from the master show file. More sophisticated instructions may include replacing broadcast style graphics with something more suited to the target device. Finally the tags should specify the exact transrating or transcoding requirements needed by the client.

MUCH ABOUT METADATA

There is a lot of talk in the industry about metadata, and most of it revolves around the static description of content. It is well understood and accepted by those who work with large volumes of material, that the actual Audio/Video essence is of little use without a detailed description that enables the content to be indexed, stored, searched and retrieved as needed. Because the detailed description of content is so important, standardized means of formatting and transferring the data is key to its usefulness. One such standard that has helped different systems in the production pipeline effectively communicate is the SMPTE 377M Material Exchange Format (MXF)[2], although it has not gained much acceptance outside the production environment. The standard more widely used for multi-distribution on the internet has been MRSS[3], although it too has limited capability in describing temporal events.

Figure 1: Types of Metadata

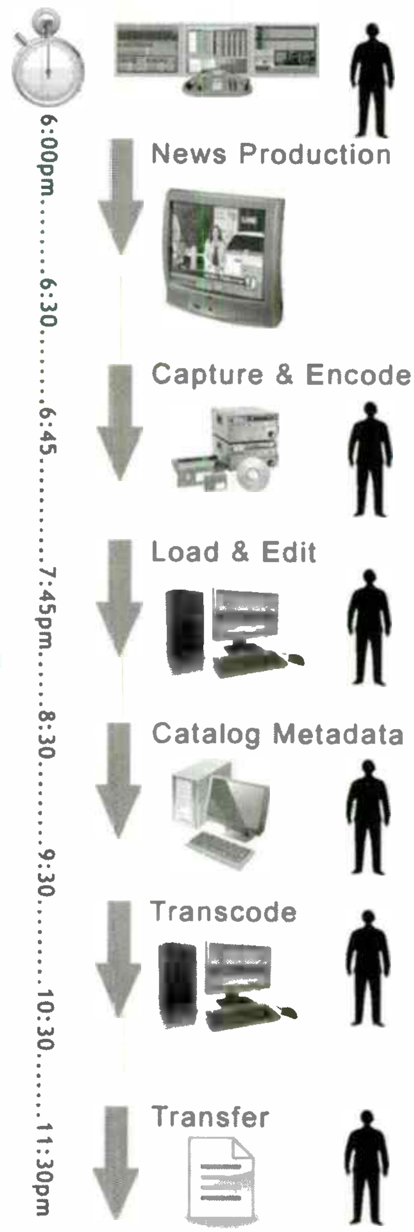
Static Metadata	Title, Subtitle, Rating, Credits, Copyright, Air Date/Time, Duration, Categories, Keywords
Temporal Metadata	Descriptive Scene Changes, Key Frames, Segment Timing, Script Text, Closed Captions, Timecode, Synchronous Auxiliary Events (Web URL, RSS, Surveys).
Distribution Metadata	Syndication Points, Expiration Date, Associated Content, DRM, Target Resolution, Bitrate, Codec

The need for describing temporal events that are synchronized with the source content is growing. Traditional examples of synchronized events in video include closed captions, Linear Timecode and cue tones. New media applications of temporal metadata include descriptive scene changes with key frame images that can help the user index to exactly the desired part of the content. For example, a tennis match could be delivered with every point, game and set identified. Temporal events also add great value to content delivered on web sites that can synchronize banner ads, RSS feeds, surveys and other web content to the source video. To help describe the interaction between audio, video, graphics, text and other elements, the W3C developed the Synchronized Multimedia Integration Language (SMIL) as a standard way to specify temporal events.[4]

FROM CAPTURE TO CONTENT DISTRIBUTION – IN HOURS OR MINUTES?

The biggest challenge in preparing content for distribution is dealing with raw source material that has little or no temporal metadata associated with it. Examples include live news, talk shows, sporting events and other dynamic sources that by their nature cannot follow a rigid timing sequence. Although there are automated tools to detect scene changes in raw video, the actual produced segments typically have many such transitions as part of the content itself.

Figure 2: Time Study of a Traditional Content Repurposing Workflow



Therefore the work becomes differentiating between basic scene changes in the source content and the actual start and end to a desired segment.

The traditional workflow to get unstructured content prepared for multi-distribution is laborious and very time consuming as illustrated in Figure 2. Once the live feed is captured, the process begins:

1. Load & Edit: The master show file is loaded into a Non-Linear Editor like Edius or Final Cut. Edits are made to optimize the content for target devices including logo insertion and aspect ratio corrections.
2. Head & Tail: The start and end of each segment is searched, marked and exported into cut files.
3. Thumbnails: Keyframes are searched and thumbnails generated to visually describe the content.
4. Metadata: Once the content is created, it must be described as needed for the distribution platform.
5. Bookends and Ads: To generate revenue, the post-production process typically includes inserting promotions in the front, middle or end of segments targeted for web or mobile applications.
6. Transcode: Each publication point typically has very specific requirements to the target codec (H.264, Flash, MPEG2), transport wrapper (3GP, QuickTime), resolution (CIF, D1, 1920x1080) and bitrate (125Kbps to 20Mbps).
7. Transfer: Once the final content file is created, it must be transferred, typically by FTP, along with the associated metadata and content.

The result of the traditional workflow is a show that has been cut into all the necessary segment files after hours if not days of very tedious work.

ENTER AUTOMATED CONTENT PRODUCTION

In order to dramatically improve the efficiency of producing live content, news departments are implementing sophisticated control room automation systems like Ignite.[5] The automation system can control virtually everything needed for news production including robotic cameras, video switchers, audio switchers, CGs, video servers and closed caption encoders all at the hands of a single technical director. The results are a tighter and more consistent newscast with all the production elements and features of the original product. The benefit for repurposing news content is that the automation system becomes the engine that drives the efficient creation of content with time accurate results.

Figure 3: A Single Technical Director Controls The News Production With The Ignite Control Room Automation System.

PRE-PLANNING PREVENTS POST-PERFORMANCE

Many newscasters today spend most of their time preparing and repurposing content in the postproduction stage. Not only is content frequently segmented and edited by hand in a non-linear editor, but metadata is typically entered after the content is created. The problem with this workflow is that time-consuming tasks must be performed to time-critical news segments before they can be distributed to the viewer. Also extensive postproduction requirements effectively limit the quantity of content published for web and mobile devices and therefore limiting advertising availabilities to generate more needed revenue.

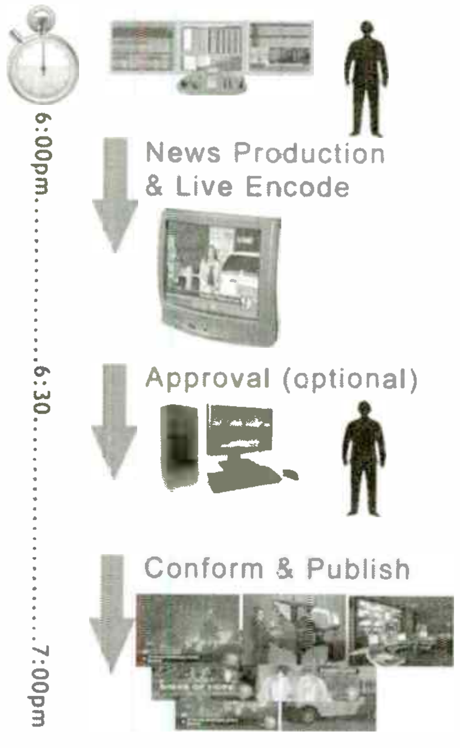
Instead of staffing up the postproduction process to repurpose content faster by brute force, it makes sense to complete some of the tasks before the newscast is even shot. The preproduction of live news in television stations is typically handled by a Newsroom Control System (NRCS) like iNews or ENPS, where journalists enter their stories and associate content as needed. By adding a special markup tool specific for repurposing content, the journalist or web producer can easily use their existing NRCS system to specify the static, temporal and distribution metadata needed in the production process.

CREATE ONCE - TIMING IS EVERYTHING

Once a news show is marked up with all the necessary production metadata, the control room automation system can import the show from the NRCS and run it with time accurate results. As the show is running, the uncompressed audio and video is captured and encoded into the high resolution master show file needed for repurposing the content in postproduction.

The static and distribution metadata entered in the preproduction process for the show and each segment can be reviewed and ultimately carried through to postproduction for a seamless workflow. However large efficiencies in work is gained by the addition of accurate temporal metadata inserted into the workflow by the control room automation system. The start and end of each segment is registered as those rundown elements are executed by the automation system. Temporal events within each segment are also accurately recorded with the desired URL, RSS or survey specified up front by the web producer. The result is one copy of the content stored in a master file with all the static, distribution and temporal metadata to accurately and automatically repurpose the content.

Figure 4: Timely And Efficient Production And Multi-Distribution In A COPE Workflow.



OPTIONAL POST-PRODUCTION ?!

By moving the metadata tasks into the preproduction stage and creating a time accurate Edit Decision List (EDL) during production with a control room automation system, the necessary work in post production is virtually eliminated. In fact content that is flagged for automatic distribution could skip any review step altogether and get sent out within minutes of creation. Think of the possibilities this type of enhanced workflow can provide: timely weather, traffic and late breaking news can actually be meaningful to the viewer.

The postproduction process in the COPE framework is essentially reduced from a traditionally labor intensive task to one of optional review and approval as shown in Figure 4. The mark in and out of each segment can be inspected and adjusted as needed. Also the thumbnail associated with the content can be updated to best reflect the desired subject. In practice these minor tasks, along with a quick review of the relevant metadata, can be done for a whole show within minutes given the right tools and workflow.

CONFORM AND PUBLISH EVERYWHERE

As soon as a segment is either automatically or manually approved, the distribution and temporal metadata associated with the content determine the final steps. The distribution metadata includes specific information to the resolution, bitrate and codec needed by each publication point. By applying the time accurate EDL to the master show file, the conformance step can automatically generate content that is ready to publish and satisfy the specific requirements of the distribution platform.

Transferring the content to each Syndication Service, Application Service Provider or Content Distribution Network can also be automated from the distribution metadata. A single segment may be published to one or more clients as originally specified in preproduction or augmented in postproduction to meet last minute demands. Each publication point references the actual transfer method and credentials needed to efficiently get the finished content, metadata and thumbnails to where they need to be. The actual transfer methods can range from a basic FTP site to transactional APIs that can efficiently process the content and extended metadata. Higher level publication interfaces can also execute content management transactions for expiration, recall and republishing content as needed.

Once the content has been successfully transferred, the many derivative files created for each publication point can be safely purged from the local production system. However the master show file and all the associated static, temporal and distribution metadata should be archived for future use. For example if a piece of distributed content needs to be recalled or replaced, the master show file can easily be used once again to update the metadata and generate the needed files for the client.

CONCLUSIONS

In live news where time is of the essence, real time traffic, breaking news, weather, financial data and sports scores can all be published quickly and efficiently leveraging an automated single workflow; multi-distribution process. Broadcasters have important and valuable national and local video based content that is in high demand for advertisers seeking to create campaigns and awareness with precision to a target audience on web and mobile devices. This same requisite will follow through to the on-demand world for digital television. Today, advertisers have plenty of inventory but not enough relevant and available content to satisfy the demand. Leveraging automated workflows provides a cost effective means to produce significantly more content, creating more advertising availabilities while keeping costs and timeliness under control. Broadcasters can now realize the "full" value of their content on multiple platforms and distribution end points to maximize their return on investment. So the next time a customer comes to you for content their way, can you satisfy the demand?

ACKNOWLEDGEMENTS AND REFERENCES

[1] http://www.baskinrobbins.com/About/OurHistory.aspx

[2] http://www.smpte.org/standards

[3] http://search.yahoo.com/mrss

[4] http://www.w3.org/AudioVideo/

[5] www.thomsongrassvalley.com/products_broadcast/

ABOUT THE AUTHORS

Mark Bishop manages product development for the Integrated Production Solutions (IPS) business unit of Thomson Grass Valley that produces the Ignite Control Room Automation System, the Ignite Robotic Cameras and the MediaFUSE Content Repurposing and Multi-Distribution System. Mark has over 20 years of experience developing new products in television broadcast automation, IPTV, DTV and digital imaging. Mark holds a Bachelor of Science degree in Computer and Electrical Engineering from Purdue University and an MBA from the University of Florida.

Alex Holtz is the managing director for the IPS business unit of Thomson Grass Valley and has been a true pioneer in the fields of News Control Room automation, robotic cameras, live content production and multi-distribution systems for over-the-air, cable, web and mobile devices. He originally worked on many of the methods needed for file based repurposing and multi-distribution in the mid to late 1990's and has been granted 6 patents with another 12 pending. Alex has over 30 years of experience developing new products and application innovations and holds a Bachelor of Technology degree from the University of North Florida.

LEVERAGING FEC ADVANCES TO OPTIMIZE DVB-H NETWORKS

Marshall Porter, Jerry Miller, and Thomas Stockhammer
Digital Fountain, Inc, Fremont, CA

INTRODUCTION

In the past six months, streaming-related concerns in DVB-H have come to a head. Urban and building penetration continue to be an issue and video quality has been unacceptable. Yet the tools to combat these issues, as defined by the standard, don't solve the issues; in fact, some operators have removed MPE-FEC, the FEC specified in the DVB-H standard, from their networks as the bandwidth overhead does not justify the minimal gains in quality that have been experienced. Thus, operators have been exploring alternatives that will enhance the DVB-H experience for streaming broadcast television. In their collective search of the alternatives, operators have four specific concerns that any proposed supplement to DVB-H must address:

- Infrastructure costs: DVB-H networks are expensive to deploy, so operators are particularly interested in advances delivering a link margin gain that would mitigate some of the required infrastructure build out. Operators' specific pains in this instance revolve around covering urban areas (indoor penetration) and extending coverage to more rural areas.

- QoS: Any alternative must be able to deliver a better consumer viewing experience, both at high rates of speed and indoors.

- Legacy equipment: The introduction of an alternative must co-exist with existing deployed DVB-H infrastructure and handsets. Operators that have already deployed a DVB-H network simply are not willing to start anew.

- Migration: Any alternative should be able to be introduced gradually into existing deployments to minimize or eliminate any transitional pains.

While the four criteria cited above establish a difficult environment, there are options available to DVB-H operators that will allow a higher-quality consumer experience, create link margin gains, and meet the requirements related to existing systems. The most promising alternative that meets the above requirements is Multi-Burst FEC, also called iFEC, which itself has multiple available flavors, including both Reed-Solomon based codes and Raptor codes. Raptor codes are already "strongly recommended" for file delivery in IP Datacasting services in DVB-H, so there are benefits to considering it for streaming as well. The file delivery benefits of Raptor codes are compelling once a file to be sent exceeds 32kB: increased link margin, faster file delivery times, and substantial power savings.

But before understanding iFEC, an examination of the present role and function of FEC in DVB-H is required.

BACKGROUND ON DVB-H

A key to the success of DVB-H in Europe is that it is introduced almost exclusively as a link layer on top of DVB-T to enable mobile reception. For this purpose DVB-H includes time-slicing for power savings, Multi-Protocol Encapsulation (MPE) to transport IP data over MPEG-2 Transport Streams, and optionally MPE-FEC to protect the IP datagrams. During the definition phase of DVB-H, it was recognized that some kind of forward error correction (FEC) above the physical layer is necessary to support mobile reception – hence the option of MPE-FEC was introduced. MPE-FEC relies on Reed-Solomon Codes (RSC). But MPE-FEC has limitations on the amount/duration of data it can protect because of its high cost to terminal memory and processing as well as the restricted code word size of RSC. As asserted today, MPE-FEC does not solve the job satisfactorily and is an unused option in many deployments because of signal variations at the receiver: signal outages due to penetration losses, blockage from buildings, signal fading, body shadowing, and other disturbances are unavoidable.

When the duration of the outage is only a fraction of the time-slice burst duration (usually in the range of at most a couple of hundred milliseconds), MPE-FEC may help, but outage durations generally exceed fractions of a time-slice burst. Cases where a significant amount of a burst, or even the entire burst, is lost happen quite often and are the major impact to coverage. To achieve sufficient coverage, signal reception even in very adverse conditions must be ensured, and as MPE-FEC cannot provide a resolution to this problem, transmit power must be increased and additional transmit sites must be installed resulting in significant additional investments and operational expenses. Clearly some other solution is necessary.

MULTIBURST FEC PROTECTION

In order to overcome the problem of lost bursts, an FEC scheme that spans multiple time-slice bursts has been introduced. The benefits of this have already been recognized in the download delivery service in DVB-H by using FLUTE together with application layer FEC (AL-FEC) based on Raptor codes [2][4]. In addition, some research work in this direction proposes similar concepts for streaming delivery over DVB-H [8]. In standardization efforts, similar concepts are currently investigated in DVB within the satellite services to portable (SSP) group and a detailed and backward-compatible extended MPE-FEC proposal has been endorsed by the SSP group in [1] for further considerations in DVB. Detailed results on the performance of this multiburst FEC are provided in [5].

The idea of multiburst FEC generation is shown in Figure 1: A continuous data stream (yellow) is chopped into pieces of certain size. The data and the generated FEC from a single block encoding process is then distributed over multiple time-slice bursts. Two different sending arrangements are shown. Sending Arrangement 1 distributes the data and the FEC sequentially over the slice bursts—this generally results in bursts that either only contain data or only contain FEC. Sending Arrangement 2 distributes the data and the FEC so that each burst contains data and FEC. While either arrangement is significantly better than the current standard, both arrangements have advantages and drawbacks as outlined below.

Sending arrangement 1 separates FEC and time-slicing to the most extent. In case of good channel conditions, one might be able to ignore entire FEC bursts leading to power savings. However, this would require new signaling of time-slice bursts. Sending Arrangement 1 also might result in increased tune-in delays, since when tuning during a FEC burst. no data is available to be decoded and displayed.

Sending Arrangement 2 always sends data along with a time-slice burst in a very similar manner as DVB-H. This sending arrangement supports fast channel switching and allows reuse of the legacy time-slice burst signaling, but it suffers some slight end-to-end latency degradation. Nevertheless, due to the backward-compatibility to DVB-H MPE, sending arrangement 2 is generally considered more attractive.

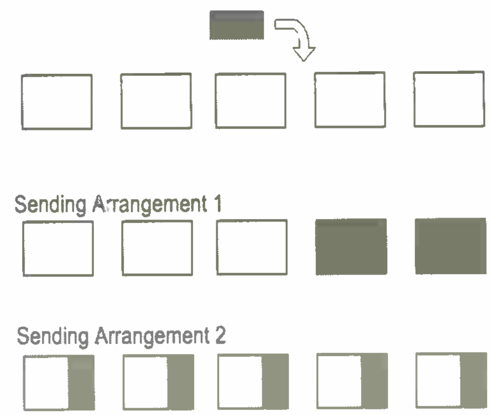


Figure 1. Multiburst FEC schemes in DVB-H

The introduction of multiburst FEC scheme may be done either

- above the IP layer as application layer FEC following the FEC framework currently defined by the IETF and the MBMS streaming framework (see [6]), or

- below the IP layer as link layer FEC following the MPE-FEC principle as well as MPE iFEC (Inter-burst FEC) concept [1] as endorsed by the DVB-SSP group.

The performance in both cases is expected to be identical, only some implementation and signaling specific aspects may differ.

MPE iFEC

The MPE iFEC framework defines a new MPE section type that can either replace the regular MPE-FEC or be sent in parallel. Therefore, the protocol stack is unmodified to DVB-H, and the new FEC is on the same layer as the MPE-FEC. Some features of this MPE iFEC specification are backwards-compatibility with MPE and MPE-FEC, operation on a time-slice burst basis, support for fast channel zapping, support for recovery of burst losses or bursts with high loss rates, full support of variable burst sizes and variable bit rate streams, and support of Raptor codes as specified in DVB-H IPDC and MBMS.

The overview of a sender supporting MPE iFEC is shown in Figure 2. The basic steps are as follows:

1. Datagram bursts of variable size are processed

2. Each of the burst is mapped to encoding matrices

3. Each encoding matrix periodically produces parity information

4. The parity data from the different encoding matrices is mixed with an original datagram burst to generate a burst.

A detailed presentation of the framework is available in [1].

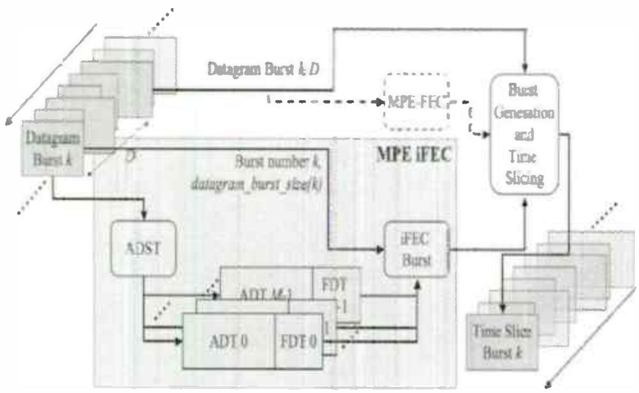


Figure 2. MPE iFEC framework at sender

The sender architecture is such that the new FEC scheme could be added to the IP encapsulator, since full time-slice awareness is also required for the extended FEC to provide all the benefits of DVB-H.

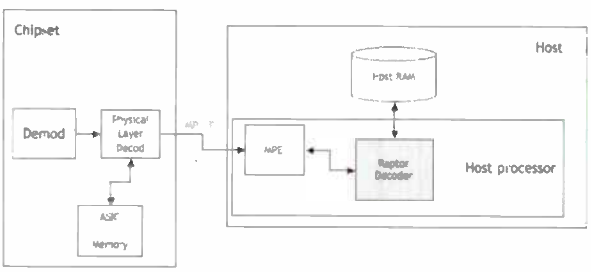


Figure 3. Potential receiver architecture for multiburst FEC on MPE level

One potential receiver architecture is shown in Figure 3. As the FEC processing is done in the host processor, it is more suitable to do the MPE processing on the host. The interface between the host chipset and the host can also be MPE sections corresponding to IP packets. Note that the introduction of such a scheme will not require any modifications to DVB-H chipsets or hardware and can therefore be introduced as a software update at the receiver.

Raptor Codes in Multiburst FEC

As already mentioned, the MPE iFEC framework supports the use of Raptor codes. The flexible and low-complexity Raptor technology enables overcoming the above mentioned signal outages by channel coding. Due to the flexibility of the solution, the operator is provided with the option to design each of the services individually, according to service requirements in terms of priority, reliability, resource consumption, and latency, as well as take into account its measured and experienced conditions. The flexibility of fountain codes provides a unique possibility to adapt a deployed system to the specific conditions, which is particularly important since operators are concerned about channel switching times and their effect on the consumer experience. Recent innovations include new methods to minimize any effects on channel change time with MPE-iFEC.

Link Margin Gains with Multiburst FEC and Raptor Codes

To provide insight into the potential performance gains from Raptor codes, some initial simulations have been performed. A simplified channel model for the evaluation of the potentials of multiburst FEC has been proposed by Teracom in [3]. A brief overview is provided. With DVB-H the error behavior can be modelled so that as long the average field strength (FS) during a time-slice burst is above a certain threshold, all IP data in the burst will be error-free after MPE-FEC error correction, whereas if the average FS (or C/N) is below the threshold, a significant amount of data is considered lost. Shadow fading is added to this model. The average C/N of a burst will vary from burst to burst with independent probabilities following a Gaussian distribution around the average FS value. The Gaussian distribution is characterised by its standard deviation, sigma, which is often assumed to be 5.5dB.

The MPE-iFEC framework is applied. For Raptor codes, the framework is most suitably used with B=S=1 resulting in basically only one encoding matrix to host source data from EP datagrams (B=1). In addition, the generated parity data is also distributed over EP time-slice bursts (S=1). For D=0, the interleaver spread is 2*EP as the parity data is sent after the source data. For D=EP, the sending of the data is delayed by EP bursts and therefore, the interleaver spread is only EP. The interleaver spread influences the performance and the channel switching time. Typically, time-slice burst repetition periods are in the range of one to several seconds, such that a spread over *EP* bursts results in a delay of around *EP* seconds.

Figure 4 provides simulation results for ESR5(20) (Erroneous Second Ratio measured over a window of 20 seconds and at most 5% erroneous seconds) over link margin in dB for FEC code rate, 3/4, and different FEC spreads *EP*=4,8,12 compared to MPE-FEC at same code rate (CR) of 2/3. Also shown are the gains in dB for different schemes. Two figures are presented, one corresponding to D=EP, and one corresponding to D=0. Typically, operators want to achieve an ESR5(20) of below 1%.

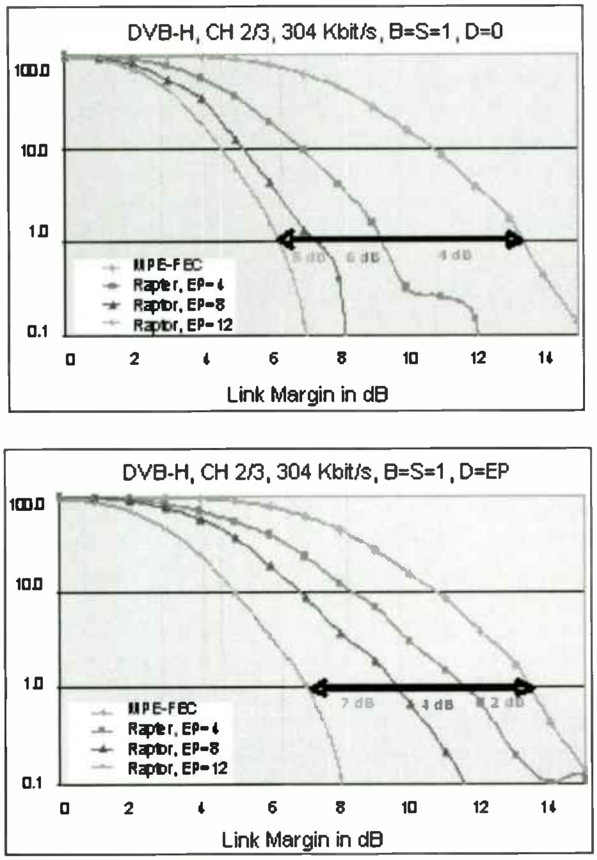


Figure 4. Simulation results: ESR5(20) over link margin in dB for Raptor MPE-IFEC for different interleaver spreads EP and different D compared to MPE-FEC at same code rate 2/3.

From the figure it is observed that gains between 2dB for low EPs and up to 8dB for high EPs and D=0 can be achieved compared to the MPE-FEC by the use of multiburst FEC with Raptor. The parameter variations show the flexibility of the solution, each service can be optimized towards delay, quality, and efficiency. More comprehensive results for different code rates, interleaver depth, etc. show very similar tendencies and are available on request.

Channel Switching and Latency

Channel switching time is an important parameter for DVB-H video services. Generally, the use of multiburst FEC increases channel switching times. However, with sending arrangement 2 (mentioned above), each time-slice burst contains useful data. If the time-slice burst is not lost and the burst contains a random access point to the media stream, the data can be immediately decoded and displayed, and channel switching times are as good as for regular DVB-H. Switching then happens on non-protected data and only relies on in-burst FEC.

Once tuned to a service, the multiburst FEC is quite likely required at some later time. In a simple receiver implementation, the video decoder would then need to rebuffer once any iFEC is required. However, the video and audio decoders can easily and without perceptual degradation slow down the media playout. This concept is known as adaptive media playout (AMP), see for example [7], where after switching, the media decoder slows down the playout by, for example, 25% such that the buffer needed for iFEC decoding can be built up over time. With a slowdown of 25% and an iFEC delay of 8 seconds, iFEC could be fully exploited in 32 seconds. If iFEC needs to be available sooner, more aggressive strategies might be used for shorter iFEC delay or more aggressive slowdown of playout. In addition, for services with no delay constraints, the data may be delayed even further such that the FEC is sent ahead of the corresponding data. In this case, less buffer slowdown is necessary. The configuration of the receiver playout strategy is a quality versus efficiency tradeoff and may be adjusted according to the previously experienced reception conditions.

THE BUSINESS CASE

Based on the results in the simulations above, the implications to the business case of DVB-H are nothing less than substantial. Link margin gains using Raptor technology (compared to the status quo) were 2dB – 8dB. To translate those dB gains into meaningful numbers, note the following table that equates link margin gains to area gains in the service area:

dB Gain with Raptor iFEC	2	3	4	5	6	7	8
Coverage Area Increase	30%	49%	70%	95%	123%	155%	191%

Table 1. Example dB gain translation to coverage area increase.

A realistic link margin gain of 5dB with Raptor codes would nearly double the service area, allowing an operator to deploy a broadcast network with half the investment in transmission towers or to achieve a substantially better QoS with the same number of towers. Further, for urban areas, a 5dB gain would dramatically improve indoor reception, again improving the customer experience though better QoS. In an instance of a medium-sized country with a moderately dispersed population where a DVB-H rollout has cost upwards of 300m Euros, Raptor codes could yield a start-up infrastructure savings of 150m EUR. Understandably, this has gathered the attention of a number of operators developing DVB-H solutions.

So the business case is clear and affects both sides of the coin: operators can leverage FEC advances to:

- Reduce costs and infrastructure investment
- Increase service revenue through:
    - Reaching a broader area, bringing the broadcast network to more subscribers
    - Improving QoS for all receivers by strengthening the broadcast signal and reaching previously unreachable urban areas, creating a more seamless, high quality network
    - Lower churn, as the service improves and customer frustration over network quality is addressed

CONCLUSION: A NEW DAY FOR DVB-H

Much has been made of DVB-H technology and the role it will play in spurring mobile multimedia further. Operators globally are evaluating the network or making massive investment in the buildout of a DVB-H network. But rollouts to date have been plagued by high cost, wasted bandwidth, flawed business models, and lackluster quality. Indeed, alternative mobile broadcasting models have used the troubles of DVB-H to highlight the relative strengths of alternative broadcast networks.

The aforementioned shortcomings in DVB-H require an analysis, if not a rethinking, of DVB-H in order to find areas where recent technological improvements can measurably impact DVB-H. It is clear that the advancements in FEC create this opportunity. The adoption of iFEC, and particularly Raptor-enabled iFEC, has several advantages for DVB-H networks:

1. Link margin gains that create substantial infrastructure savings, allowing operators to reach broader service areas and reduce the investment required for a network buildout by as much as 30-50%
2. Quality of service improvements from link margin gains enabling perfect quality even at high rates of speed
3. Bandwidth savings, so operators can deploy more services and create a more favorable business model for operators
4. Seamless backwards compatibility with existing networks, allowing operators to leverage existing DVB-H investments

The business impact of these advantages is clear: lower costs and greater revenue while reducing churn, making the service more attractive to consumers as network quality improves.

There's little questioning the promise of DVB-H: the technology has been adopted by operators globally and endorsed by the European Commission and investment by the entirety of the mobile ecosystem—from operators to OEMs to content providers—continues to grow with an eye toward a mobile multimedia explosion. But the existing issues must be solved for DVB-H to be successful.

The solution is clear—by leveraging the advancements in FEC over the past five years since specification, DVB-H can offer a more attractive business model, both on the cost and the revenue sides, for operator and consumer adoption while maximizing the consumer experience.

References

[1] DVB TM3880: "MPE IFEC Specification", Alcatel-Lucent, UDCast, Digital Fountain, September 2007 (see also DVB TM-SSP0331,slightly revised version).

[2] DVB-CBMS, "IP Datacast over DVB-H ; Content Delivery Protocols (CDP)", DVB Document A101, December 2005.

[3] DVB TM-CBMS1209: "Simplified Model to assess Gain by using Application Layer Generic Forward Error Correction (ALG-FEC) for File Delivery", Teracom, Digital Fountain, February 2005.

[4] DVB TM-CBMS1397: "AL-FEC Summary," Digital Fountain, August 2005.

[5] DVB TM-SSP240: "Upper Layer FEC Simulation Results," Digital Fountain, May 2007.

[6] DVB TM-SSP286: "DVB-SH Simulations with Upper Layer FEC," Digital Fountain, September 2007.

[7] 3GPP TS26.346: "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs," July 2005.

[8] M. Kalman, E. Steinbach, and B. Girod, "Adaptive Media Playout for Low Delay Video Streaming over Error-Prone Channels," IEEE Transactions on Circuits and Systems for Video Technology, June 2004.

[9] D. Gomez-Barquero and A. Bria, "Application Layer FEC for Improved Mobile Reception of DVB-H Streaming Services", in Proceedings IEEE VTC Fall, Montreal, CA, September 2006.

New Techniques for Mobile TV Broadcasting Based on ISDB-T

~ Automatic Activation of Mobile TV Receivers by Emergency Warning Broadcasting System and a New Method of One-Seg Retransmission for Mobile TV Receivers ~

MASAHIRO OKANO, YOSHIKAZU NARIKIYO, MAKOTO TAGUCHI, KENICHI TSUCHIDA, MASAYUKI TAKADA, HIROYUKI HAMAZUMI, YASUHIRO ITO, and KAZUHIKO SHIBUYA

NHK Science & Technical Research Laboratories, Tokyo, Japan

INTRODUCTION

Digital terrestrial television broadcasting (DTTB) called Integrated Services Digital Broadcasting - Terrestrial (ISDB-T) [1][2] was launched in Japan in December, 2003. The ISDB-T system is designed to provide high-quality video, sound, and data broadcasting not only for fixed receivers but also for portable/mobile receivers.

One-Seg, which is an ISDB-T service for portable/mobile receivers, began operating in Japan in April 2006. One-Seg has become popular for its interactivity and for its providing broadcasts to viewers anywhere, anytime. The majority of portable/mobile One-Seg receivers are Internet-enabled cellular phones, and domestic shipments of cellular-phone-type One-Seg receivers had reached about 15 million by the end of Oct, 2007.

In this paper, we describe a new technique for One-Seg service to provide an emergency warning broadcasting service and a technique to retransmit the One-Seg signal.

One-Seg can send information to a large number of One-Seg receivers at the same time without traffic congestion. It is an effective means of alerting the public about an impending natural disaster and of informing them about preventive measures and rescue measures after a disaster happens. An emergency warning broadcast signal automatically activates warning receivers when they are in stand-by mode. We developed a remote activation system with low power consumption[3][4], and here, we describe its receiving capability and the effect of reducing its standby power.
In addition, to achieve full One-Seg coverage, we propose a method of One-Seg combined retransmission[5] as an efficient means of retransmission in locations where broadcast waves from transmitting stations cannot be directly received. This method extracts One-Seg signals from the individual broadcast waves, combines them, and retransmits the combined signals. We describe the characteristics, functions and performance of this system.

EMERGENCY WARNING BROADCASTING SYSTEM

Reception of Emergency Warning Broadcast Signal

An emergency warning broadcast service for One-Seg receivers would be an important tool for disaster prevention. In this section, we discuss the technique by which an emergency warning broadcast automatically activates One-Seg receivers.

The Emergency Warning Broadcasting System (EWBS) is a disaster information distribution system that can remotely activate receivers by using radio and television signals in times of earthquake, tsunami, or other disaster. It has been operating in Japan since September, 1985, and it is used by national and municipal governments to issue notifications such as large-scale earthquake or tsunami warnings. It has been used 15 times since its inception, as shown in Table 1.

Table 1 Emergency warning broadcast history

Incident name	Date
Miyazaki Earthquake	1987/3/18
Sanriku Earthquake	1989/11/2
Hokkaido Nansei Earthquake	1993/7/12
Hokkaido Tohou Earthquake (1)	1994/10/4
Sanriku Haruka Earthquake	1994/12/28
Earthquake near Amami Ooshima	1995/10/19
New Guinea Earthquake	1996/2/17
Hyuganada Earthquake	1996/10/19
Earthquake with epicenter in the sea south of Ishigakijima, Okinawa (1)	1998/5/4
Earthquake with epicenter in the sea south of Ishigakijima, Okinawa (2)	2002/3/26
Earthquake with epicenter near Taiwan	2002/3/31
Hokkaido Kushiro Earthquake	2003/9/26
Tokaido Earthquake	2004/9/5
Hokkaido Touhou Earthquake (2)	2006/11/15
Hokkaido Touhou Earthquake (3)	2007/1/13

(All of the above are tsunami warnings.)

However, in order for One-Seg receivers to be able to receive emergency warning broadcasts, the emergency warning broadcast activation flag within the Transmission and Multiplexing Configuration Control (TMCC) data must be monitored. If power consumption for monitoring the activation flag were too high, it would be difficult to maintain stand-by for a long time. This is a significant issue that must be resolved.

We developed a method for demodulating the TMCC carrier that does not require a fast Fourier transform (FFT) and uses a simple circuit architecture. In addition, we investigated a time-sliced reception approach that can continuously monitor the emergency warning broadcast activation flag and reduce power consumption. We also developed prototype test equipment using a field programmable gate array (FPGA) (Figure 1 beneath cellular phone). We evaluated the characteristics of the prototype in laboratory tests.

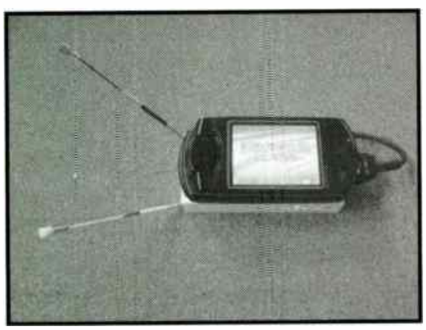


Figure 1 Emergency warning broadcast activation flag receiver

Low-power Activation flag Receiver Device for Emergency Warning Broadcasts

As shown in Figure 2, in normal standby mode, the frequency of the signal received by the antenna is converted by the tuner block and the emergency warning broadcast activation flag in the TMCC data is monitored by the activation flag detector block. When a change to "ON" is detected in the activation flag, the demodulation block is automatically activated so that the emergency warning program can be displayed on the receiver. Methods for receiving the TMCC data and for time-sliced operation are described below.

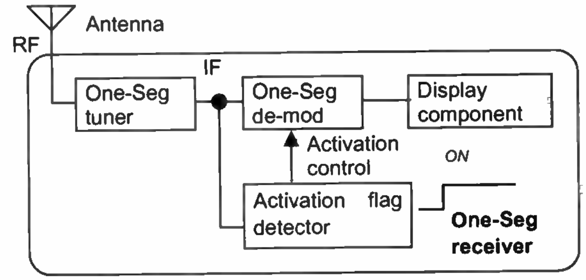


Figure 2 Block diagram for receiving the emergency warning broadcast activation flag

TMCC Reception

As shown in Figure 3, there are four TMCC carriers within the One-Seg band. TMCC data bits are mapped into DBPSK, and each TMCC carrier is modulated by the same data. The circuit architecture for the receiver is shown in Figure 4. Two of the TMCC carriers are demodulated by differential detection using a shared circuit. The other two are demodulated using independent differential detection circuits. The resulting four TMCC data are combined using diversity to increase the reception reliability. Then, the value of the activation flag is decided by majority decision between multiple OFDM frames. In the prototype receiver, we use a simple circuit to demodulate the TMCC data instead of applying an FFT.

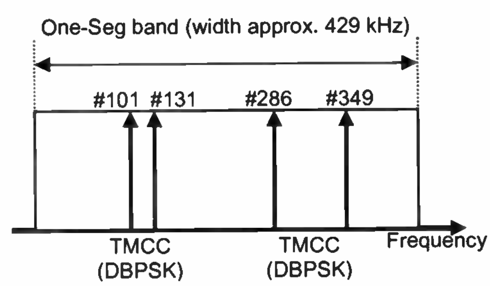


Figure 3 Positions of the TMCC carriers within the One-Seg band

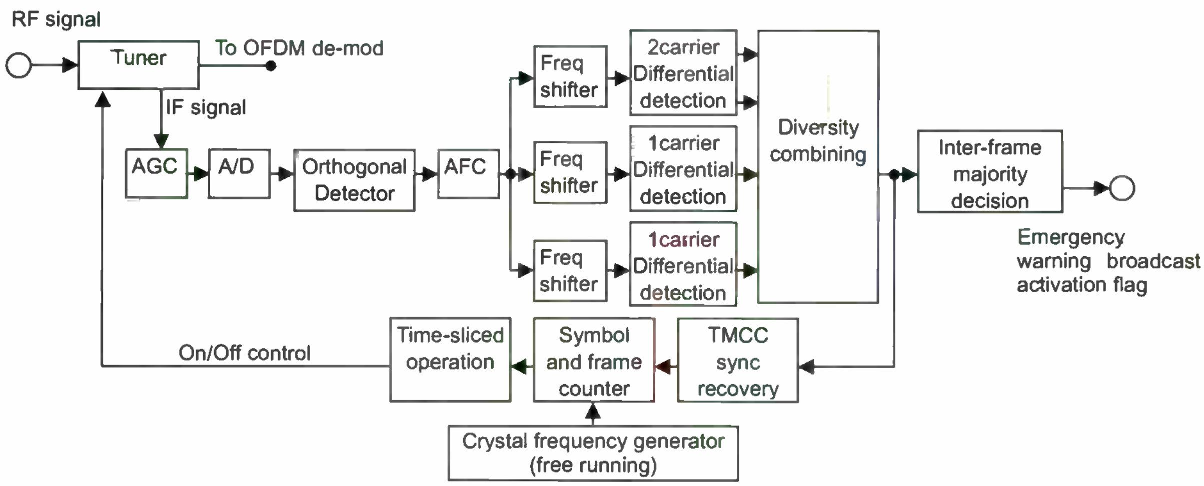


Figure 4 Prototype emergency warning broadcast activation flag receiver

Time-sliced Operation

We studied ways to activate and deactivate the tuner block that would reduce power consumption. A schematic of this procedure is shown in Figure 5. In order to periodically receive the 27 symbols from the OFDM frame header to the emergency warning activation flag (hereafter called symbol time slicing), automatic gain control (AGC) operation must be stable before the OFDM frame header arrives. As a result of experiments on the prototype circuit, we found that AGC operation stabilized seven symbols after the tuner power turned ON. Accordingly, the tuner activation duration must total 34 symbols, so symbol time slicing can reduce power consumption to 34/204 (approximately 17 %) for a single frame.

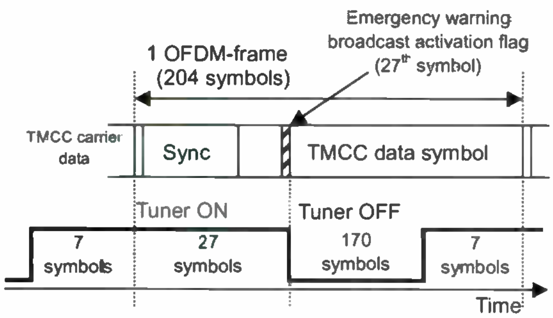


Figure 5 Timing diagram for TMCC receiver power control

To further reduce power consumption, we thinned the OFDM frames (frame time slicing). If the circuit is activated for only one in eight OFDM frames, 34/204 symbol time slicing is performed only on the first frame and the tuner is not activated for the following seven frames. In this case, a long detection time is needed, but the number of times the tuner is activated is reduced by a factor of 8. Thus by using both techniques, tuner power consumption can be reduced to 1/48th of the original amount.

Reception Characteristics

CN Ratio vs. bit error rate Characteristics

Figure 6 shows the TMCC reception characteristics of this circuit under the additive white Gaussian noise environment for different numbers of TMCC carriers used for diversity combining. The horizontal axis is the input-signal carrier to noise ratio (CNR), while the vertical axis is the bit error rate (BER). At a BER of around 2.0 x 10^{-4}, diversity combining provides improvements of 2.4 dB (two carriers) and 4.5 dB (four carriers).

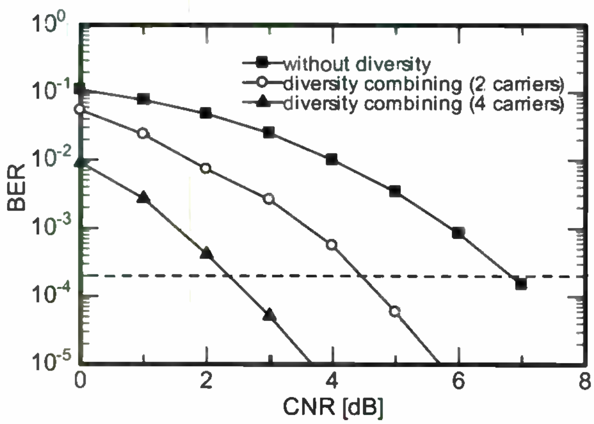


Figure 6 CNR vs. bit error rate

TMCC Reception Characteristics

The emergency warning broadcast activation flag must be detected accurately within the One-Seg coverage area, so the TMCC reception characteristics must be at least as good as the One-Seg reception characteristics. The reception characteristics for four-carrier diversity combining using majority decision between frames is shown in Figure 7. The characteristics improve as the number of frames used for majority decision increases. The bit error rate for a single false detection of the activation flag in one day, as calculated by the following equation, is 2.7 x 10^{-6}.

$$\text{BER} = \frac{(Tu+Tg)\times n[\text{sec}]}{\text{T[sec]}} \quad - - - \quad (1)$$

Tu: Useful symbol period (s)
Tg: Guard interval period (s)
n: Number of symbols per frame (204)
T: Period allowing one error (one day)

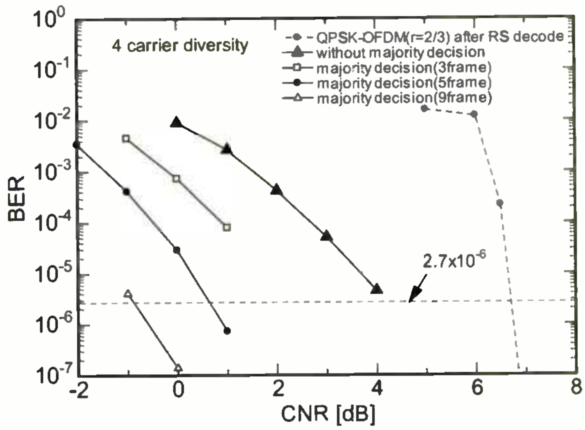


Figure 7 TMCC receiving characteristics after majority decision

When the emergency warning broadcast is actually received by a One-Seg receiver, the One-Seg demodulator circuit is first activated by the activation flag, and then the emergency warning descriptor within the transport stream (TS) is confirmed. This process can cancel a once-per-day false detection from the time-sliced detection circuit, so receiver does not alert the user by mistake. This characteristic, after applying the majority decision, is better than that of a Reed-Solomon decoded signal according to the transmission parameters of the One-Seg service (QPSK, encoding rate, r=2/3). Hence, the proposed method is sufficiently high reliability.

One-Seg COMBINED RETRANSMISSION SYSTEM

Retransmission of Terrestrial Digital Broadcasts for Mobile Receivers

Even within the coverage area of terrestrial digital broadcasting, the signals from transmitting antennas cannot directly reach, or is extremely weak, in areas such as underground malls, subways, or behind buildings. In these areas, viewers would likely try to use a portable/mobile receiver, so a system to efficiently retransmit the One-Seg signal is required. We have developed retransmission equipment to fill such gaps in One-Seg coverage.

There are various approaches to retransmitting the One-Seg signal, including simply retransmitting the entire received broadcast wave at the same frequency, or extracting only the One-Seg band from the received signal and retransmitting it at the same frequency. With these methods, the signal is retransmitted at the same frequency, creating a single frequency network (SFN), and it becomes necessary to consider how reception will be affected in areas where the originally broadcast signal overlaps the retransmitted signal.

To avoid interference with existing broadcast waves, retransmission using another channel is effective. However, to retransmit all ISDB-T signals only for the purpose of expanding One-Seg coverage is inefficient. Accordingly, we devised a One-Seg combined retransmission system that extracts several One-Seg signals from the broadcast waves and combines them into one signal to be retransmitted over a single channel.

Combined One-Seg Retransmission Method

One-Seg service uses only the center segment of the ISDB-T broadcast signal. The One-Seg bandwidth is 1/14th of the channel bandwidth (6 MHz), so it is possible to save bandwidth by arranging several One-Seg signals within a single channel for retransmission. In particular, if signals are combined without a guard-band and transmitted, bandwidth usage will be as efficient as the ISDB-T system. However, simply rearranging multiple unsynchronized signals and transmitting them will cause interference between the signals, so this issue must also be resolved.

The One-Seg combined retransmission system extracts the One-Seg signals by partial reception of multiple broadcast waves, rearranges the extracted signals within a single channel, and retransmits them. A diagram illustrating the system is shown in Figure 8. As an example, let us consider the case in which several broadcasters transmit ISDB-T signals using UHF channels 20 to 27. The signals from each channel are extracted by partial-reception of the One-Seg segment and demodulated. These signals are combined together and re-modulated into a 13-segment signal similar to ISDB-T and retransmitted, for example on UHF channel 34. Using re-multiplexing and re-modulation together makes it possible to retransmit stably without a guard band. To perform re-multiplexing, the reception signal is changed in two ways. One is to modify the program clock reference (PCR) arrangement, and the other is to insert or delete Null packets for synchronization of the bit rate.

Up to 13 One-Seg signals can be combined on a single channel with this method, so if fewer than 13 channels are being retransmitted, the remaining segments can be used for community broadcasting (e.g. public or commercial content).

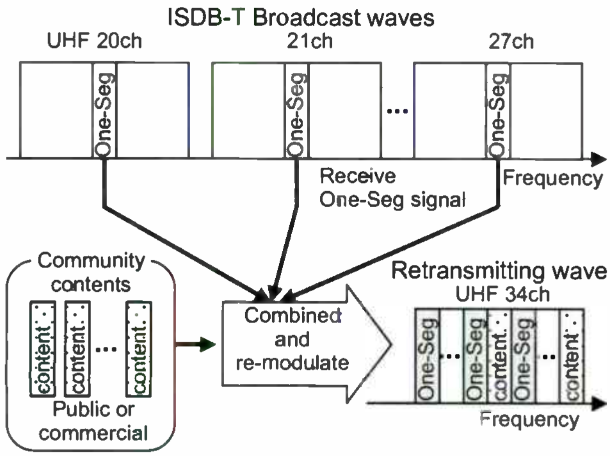


Figure 8 Outline of combined One-Seg retransmission method

Current One-Seg receivers scan channels in 6-MHz steps according to the placement of One-Seg signals. With the system, One-Seg signals can also be transmitted in segments other than the one in the center of the 6 MHz band, so the channel-scanning step size on receivers supporting this system will have to be changed to the bandwidth of a single segment, or 432 KHz (6 MHz/14).

Characteristics of The One-Seg Combined Retransmission System

The characteristics of the system are summarized in Table 2. The system requires the use of a further channel in addition to the received channels, but causes no interference with the received channels, and it utilizes the channel efficiently by combining several One-Seg signals. It is also able to transmit community content on unused parts of the channel so that it both expands the coverage and delivers more localized One-Seg service.

Table 2 One-Seg combined retransmission system characteristics

	Requirement	Proposed system
Efficient spectrum	Must be high	A single channel is required for retransmission 13 One-Seg services delivered per channel
Power consumption	Must be low	Efficient, as only One-Seg is retransmitted
Signal degradation due to retransmission	None (Minimal)	A type of regenerate relay, so high-quality retransmission signal is possible.
Interference with existing broadcast signal	None (Minimal)	No interference with existing signal (retransmitted on a different received channel)
Content consistency	Consistency is preserved	Consistency at the TS level (Only the PCR is replaced)
Changes to receivers	None (minimal)	Skip frequency for scanning must be changed
Compatibility	No effect on existing receiver equipment	The center segment of the retransmitted signal can be received by existing receiver devices.
Service extensions	Possible in the future	Community content can be added easily

Prototype Equipment

We developed prototype One-Seg combined retransmission equipment to evaluate the system. A block diagram of the prototype is shown in Figure 9. The device is composed of a receiving block, a TS re-multiplexing block, and a retransmitting block.

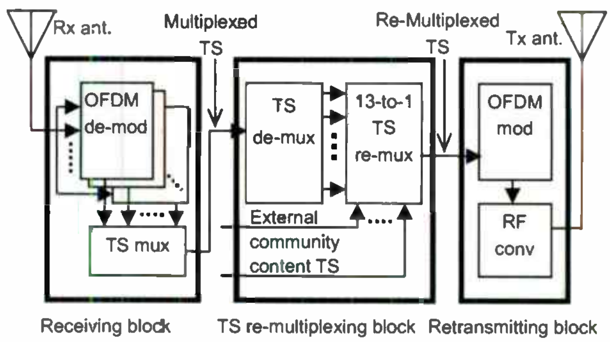


Figure 9 Prototype equipment architecture

The receiving block demodulates eight One-Seg signals into MPEG-2 transport streams (TS), and multiplexes them before output. The TMCC data from each of the signals is packetized and multiplexed as well.

The TS re-multiplexing block takes the multiplexed output from the receiving block, temporarily de-multiplexes it, and then re-multiplexes it, adding any external community content TS up to a maximum of 13 TSs. The system synchronizes each TS signal by re-multiplexing.

The retransmitting block assigns each TS signal to a specified segment and processes each, by adding error-correction coding, interleaving, delay-correction and carrier modulation, and then performs OFDM modulation on all at the same time by using an inverse fast Fourier transform. The modulated signal is transmitted by the transmitting antenna via the frequency conversion and power amplifier circuits. The arrangement of the pilot carriers of each segment is done in the same way as that of the One-Seg segment. The TMCC carriers are re-modulated according to the received TMCC data and retransmitted. Note that if the transmission parameters are modified, this change is reflected in the TMCC data.

The received One-Seg TS and the community content TS are allocated to segments, and various transmission parameters for retransmission signal are set in the TS re-multiplexing block.

Evaluation of Combined Retransmission

To confirm that no degradation occurs in the signal received from the combined retransmission equipment, laboratory tests were performed using the prototype One-Seg combined retransmission device. The test system is shown in Figure 10. The TS packet error rate (PER) was measured with and without the retransmission device. The case without retransmission device indicates that the OFDM modulator output was fed directly into the One-Seg receiver, whereas the retransmission case indicates that the output of the combined retransmission device was fed to the One-Seg receiver. In both cases, the receiver input level was adjusted by attenuator (ATT) before being fed to the One-Seg receiver.

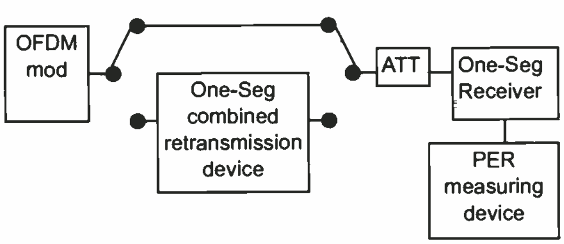


Figure 10 Laboratory test system

The results of these measurements are shown in Figure 11. The PER for the receiver input level characteristics were essentially the same, confirming that there was no degradation of characteristics due to the combined retransmission.

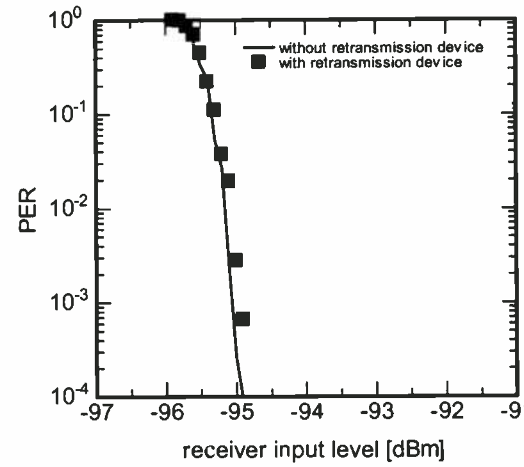


Figure 11 Packet error rate vs. receiver input level

CONCLUSIONS

In this paper, we discussed automatic activation of One-Seg receivers by emergency warning broadcasts and a technique for implementing this. We described the technique to conserve power on the mobile receivers. After that, we demonstrated the prototype receiver's excellent reception characteristics in laboratory tests.

In addition, we discussed a One-Seg combined retransmission system that can be used to achieve full One-Seg coverage. The system provides both stable retransmission of One-Seg signals and can also transmit community content; hence, besides expanding One-Seg coverage it also may supports localized services.

REFERENCES

[1] "Standard regulations for transmission of Digital Terrestrial Television broadcasts", Association of Radio Industries and Businesses, ARIB STD-B 31 (http://www.dibeg.org/aribstd/ARIBSTD-j.htm).

[2] M. Takada, "Transmission System for ISDB-T", IEEE Proceedings, Vol. 94, No. 1 (Jan. 2006).

[3] Y. Ito, "Study on remote activation of One-Seg handheld receivers with Emergency Warning Broadcasting System" ABU Technical Review No. 231 (pp. 12-14) (Jul.-Sep. 2007).

[4] M. Taguchi, "A Study of automatic activation of portable One-Seg receivers with Emergency Warning Broadcasts", IEICE The Technical Report, vol.106, no.579, EMCJ2006-120, p.29-32 (Mar. 2007).

[5] Y. Narikiyo, "Proposal of Connected Retransmission Method for One-Seg and its Gap Filler," IEICE The Technical Report, vol.106, no.349, OCS2006-63, OFT2006-48, p.11-14ocs2006-63, OFT2006-48 (Nov. 2006).

Preserving SFN in a Broadcast Network using IP distribution

Nicolas FANNECHERE
ENENSYS Technologies
France

ABSTRACT

Present paper takes a closer look at the technologies used today for synchronizing transmission of a Digital Terrestrial TV or Mobile TV broadcast signal countrywide using SFN synchronization means. Paper will detail as well the means to transport compressed digital video over a long distance to introduce fundamentals about video over IP transport. Conclusion will be made that it is possible to date to offer a distribution of a MPEG2 compressed video stream over IP networks while preserving SFN networks, applicable to Terrestrial or Mobile TV broadcast standards.

WHAT IS SFN ?

SFN means Single Frequency Networks. Migration to SFN networks is ramping up today because broadcasters and network operators are facing frequency shortage. Even with analogue TV switch off and the release of frequencies for a use in digital, the fast-growing commercial deployment of new terrestrial and mobile TV services most of time lead operators to optimize the use of granted frequencies. To do so, they have to minimize the number of channels on-air that will broadcast the same content over many frequencies. Such a reduction of frequencies is called SFN. As a contrary, MFN (Multiple Frequency Network) transmission exists, where every single transmitter broadcasts over a frequency that is different from others. Overlapping areas of transmission will thus not be an issue for any receivers since there is no risk of signals perturbing each other, receivers being locked on separate frequencies. In SFN (Single Frequency Network), transmitters broadcast over the same frequency to provide expected spectrum and bandwidth optimization.

WITHIN SFN CELLS

A SFN network is made of at least one SFN cell. For instance, an SFN cell can cover a nation (with national content, broadcasting on frequency #1), or a region (with regional content, on frequency #2), or another region (frequency #3). Transmitters belonging to the same SFN cell shall radiate the very same RF symbols, over the same frequency, at the very same time, using the very same RF modulation. Any receiver located in the area covered by the SFN cell will receive signal from one transmitter, or another, without any difference, and moving from an area covered by one transmitter to another within the same SFN cell will not lead to any macroblock or black picture. Even in overlapping areas, where a receiver gets signal from two transmitters from the same SFN cell, no perturbation can occur.

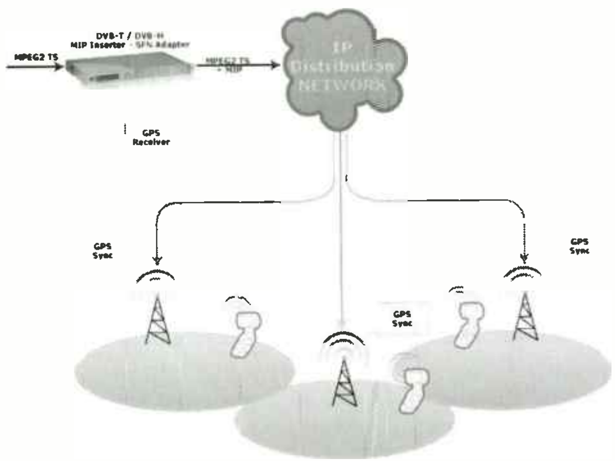
SFN SYNCHRONIZATION MEANS

Such a perfect synchronization between all transmitters within the SFN cell is facilitated by the use of a SFN adapter at the head end side. SFN adapter adds extra information to the MPEG2 TS stream aiming at synchronizing every transmitters with each other within the same SFN cell: Synchronization information and Transmission (ie modulation) parameters. SFN adapter can also add optional data to address individual transmitters: Transmission delay, Frequency offset, transmission power...

SFN adapter is centrally located at Head End side to build a unique MPEG2 TS source that all transmitters will have to broadcast in a synchronized way.

Those synchronization information are contained in MIP packets (Megaframe Initialization Packet) inserted in the MPEG2 stream by the SFN adapter, at the Head End. SFN adapters sometimes is called a MIP Inserter. In DVB, MIP packets have a normalized with value: 0x15. and do embed the following data: Transmission parameters (Transmission Parameters Signaling bits – aka TPS bits, Modulation settings, guard interval, bandwidth, FFT mode, etc.), Synchronization information (Timestamp : Synchronization TimeStamp - aka STS, maximum_network_delay).

Optional functions can be carried as well, aiming at addressing transmitters individually according to their TxID (Transmission delay, Frequency offset, transmission power, cell identifier...).



SFN adapter's duty is to build a megaframe, made of several TS packets. Megaframe size depends on code rate and constellation while its duration depends on bandwidth and guard interval used by modulation .

Using MIP packets, SFN adapter device located in Head End is able to control every single transmitter independently on all critical parameters: configuration, frequency and time.

MIP PACKETS

SFN adapter inserts MIP packets and its output bitrate must follow the exact modulation bitrate mandated by the transmission modulation. Actually, SFN adapter output is to feed all transmitters of the target SFN cell, and exciters (ie modulators) inside transmitters do not have an infinite input buffer size that may crack if data rate in input is higher than what modulation can transport. On the other hand, input buffer may dry run if data rate is lower than expected.

As said before, MIP packets contain the transmission parameters that transmitter must extract to auto-configure accordingly. MIP packets also aim at frequency synchronization. Temporal synchronization makes Transmitters broadcast synchronously, at the same time, through the use of GPS timing reference. Indeed, since SFN adapter and transmitters are not located at the same place, the only common timing reference available in the world is GPS satellites. They need as well to have a frequency synchronization to broadcast exactly the same set of sub-carriers. This is why in SFN networks, SFN adapters and transmitters are connected to GPS receivers providing them with 1 PPS and 10MHz signals. 1 PPS for temporal synchronization, 10 MHz for frequency synchronization.

NETWORK DELAY

If 1 PPS signal provides with synchronized pulse information on every place on earth, 1 PPS signal brings one constraint to a SFN broadcast: content distribution from head end to transmitters MUST be done within one second. If for any reason it takes longer than 1 second (ie between 2 and 3 seconds), then distribution duration should be the very same for every transmitter. This is where time budget of distribution enters in the discussion: to setup a SFN network, network operator must assess the longest path (that will take the longest duration to transport content) between head end and every transmitter. The longest path will correspond to the network delay (ie the worst case time budget), and during that time, some of the transmitters already got the data to transmit because their distribution duration was shorter. That explains the interest of carrying the network delay information in MIP packets: transmitters parsing that data know how long they have to retain information, and deduce the exact time they have to broadcast it. Synchronization of all the transmitters of the SFN cell will be synchronized since every transmitters, even the ones with the longest transport duration, will have the data in buffer, ready to broadcast. For instance, we can imagine transmitters receiving distribution by satellite and some transmitters get distribution via IP or microwave links. Transport delays can vary from 100ms to 500ms in average, that clearly explains that some transmitters have to bufferize data during the network delay time, ie the maximum transport time on distribution network.

FREQUENCY CONSTRAINTS

In DVB-T/H, DVB-SH, etc. modulation scheme used is COFDM modulation, which is the the ideal for multi-path echoes (multi carrier modulation) such as urban environment. It also fits with moving reception and enables diversity reception. A large number of carriers used (2000, 4000 or 8000) within the available bandwidth, and the best receiving conditions can be ensured if all transmitters broadcast the very same sets of carriers, at the same time and the same frequency. Obviously, frequency shifts from one transmitter have to be avoided, and the accuracy of the 10MHz signal used by transmitters is a key. Such an accurate 10MHz reference clock is provided to exciters, in charge of modulation inside transmitters, from GPS signal.

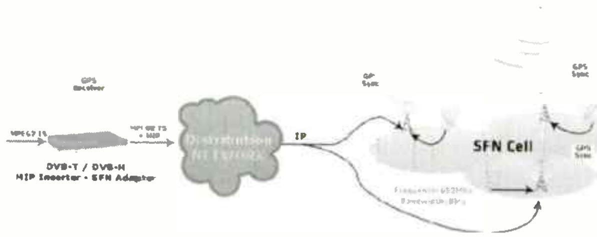
SFN NETWORK OPTIMIZATION - CONCLUSION

SFN networks are not only the key for frequency spectrum allocation issues. SFN networks enable transmission areas to be covered by several low power transmitters instead of a very high power, unique broadcast point. It improves overall coverage, allows to build large SFN cell (nationwide, for instance). SFN synchronization between head end and transmitters have to be secured with Temporal and Frequency synchronization through the use of (1PPS + 10 MHz) signals.

The more accurate the SFN network is, the better reception quality will be !

CONTENT DISTRIBUTION

Content distribution is the way to carry multiplexes from head end premises (eg located in main city) to transmission sites (eg in regions). There are many ways to distribute content, according to historical technologies used by a country or an operator, led by environmental and geographical constraints.



The following non exhaustive list provides an overview of pros and cons of some of these transport methods: MPEG2-TS over ASI cables (Very expensive, useless for long reaches unless repeaters, cable equalizers ... are required), MPEG2-TS over ATM (high operating costs, needing expensive adapters as well, complex to setup and not that flexible), MPEG2-TS over Satellite (not cost efficient if limited number of transmitters, Limited bitrate in DVB-S, presence of shadowed areas, high operating costs), microwave links (rand last but not least: distribution using IP back haul.

WHY USING IP ?

IP is really easy to setup, there is Ethernet connectivity in every building now, and IP prices dropped down compared to other distribution solutions. To date, IP networks are replacing standard distribution solutions: leased lines prices are decreasing, Bandwidth capability keeps on increasing, CAPEX and OPEX expenses drastically decrease over time.

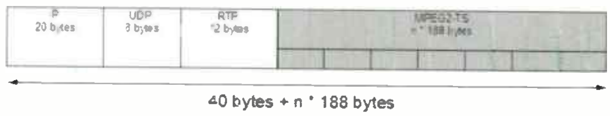
Back to compressed video transport, RFC 4259 (« A Framework for Transmission of IP Datagrams over MPEG-2 Networks ») described the way to encapsulate MPEG2-TS packets in IP datagrams, reused by DVB standardization body (DVB-IPI working group). On top of that, Pro MPEG Forum Code (an association of professionals working around broadcast) released Code of Practice #3 specifying the way to protect MPEG2 over IP data with appropriate Forward Error Correction schemes. Objective of that standard pushed by industrial consortium is to ensure interoperability between products on the market.

IP ADVANTAGES

IP Distribution is a one to many distribution based on the data multicast principle. Multicasting MPEG2 TS over IP makes every transmitters receives the very same content. But IP networks have some drawbacks (delay, jitter, packet loss, packets dis-ordering...), anyway, in the following sections, we will explain how MPEG2 over IP distribution can be suitable as well for SFN operation that require a steady and stable distribution relying on accurate 10MHz and 1 PPS signals.

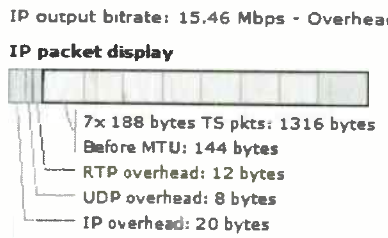
MPEG2 TS TO IP ADAPTATION

Meant as well by MPEG2-TS encapsulation over IP, MPEG2-TS packets are encapsulated into IP packets using RTP. Unicast or multicast addresses can be used according to requirements. Transport service is thus provided by UDP (Checksum) and RTP (Sequencing, Time stamping). Then, IP payload is made of 1 to 7 MPEG2 TS packets.



Overhead of 40 bytes per RTP packet is relatively low (for instance, 3% with 1316 byte payload as shown on next picture).

Pro MPEG Forum Code of Practice #3 adds protection against errors with Forward Error Correction (FEC) scheme. COP3 enables compensation for packet errors and out of order packets using a FEC matrix that is generated by ASI to IP adapter and transmitted on two separate UDP ports. According to FEC settings choice, presence of redundant data can lead to a 10% to more than 50% overhead.



MPEG2 TS FROM IP PACKETS RECOVERY

MPEG2-TS dis-encapsulation from IP requires out of order packets re-ordering, and use of Forward Error Correction (FEC) restore lost/corrupted packets. As said before, as FEC data are transmitted separately from data port (same address but separate ports), a receiver may ignore FEC data if not required.

IP TO ASI DIS-ENCAPSULATION CONSTRAINTS

Gateway must be able to reliably identify target output bitrate and have a straight convergence to the right bitrate value since output stream is to feed target device (eg a transmitter) with the very same bitstream as exactly was at the input of the IP network (ie output of head end). Gateway must compensate network jitter in dynamic, continuous way, which can be achieved throughout the use of a data input buffer in charge of rate analysis, adapting and smoothing. Buffer size used in input usually can be expressed in time: for instance, from 10ms to 500ms. The shortest the buffer size is, the less precise the bitrate estimation will be. The highest the buffer size is, the more exact assessment will be... but with static delay in output.

That's for MFN...

SFN is another matter

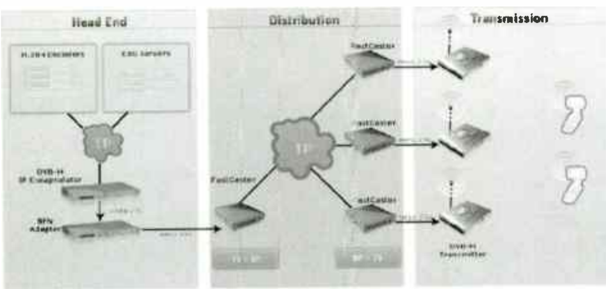
THE KEY FOR A SUSTAINABLE SFN NETWORK WITH IP DISTRIBUTION

SFN needs full control over distribution, no room for inaccurate process !

In SFN, modulation bitrate at the input of the transmitter must be the very same as the one that outputs from SFN adapter, meaning that IP to ASI gateway has to assess bitrate and feed modulators with a steady rate. Gateway's output bitrate must be the same as modulation bitrate, and extremely stable: Transmitters in SFN are not allowed to do bitrate adaptation. The only means to achieve that is to use buffering tricks in gateway, adjusting the input buffer size.

One of the constraints brought by SFN is first the 1 second rule as a maximum time for content transportation from head end to transmitter. Whenever IP or satellite (or both) distribution methods are used to cover a complete set of transmitters in the field, all distribution data MUST be received by every transmitter withing that second. We quoted some typical time for network delays: between 10 to 500ms, sometimes even more.

As a conclusion, input buffer must be used to have the most accurate estimation of bitrate and restitution of MPEG2 stream, but it should take the shortest time since some more room is needed to fit into the network delay value.



To synchronize precisely together, SFN adapter and transmitters have the same clock reference (10 MHz from GPS receiver). A solution can be to have the IP to ASI gateway use the same clock reference (ie the external 10 MHz clock reference from GPS receiver). If not available, an internal 10 MHz clock reference has to be regenerated by gateway with clock drift correction. Such a process requires a lot of know how to enable the same accuracy as an external clock reference, consistent with the clock reference that was used in the head end side... Otherwise, a clock drift will appear and break down SFN synchronization on the transmitter after a while: if this clock drift is constantly increasing/decreasing (even if this is just a few bits per day or week), transmitters' buffer will irremediably run dry or overflow after a while. One should keep in mind that once a transmitter lost synchronization and put RF signal OFF, it can take up to 30 seconds after re-synchronization to get RF signal back again on air...

Thus, gateway must continuously evaluate its internal clock drift and correct it accordingly. Such a complex algorithm requires high grade tests beds and very long-run tests.

Moreover, to avoid SFN network “falling down” after some weeks of running, that feature of internal 10MHz signal re-gerenation is mandatory in case the external 10MHz reference is lost: a switch from external to internal 10 MHz clock reference is mandated to preserve SFN. Gateway must also be able to switch back to external 10 MHz clock reference when it comes back.

CONCLUSION

To date, it is possible to succeed in “SFN over IP”, and this can be achieved with high grade IP gateways featuring the support of both external and internal clock references locking capabilities. Developing those gateways requires high skills in SFN network, distribution over IP and transmitters, complex, expensive and dedicated test beds, very long-run tests (several months on stability and reliability). Thus, only recommendation for that technology is to use mature products that already are validated by operators. As a matter of fact, preserving SFN networks using an IP distribution is already operational in DVB-T/H networks today, already running at broadcast network operators. IP appears to be the solution for the next coming years in SFN content distribution.

IMPROVING MOBILE TV STREAMING: LESSONS LEARNED FOR SUCCESSFUL DEPLOYMENTS

Boris Felts & Ian Locke
Envivio Inc.

ABSTRACT

Mobile TV has become a key, if not the key. application that mobile operators need to deploy to attract new subscribers, or simply retain their existing base. While the business models and killer applications for mobile video are still being researched, mobile operators and broadcasters already face a wide variety of mobile video distribution standards, devices, and applications that deeply affect how those services are deployed.

This presentation will focus on the knowledge gained from early trials of MPEG-4 encoders for Mobile TV. Featured highlights include case studies on the Orange 3GPP and Doordarshan DVB-H deployments, their technical challenges as well as current requirements for successful mobile TV deployments.

Also presented in this session is an analysis of the different standards that are being deployed, their pros and cons, and their impact on how and what services can be offered.

The Big Lesson

As with most new consumer services, while Mobile TV services are enabled by advances in technology, the technology itself is not what the user values although video is still perceived to be a high value driver of premium service adoption. Instead operators deploying commercial Mobile TV services attribute their success to understanding and optimizing the quality of experience, and by focusing on delivering services that meet or exceed the customers' expectations at a price that is perceived to be acceptable.

The Second Lesson

Mobile TV deployments can only be considered a success if they are a net positive to an operator's business. Most operators are looking beyond user subscription fees as the sole revenue source of the service with the addition of targeted advertising revenue and other paid content. Operators also see tangential benefits of improved subscriber adoption, increased penetration of premium services and handsets and overall brand visibility.

Successful Operator Experience

Operators who optimize their video headends to improve the end user Quality of Experience see advantages of user adoption, revenue potential and positive brand awareness. While there is ongoing discussion of the potential for Mobile TV, it is clear that operators who create attractive video services are able to generate substantial revenues and establish brand leadership.

This paper details the encoder specific features and technologies that enhance a MobileTV service by focusing on improving the consumer Quality of Experience.

1. QUALITY OF EXPERIENCE

Except for early adopters of technology, most users are blissfully unaware of the technology that enables the services they consume. To each user the service is judged on how well it enhances their lifestyle in balance with the cost or inconvenience of the service.

Operators offering Mobile TV need to understand these lifestyle benefits, which are predominantly information and entertainment delivery to improve personal productivity and alleviate boredom. There is also a measure of prestige associated with the early adoption of premium video services.

Any benefits for a Mobile TV service will be weighed in balance with the costs or negative users experience associated with the service. If the selection of content is limited, the picture quality is unacceptable, the user experience is unreliable or confusing or the service is perceived to be too costly, then the service will not have good customer traction.

1.1 Video Quality

Given the very low bit rates attributed for mobile services, compression quality and bandwidth constraints are key parameters to take into account.

Although the explosion of successful video services for PCs has proven that consumers will watch video that is less than traditional broadcast quality, it is also true that in a direct comparison between service providers that users appreciate improved video quality. Also, since many operators are rebroadcasting commercial television channels, user expectations for quality are higher than for user-generated content.

Video encoders are the most important component to contribute to the audio and picture quality the user experiences, although it can be negatively affected by the network and the decoder/player in the mobile device. For an encoder to provide maximum video quality it should combine the following features.

- Highest quality signal interfaces for broadcast content ingest
- Optimized audio and video prefiltering and resampling
- Best audio and video codecs (MPEG-4 HE-AAC and H.264)
- Efficient codec with pre-analysis, look-ahead, advanced toolsets and rate controls

Lip synchronization, or audio/video synchronization is a common problem in mobile streaming. Indeed, because of the large variety of audio sampling and video frame rates, the perfect synchronization of the two media is difficult to achieve. Furthermore, the inter-stream synchronization relies on IETF protocols (RTP and RTCP) which are poorly implemented in most of PC/software products

Encoders should optimize the audio samples and video samples with high precision time-stamps as they are captured, thus insuring perfect stream output synchronization.

1.2 Network Performance

Optimal Packetization and Jitter

Mobile networks are extremely sensitive and sending real-time media can be challenging. Any bandwidth variation or jitter in the source streams has immediate effect on the quality of the service.

In Mobile Streaming, the audio and video data is packetized according to the RTP, UDP and IP specifications. The packets transmitted by the encoder are composed of a media payload and a packet header. This has a fixed size and usually represents a small portion of the information transmitted. However, at very low bit rate it becomes more significant.

Also, in most encoders available on the market the media is encoded and then packetized without knowing how the media will be fragmented and encapsulated. This results in bandwidth variations which can generate packet loss, especially under tight bandwidth constraints (see Figure1)

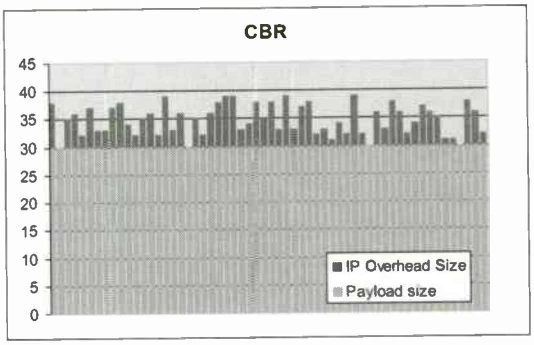


Figure 1: Packetization variations in Constant Bit Rate

In order to solve these problems, *network-aware rate control* provides the following functionalities:

- optimally fragments the data in order to minimize the amount of packet headers
- takes into account packet fragmentation in the codecs to produce optimally-sized packets with improvement of picture quality

Hence for a given bandwidth, more bits are spent towards audio and video streams than packet headers Figure 2 illustrates this in the example of a constant bit rate stream.

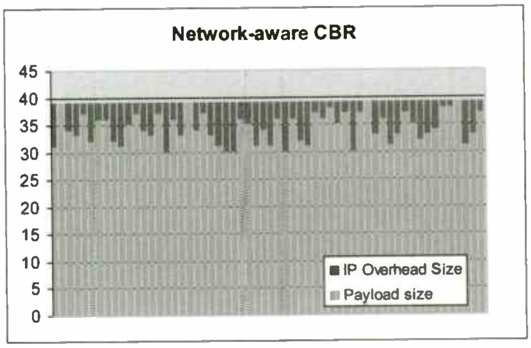


Figure 2: Network-aware CBR

Finally, the source jitter (potential variation of the packet emission times) should be controlled within the encoder. With software encoders on standard PC and operating systems, the packet emission time depends on the scheduling of the operating system. The jitter then increases according to the operating system version, the CPU load, the type of platform and the other traffic.

Encoders should support advanced packetization and source jitter controls to present two significant advantages:

- the video is smoother, because more bits are spent towards quality than towards IP signaling
- the packet loss is reduced, leading to fewer video freezes

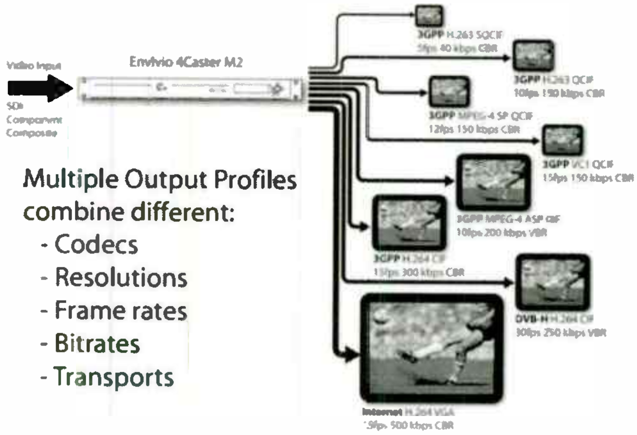


Figure 3: Multi format output

1.3 Multiple Profile Outputs

Due to the large and heterogeneous variety of mobile devices deployed, it is often necessary to output video in various formats, in order to make it accessible on networks such as

- GPRS
- EDGE
- UMTS
- HSDPA

The same video channel needs to be encoded multiple times, with various codecs, resolutions and bit rates. Also, as we have seen in the introduction using a hybrid approach to deploy mobile video services would allow the operators to take advantage of the qualities of the various networks available. This could be for example:

- DVB-H in large network areas
- 3G / HSDPA in more rural environments
- Wireless networks at home

- Live Bit Rate Switching

3GPP PSS release 6 and 3GPP2 "3G Multimedia Streaming Services" specifications have defined a way to communicate network bandwidth and device capability in real-time back to the video server. This new function allows real-time adjustments of the content delivered to the player.

In this case, multiple "elementary" or single rate profiles are contained within one multi-rate profile and sent to the streaming server. According to the bandwidth available between client and server, the latter picks the most suitable profile and forwards it to the client.

Although it is a relatively new feature for MobileTV, service providers should deploy encoders capable of supporting this advanced streaming mode. The video quality is automatically adapted to the network and device capabilities, allowing the user to benefit from the best user experience, regardless of his location and device.

This feature is further discussed in section 1.6.

1.4 Handset Compatibility & Optimization

Advanced interoperability with mobile video players is a key factor for large scale deployments. Vendor participation in industry interoperability events (such as the ones organized by IMTC, ISMA, BMCO Forum) is a good method to improve end-to-end solutions.

However as player implementations may vary among vendors (some vendors implement multiple players), encoders need to support the specificities of each player and cope with them. Operators can then create a set of profiles for encoding the same content to optimize the viewer experience on each device using a combination of:

- Maximum Screen Resolution
- Optimal Codec support
- Player/Hardware Compatibility
- Device/Network Protocol handshaking
- Network Bandwidth

In reality, many of these compatibility issues can only be resolved with field experience where the idiosyncrasies of each device have been addressed.

1.5 Service Reliability

High availability

The reliability of some solutions used for live streaming is extremely limited. This is particularly the case for software encoders installed on PC platform with various capture cards. In most deployments, this type of platform shows unstable behavior over time and leads to frequent and long service interruptions: the unit requires manual intervention to reboot, takes a long time to launch and the configuration might be lost or corrupted.

In comparison, dedicated hardware appliances are highly reliable platforms. The units are designed for 24/7 broadcast operation with a high level of availability. In case of power failure, the reboot time is minimal and the encoder automatically returns to its normal encoding state, using the configuration and parameters previously set.

- Management

Operators need to deploy Mobile TV head ends with an advanced level of failure detections and management. with the management system should integrate error detections (I/O output loss, hardware platform issue), SNMP alarms, logs and redundant IP output. The system should also provide automatic redundancy to decrease the service downtimes and on-going operational costs. Operations such as firmware upgrades should be easily scheduled and require only a few mouse clicks, compared to PC/software upgrades which can be extremely problematic in large deployments.

To further improve service availability, the management system should support multiple redundancy schemes:

- N+1 (N encoders and one automatic backup)
- N+M (N encoders and M automatic backup)
- 1+1 (each encoder is backed up)

Most mobile solutions available on the market do not have this level of redundancy and management. In some cases, only 1+1 redundancy is available, leading to quite large and expensive solutions.

1.6 Ecosystem Integration

Servers

In the case of live streaming, several core functionalities of the streaming server can greatly contribute to improving the quality of service:

- Fast channel switching

The initial player-server session setup is done with RTSP and requires quite a few exchanges. Once the session is established, the player needs to buffer the incoming stream before starting playback. All this can lead to fairly long interruptions when switching channels. In order to reduce this delay, two strategies can be used:

- "caching" of the RTSP session negotiation by intercepting the first RTSP requests and switching source streams
- "acceleration" of the player buffering by sending the first packets in a burst mode. This is sometimes called "instant-on playback"

Advanced architectures use a combination of both, so that the delay for channel switching is reduced.

- Live Bit Rate Switching

As explained previously, 3GPP release 6 defines a way for the player to signal network capacity and device capabilities. According to this information, the server dynamically selects the adequate streams for the player. During a live TV session, if the mobile user travels from an area with high bandwidth availability to an area with lower bandwidth available, he can still watch the same video, at the expense of a slight decrease in video quality. Without this mechanism, the viewer would only be able to watch good quality video in few areas and the service would be interrupted in others.

- Reliable UDP

UDP is an efficient protocol for streaming, but is a "best effort" mechanism: streams are continuously sent from the server without checking if the player received them. On the other side, TCP offers the possibility to check that each packet has been received correctly and to request the missing information. This insures a guaranteed quality of service, but significantly increases the exchanges and network traffic between the server and the player.

A trade-off solution, called *reliable UDP*, aims at increasing the quality of service while keeping the amount of traffic relatively low. With this mechanism, the server sends packets in UDP, and the players sends back information about the potential packet loss. Based on this information, the server can then resend the packets previously lost, or decide to ignore the request because it came too late. While this mechanism does not guarantee the service, it helps reducing the amount of packets lost.

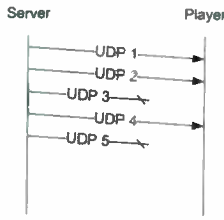


Figure 4: Player-Server exchanges in UDP: packets 3 and 5 are lost

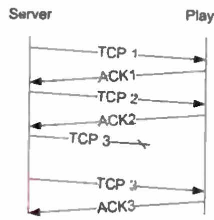


Figure 5: Player-Server exchanges in TCP: all packets are acknowledged

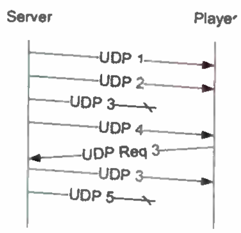


Figure 6: Player-Server exchanges in reliable UDP: lost packets are requested by the player and can be recovered

Conditional Access and DRM

Another important ecosystem partner to integrate is the content protection infrastructure. This includes authentication systems, rights management systems, key generators, stream encryption and client decryption.

In an effort to simplify the head end components some encoders for Mobile TV now offer integrated stream encryption, which has the added advantage of providing additional content security by ensuring that content is protected when it leaves the encoder.

These third party components interoperate using the SimulCrypt standard.

The operator must also be aware that any encryption system deployed in the head end must be supported by the mobile handsets. This usually requires device support for the OMA ISMACryp, which is the predominant standard for mobile content encryption.

1.7 Other QoE Considerations

In addition to the head end specific features discussed above, the operator should seriously consider the valuable lessons learned from successful operators.

Client Navigation & Workflow

The portal through which the end user selects and watches the video is an important contributor to the overall customer experience. Making sure that this user interface is easy to use and does not interfere with the viewing habits of users will significantly improve the consumer's perception of the video service. The portal can also be used to influence viewing behavior towards more profitable content and services.

Attractive Content

It cannot be underestimated that most "fast growing" Telco video services can attribute their success to securing rights to unique and valuable content. Whether it is a library of unique movies or premium sporting channels, if an operator can promote content that users find desirable then they will be more likely to adopt Mobile TV services.

Suitable Pricing

Although users have shown they are keen to consume Mobile TV content, they perceive its value to be less than premium broadcast TV and pricing models offered by service providers should reflect this. There are numerous examples of ingenious and flexible pricing schemes some of which are similar to cable tiered packages and some of which are "all you can eat". What has proven universally unsuccessful is to price video as an extension of existing data plans where the high bandwidth video content is quickly seen as unaffordable.

2. CUSTOMER SUCCESS

Envivio has successfully deployed more than 100 video head ends with mobile operators now delivering Mobile TV services. These cover both 3GPP network deployments as well as DVB-H and ISDB-T mobile broadcast services. We will look in detail at the success of two operators, Doordarshan in India with their DVB-H deployment and Orange in Europe with their 3GPP deployment.

2.1 Doordarshan

Doordarshan, the national television service of India, is devoted to public service broadcasting and runs the largest free-to-air satellite services covering the entire country. Its analog territory broadcast network covers more than 92% of India providing free news, sports and entertainment.

Service Description

The Mobile TV pilot, launched in early 2007 with eight channels to test reception quality of coverage, assess service schemes including advertising and interactive services, as well as gauge consumer expectations. Upon the successful completion of the trial, the channel lineup was expanded to 12 channels for the live deployment. This service delivers the video to Nokia DVB-H enabled handsets.

The current service is deployed and broadcasting in the Delhi market with plans to expand this deployment to three additional cities with broadcast services in Mumbai, Chennai and Kolkata.

2.2 Orange

Orange is the leading integrated operator in Europe offering converged video services. Since 2006, Orange has been the France Telecom group's single brand for mobile, Internet and TV services. As of December 2006, the Group provided services to almost 160 million clients worldwide, of whom two thirds are under the Orange brand.

Key dates

- 2002 : Launch of VOD on PC: First on the French market
- 2003 : Launch of TV/VOD via ADSL: First in French market
- 2004 : Launch of mobile TV: First on the European market
- 2005 : Launch of the first digital recording service on ADSL
- 2006 : Launch of high definition channels on ADSL: First worldwide
- 2006 : Launch of high definition TV via mobile: First on the European market

Service Description

- Rich and varied content: 60 TV channels, 3000 videos, exclusive sports coverage (football, rugby).
- The first operator to offer unlimited access.
- The only operator to offer high definition quality.

Access

TV/Video on Orange mobiles is accessible by over 95% of the population in France. It is available from the Orange World portal, Video/TV section, or by downloading Orange World TV. The viewer just has to have a mobile device equipped with EDGE or 3G technology. 3.6 million Orange customers are equipped with TV-compatible handsets. A high definition TV-compatible handset has been available on the French market since Christmas 2006 (Samsung Z560); the range will continue to grow.

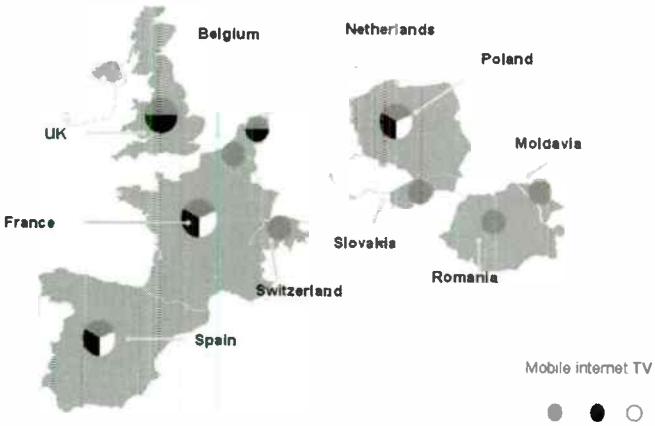


Figure 7: Presence of Orange Video Services

Content

Orange is both a producer and a distributor of the TV/video mobile content broadcast on the Orange World Portal. Three formats are available: classic streaming, TV on demand, or "mobile loop" (e.g. LCI mobile)

- 60 live TV channels
- 3,000 videos
- 8 themed TV categories: General, News, Sport, Entertainment, Music, Youth, Practical, Discovery
- 9 themed video categories: News, Sport, Cinema, Music, Humor, Cartoons, TV/series, Live Cam, and Erotic
- New types of content, created specially for mobile broadcasting: LCI Mobile, "mobisodes" and summaries of the "Plus Belle La Vie" series, Johnny Halliday videos, and "Arthur and the Minimoys" miniepisodes
- Exclusive programs with the Sport option include access to all sports news via continuous programming, sports news and sports TV channels (Info Sport, TV foot, etc.), videos, commentary and text alerts.

Continuous Improvements:

- In terms of quality: access to Orange mobile HD TV since November 2006.
- To access the content rich media TV player available by downloading on compatible handsets
- Increasing personalization, for example by integrating SMS TV/video alerts according to categories chosen by the customer.

Pricing with unlimited options

- Unlimited on weekends for 1 year with Orange Intense contracts
- TV option (€6/month): access to all channels, including unlimited access to 20 and unlimited video on demand
- Total TV option (€10/month) – unlimited access to all 60 TV channels: TV option + over 30 themed channels + unlimited video on demand (whole VOD catalogue)
- TV-music-web option (€12/month): TV option + all music-oriented TV channels and videos + music services and web browsing
- Sport option (€9/month): Sports news + sports TV channels + videos, commentary, sms alerts.
- Total mobile HD TV option (€12/month with no commitment)

The viewer just needs a TV compatible mobile handset to access the 60 channels and the 3,000 video programs. Two pricing methods are proposed: for occasional use: price per volume of data consumed (Kb); for regular use without limit: unlimited options.
(tariffs avalable on www.orange.fr "L'offre")

Statistics

- 51 million videos/TV programs watched in 2006
- Average use of 43 minutes per month and up to 2 hours per month on unlimited offers
- Predominantly indoor usage, notably at home (40%) and between 20% and 30% in public transport or in public places (restaurants, cafés).

3. CONCLUSIONS – KEY LESSONS LEARNED

Operators are now deploying both technically and commercially successful Mobile TV services and these operators have learned to manage a broad range of considerations that will contribute to their ability to deliver, manage and scale video services.

Most important in planning for a successful deployment is to prioritize the overall client interaction and the value users perceive from using the service. This is defined as the Quality of Experience and is measured by more than just picture quality.

Although there are several business and workflow decisions that contribute to the Quality of Experience, the operator should look for hardware vendors that consider not just compression optimization but also address issues of reliability and interoperability to maximize the value users receive from their Mobile TV service.

If users see value in the video services then the operator is able to generate direct commensurate revenues and indirect benefits to the operator's brand. Only if the service provider can recognize these direct and/or indirect benefits can a Mobile TV service be considered a success.

Beyond Mobile TV

Operators are already moving to provide video services to not just their mobile clients but to any subscribers they have on other video enabled platforms such as PCs and residential TVs. This move to the three-screens is seen as the logical way to leverage premium services to new customers or to offer additional premium services to existing customers.

In this move to migrate video services to all client platforms, operators are looking for vendors that can bring simplified head end architecture able to simultaneously address not just Mobile TV but also Internet TV and IPTV. As experts in multi-platform advanced video compression, Envivio sees this demand for a "convergence head end" as the next essential requirement for operators looking to provide ubiquitous IP video services.

Radio Technology Advancements

Sunday, April 13, 2008
1:00 PM – 5:30 PM

Chairperson: Milford Smith
Greater Media, Inc., Lawrenceville, NJ

Radio Broadcasters: Building File-Based Networks

Eric Wiler, Jones Radio Networks, Centennial, CO
Gary Pelkey, Wegener, Duluth, GA

HD Audio Quality and Netcasting

Greg Ogonowski, Orban/CRL Systems, Inc, San Leandro, CA

From ITM to ITWOM: Correcting, Completing, and Updating the Longley-Rice Irregular Terrain Model

Sid Shumate, Givens & Bell, Inc., Haymarket, VA

Can the Public Internet Be Used for Broadcast?

Simon Daniels, Audio Processing Technology, Belfast, UK

A New Approach to Peak-to-Average-Power Reduction for FM+IBOC Transmission

Philipp Schmid, Nautel, Inc., Bangor, ME

Field Tests for Service Area and Handover Service in T-DMB

Sang-Hun Kim, Korean Broadcasting System, Seoul, Republic of Korea

An Improved Coverage Prediction Method for HD Radio

John Kean, NPR Labs - National Public Radio, Washington, DC

*Brazil's Digital Radio Technology Choices

Acacio Luiz Costa , Mix TV Network, Sao Paulo, Brazil

Advances in Digital Measurement Techniques for FM Broadcast

Frédéric Allard, Audemat, Bordeaux, France
Tony Peterle, Audemat, Bordeaux, France

*Paper not available at the time of publication

Radio Broadcasters: Building File-Based Networks

Eric Wiler
Jones Radio Networks
Centennial, CO 80112

Gary Pelkey
WEGENER
Duluth, GA 30097

ABSTACT:
The challenge for national radio broadcasters is to provide programming to multiple locations while at the same time giving the (local) listener the impression of local presence of the station.

Many radio networks today integrate multiple devices at each affiliate station for local ads, station identifiers, local traffic and weather. A more efficient means of achieving these goals is to leverage a centralized solution that uses addressable devices to receive live audio broadcasts, store regionalized to local content and seamlessly combine them to customize broadcasts. This solution can be achieved through a seamless combination of network control, content management and media server technology. Jones Radio Networks is an example of a large radio network utilizing this technology.

In this paper Jones Radio Networks and WEGENER will discuss how an intelligent centralized solution allows radio networks to support the technical and business needs of today's radio networks. The technical portion will discuss how one-time distribution of repetitive material to addressable media servers can be used to enhance national live programming.

It will also discuss how the same technology supports the business aspect of national radio networks to help increase ad revenue through regional to local programming; timeshifting and low cost multicasting.

Consolidating radio services from multiple end broadcast sites to a single centralized location promises dramatic economies of scale, but also presents unique challenges. Success in this space requires replicating all of the local *services* provided by in station personnel, without *replicating those personnel* at the central control facility. The end result should be that the local listener remains convinced that the station he/she is listening to is 'locally derived'. Integrating **File Based** workflows throughout a network and its affiliates is the key to providing such a 'hyper-localized' listening experience, while at the same time dramatically reducing satellite and human bandwidth requirements.

Jones Radio Networks®, Inc. (JRN) is America's leading independent radio programming company with offices and production studios in New York, Los Angeles, Chicago, Washington, D.C., Seattle, Denver, Nashville, and Florida. Through its subsidiaries, the company serves over 5,000 radio stations with a full menu of radio programming and services. Jones Radio Networks began as a 24-Hour music format company serving commercial radio stations. They are now the largest provider of satellite-delivered live formats. Jones Radio Networks uses satellite delivered programming because its ability to cost-effectively provide a choice among highly reliable programming feeds (full-time music services, ad-hoc shows, features, etc.), with essentially hands off operation at the station level to thousands of locations.

In 2004 Jones Radio Networks identified the need to enhance their products and targeting Wegener as their technology partner to develop their next-generation delivery platform. This next generation network architecture incorporates elements of file based workflows and robust management tools in both the centralized satellite uplink location and at affiliate stations.

Stations may choose to receive programming from one of several music formats. Stations also can choose to subscribe to sets of short form programming (such as talk and music shows), which are not part of any particular, dedicated format.

Both of the above are easily achieved with a traditional linear digital receiver and addressable control system. To keep the management of the program switching from exploding in complexity, while still allowing flexible end user selection of formats and long form programs, requires a control system that provides a 'funneling' of user requests into manageable *packages*, which can then be implemented via *switch groups*. In many cases such switching is desired to take place at a specific *local time*, independent of time zones.

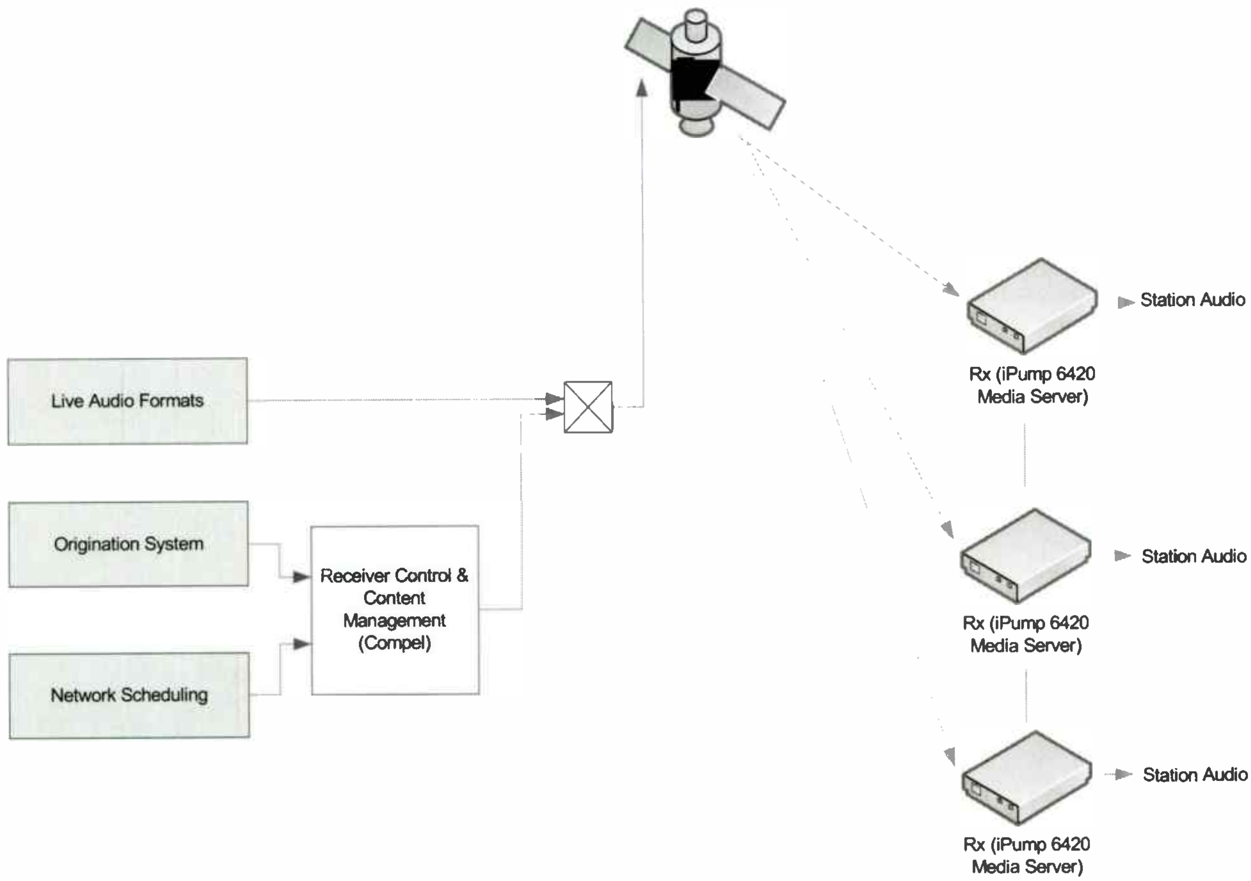


Figure 1: Satellite Distributed Programming

Jones Radio Networks' updated network control system, provided by WEGENER, provides a mechanism to, for example, switch a set of receivers, on Monday through Friday from 'The Clark Howard' show to the 'Neal Boortz' program, at precisely 3:00 <u>local</u> time. Once receivers are moved into or out of their correct switching and time zone groups, all format and switching takes place automatically, without further operator intervention. Thus thousands of stations may receive essentially custom programming without any local personnel, and without an army of centralized personnel.

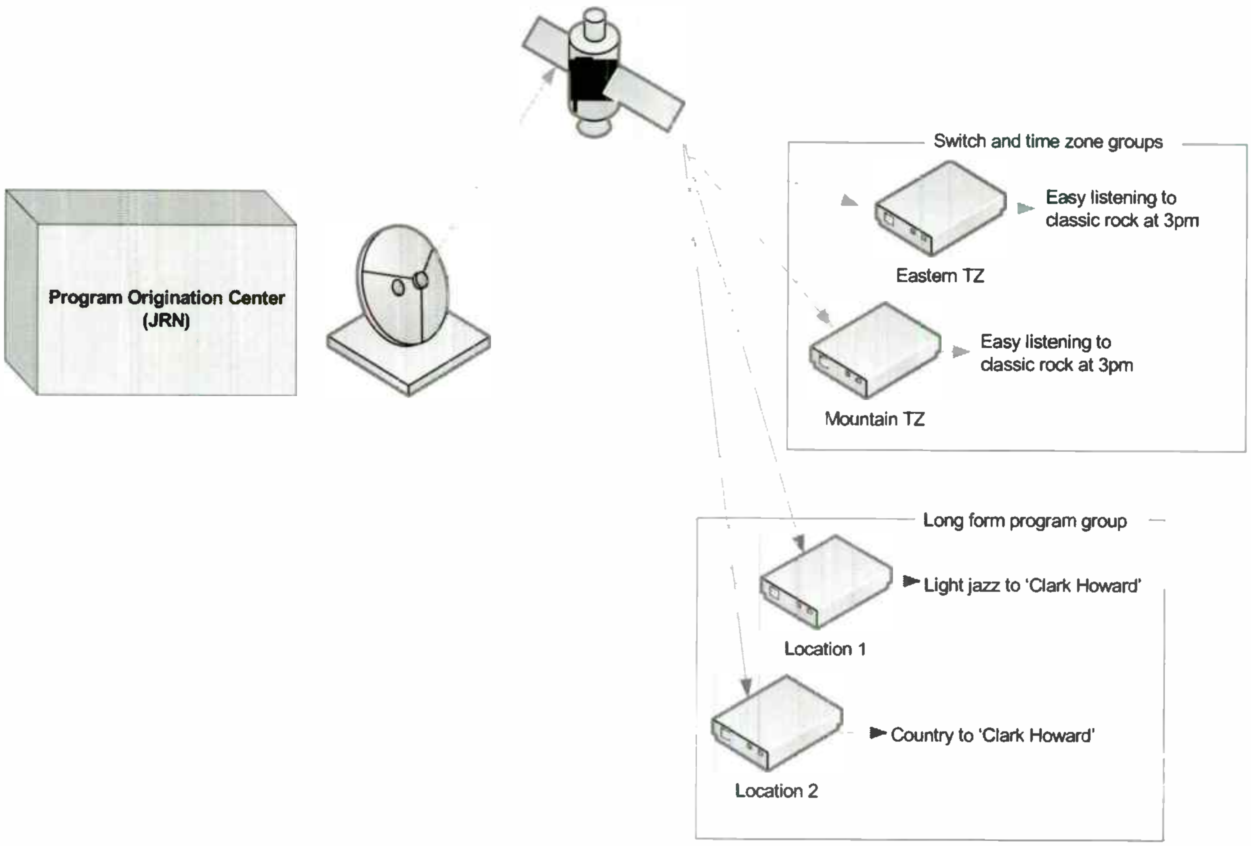


Figure 2: Standard Format and Long Form Program Switching Control

Moving beyond this first set of services requires file based operations. File storage at the local level allows for the introduction of *hyper-localizing* features including liners, regional or demographic based ad insertion, time zone shifting, Show Shifting™ to customized music programming at the *affiliate* level.

Let's discuss Liners first: Liners come in various flavors and styles, from station call letters to promotional inserts to fully produced promotional montages. Liners are the primary means for adding local sounding flavor to an otherwise generic stream in addition to identification of the station to the listener. Previously stations triggered the start of these Liners by utilizing a relay closure to their local automation system. In the Jones Radio Networks system the ability to play the liners from the satellite receiver was a prime target for inclusion in the delivery platform. However, liners must 'play out' of the receivers at the same time, so it is not feasible to provide hundreds of live linear satellite channels to dynamically switch to just so the local user hears his own call letters and city name. That's where storage comes in.

In the Jones Radio Networks system, hundreds of individual liners may be recorded and edited into files at the studio and then downloaded individually into new satellite receivers with on-board media storage (receiver/media server). The correct liner can then be commanded to be played out based on a common trigger from the network studios. These liners can be organized into 'styles' (5 second dry voice, 8 second fully produced, legal ID, etc.), formats and DJs, such that the appropriate liner is played on the right format at the right time. The receiver/media server may even insert the file using an audio 'mix' with the live format audio. Thus when the local Pittsburgh listener hears "This is Phil McCoy playing all Pittsburgh's Classic Rock here on WKRP" over the end of a fading song, the illusion is complete (see figure 3). They do not picture Phil McCoy recording that seven second segment in a recording booth in Denver four months ago.

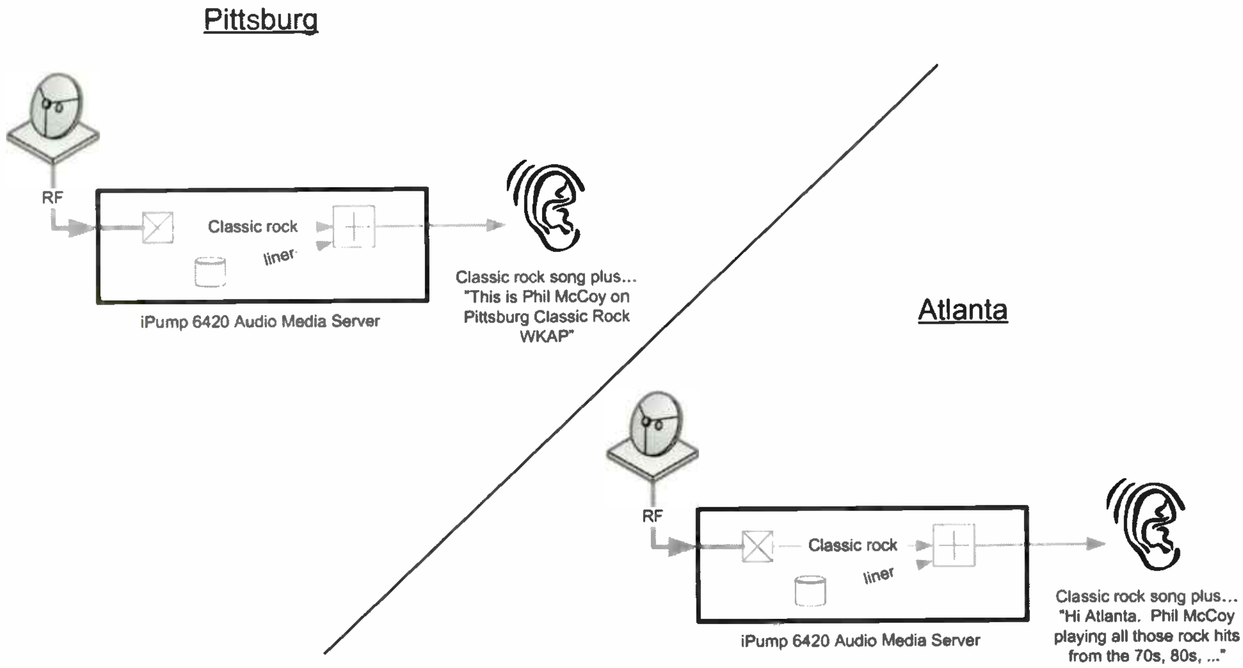


Figure 3: Liner Control

Several elements come into play to make these personalized liners work. For one thing, the network control system and the receiver/media server must have an addressing mode which says, in effect, 'If you are currently playing this audio channel/format, accept this Liner Insertion command'. This is implemented using Wegener's patented "Parameter Addressing" mode, which can be combined with other group addresses (such as time zone groups, mentioned above). Parameter addressing allows addresses to be commanded on a national level which utilizes the receiver's groups to create complex logic. For example, a liner which would say "Good Morning it's 6AM" in the Eastern Time zone would be inappropriate in the Pacific Time zone. By sending a command to a FORMAT AND EASTERN TIME ZONE GROUP address and a completely different trigger to FORMAT AND PACIFIC TIME ZONE, the advantages of a *hyper-localized* system become obvious.

Second, the control system must be able to direct the receiver/media server units to play a liner that was previously recorded *by the DJ that is now on the air.* This is done via a DJ Schedule Table, which is created and maintained at the network operations center. When the network origination system closes a relay to signify to the control system that a 7 second break is taking place on, say, the Country channel, the control system does a quick look up of which DJ is now on duty, and sends this information to the receiver/media server. The receiver/media server uses the DJ information to select an appropriate directory from which to play the liner. This is another example of minimizing network personnel – decisions that can be made ahead of time (which DJ is on duty) are absorbed into the control system and used intelligently, rather than having network personnel make time-critical decisions (what button do I push for Phil McCoy?!) on the fly.

Finally, to keep the presentation 'fresh' to the listener, software is incorporated within the receiver/media server to support a RANDOM PLAY feature. Using this feature, the control system may merely 'point' to a directory (for the correct format and DJ) and the receiver/media server will randomly select one of several liner files from the directory.

Another important 'localizing' feature is the ability to play local or regional advertising. Local stations sell advertising, and require cueing information to properly trigger downstream insertion equipment (see figure 4). The receiver/media server must support this with a bank of relays for each audio output. Select relays are activated, using the above mentioned Parameter Addressing feature, such that a relay will only close if the associated decoder is outputting an audio format that is currently undergoing a 'break'. These triggers can be further restricted to only those sites that have contracted for this feature with Jones Radio Networks, by selectively including them in special 'Local Avail' groups maintained by the control system.

Alternately, ads may be inserted as files directly by the receiver/media server itself, under complete control of Jones Radio Networks. This is done by first loading a spot as a file onto the receiver/media server hard drive. Spots with different content (but with a common name) may be loaded regionally or individually into receiver/media servers. Then at the spot avail time, a single control system command is transmitted, resulting in highly 'tailored' ad output. Using the Internet based return path, spots are verified to have 'landed' successfully on the receiver/media servers, and 'As-Run' logs may even be later retrieved to verify the spots played at the station level.

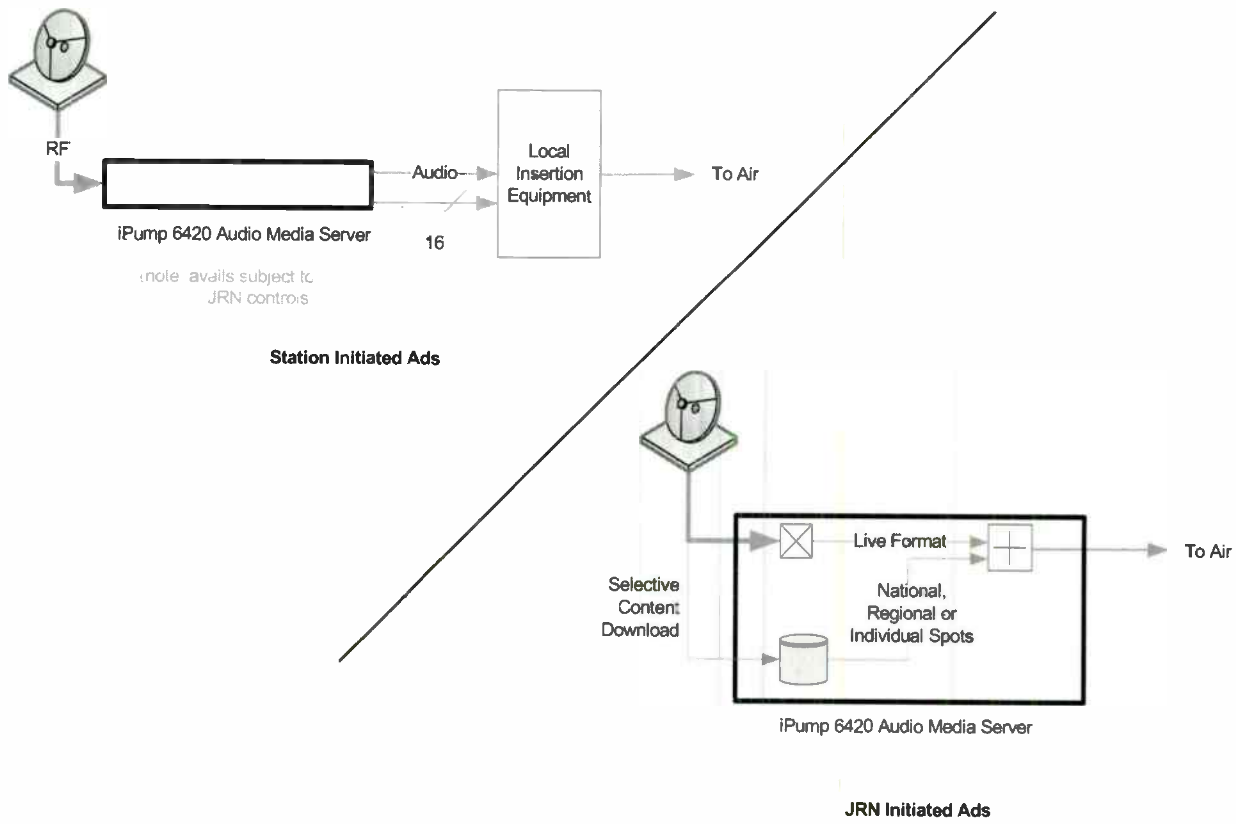


Figure 4: Ad Insertion

The most obvious example for this feature is *regional advertising*, such as playing a snow tire ad in the North, while playing a rain tire add in the South. However, the addressing flexibility and distribution infrastructure allows other types of selective insertions to be performed. For example, ads may be loaded into hard disk 'slots' with content that is based upon end location demographics. So while an ad insertion is taking place in a market where the average listener's age is 28, an entirely different ad is playing out for the market where the average listener is 35, even though both stations may be playing the same music format.

Like liner insertions, listeners at hundreds of different locations can hear different content, highly tailored to their locale, ***without using any additional satellite channel bandwidth*** (other than the one time bandwidth penalty of loading the files onto the hard disk.

The ultimate embodiment of *hyper-localizing* national music formats is the shift of music programming from the network's studios playback system to the receiver/media server. With the ability to group receivers on many levels it becomes possible to vary the programming by regional or local metrics. A receiver/media server can be placed in a group based on the network's own individual requirements. Based on these groups, a playlog of music files may be distributed only to the stations which match the needs of the music programming. The music programming may be adjusted to suit the needs of the particular market and level of service being subscribed to, creating a total radio station solution programmed from another state by Jones Radio Networks.

Two other features made possible by file based networks are *Timezone Delay* and *Showshifting* ™

Time zone Delay allows an end station to completely shift the network broadcast format such that the audio plays out of the unit in exactly the same time relationship (to the local time zone) that it had when it originated (relative to the network's time zone). Thus a typical audio format (e.g. 'jazz') may have all of its songs, DJ banter, liners, PSAs, and even local breaks and PAD (Program Associated Data) shifted forward in the day by a set number of hours.

Note: to do this the receiver/media server records not only the live audio program, but also the PAD data and the local insertion commands, with their correct chronological relationship to the audio. This is all accomplished by a simple one-time setting of the receiver/media server, with no ongoing maintenance required by either station or network personnel.

Showshifting™ is another highly customizing feature of the system. Using it, the station end user (or network operator) may 'point' to a show (such as 'Neal Boortz') and specify that program to play out of the receiver/media server at 9:00 P.M. every evening (rather than the 10:00 A.M. time at which it is transmitted). This may even occur if 'Neal Boortz' is transmitted on another channel than what the receiver/media server is statically set to decode. As with Timezone Delay, all activity associated with that program (internal or external ad insertions, liners, PAD data, etc.) are captured and used in the correct manner as the show is later played out. Showshifting™ yields tremendous satellite bandwidth savings over linear operations. Without it, the network is forced to rebroadcast a show multiple times to allow affiliates the ability to air it at a more desired time slot.

By the creation of a national file based satellite network, it is also possible to provide programming and content to the station, previously delivered via CD or Internet downloading. Quick turn around distribution to the local station's storage contained within the receiver eliminates the trouble associated with overnight or mail delivery of optical media. Internet file downloading, while practical in many cases, does suffer from speed limitations for stations without adequate bandwidth. Weekend programs, commercial inventory, network memos and text files can all be easily recovered from the receiver using a PC and drop-and-drag functionality.

Advanced file based network features are presently at various stages of deployment. With all of the software in the receiver/media server fully downloadable via the satellite link, Jones Radio Networks are free to evolve the system at a deliberate pace, introducing new features as well as optimizing the system and correcting bugs. With the addressable control and store and forward infrastructure in place, further 'localizing' features can be contemplated which meet the needs of an ever changing radio market.

This satellite delivered, file based network architecture can offer affiliates a 'hyper-localized' listening experience. It also offers robust mechanisms at the uplink and affiliate location to continually monitor network performance 24x 7. To maximize reliability, Jones Radio Networks has built-in various playout features such as a hardware *watchdog timer* and LOS (loss of signal) procedures. LOS inserts 'evergreen' audio whenever the desired audio is unavailable, including interruptions to the delayed content (see above Timezone and Showshifting delay modes), caused by problems with the satellite signal or even AC power. A prime example of LOS functionality allows the receiver to locally cover the normal spring and fall sun outages… While only a few minutes per day for 3-4 days per season, sun outages have always required station resources to correct. LOS allows station personnel to focus on other issues. These tightly integrated file based network features are hard to duplicate and even harder to support using 'homegrown' PC solutions.

In conclusion, Jones Radio Networks has identified the need to leverage technology to meet the new demands of the radio industry and deployed a scalable solution which allows Jones to meet the current needs of its advertisers and programming customers but also to grow the capabilities for future technology driven products and services. By partnering with Wegener during all stages of development, Jones Radio Networks' file based network continues to "push the envelope" beyond anybody's expectation and is opening the door to new revenue opportunities.

HD AUDIO QUALITY AND NETCASTING

GREG J. OGONOWSKI
Orban/CRL Systems, Inc.
San Leandro, CA USA

ABSTRACT

The way consumer audio is being consumed is changing rapidly. This includes retail music and broadcast radio. The main driving forces are the technological advances in delivery mechanisms, which range from new portable music players for downloadable music to new streaming players and wireless mobile devices. These devices have become mainstream and are now frequently overlooked features of electronic components like home theater receivers. Streaming is another broadcasting opportunity that has finally reached technical maturity, but its implementations and audio quality issues need to be addressed if it is to be taken seriously by the consumer while favorably competing with the many diverse forms of media delivery. You only have one chance to make a first impression.

BRIEF HISTORY

For more than 75 years, the business model of radio broadcasting has depended on the wide availability of receivers that are universally compatible with the broadcasts. Unfortunately, Internet radio, streaming, and netcasting do not offer such easy compatibility. Instead, while streaming technology was maturing, many software vendors, content providers and content distribution networks created proprietary streaming formats and protocols that required different player software. Except for Microsoft Windows based devices, mobile streaming uses standards-based protocols employing MPEG-4 QuickTime technology, which has its development roots at Apple Computer. This same technology can be used to deliver streams to both computers and appliances, allowing the same protocols and formats to reach multiple devices.

Another problem with Internet streaming has been requiring listeners to download and install player software on client computers. To enforce security policies, many organizations and enterprises now prohibit the casual downloading and installation of software. While Windows Media Player is typically bundled with every copy of Windows, this does not help multimedia users with non-Microsoft computers or operating systems. Compared to currently available MPEG-4 AAC or HE-AAC codecs, Windows Media Player's WMA provides poorer audio performance. MP3 is even worse at the low bit rates commonly used in streaming. Nevertheless, many content providers have chosen to stream via Windows Media or MP3 simply because the player software penetration is high. This foregoes the quality benefits that can be gained by using MPEG-4 AAC/HE-AAC, played through the Orban Windows Media Player AAC/HE-AAC Plugin, RealPlayer, or Apple QuickTime.

The Adobe Flash protocol offers wide cross-platform penetration. Recently, Adobe made Flash compatible with MPEG-4 AAC/HE-AAC. This changes the streaming protocol game again because Flash-streamed MPEG-4 AAC/HE-AAC uses yet another proprietary protocol.

By confusing consumers, incompatible streaming formats and protocols impede the success of this emerging technology while also challenging content providers. In addition to computer desktops, Internet streaming now extends to wireless mobile devices and stand-alone multimedia appliances, so content providers must stream in multiple streaming formats using multiple streaming server types if they want to reach all possible destination devices without compromising audio quality. This could include using Windows Media Streaming Server for WMA streams, SHOUTcast/Icecast2 Streaming Server for MP3/AAC/HE-AAC streams, QuickTime/Darwin Streaming Server for MPEG-4 and 3GPP AAC/HE-AAC mobile streams, Real/Helix Mobile Server for 3GPP AAC/HE-AAC streams, and now Adobe Flash Media Server for MPEG-4 AAC/HE-AAC Flash streams. Multiple format streaming smoothes the consumer experience and facilitates successful media distribution.

BUSINESS

To be successful, Internet streamers must adopt an achievable business model just as terrestrial radio broadcasters did. Although Internet streaming to the computer desktop offers features beyond the simple radio model, like a graphical display which can be associated with the audio content, this capability is not always available to or even desired by the consumer, especially when it leveraged for banner advertising. Consumers perceive Internet banners as graphical spam. In fact, research has shown that the only graphical information that the consumer wants is the stream, artist, and title. Some consumers also want the opportunity to purchase a digital download. Because Internet streaming and satellite radio can display this, many FM broadcasters have recently responded to this competition by implementing RBDS/RDS—a technology that has been available for years.

Internet streaming listeners often minimize their players while working with other applications. If listeners don't see banners, advertisers get no results. Furthermore, many content providers conceal their streaming URLs from listeners to prevent them from bypassing the banner advertising. This makes it difficult or impossible for the typical consumer to listen to these streams on multimedia players, which are starting to gain momentum because they are easy to use without a computer. Therefore, audio advertising, the model that has worked for terrestrial radio since the 1920s, should be pursued for revenue generation.

Certain broadcast commercials do not have Internet clearance. When terrestrial radio stations stream, these commercials must be replaced with other commercials or fill elements. Many content distribution networks have systems to do this but not all of them are elegant. Some of these systems damage audio and stream quality (a bit like the old cable television local insertion systems), causing audio glitches and gross level discrepancies. To smoothly control switching of the audio sources, the system must operate at the streaming audio encoder side, not the streaming server, and must be tightly coupled with the computer sound card used for encoding and audio signal processing.

AUDIO

Internet streaming audio technology has matured to the point where it is fully competitive with terrestrial and satellite radio. By using MPEG-4 AAC/HE-AAC, netcasters and podcasters can deliver broadcast-quality, full-fidelity audio bandwidth at cost-effective bit rates. This creates an HD audio experience that listeners can enjoy on computers, mobile wireless devices and high quality multimedia players. The audio quality can exceed that of terrestrial and satellite radio.

The three most important considerations for a high quality streaming audio netcast are:

- Audio source material
- Audio processing
- Audio codec

It is also crucial to employ personnel who understand both computer networking and professional audio. Many times, a team is needed to get the required results because most computer professionals have limited knowledge of pro audio. Achieving a high-performance, economical stream requires thorough knowledge of both of these technologies in much the same way that specialized expertise is required to maintain radio transmission equipment.

AUDIO SOURCE MATERIAL

A quality stream starts with quality source material. This includes both live and recorded sources. Follow professional audio practices; you can achieve best results by using uncompressed audio sources and allowing generous headroom in the signal path. Music should be sourced from uncompressed (or losslessly compressed) digital files or ripped CD transfers. Achieving clean audio is one of the most misunderstood topics in broadcasting and netcasting. More detailed information on this may be found in the publication "Maintaining Audio Quality in the Broadcast Facility" by Robert Orban and Greg Ogonowski, available for free download from www.orban.com.

AUDIO PROCESSING

Broadcasters have been accustomed to processing audio for AM and FM transmission with transmission audio processors like Orban Optimods. These processors compress dynamic range to make the signal comfortably listenable in noisy environments and to make the best use of the dynamic range limitations of the channel itself. In analog services (like FM radio), this dynamic range varies as a function of reception conditions, being poorest in the fringes. Audio processing therefore also increases the potential coverage area of analog transmissions.

Digital transmissions behave differently. The technical specifications of the transmission system determine the signal to noise ratio. This does not change with the signal strength in wireless transmission (and is even more irrelevant in a wired environment). Internet reception anomalies are typically audio dropouts rather than added noise.

Audio processing in systems with a low noise floor still has several vital functions:

- Compressing dynamic range to accommodate the signal into typical listening environments like autos and homes. In autos, the acoustic dynamic range is severely limited by wind and road noise. In most apartments and multi-family dwellings, the available dynamic range is limited by the need to avoid disturbing family and neighbors with excessive sound levels. In public spaces like busses, subways, and airports, there are a wide variety of acoustic noise sources. There are relatively few environments where the full, uncompressed dynamic range of the original program material is useable or desirable.

- Ensuring a consistent presentation. In radio, program material from different producers is constantly juxtaposed. Yet most successful broadcasters agree that achieving a "major market" sonic image requires an overall consistency of sound texture and spectral balance from source to source. Multiband compression can achieve this. By setting a target spectral balance and automatically re-equalizing program material that does not have this balance, the multiband compression helps the radio station achieve a "big-time," highly produced sound that sounds authoritative to listeners.

- Reducing the peak-to-average ratio of the signal to increase its relative loudness compared to an unprocessed signal normalized to the same peak level. In netcasting, all signals share the same "pipe." Each signal has exactly the same reach. So the only thing that broadcaster can do to stand out from his neighbors (and possible competitors) is to broadcast a louder and punchier audio signal. Experience has shown that a combination of multiband compression and sophisticated peak limiting is the most effective way to do this.

- Improving the intelligibility of substandard program material, particularly news actualities and incoming telephone calls. Properly designed multiband compression can make startling improvements in this material without need for preprocessing in a production studio.

Preprocessing each program element before it is stored on a playout system is not as effective as preprocessing the mixed audio on the program line immediately before it is streamed. The latter technique maximizes the smoothness of transition between program elements and makes voice from announcers or presenters merge smoothly into the program flow, even if the announcer is talking over music.

It is important to understand that AM, FM, or TV audio processors that employ pre-emphasis/de-emphasis and/or clipping peak limiters are highly inappropriate for use with the perceptual audio coders used in netcasting. The pre-emphasis/de-emphasis limiting in these devices unnecessarily limits high frequency headroom. Further, their clipping limiters create high frequency components — distortion — that the perceptual audio coders would otherwise not encode. It is therefore not desirable to use an FM audio processor for netcasting.

An audio processor like the Orban Optimod-PC 1100 Professional PCI Sound Card consists of several cascaded stages. These may include *input conditioning*, including defeatable highpass filtering and defeatable phase rotation; *stereo enhancement*; *multiband gated AGC*, with target-zone window gating and silence gating; *equalization*, which may include program-adaptive high-frequency enhancement; *multiband compression* in several frequency bands, depending on the processing structure; and *look-ahead limiting*.

A *highpass filter* removes low frequency noise that can contaminate some recordings and microphone chains. This noise can otherwise cause problems with the rest of the audio processing and with the codec, which should never waste its bit budget by encoding noise. The *phase rotator* makes speech more symmetrical, reducing its peak-to-average ratio by as much as 6 dB without adding nonlinear distortion as fast peak limiting does. Hence, phase rotation can be very useful for loudness processing of speech when maximally clean sound is desired.

There are a number of *stereo enhancement* technologies available. Orban prefers one based on its patented algorithm that increases the energy in the stereo difference signal (L–R) whenever a transient is detected in the stereo sum signal (L+R). By operating only on transients, this algorithm increases width, brightness, and punch without unnaturally increasing reverb (which is usually predominantly in the L–R channel). Gating circuitry detects "mono" material with slight channel or phase imbalances and suppresses enhancement so this built-in imbalance is not exaggerated.

AGC compensates for varying input levels. This is particularly important nowadays because many CDs are aggressively processed for loudness when they are mastered. Older CDs having the same peak levels can have average levels (which correspond approximately to loudness) more than 10 dB below the average levels of current product. This means that the common technique of normalizing audio files for the same peak level causes huge problems with source-to-source loudness consistency, particularly when old and new material is juxtaposed. AGC can go a long way toward smoothing out such inconsistencies.

Equalization applies processing similar to "tone controls" to the signal but typically allows the user to adjust the tonal balance in a much more detailed way. EQ has two purposes in a broadcast processor. The first is to establish a signature for a given station that brands the station by creating a "house sound" by subtly emphasizing the bass, midrange, or high frequencies. The second purpose is to compensate for the frequency contouring caused by the subsequent multiband compression and limiting. These may create an overall spectral coloration that can be corrected or augmented by carefully chosen fixed EQ before these multiband dynamics stages.

Multiband compression and limiting may occur in one or two stages, depending on the developer. If it occurs in two stages, the multiband compressor and limiter can have different crossovers and even different numbers of bands. If it occurs in one stage, the compressor and limiter functions can "talk" to each other, optimizing their interaction. Both design approaches can yield good sound and each has its own set of tradeoffs. Usually using anywhere between four and six bands, the multiband compressor/limiter reduces dynamic range and increases audio density to achieve competitive loudness and impact. It's common for each band to be gated at low levels to prevent noise rush-up, and developers often have proprietary algorithms for doing this while minimizing the audible side effects of the gating.

A *look-ahead limiter* controls the peak level at the output of the processor to prevent clipping the codec. The limiter prevents overshoots by examining a few milliseconds of the unprocessed sound before it is limited. This way the limiter can anticipate peaks that are coming up.

Any peak limiter (even a clipper) can be modeled as multiplying its input signal by a gain control signal. This is a form of amplitude modulation. Amplitude modulation produces sidebands around the "carrier" signal. In the case of a peak limiter, each Fourier component of the input signal is a separate "carrier" and the peak limiting process produces modulation sidebands around each Fourier component.

Considered this way, a hard clipper has a wideband gain control signal and thus introduces sidebands that are far removed in frequency from their associated Fourier "carriers." The "carriers" hence have little ability to psychoacoustically mask the resulting sidebands when compared with the sidebands that a look-ahead limiter introduces because the look-ahead limiter's gain control signal has a much lower bandwidth. Therefore, compared to a hard clipper, a look-ahead limiter produces considerably less audible modulation distortion. This is particularly important when one is driving a low bitrate codec because one does not want to waste precious bits encoding this distortion.

Simple wideband look-ahead limiting can still produce audible intermodulation distortion between heavy bass and midrange material. Advanced-technology look-ahead limiters use sophisticated techniques to reduce such IM distortion without compromising loudness capability.

Beware of codec pre-processor potions or tonics that claim to remove codec artifacts. Codec artifacts are caused by fundamentally very different signal processing than that used to control gain, equalization, and peaks. Therefore the only way to keep codec artifacts to a minimum, especially with low performance codecs, is to insure that the audio signal processing is accurately aware of all the frequency content being encoded, and to not exaggerate any frequencies that would irritate the codec. Most current generation multiband audio processors do not meet these criteria. The Orban Precode™ technology used in Optimod processing does, by using a patented high-order crossover to maintain audio control consistency while being aware of the entire audio spectrum. Without this, the subsequent final peak limiter is required to do more work, further reducing audio quality.

The best modern audio codec design has taken into account the fact that many modern recordings are highly processed with a very low peak to average ratio because of "loudness wars" mastering. If the audio codec in use cannot meet quality expectations when fed by such material, then the only way to improve codec performance is to compromise the audio fidelity of the source feeding the codec by attenuating frequencies that cause codec artifacts, usually high frequencies. This reduces the quality of the transmission chain even further compared to a chain using a more advanced codec.

AUDIO CODECS

When trying to achieve the best audio quality, it is crucial to select the audio codec carefully. Some codecs sound so inferior that we believe they insult the artists who have created the content that plays through them.

The basic principle of perceptual coding is to divide the audio into frequency bands and then to code each frequency band with the minimum number of bits that will yield no audible change in that band. Reducing the number of bits used to encode a given frequency band raises the quantization noise floor in that band. If the noise floor is raised too far, it can become audible and cause artifacts.

A second major source of artifacts in codecs is pre- and post-echo caused by ringing of the narrow bandpass filters used to divide the signal into frequency bands. This ringing worsens as the number of bands increases, so some codecs may adaptively switch the number of bands in use, depending on whether the sound has significant transient content. This ringing manifests itself as a smearing of sharp transient sounds in music, such as those produced by claves and wood blocks

.

Psychoacoustic Models

Perceptual coders exploit complex models of the human auditory system to estimate whether a given amount of added noise can be heard. They then adjust the number of bits used to code each frequency band such that the added noise is undetectable by the ear if the total "bit budget" is sufficiently high. Because the psychoacoustic model in a perceptual coder is an approximation that never exactly matches the behavior of the ear, it is desirable to leave some safety factor when choosing the number of bits to use for each frequency band. This safety factor is often called the "mask-to-noise ratio," measured in dB. For example, a mask-to-noise ratio of 12 dB in a given band would mean that the quantization noise in that band could be raised by 12 dB before it would be heard. (That is, there is a safety margin of two bits in that band's coding.) For the most efficient coding, the mask-to-noise ratio should be the same in all bands, ensuring that the sound elements equitably share the available bits in the transmission channel.

Coding Efficiency

Different sounds will vary greatly in the efficiency with which a perceptual coding system can encode them. Therefore, for a constant transmission bitrate, the mask-to-noise ratio will constantly change. Pure sounds having an extended harmonic structure (such as a pitch pipe) are particularly difficult to encode because each harmonic must be encoded, the harmonics occupy many different frequency bands, and the overall spectrum has many "holes" that are not well-masked, so that added noise can be easily heard. The output of a multiband audio processor that uses clipping is another sound that is difficult to encode, because the clipper creates added distortion spectrum that does not mask quantization noise well, yet may cause the encoder to waste bits when trying to encode the distortion.

AAC and HE-AAC Codecs

AAC is intended for very high quality coding with compression up to 12:1. The AAC codec is about 30% more efficient than MP3 and about twice as efficient as MP2.

With common program material, the AAC codec can achieve "transparency" (that is, listeners cannot audibly distinguish the codec's output from its input in a statistically significant way) at a stereo bitrate of 128 kb/sec, while MP2 requires about 256 kb/sec for the same quality. The MP3 codec cannot achieve transparency at any bitrate, although its performance at 192 kbps and higher is still very good and has been progressively improved over the last decade.

AAC stands for Advanced Audio Coding. Intended to replace MP3, AAC was developed by the MPEG group that includes Dolby, Fraunhofer (FhG), AT&T, Sony, and Nokia—companies that have also been involved in the development of audio codecs such as MP3 and AC3 (also known as Dolby Digital™). (AAC does not stand for Apple Audio Codec, although Apple was one of the first to implement this technology with the introduction of Apple iTunes, the most successful downloadable music source, and QuickTime.)

The Dolby/Coding Technologies "Spectral Band Replication" (SBR) process can be added to almost any codec. This system transmits only lower frequencies (for example, below 8 kHz) via the codec. The decoder at the receiver creates higher frequencies from the lower frequencies by a process similar to that used by "psychoacoustic exciters." A low-bandwidth signal in the compressed bit stream provides "clues" to modulate these created high frequencies so that they will match the original high frequencies as closely as possible. Adding SBR to the basic AAC codec creates aacPlus, which offers the best subjective quality currently available at bitrates below 128 kbps. At bitrates below 128 kbps, full subjective transparency cannot be achieved at the current state of the art, yet the sound can still be very satisfying. (In the phraseology of the ITU 1 to 5 subjective quality scale, this means that audible differences introduced by the codec are judged by expert listeners to be "detectable but not annoying.")

Dolby/Coding Technologies HE-AAC/aacPlus v2, the latest in MPEG-4 Audio and previously known as "Enhanced aacPlus," is aacPlus coupled with the MPEG Parametric Stereo technique created by Coding Technologies and Philips. Where SBR enables audio codecs to deliver the same quality at half the bitrate, Parametric Stereo enhances the codec efficiency a second time for low-bitrate stereo signals. Both SBR and Parametric Stereo are backward- and forward-compatible methods to enhance the efficiency of any audio codec. As a result, aacPlus v2 delivers streaming and downloadable 5.1 multichannel audio at 128 Kbps, near CD-quality stereo at 32 Kbps, excellent quality stereo at 24 Kbps, and great quality for mixed content down to 16 Kbps and below. MPEG standardized Coding Technologies' aacPlus as MPEG-4 HE-AAC (MPEG ISO/IEC 14496-3:2001/AMD-1: Bandwidth Extension).

With the addition of MPEG Parametric Stereo (MPEG ISO/IEC 14496-3:2001/AMD-2: Parametric coding for high quality audio), aacPlus v2 is the state-of-the-art in low bitrate standards-based audio codecs. The Dolby/Coding Technologies codecs provide the absolute best possible sound per bit the current state-of-the-art will allow, without the typical resonant, phasey, watery character of older-technology codecs.

Streaming and file audio encoders are available that use the genuine Dolby/Coding Technologies HE-AAC/aacPlus codec. Because some such encoders support both RTSP/RTP (Real and QuickTime) and HTTP/ICY (SHOUTcast and Icecast2) standards-based streaming servers, and because free player clients are available from Orban, Real, Winamp, and many mobile phones, there is little reason for netcasters not to consider using HE-AAC/aacPlus to deliver quality audio to the increasing sophisticated Internet streaming audio audience while saving on bandwidth costs thanks to the HE-AAC/aacPlus higher efficiency than MP3 and WMA. AAC/HE-AAC is changing the way streaming audio is perceived and consumed.

FROM ITM TO ITWOM: CORRECTING, COMPLETING, AND UPDATING THE LONGLEY-RICE IRREGULAR TERRAIN MODEL

Sid Shumate
Givens & Bell, Inc.
Haymarket, VA

INTRODUCTION

In 2006, at the National Association of Broadcasters (NAB) Broadcast Engineering Convention, we presented a new, sophisticated yet simple to use method of predicting and displaying FM radio (including HD Radio) reception. This new reception presentation system report provided the background, including adjacent channel interference and noise criteria thresholds, and a sophisticated methodology to use, in order to create simple-appearing FM reception quality maps. While deceptively simple in final appearance the automated calculation and analysis underlying the new mapping system required a high-accuracy, high quality, specialized analysis system incorporating Longley-Rice calculations, to analyze potential co-and adjacent channel interference for all stations in a market, and then provide resulting displays on a individual station basis. This analysis includes the situations most critical to HD radio, the various single and dual adjacent signal interference scenarios, as well as noise floor considerations.

To achieve this as a viable, reliable product, we found it necessary to create our own, customized Longley-Rice predictive software system. We upgraded an already sophisticated, all-points considering, open source Linux wrap-around software package to operate with the official standard Longley-Rice core subroutine set, and with a dual database, and distributed free copies of this software (SPLAT with PLOP) at the 2007 NAB BEC presentation.

But we were very disappointed with the results; the SPLAT wrap-around software, with the dual-database PLOP modifications, worked as advertised; but the old core software did not respond the way the NTIA ITM documentation stated that it should. The unexpected, but necessary next step was to dig into the core.

THE CORE, AND THE WRAP-AROUNDS

The Longley-Rice Irregular Terrain Model used by the FCC and by radio engineers worldwide consists of two main components. The first is the core; the National Telecommunication and Information Agency's (NTIA) Irregular Terrain Model. (ITM). This over-quarter-century-old set of computer subroutines, primarily based on the original Longley-Rice methodology in Tech Note 101, was originally written in FORTRAN IV (ANSI 66) and is now available in a Windows DLL-compatible c++ version. It has recently been slightly updated to version 7.0; however, the official version used by the FCC is version 1.2.2.

The second component is wrap-around input-output and graphic display interface software that processes input information and prepares a terrain database to input to the core. After the core is run, the wrap-around software takes the output from the core and processes it for display and print out. Many versions of this wrap-around software exist, including sophisticated commercial software sold by several companies, freeware Windows software, an open-source Linux package made available under the GNU GPL license, a very basic Windows freeware application from the NTIA, and, of course, custom software written by or for end users. But all of these are wrapped around the same old core that the NTIA's annual reports have described since the turn of the century as having "remained essentially unchanged since 1982".

THIS IS THE THIRD OF TWO REPORTS

At the 2007 BEC, we presented and distributed free copies of an expanded open-source wrap-around input-output software package. Our intention and schedule at this time was that this was to be the unveiling of the commercial availability of the new reception prediction calculation and display methodology. However, we found that the core subroutines were not responding properly to the input data, and would not produce good results for 3-arc-second terrain databases, and produced no better results using dual terrain databases than with single databases. We presented little more than a progress report, as we realized that it was necessary to analyze and repair the NTIA core.

THE ROTTEN CORE

Not just one, not just a few, but many problems were found. Here are a few of the more significant problems:

1. Dysfunctional least-squares line fitting; the zlsql subroutine can't even draw a straight line correctly.
2. Subroutines that were written based on a manual calculation method. Direct conversion of the manual calculation method in Tech Note 101 into computer code, resulted in a small, fixed number of intervals on a radial that can be considered before a computation drifts off or fails. This ranges from a (hidden) fixed limit, to an algorithm that gradually gets worse as the number of intervals increase. As a result, the more detailed the terrain database used and/or the longer the path, the worse the results.
3. Multiple errors were found in calculating the take-off angles, terrain irregularity factor, and terminal effective heights.
4. Confusion over the meaning and coding of "log". In common notation today, the use of "log" denotes a common logarithm, a logarithm function to the base 10, used for calculating in decibels. To the user of a scientific calculator, the "log" button also calculates a base 10, or common logarithm. To a FORTRAN coder in 1966, the only logarithm function available was ALOG, which calculates a natural logarithm, what the "ln" button on a calculator does. To convert to a base 10 logarithm, the FORTRAN coder had to multiply the results by 0.434. Later, the function ALOG10 was added, but the original conversions remain in the code. The original FORTRAN code was later ported (converted) to c++ code, including all the original errors, with only one new one added. In c++, however, as in FORTRAN, a natural logarithm is written as log, and a base 10 logarithm is coded as log10. So the confusion continues; the radio engineer says "make that a twenty times log of the frequency function", thinking of a common logarithm, and the computer engineer and c++ coder writes: 20*log(f).
5. Two key subroutines were found to have never been completed to properly calculate path loss when obstructions occur in the path prior to the sum of the two estimated maximum horizon distances, due to failure to recalculate the horizon distances when an obstacle is found. This is one of two problems that trace back all the way to omissions in the original Longley-Rice methodology in Tech Note 101.
6. The other omission is that the cause of the additional attenuation found in the line of sight range, in addition to free space loss and two-ray multipath cancellation, is not identified, nor is a means provided for its calculation. The computer code incorrectly substitutes a diffraction calculation to estimate this attenuation; a methodology of questionable validity that becomes absurd when an obstacle blocks the line-of-sight path. The correct method to use to determine and calculate this attenuation is based on Radiative Transfer[3]. Johnson and Schwering[2] revealed this in 1985, but that was after development of the ITM methodology had ceased. In addition, Donald Barrick's observations[1] about the behavior of reflected signals were not considered or incorporated.
7. While the ITM model documented in Tech Note 101 considers large obstructions that end the line of sight path, the computer implementations in ITS-67[6], and ITM versions 1.2.2 through 7, never fully have. Despite the fact that the point-to-point subroutine in ITM computer implementation versions 1.22 through 7 reports out that obstructions are being considered prior to the horizon, and that the mode has switched to (post-obstruction) diffraction, in fact, the instructions in the lrprop subroutine, where the work is actually performed, did not fully incorporate the new concept instruction set when lrprop was assembled from the subroutines in the earlier ITS-67 smooth-or-slightly-irregular-terrain code,

and does not allow full consideration of obstructions until the distance "dlsa" is passed. The distance 'dlsa" is the distance from the transmitter to the calculated smooth-earth transmitter horizon, added to the distance from the receiver to the calculated smooth-earth receiver horizon.

8. The computations in subroutine lrprop originally generated averaged and truncated results, perhaps because phenomena that produce results that are not contiguous and cumulative make statistical analysis more difficult, and the original goal of the Longley-Rice solutions were to produce statistically valid average results. The two-ray calculation results were modified to eliminate the comb-filter effect nulls found between the transmitter site and the point where the two signal paths lengths are within half a wavelength. For the line of sight, solutions were computed for locations at three points in the line of sight range, using a weighted combination of two-ray and diffraction computations. A straight-line formula, y = ax + b, was then fitted to the three data points. The distance to the location was then used to solve this formula for the location being considered. So it was never possible to accurately pick a spot and make a reliable prediction of received signal strength at that specific location; the results were always greatly averaged to produce results valid for reception in the general neighborhood. Today, we have engineers in the field, holding a GPS unit in one hand and a DTV field strength meter in the other hand, wanting an accurate prediction of field strength. HERE.

WHY, IN MORE THAN A QUARTER OF A CENTURY, WERE NONE OF THESE PROBLEMS DEBUGGED?

First, this is a more than 1,500 lines of old, dense code, much of it written in the days of punch cards and 64kilobit memory, when every instruction was gone over and over to see if it could be consolidated or if other tricks could be used to multitask and consolidate each line of code. There is one line of code, for example, in subroutine alos that combines at least three equation functions, and takes fourteen typewritten pages to explain. And, it has an incorrect implementation of a logarithmic function in it. The engineers who originally understood and wrote this software are long since retired and gone. When this software, originally written and documented in ITS-67 for the army in 1967, was rewritten over a decade later, in the late 1970's and early 1980's as the Longley-Rice Irregular Terrain Model, it appears that subroutine lrprop, which was assembled from several earlier subroutines found in ITS-67, was assembled without benefit of the understanding of the changes that needed to be made to convert it to a Irregular Terrain with Obstructions model.

The ITM software, despite its many shortcomings, has become a standard by which other radio reception prediction software is judged; clearly, its caretakers were cautious and conservative in considering any changes to what has now become a FCC-sanctioned (with reservations) standard tool for predicting DTV reception and solving relocation and move-in disputes over city grade coverage.

WHAT HAS BEEN DONE TO CORRECT THE PROBLEMS?

The ITM software has been analyzed from end to end. Each problem found has been documented and addressed. Most could be corrected with relatively minor changes to the existing software, with two exceptions: the alos and lrprop subroutines. These two subroutines required a major rewrite, and the addition of a new subroutine, saalos, to address problems 6 and 7 mentioned above.

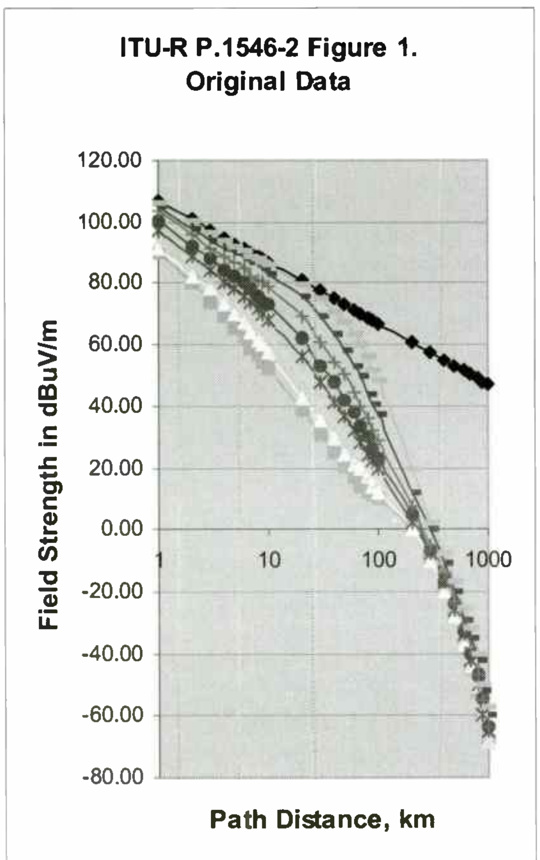
The main problem in rewriting alos and lrprop was that there was no good existing fast-calculating set of approximation equations for the Radiative Transfer Functions, which were needed to replace the diffraction calculations used in the line-of-sight range.

THE NEW RADIATIVE TRANSFER LINE OF SIGHT EQUATIONS

To develop and test the needed equations and methodology, as a side project, we undertook an analysis of the massive amount of empirical data embodied in the over-land Figures 1, 9, and 17 of ITU Recommendation P.1546-2. From this we developed a set of deterministic approximation equations that, when assembled on a computer spreadsheet, adequately duplicate the P.1546-2 results. This provides both a practical and useful tool, and a proof of concept and accuracy for the equations.

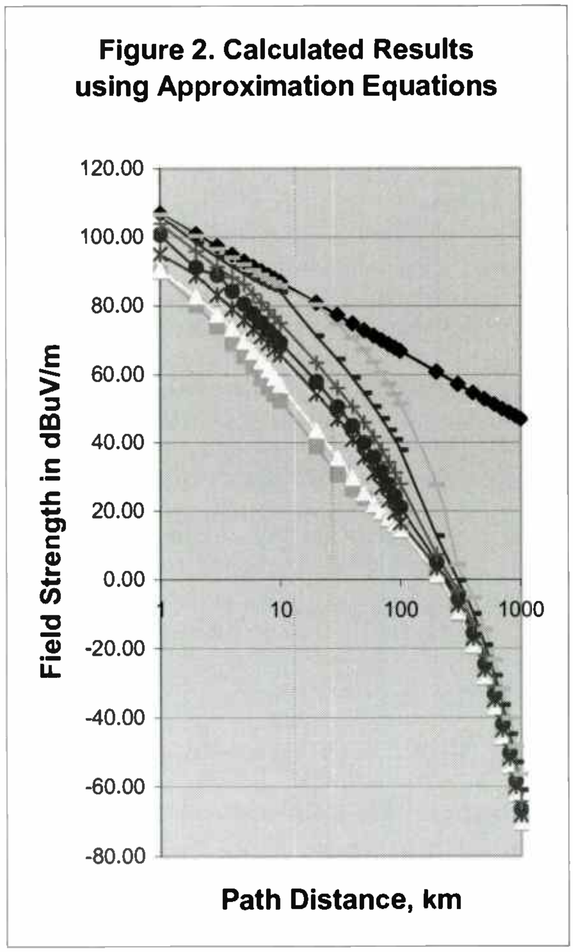
I have prepared a paper describing and explaining these new equations and methodology, "Deterministic Equations for Computer Approximation of ITU-R P.1546-2", this paper has been proposed for presentation before the 10th annual International Symposium on Advanced Radio Technologies, at the NIST Boulder Labs, in Boulder, CO, June 2-4, 2008. Members of Working Group 3 of the ITU, which oversees recommendation P.1546-2, are expected to be in attendance, as they are meeting in session at Boulder Labs, following the Symposium.

A comparison of the match between the following two charts demonstrates the accuracy of the results from these new equations. Figure 1., below, shows the original ITU data from P.1546-2 Figure 1 on a computer spreadsheet-generated chart. These are the curves of the field strength, for various transmitter heights, for over-land paths at 100 MHz, 50% of the locations, 50% of the time, taken from empirical (measured test) data.



The top, straight line is the free-space line, the minimum loss possible in a line-of-sight radio path. Each line below represents the field strength along a path, each for a different transmitter height. The lowest line represents a transmitter height of 9.35 meters AGL, the next line up is 18.75 meters, proceeding upward to 37, 75, 150, 300, 600, and 1,200 meters AGL.

Figure 2 shows the equivalent curves calculated using the new set of deterministic approximation equations (Shumate's Approximations). The point of showing both charts is this; the closer the two charts are to each other, the better. The new approximation equation results match up very well with the original, empirical data.



The original ITM subroutines have been reviewed in depth, both basic and advanced mathematical errors have been corrected en masse, and the subroutines have been revised and completed to properly consider major obstructions. The misapplication of diffraction loss calculations prior to an obstruction have been replaced, to properly calculate Radiative Transfer Engine attenuation due to the presence of clutter in the line-of-sight range, using Beer's Law as a foundation for propagation calculation, and Snell's Law-based trigonometry to calculate the radio signal path geometry.

Figure 3, below, shows the usual results for a Longley-Rice field strength mapping for WETA (FM), in the Washington, DC area, using SPLAT with a single database and the ITM core:

Figure 3; WETA (FM) using SPLAT wrap-around sofware with old ITM core: (images reduced to 340x340 pixels from 2400x2400);

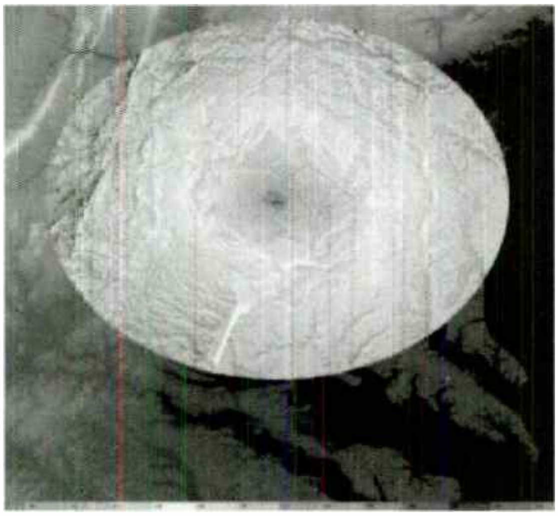


Figure 3.
Results with ITM versions 1.2.2 through 7.0 cores. (1982 to 2007)

The increase in the amount and depth of signal variation detail, especially in the line-of-sight range, is visually noticeable, even at this reduced size.

This presentation will be the first unveiling to the general public of the results of the Irregular Terrain with Obstructions Model (ITWOM), in operation. Fully operational results are to be presented at the NAB Broadcast Engineering Conference Presentation on Sunday, April 13, 2008.

Figure 4; WETA (FM) using SPLAT wrap-around software with new ITWOM core; preliminary results for the same exact run with the ITM core replaced with the new Irregular Terrain with Obstructions (ITWOM) core subroutines:

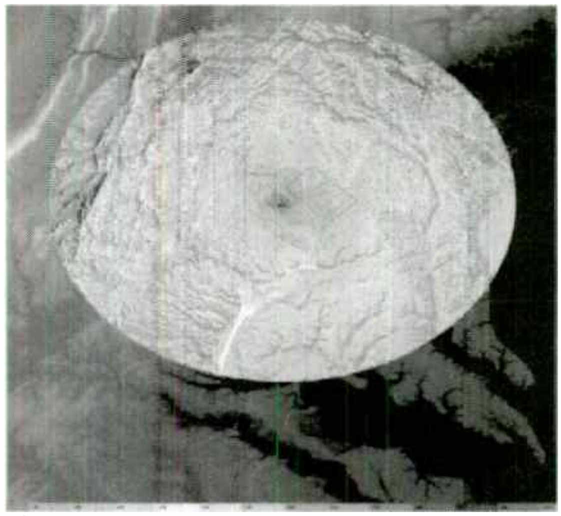


Figure 4.
Early results with ITWOM ver. 0.3 core. (2008).

REFERENCES:

1. "Grazing Behavior of Scatter and Propagation Above Any Rough Surface" by Donald Barrick. IEEE Transactions On Antennas and Propagation, Jan.1998.
2. "A Transport Theory of Millimeter Wave Propagation in Woods and Forests", by R. A. Johnson and F. Schwering, Center for Communications Systems. Research and Development Technical Report CECOM-TR-85-1, U. S. Army Communications-Electronics Command, February 1985.
3. "A Generic Model of 1-60 GHz Radio Propagation through Vegetation – Final Report" by N.C. Rogers, A Seville, J Richter, D Ndzi, N. Savage, RFS Caldeirinha, A.K. Shukla, M.O. Al-Nuaimi, K. Craig, E. Vilar and J. Austin. Prepared by Dr. Neil Rogers, Principal Investigator, OinetiQ, for a consortium comprised of QinetiQ, the Rutherford Appleton Laboratory, and the Universities of Portsmouth and Glamorgan, for the U. K. Radiocommunications Agency. First Issue, May 2002.
4. "Transmission Loss Predictions for Tropospheric Communications Circuits", Volume I, National Bureau of Standards Technical Note 101, Revised January, 1967, by P.L. Rice, A. G. Longley, K.A. Norton, and A.P. Barsis, Institute for Telecommunications Sciences and Astronomy, U. S. Dept. of Commerce National Telecommunications and Information Administration.
5. ITU Recommendation P.1546-2, 2005.
6. "Prediction of Tropospheric Radio Transmission Loss Over Irregular Terrain, A Computer Method –

1968" ESSA Technical Report ERL 79 – ITS 67, by A. G. Longley and P.L. Rice, Institute for Telecommunications Sciences, U. S. Dept. of Commerce National Telecommunications and Information Administration.

CAN THE PUBLIC INTERNET BE USED FOR BROADCAST?

Simon Daniels
Audio Processing Technology, Belfast, UK

For the forthcoming Olympics in Beijing, several European Public Service broadcasters will take home audio via IP. Some will go over the well managed, low-jitter, low delay EBU Fine network with Net Insight Nimbras in combination with satellite feeds, others will use the public internet.

With IP Audio Networking moving into the mainstream of audio broadcasting, radio station engineers are faced with a wide choice of IP links over which to transfer their content. Options include dedicated IP LANs, contended lines and the public internet using DSL, FiOS, WiMAX, 3G. Issues such as network latency, connections and protocols arise and examples of various approaches in different circumstances can help us make informed choices.

The foundation of the broadcast industry for many years, synchronous networks have been considered the industry standard for audio transport worldwide. Balanced analog copper circuits, microwave and synchronous based systems such as V.35/X.21 or T1/E1 have been the traditional choice for studio transmitter and inter-studio links in professional audio broadcast networks. However, the reign of synchronous links as the preferred choice is currently coming under threat from a new challenger in the form of IP based network technology. While IP technology does have some disadvantages for audio transport, the benefits over existing synchronous networks are increasingly proving too persuasive for broadcasters to ignore: Cost , Greater flexibility, Greater scalability, Widespread availability, Network efficiency and Resource efficiency.

While common broadcast practice dictates the need for dedicated lines, broadcasters are realizing that remote broadcasts can be conducted using contended services such as the public internet at substantially less cost and effort. Historically, ISDN has been readily available from all major service providers, it's popularity has been due to the fact that it offers reliable, point-to-point and bidirectional communication at guaranteed data rates. Broadcasters are seeking IP alternatives and in the absence of temporary network contracts the Public Internet is an attractive option. Long term commitments to ISDN vary from country to country, and these need to be considered in planning alternatives. We also need to consider the different stages of development of secure use of the Public Internet in various countries.

Increasingly Radio Broadcasters are looking at making the transition from synchronous networks to IP networks for their distribution and contribution links. As a completely different transport mechanism presenting different advantages and disadvantages to the broadcaster, IP audio transport poses many new challenges. Despite the scalability, cost-efficiency and flexibility of IP networks, packetized audio transport is not a perfect solution and broadcasters must take care that the quality of their audio and reliability of their audio delivery is not sacrificed in the interest of cost savings. There are many important issues that broadcasters need to consider before embarking upon IP migration and the Public Internet can be the ultimate example of the problems associated with IP.

This paper will look at a number of examples of remote broadcasts over various IP links and examine the key points in ensuring their success. We will talk about issues such as jitter and latency and provide recommendations regarding essential features on IP codec equipment. The experiences of broadcasters trialing audio over the Public Internet can inform future choices and generate awareness of the pitfalls and possibilities associated with using the Public Internet for essential broadcast links

Why Use The Internet for Broadcast?

The technology is there - so of course we are going to use it! As any technology presents itself, broadcast engineers will consider how it can be used. What problems does it solve, what opportunities does it open up and what less convenient technology can it replace? This is the same for the Public Internet as for the use of analogue, POTS, T1, ISDN or IP.

Transporting broadcast audio across the internet is certainly a challenge, but it is also an opportunity and in some cases may be a necessity. Traditional resources for remote broadcast are not always available. POTS may not be convenient or the infrastructure may be old. ISDN has long been the most suitable resource, with guaranteed bandwidth and bidirectional communication, but it is not always available. It is becoming increasingly difficult to obtain new ISDN installations in many parts of the world and in some countries there are scheduled dates for the shut down of ISDN services.

There may also be positive opportunities to broadcast higher quality audio or from new locations. IP links may provide enough bandwidth for stereo or even surround sound where previously ISDN could only deliver mono. A 3G link into the internet can be used to broadcast from previously inaccessible locations.

IP is attractive to broadcasters as it offers a number of advantages over other mediums. It can be easily scaled to meet requirements, is readily available and offers a good ratio of bandwidth to buck.

In many respects the public internet is the most extreme variance of IP as a solution. It offers a low cost and high availability solution, but it is also high risk because there can be factors outside our control. We are faced with the same challenges as on any IP network but with the additional question - what's behind the network cloud?

CHALLENGES IN IMPLEMENTING IP NETWORKS

Connections and Protocols

While not attempting to cover anything like all the connection types and relevant protocols, it is worth taking a minute to review some of the most commonly available, and some of those that have been used in the case studies in this paper.

DSL was originally implemented as part of the ISDN specification, giving it an often forgotten historical link to an accepted broadcast format. Domestic variants often run from as little as 512 kb/s but much higher data rates are possible. The variant ADSL, Asymmetric Digital Subscriber Line, offers a high down load speed but a low up load speed, this may be ideal for a home user down loading (legal) mp3 files but is rarely useful for broadcast applications. At the broadcaster's end of the connection, a DSL modem or DSL router converts digital signals into a voltage signal of a suitable frequency range which is then applied to the phone line and connects to a synchronised DSLAM at the phone company site. Attenuation can be a problem if the distance from the DSL modem is too great. If a DSL router is used it is likely to combine fire wall functionality and the broadcaster may need to enable communications to and from specific ports and IP addresses.

FiOS (Fiber Optic Service) will be familiar to many US broadcasters. It is a telecommunications service providing fibre optic connections direct to a location and comes in both domestic and business oriented packages. It typically offers better data rates than DSL, generally 2MB/s, a figure that will be very familiar to anyone who has used the European synchronous standard E1.

Satellite and Microwave technology can be harnessed for IP and internet connections. WiMAX is a standards-based technology enabling the delivery of last mile wireless broadband access as an alternative to cable and DSL. It can be used to provide bidirectional high data rate links across distances in excess of 50 km and includes QoS and security systems.

The ITU mobile telephony standard 3G has been widely adopted by phone companies in the USA, Europe and South East Asia, and its availability is growing in other parts of the world. 3G uses .6 MHz channel carrier width to deliver higher data rates tan the old 2G standard. The motivation for the telecommunications companies has been the delivery of mobile web content and various charged services, but 3G can provide internet connections for broadcast audio. 3G connections have been successfully used by a number of broadcasters including Z100 WHTZ New York.

Transport protocols

TCP, Transmission Control Protocol, uses a systems of positive acknowledgment with re-transmission. Each packet sent from Point A to Point B is acknowledged and if no acknowledgement is received at Point A then the packet is re-transmitted. This offers some security of transmission but at a cost. The procedure sometimes incurs long delays and can produce large additional amounts of network traffic causing bottlenecks and creating rather than solving problems.

UDP, User Datagram Protocol, has a "send and forget" policy. This dumb sending may not offer security but is faster and more efficient in its bandwidth use. It is also suitable for multicasting which TCP is not since it requires packet acknowledgment between points.

RTP, Real-time Transport Protocol, provides headers with payload identification, sequence numbering and time stamping.

Most practitioners have come to the conclusion that using a combination of UDP and RTP is most effective for broadcast audio links over LAN and WAN. This is also likely to make sense on public internet connections but TCP is also possible and we shall see examples of both.

Network Protocols

IP, Internet Protocol, provides a unique global addressing system that works across a mix of networks including Ethernet, ATM, Wi Fi, etc. ARP, Address Resolution Protocol, works in conjunction with IP, providing IP to Mac address resolution, in other words tying the IP address to the physical identity of a particular piece of equipment. The IP information is necessary for Routers and Switchers to operate on a network and deliver the correct packets to the correct destination.

Systems such as SIP provide a user friendly front end to IP addressing, allowing the user to work with names of locations or connections rather than worrying about twelve digit IP addresses. Other times, we have to set IP addresses for connections. We may need to check in advance the IP address of units we need to connect to. We may also need to be aware of rules relating to IP addressing, for example the ranges of IP addresses designated for use in multi-casting.

Data Rates, Packets & Bandwidth

Looking at IP connections the headline data rate figure can appear high but the reality can be different. On any contended network there is no guarantee of a consistent data rate throughout the connection, and lower data rates mean less lost packets. To achieve a safe data rate audio compression such as MPEG I/II Layer II or Enhanced apt-X is generally required. We also need to consider the impact of packet size.

In synchronous networks the audio bandwidth equals the bandwidth required to transport that audio but in IP we must add a packetisation overhead. Each packet of data contains Ethernet header bytes containing information relating to routing as well as header bytes for UDP or TCP and RTP. Packet headers are examined by routers directing data as well as by the receiving unit.

Most codecs will allow the user to set the packet size. Using a large packet size reduces the overall bandwidth requirement because the proportion of header data to audio data is small. but this has the adverse effect of increasing latency because the time for each packet to be filled with data before transmission is increased. There are also risks involved in using larger packet sizes, some network routers may fragment large packets and the impact on the listener of a dropped packet increases as the packet size increases.

Using a small packet size reduces the impact of a dropped packet and reduces latency but adds to the overall bandwidth requirement because the proportion of header data to audio data is relatively high.

it is worth considering that the lower our bandwidth requirements the less risk of lost data! Experimenting with packets size may help.

Network Latency

A number of factors contribute to latency across a network or internet connection. Encoding time, packetisation delay, network transmission latency, and the buffer time included in case of jitter.

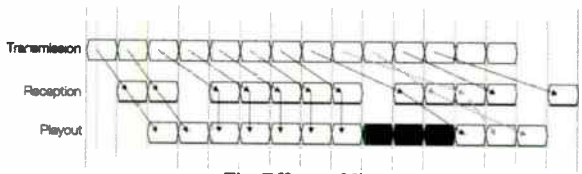
We can minimise some of these factors, for example by using non-frame based audio compression algorithms we minimise encoding and decoding time, small packet size minimises the packetisation delay. If we are using the public internet we cannot control the delay across the network but we may be able to avoid the worst cases with a combination of experience and planning how to connect. A jitter buffer is always necessary but as we shall examine, it only needs to cover variations in latency, not the total latency. It is good practice to measure network latency at various times and under various conditions, if we experiment and keep good records we can develop a picture of the minimum jitter buffer settings that can be used.

Sometimes using the public internet for broadcast audio will only be appropriate for simplex connections because we will not be able to reduce the latency to a point at which it is comfortable to use for duplex communication.

Jitter

A feature of any packet switched network is the fact that a packet can take any route from source to destination. On a simple LAN there may be only one route from our transmission point A to our reception point B. On a larger LAN or a WAN there can be several possible routes. If we use the public internet, we cannot see behind the network cloud, we don't know how many possible routes can be taken. More significantly, we don't know how long each route will take.

In an ideal model network, the time a packet takes to go from point A to point B will be consistent. In practice, if one packet takes one route and the next packet takes a different route they are likely to take different amounts of time for the journey. This can lead to packets arriving out of sequence or a gap between one packet and the next.



The Effects of Jitter

It is necessary to use a jitter buffer so that the play out of packets can be consistent and in order even if variable latency across the network causes packets to arise out of order or with gaps. If jitter increases above level allowed for by the jitter buffer, then packets are dropped resulting in corrupted audio. Experimentation and experience will determine the amount of buffering required.

Packet Loss

All packet based delivery systems are susceptible to dropped packets resulting in dropped audio. We can negate the effects of dropped packets by choosing smaller packet sizes but this incurs greater packetisation overheads and therefore bandwidth requirements. Re-sending packets using TCP has consequences with regard to bandwidth and latency. FEC appears attractive but adds significant extra bandwidth and is not 100% effective even with a full implementation.

Equipment Considerations

Given the number of variables that can come into play using the public internet for broadcast audio, and the constantly changing network conditions, it is very important that our equipment gives us effective network monitoring tools. We need at a glance status of codecs throughout the network and clear monitoring of network performance through parameters such as buffer status and sequence errors. Alarms are as important now as ever, and Alarm and Event Logs can help diagnose persistent network problems. Depending on the circumstances in which we need to work other considerations may come into play such as equipment size, ease of configuration through the use of SIP or Speed Dials. Of course reliability will always be important and we will have our own views on what is appropriate, to quote Bernie Courtney of Z100 New York: "I didn't want a computer disguised as a broadcasting device".

CASE STUDIES

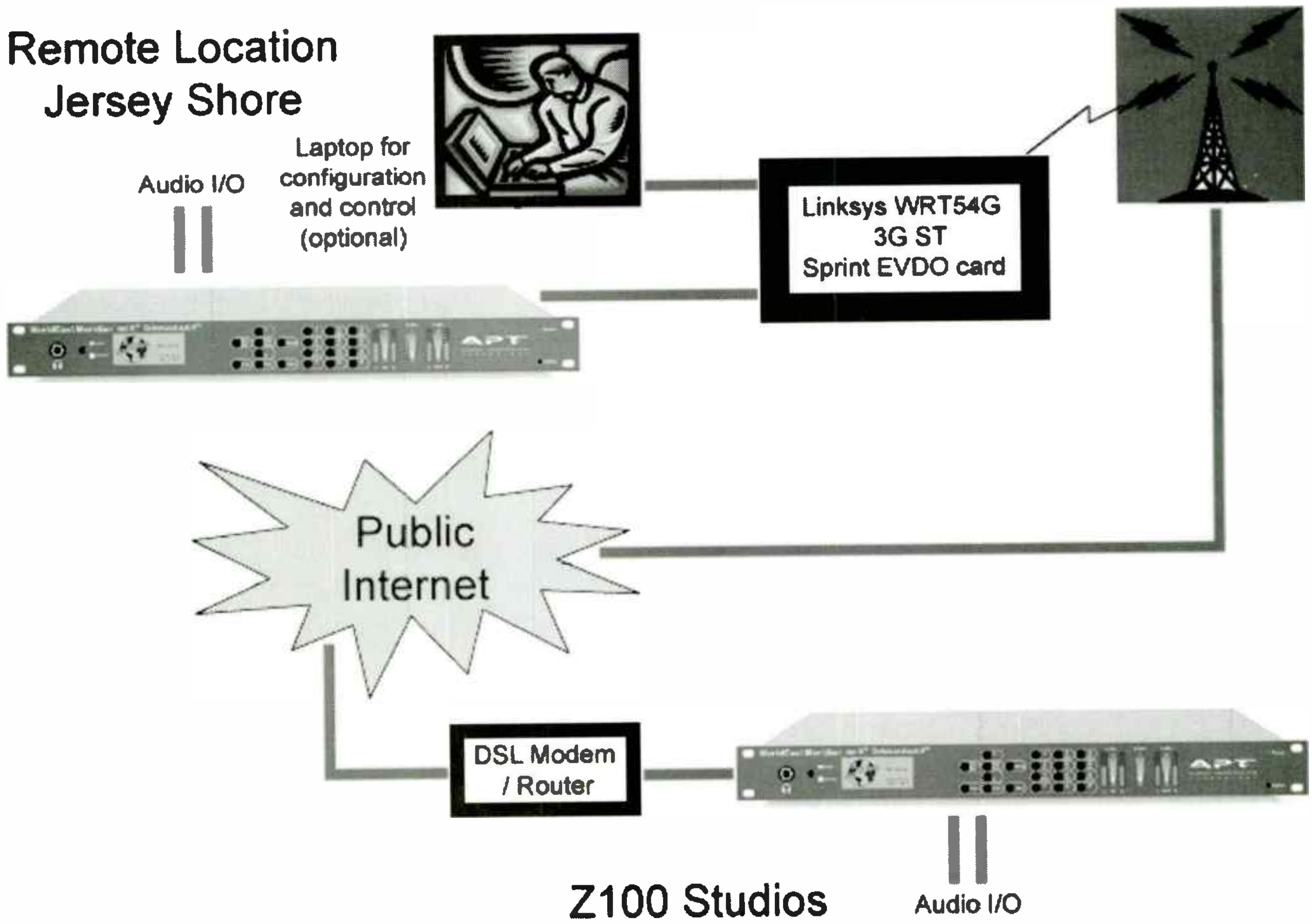
Z100/WHTZ New York

Bernie Courtney, Metro Broadcast Services

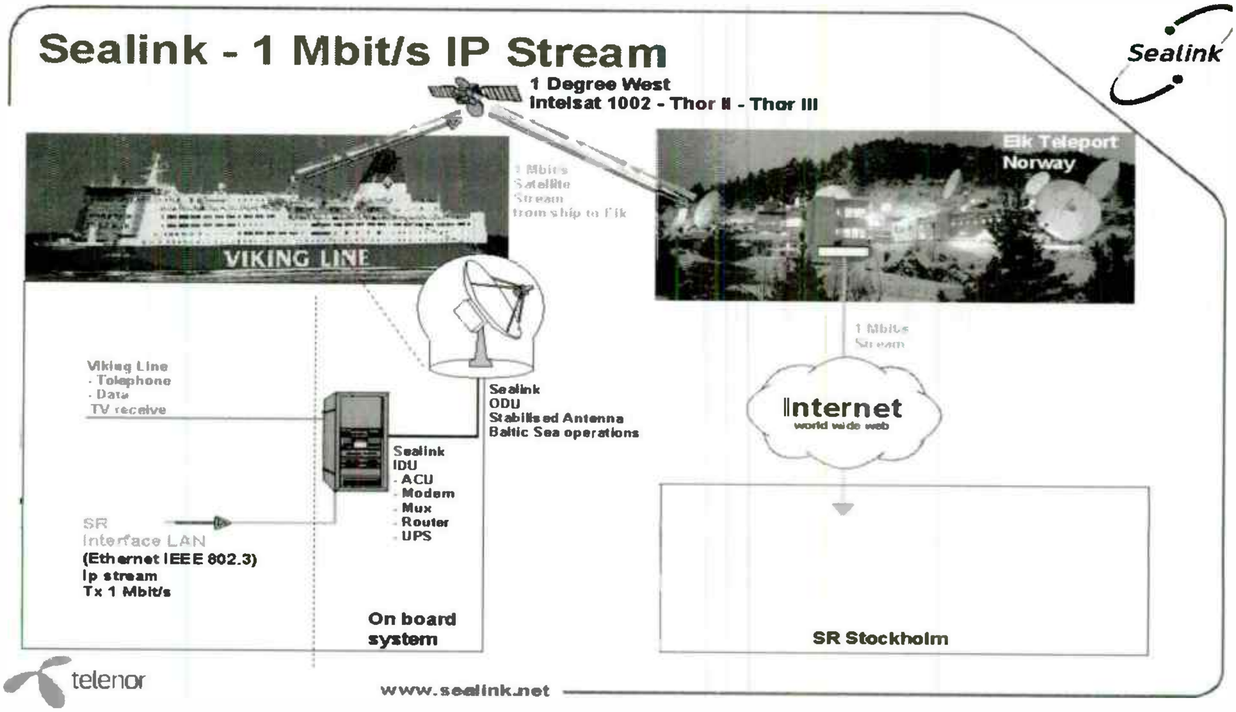
Happy Hour broadcast from a nightclub at the Jersey Shore for the past two years. Most of the copper infrastructure is old, years of salt air takes it's toll on the copper wiring and results is noisy lines that do not mesh well with POTS codec's. The local telephone company recently deployed FiOS, optical fiber. With ISDN becoming harder to get, Bernie decided a fireworks display would be an ideal test bed for evaluating IP technology in 'prime time' using a 3G connection from location to the network.

Bernie ran tests on the codecs between cable modem connection at home and the studio before they were installed on the shore the following Saturday. The installation process was straight forward, after setting up the appropriate port forwarding on the router on each side of the connection they were up and running. In subsequent visits, Bernie just plugged in, pressed the Speed Dial he had created, and they were on the air. The POTS gear was taken along just in case but never left the back of the van.

After the EVDO was set up, Bernie pinged the studio codec's IP address to check the latency on the link. Wireless data latency can vary widely from cell site to cell site. Latency on the EVDO is usually between 250 to 400ms, a 500ms buffer was used to give a margin of error. Enhanced apt-X compression was used running in mono with a data rate of 128kb/s.



Z100/WHTZ New York



Swedish Radio - Sealink Broadcast

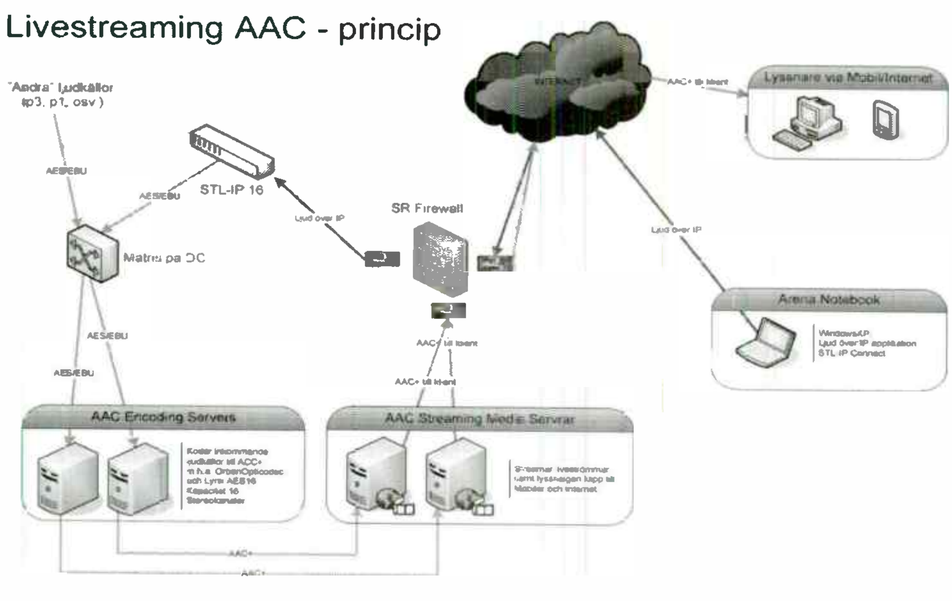
Swedish Radio

Lars Jonnson & Björn Westerberg

A live transmission was made from a Viking Line Ferry, a software encoder on the ferry compressed the audio using Layer III stereo 256 kb/s and UDP over a satellite link to a Norwegian downlink station and from there back to SR Stockholm via the internet.

The buffer size was set to 300 ms but the overall latency was greater. The satellite hop created about about 300 ms delay and the internet link to Stockholm added about another 50 ms. Layer III had probably 70-100 ms encoding time. It was a one way transmission so the delay was not critical.

The duration of transmission was one hour, there were some minor drop outs of unknown cause.



Swedish Radio - Moscow Stockholm Link

Swedish Radio also used the internet to broadcast 20 hours of Ice Hockey live from Moscow (diagram on preceding page). For the Moscow transmission SR used MDO software on PCs.

The buffer size was 300 ms and the coding used was Layer II at 256 kb/s, stereo. The delay was probably around 200 ms with the internet and the coding added together. It was only used one way, with no conversation, so the latency was acceptable.

SR transmitted 10 Ice Hockey matches, more than 20 hours of non-interrupted audio. ISDN was kept as a back up, but 20 hours of ISDN, Moscow-Stockholm, would have been a considerable sum. AAC was only used for streaming to the internet or to mobile phones.

Radio Sunshine
Lance Eichenberger

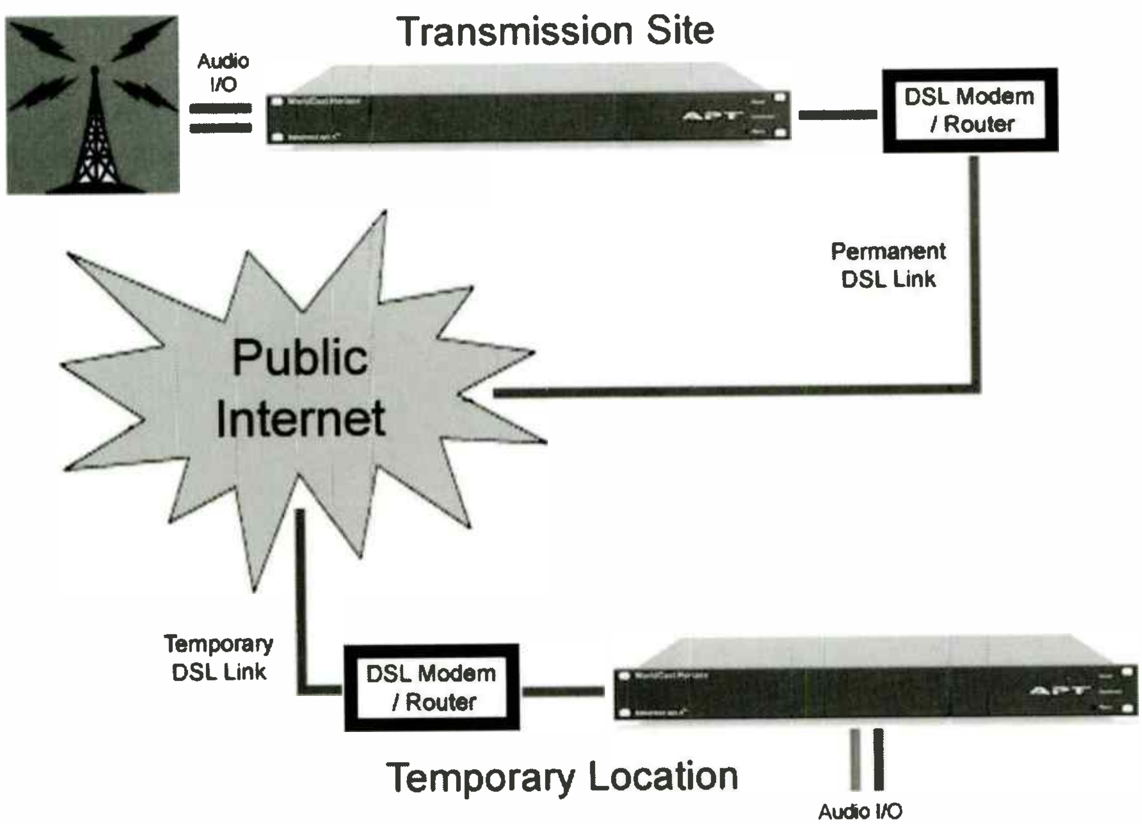
Radio Sunshine conducted the first ever public field trial of HD-Radio® in Europe and see IP as a natural choice to use for HD Radio programming. Radio Sunshine also broadcast on FM and provide special broadcasts to cover events in the locality.

Zurich City Council provides a fixed transmission installation to deal with Outside Broadcast events. The facility has links to one or two FM transmitters and is rented out on a per week basis to various organisations. The site has a permanent DSL line with a fixed IP address.

Radio Sunshine installed one of their WorldCast Horizon codecs at the city council site. The other WorldCast Horizon is a roving unit that can be placed at any location where an event needs to be broadcast, and can communicate back to the first codec using a simple DSL connection, installed on a temporary basis if necessary. This enables Radio Sunshine to broadcast from concerts, fairs and other one off events.

Connection is generally duplex with the return feed used either for talk-back or confidence monitoring. When broadcasting in HD there is a significant delay, and live two way conversation is not possible. But when working in FM mode, Radio Sunshine do use the links for duplex communication. Low delay coding (under 2ms) is important given the inherent delay in IP networking. This enables Radio Sunshine to conduct live interviews over IP links.

Aux data is used to transport RDS-Data to the transmitter site or for control of remote equipment from a lap top via RS232 or USB.



Radio Sunshine - Location Broadcast

Broadcast Associes
Sylvain Cavet

Broadcast Associes have been experimenting with using WiMAX technology to provide a digital and interactive wireless IP link and have been testing the set up on DAB Broadcasts in Paris.

The connection is point to point, operating in the 5GHz frequency range. Exterior WiMAX modules use a simple CAT 5 connection to the interior studio. Link coverage can be up to 50 km and bandwidth up to 50 Megabits. The WiMAX protocol supports QoS and various security measures including WEP encryption and Mac Address protection.

Latency on the Broadcast Associes link is 200 ms in 64 QAM Modulation mode with MAC Address protection. Buffering is set to 20 ms as the latency on the link is consistent. The tests have been running 24Bit 48kHz audio using Enhanced apt-X to achieve a 576kb/s data rate. Listening tests have been very positive and test broadcasts indicated no packet loss.

Acknowledgements

Many thanks to the broadcasters who have contributed their experiences and allowed me to reproduce their diagrams and descriptions:-

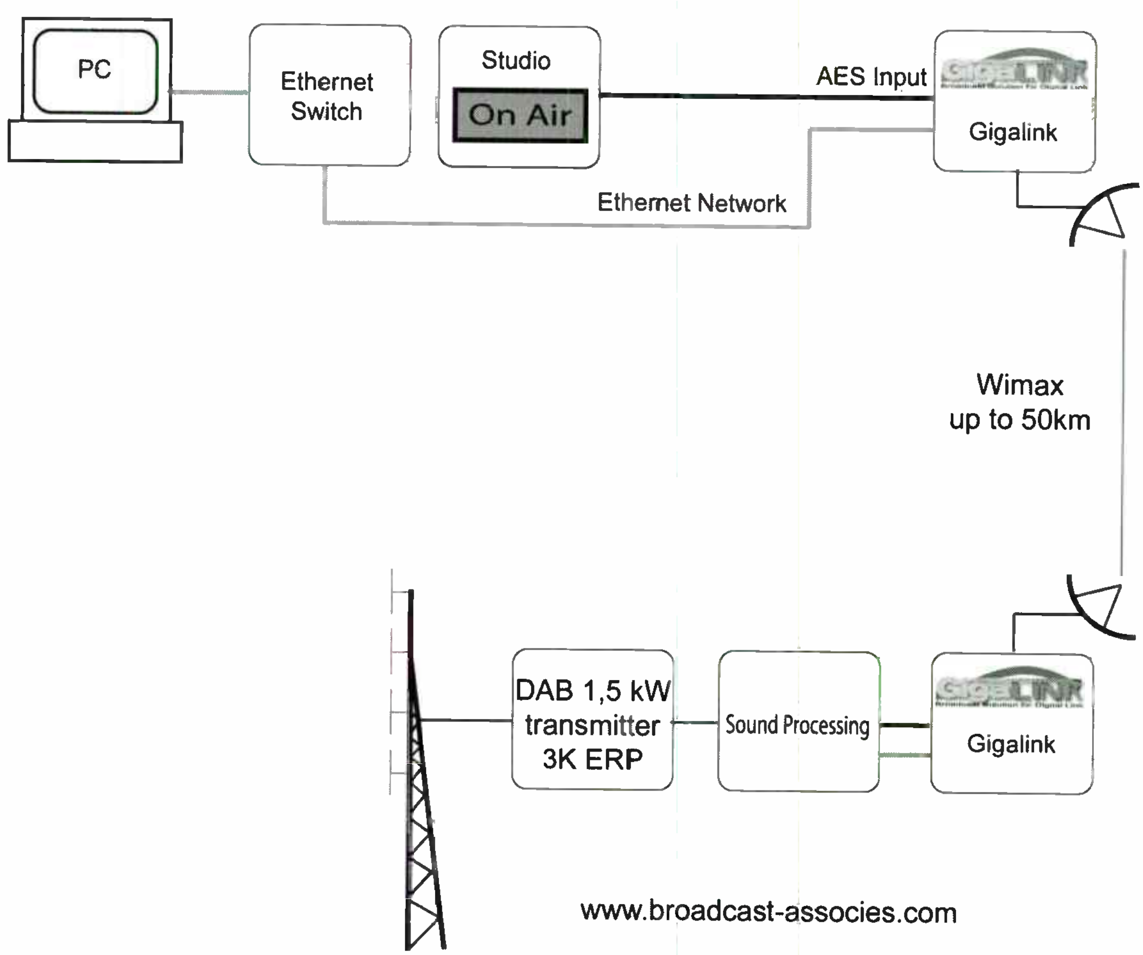
Bernie Courtney, Metro Broadcast Services, Z100 WHTZ

Lars Jonnson and Björn Westerberg, Swedish Radio

Lance Eichenberger, Radio Sunshine

Sylvain Cavet, Broadcast Associes

I would like to hear about other broadcasters' experiences using the internet for broadcast applications and can be contacted via email: sdaniels@aptx.com



Broadcast Associes - WiMAX Broadcast Link

A New Approach to Peak-to-Average-Power Reduction for Hybrid FM+IBOC Transmission

PHILIPP SCHMID
Nautel Inc.
Bangor, Maine

ABSTRACT

New developments in HD Radio™ broadcast transmission such as the move to increase the signal power allocated to the HD Radio™ sub-carriers has prompted Nautel to take a closer look at the peak-to-average-power (PAPR) reduction of the standard IBOC solution and its applicability within this new reality. While the standard PAPR reduction effectively reduces signal peaks from over 12 dB to under 8 dB, Nautel introduces a novel approach to PAPR reduction specifically for the hybrid FM+IBOC waveform. Many principles equally apply to all-digital transmission, but significant gains are realizable only in hybrid transmission. Research conducted at Nautel to date indicates that a savings of 30% or more in required transmitter power can be obtained. To take full advantage of the new sub-carrier power levels additional transmitter overhead will be required, however, even existing hybrid transmitter installations can benefit from this innovation and effectively boost the power in their digital sub-carriers.

IBOC SIGNAL CHARACTERISTICS

As HD Radio™ gains more and more momentum many radio stations have embraced HD Radio™ and are already transmitting the In Band On Channel (IBOC) signal. However, many stations cite high HD conversion costs with as of yet little realizable gains as the major factor for not adopting IBOC at this time. While IBOC broadcast equipment cost is partly to blame, the chief reason for high conversion costs is the fact that the IBOC signal provides significant challenges to broadcast transmitter designs that often require new IBOC capable transmitters to be installed.

IBOC employs orthogonal frequency division multiplexing (OFDM) to broadcast the digital HD Radio™ signal. IBOC employs time diversity through the use of data interleaving in order to deliver a more robust signal. OFDM furthermore provides frequency diversity through the use of multiple simultaneously transmitted data carriers that combat frequency dependent fades in multi-path environments. FM-IBOC further leverages frequency diversity by placing upper and lower sidebands on either side of the FM modulated signal with more than 200 kHz of frequency separation. For more information on the organization of the FM-IBOC signal refer to [1].

While multiple carriers in an OFDM signal provide for a robust signal, they require highly linear signal amplification in order to minimize carrier inter-modulation and ensure spectral compliance. Secondly, the amplifier requires a significant amount of input back-off (IBO) in order to handle large power peaks inherent in the IBOC signal. Illustration 1 shows how

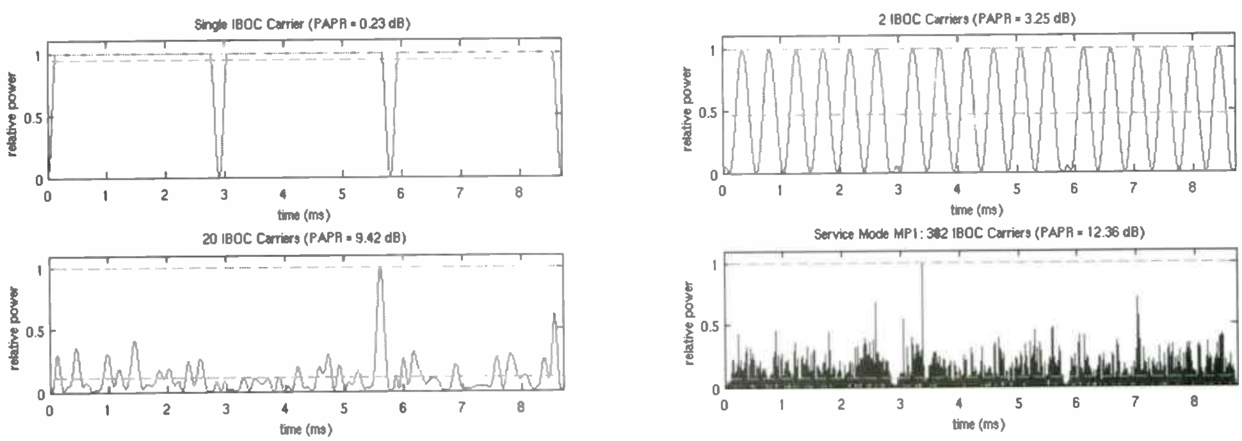


Illustration 1: Instantaneous Power Requirements over 3 IBOC Symbols

multiple carriers constructively and destructively add due to varying phase information in the quadrature phase modulated (QPSK) data carriers. Presented are three consecutive FM-IBOC symbols. The first plot is the power envelope of a single IBOC carrier, which similar to FM modulation has a constant power envelope at baseband prior to channel modulation. In this case, the IBOC pulse shaping function to smooth out the spectral impact of symbol transitions is apparent. The addition of a second carrier drops the average power of the signal by 3 dB, while maintaining the same power peaks, which can be expressed as the peak-to-average-power ratio (PAPR).

With 20 carriers, the random nature of extreme peaks is already visible and the PAPR increases to more than 9 dB. However, FM-IBOC uses a minimum of 382 carriers further increasing the PAPR to over 12 dB. While in theory all carriers could add constructively, in practice peaks of much greater than 12 dB are rarely encountered.

While the IBOC signal power is considerably less than the FM signal power in a hybrid FM+IBOC signal, requiring broadcasters to install a significant amount of additional transmitter power would present a significant hurdle to the adoption of HD Radio™. Therefore, iBiquity Digital Radio has provided an optional PAPR reduction algorithm as part of the standard IBOC modulator effectively reducing IBOC signal peaks.

Note: Some literature refers to the signal's PAPR at the RF level after channel modulation. So for example, an FM modulated signal with a constant power envelope at baseband would then have a PAPR of 3 dB at the RF level due to the fact that a sine wave has a crest factor of √2 [2]. The same principle applies to IBOC, as well. Provided we understand this relationship, we will continue to express the PAPR at baseband, which directly translates into the required IBO.

A more meaningful way of quantifying peak performance compared to the PAPR, is the complementary cumulative distribution function (CCDF), which describes the signal in a statistical way. The CCDF is defined as follows:

$$CCDF(x) = P(X \geq x) = 1 - \int_{-\infty}^{x} f(t)\,dt$$

where f(x) is the probability distribution function of the signal x.

If we take the random variable X to be the signal's power fluctuation, then the CCDF gives us the practical interpretation of the probability of clipping the signal at a given maximum power level.

Illustration 2 contrasts the CCDF of an IBOC signal to the CCDF of a peak reduced version of the signal at 1W average power. Note that the X-axis in this illustration represents a linear power scale. With a peak duration of around 1µs, a clipping probability of 10^{-6} would suggest around 1 clip per second and would require an IBO of 8 dB (6x above average power), while without peak reduction an IBO of 12 dB (15x above average power) would be needed.

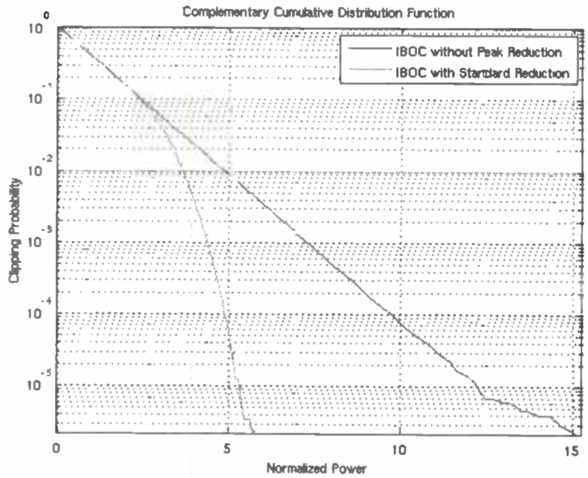


Illustration 2: Standard IBOC Peak Reduction

The CCDF not only provides a comparative measure, but it also provides us with an idea of how far a transmitter can be driven into saturation and introduce clipping. It does not, however, detail the spectral effects introduced by this clipping operation, which greatly depends on the transmitters characteristics in the saturation region. Illustration 4 below shows the effects of hard clipping on the spectrum and signal constellation.

While standard PAPR reduces signal peaks from 12 dB to 8dB, in practice, it has been found that peaks can be reduced further by driving the signal into compression. Depending on the transmitter, the signal can often be driven into compression to yield a final PAPR of 5.5 dB. What this means to the broadcaster, is that in order to achieve a 3 kW digital transmitter power output, a transmitter capable of delivering 10.6 kW of instantaneous power must be installed. Without standard PAPR reduction a much larger transmitter would need to be installed depending on how much that signal could be compressed in the transmitter.

As the standard PAPR reduction provides significant gains we will take a closer look at the operation of the standard PAPR reduction algorithm as implemented by iBiquity Digital Radio.

STANDARD PAPR REDUCTION ALGORITHM

The following is a basic description of the standard PAPR reduction algorithm as described in [3], which should be considered the authoritative source. Illustration 3 reproduces a basic version of the algorithm's flow chart presented in [3].

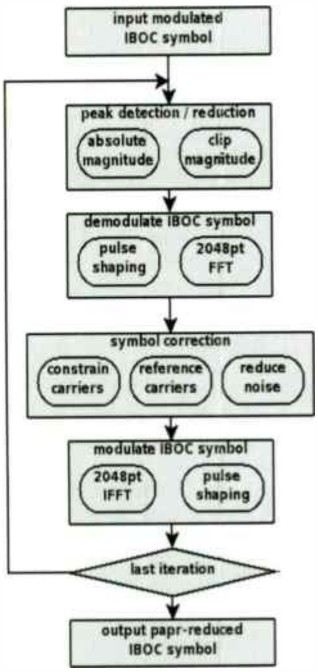


Illustration 3: Standard PAPR Reduction

Peak Detection And Reduction

The standard PAPR reduction algorithm inputs a single modulated IBOC symbol at a time. Peaks are detected by computing the absolute value of each sample point and comparing it against a pre-defined threshold value.

Once a peak is identified, the standard PAPR reduction algorithm clips the magnitude of a sample point associated with this peak to a given threshold while maintaining the sample point's instantaneous phase value. Regardless as to whether hard or soft clipping is applied, the act of clipping effectively introduces a delta function $\delta(t)$ to the digital signal.

The non-linear effect of adding a delta function to the signal is to add frequency content across the entire discrete frequency spectrum that is related to the magnitude of the peak reduction which may violate the spectral emission mask. It also introduces error in the signal constellation, which degrades the noise performance of the IBOC signal. Illustration 4 depicts the effects of clipping the IBOC signal, where the blue plots represents the original inputted symbol, and the red plot represents the clipped signal. A scattering of constellation points is observed that tends to move to the origin, as peak reduction tends to reduce the signal's power. This is not an issue, as the signal can easily be scaled back up in order to maintain the same output power. The impact on the noise floor is clearly visible and often is the limiting factor compared to the impact on the signal constellation.

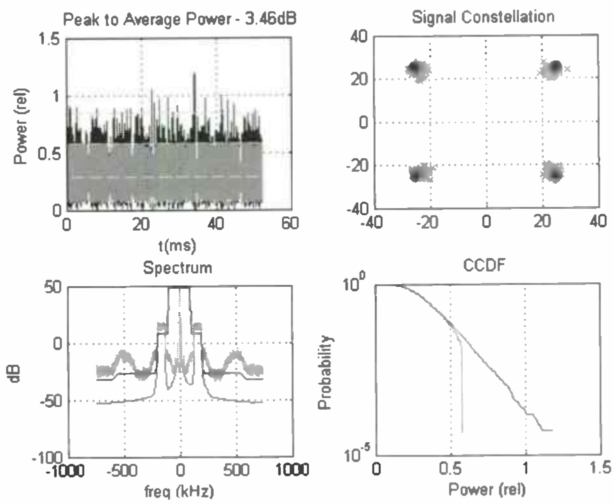


Illustration 4: Effects of Clipping the IBOC Signal

Symbol Correction

Because of the signal distortion introduced in the previous step, the signal must be cleaned up. To do so, the standard PAPR reduction algorithm performs an IBOC demodulation of the distorted signal down to the individual carrier level. This involves removing the IBOC pulse shaping from the OFDM symbol and a consequent FFT operation at a sampling rate directly related to the original symbol creation, such that each frequency bin at the output of the FFT perfectly describes the information in a single carrier.

The first part of the correction process limits the amount of error that is allowed in a single carrier by pushing the constellation point of the carrier back toward its ideal constellation point. As all carriers are QPSK modulated, this is accomplished by simply pushing all points away from the XY axes to a desired threshold but not all the way back to the ideal QPSK point. ***Pushing back the constellation points toward***

the ideal QPSK point brings back the same peaks we eliminated in the previous step. By only going part way, we increase the carrier's bit energy, but the peaks are only partially restored. Illustration 5 provides a clear example of this effect.

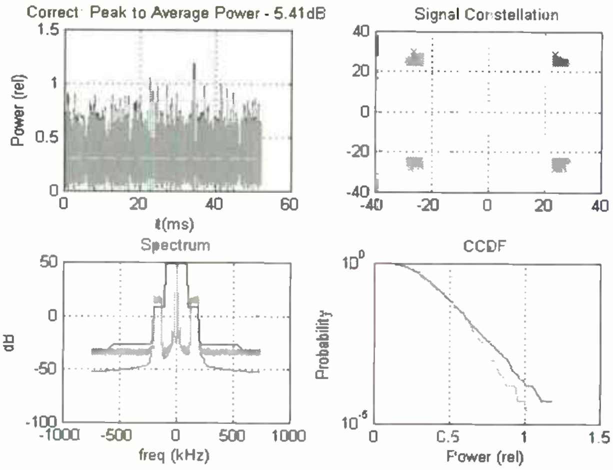


Illustration 5: Effects of Correction on a Hybrid Signal

Particular attention must be paid to reference carriers that allow a receiver to lock on to the IBOC signal. While the amplitude is not significant, it is important that the phase of the reference carrier is preserved. Therefore, the reference carrier's phase is restored to its original value while the corrected amplitude is maintained.

As a third part of the correction process, the error in the non-carrier frequency bins must be suppressed. The same principle applies here; as we correct the signal back to its original spectrum more of the peaks are starting to come back. A mask is applied in correcting this signal content, allowing varying amounts of noise to subside in the IBOC signal without violating the spectral emission mask.

Because of the opposing effect of these steps, the PAPR reduction is an iterative process, each transition through the loop yields an improved solution. However, in order to accomplish this is computationally expensive due to the iterative computation of a 2048 point Fast Fourier Transform (FFT) and its inverse (IFFT). Each additional iteration yields diminishing returns approaching a final limit that is mainly a function of our correction parameters, as well as, the frequency and magnitude of peak reductions.

Standard PAPR Reduction Performance

With the understanding of how the standard PAPR operates, we will take a closer at the output provided by the standard exgine IBOC modulator. The sample stream is directly captured at the exgine output and demodulated to determine the PAPR and signal constellation. It has already been determined that the standard PAPR reduction reduces the PAPR from 12 dB to under 8dB. This represents a significant improvement, but we need to ensure that this process has not degraded the IBOC symbol's noise performance.

Illustration 6 provides a plot of the demodulated signal constellation as captured from the exgine IBOC modulator in service mode MP1. It also highlights the reference carriers in the symbol. We observe a significant spread in the signal constellation points that indeed affects the signal's noise performance.

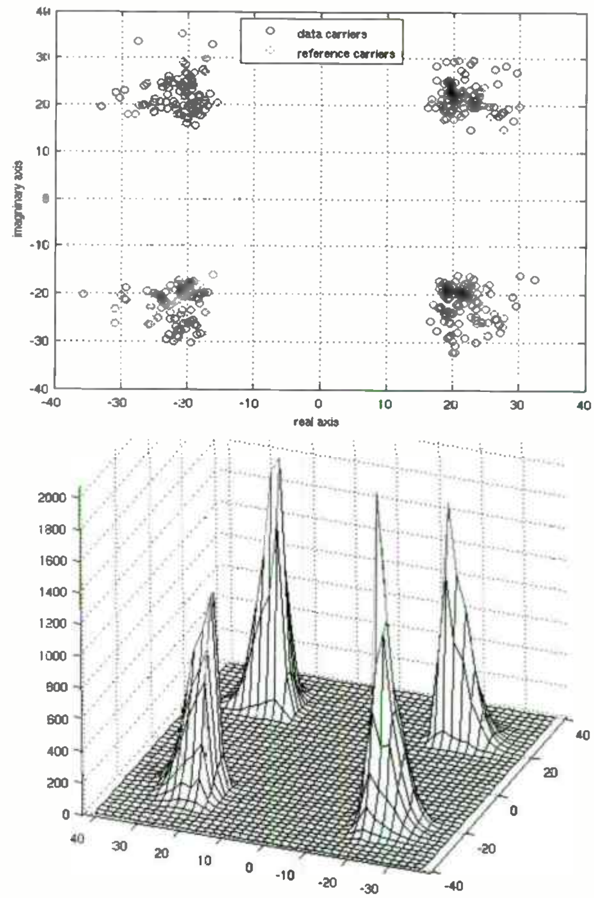


Illustration 6: IBOC Constellation and Histogram of Standard PAPR Reduced IBOC

However, looking at a traditional constellation plot may lead one to incorrect conclusions. Illustration 6 also provides a 3-D histogram that simply tallies the number of constellation points per unit area. Looking at the histogram reveals that the spread in the constellation is truly not so bad. The majority of data points are concentrated in defined clusters and only infrequent data points fall outside this region and no point approaches the bit decision boundary along the x and y axes.

A true measure of the impact of standard PAPR reduction is to look at its noise performance in an Average White Gaussian Noise Channel (AWGN). Therefore, the output of the exgine modulator is added to AWGN, the result is demodulated and the bit errors in the carriers is tallied compared to the original symbol. Depending on the tolerable bit error rate, we can quantify the power increase required to restore noise performance back to the noise performance of an ideal QPSK modulated IBOC symbol. Given a certain bit error rate, Table 1 lists the power ratio between the ideal symbol and its PAPR reduced version. This provides a figure of merit in comparing the error introduced into the constellation by different PAPR reduction schemes.

Carrier Bit Error Rate	Reduction in Noise Performance (Reduced/Ideal)
$5x10^{-2}$	0.29 dB (1.07)
10^{-2}	0.49 dB (1.12)
10^{-3}	0.57 dB (1.14)
10^{-4}	0.72 dB (1.18)
10^{-5}	0.83 dB (1.21)

Table 1: Comparative Noise Performance

Because of the substantial amount of forward error correction (FEC) inherent in the IBOC signal, IBOC can indeed operate in channel conditions with very high bit error rates. We are not really interested in the region of low bit errors, since FEC will provide us with great performance in that region regardless. We want to choose a BER at the edge of our coverage area, but which still does provide acceptable service. For the remainder of this discussion, we take this point to be at 10^{-2}, but this number may be qualified further in the future. All noise performance is to be taken with respect to the ideal IBOC symbol.

Therefore, while it is computationally expensive, standard IBOC PAPR reduction is a very effective means of peak reduction that only introduces a small to moderate degradation in noise performance. However, the argument could be made that the algorithm's parameters should be under a broadcaster's control, as it is conceptually conceivable to relax correction parameters to achieve greater gains in peak reduction and only incur a further small degradation in noise performance.

We will look at the performance and applicability of this PAPR reduction in a hybrid system next.

10DB CARRIER INCREASES: THE NEW REALITY

When discussing the effectiveness of the standard PAPR reduction in the context of low level combined FM+IBOC, we must touch on recent developments that aim at increasing digital carrier power levels by 10 dB in a hybrid waveform. This comes in an effort to more closely match the coverage area of the IBOC signal to the comparable coverage area of the simulcasted FM signal and to improve building penetration and general IBOC signal robustness.

While it is outside of the scope of this paper to discuss this development in detail, we have to look at the applicability and effectiveness of the standard PAPR reduction in the context of this new development. Increasing digital carriers by 10dB only increases the average IBOC signal power from 1% to 10% of the transmitted FM signal. However, it would be a grave

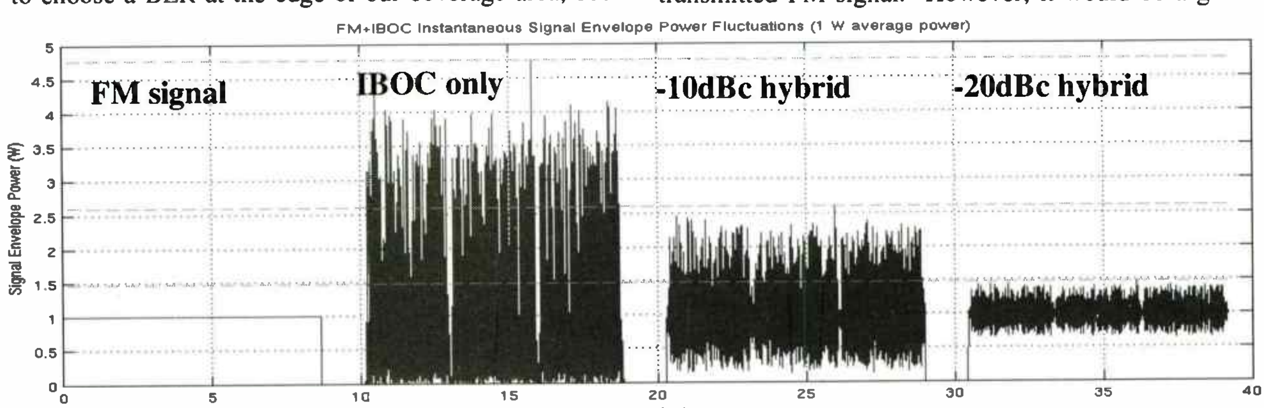


Illustration 7: Comparative Instantaneous Envelope Power Fluctuation

mistake to think that this change would have only minor implications to a low level combined hybrid transmitter. By now, it is apparent that broadcast transmitters are limited by their peak power capability and not their average power capability. So we must look at the signal peaks as shown in illustration 7, which depicts the baseband power envelope of an analog modulated FM signal, a digital only signal, and a hybrid signal at -10 dBc and -20 dBc injection levels all at the same average power of 1 W. While at -20 dBc about 40% of transmitter overhead was sufficient, going to -10dBc carriers, we now require more than 160% of transmitter power. Assuming that the spectral emission mask stays at current levels, this means that this hybrid waveform cannot be driven into amplifier compression to the same degree as in the -10 dBc case. Therefore, almost all of the signal must fall into a linear amplification region. This now means in order to achieve a hybrid TPO of 8kW, one must install a transmitter capable of handling 22 kW, while a 11 kW transmitter suffices at -20 dBc.

Illustration 8 depicts the power distribution of hybrid signals at different injection levels all scaled to the same average power. The absolute maximum point is indicated via the red dashed line and the corresponding PAPR ratios are given in the legend. This leads us to a significant observation:

As the analog component of the signal increases, the peak distribution of the resulting hybrid signal changes shape. The peaks in the digital waveform are not necessarily the same peaks in the hybrid waveform.

While it is true that a peak reduction in the IBOC signal is only but beneficial to the hybrid waveform, a peak in the IBOC signal may in fact not turn out to be a peak in the hybrid signal, if the addition of the FM signal to the IBOC signal happen to add destructively. On the flip side, a lower IBOC signal peak may entirely add constructively to the FM signal creating a notable peak in the hybrid signal. While this fact has received little attention at -20 dBc carriers, for the -10 dBc carrier case, this observation makes a significant difference.

In short, peak reduction must simply be performed on the final signal that is to be passed through the power amplifier of the transmitter. However, the standard PAPR reduction scheme cannot simply be applied to a hybrid signal without significant changes to both the algorithm operation, as well as, the implemented radio systems broadcast architecture, as defined by iBiquity Digital Radio. The remainder of this paper will detail Nautel's innovative approach to peak reduction in a hybrid signal.

PROPOSED PAPR REDUCTION

This section outlines the operation and innovation in the proposed PAPR reduction method.

Peak Detection

The major difference between the standard PAPR reduction and our proposed PAPR reduction is a difference in peak detection. Illustration 9 contrasts the difference of peak detection in the standard PAPR reduction versus the proposed reduction method at a single instance in time.

Illustration 9 depicts a complex plane, where the X-axis reflects the baseband signal's real (or in phase – I)

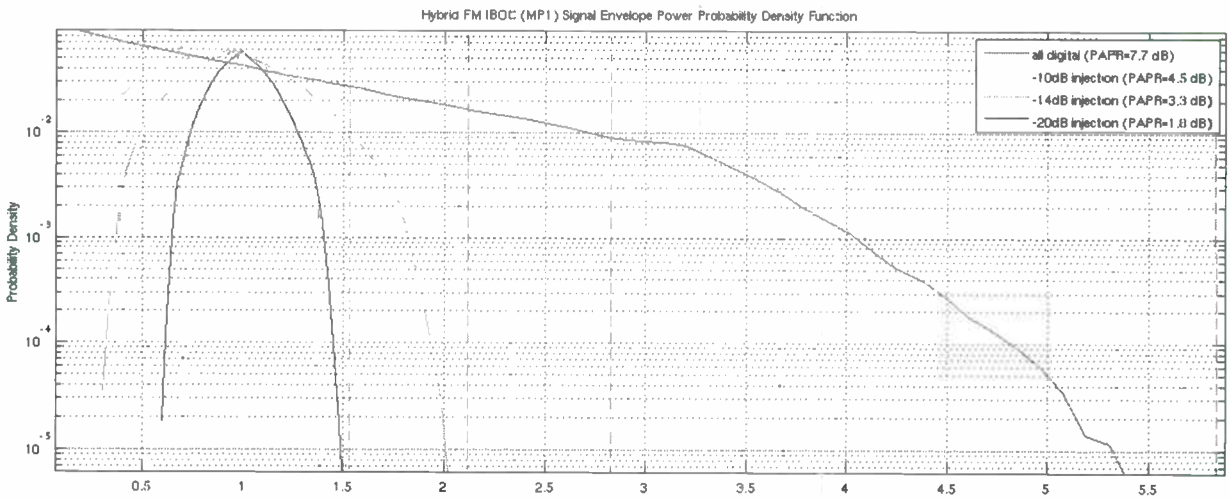


Illustration 8: Hybrid FM+IBOC Signal Distribution

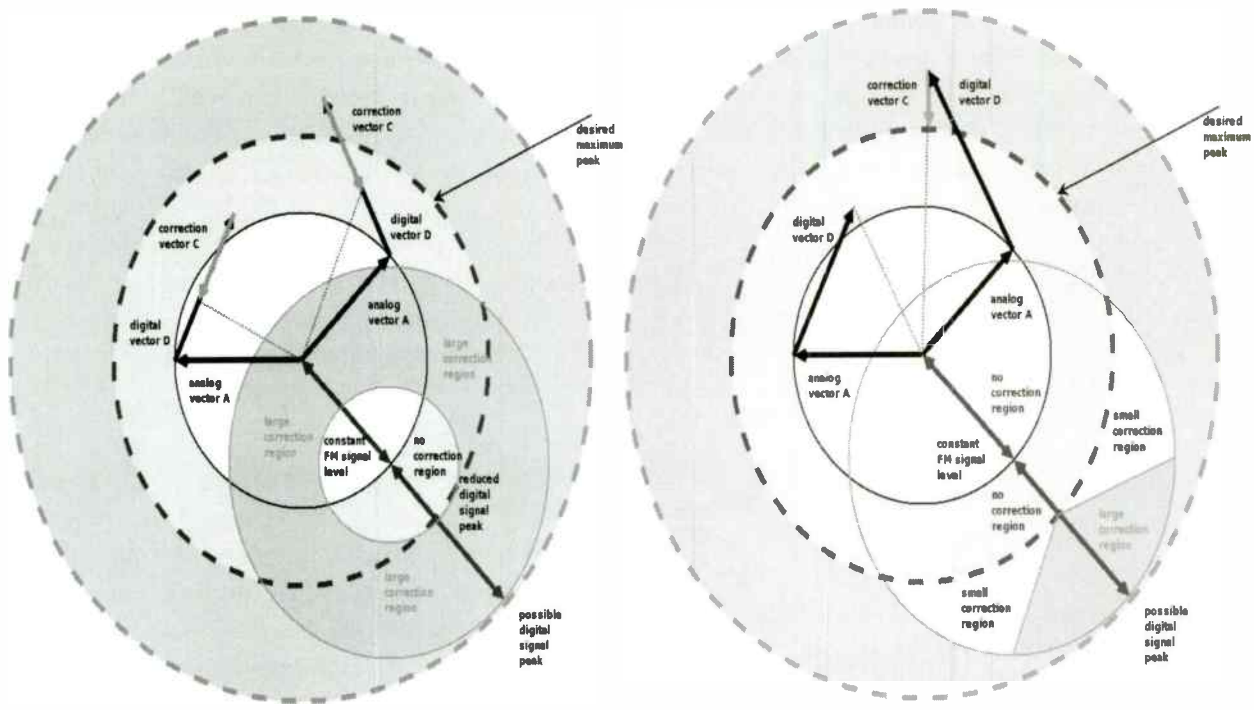


Illustration 9: Standard Peak Clipping (left) Contrasted to Hybrid Optimized Peak Clipping (right)

component and the Y-axis represents the signal's imaginary (or quadrature – Q) component. The first graphic illustrates the case of standard PAPR reduction that only operates on the digital signal and then adds the result to the analog signal. The second graphic, on the other hand, shows how the analog signal is taken into account in detecting a peak.

The output of the FM modulation process produces a constant envelope signal with varying phase. At baseband, this signal is represented as a vector in the complex plane with constant amplitude, which is represented by the white circle in our illustration.

For the sake of comparison, let us assume that both methods could achieve the same level of peak reduction. Since the standard PAPR reduction method is agnostic of the analog modulation, it can only detect a peak based on the digital signal alone and it does not know whether this peak adds constructively or destructively to the analog signal. Should the peak add constructively, then the standard reduction method performs the correct operation by introducing a large peak correction. However, if the peak adds destructively to the analog, ***the standard PAPR reduction unnecessarily performs a potentially large peak reduction***.

For demonstration purposes, if we choose an analog signal point at one point in time and perform a vector addition of all possible digital signal points, then a peak in the digital signal creates a large circle around the analog signal point. The standard PAPR reduction method reduces the peaks in the digital signal down to the radius of the inner circle, which borders the circle representing the maximum desired peak of the combined signal. This leaves a large area in the complex plane where peak reduction is performed as indicated by the red shaded area in the illustration. Hence, ***the standard PAPR reduction scheme causes many sample points being unnecessarily corrected for when they do not in fact form an actual signal peak when combined with the FM signal***.

The proposed innovation suggests a different approach for determining the correction vector C, which is used as the input to the peak correction process. When determining a peak, the analog vector A is first added to the digital vector D. The resultant hybrid vector H is then thresholded against the maximum desired peak threshold. ***Only if the digital signal adds constructively to the analog signal, is a large correction required. Only a smaller correction is needed, if the vector addition falls close to the maximum desired peak and no correction is required***

if the result is below the maximum desired peak. Our illustration comparatively shows a red shaded region where a relatively large correction is required in the same way as is performed in standard PAPR reduction and a yellow shaded region where only a smaller degree of correction is required.

In comparison, using the proposed PAPR reduction method yields a much smaller region that requires a large correction. ***By introducing a lower amount of correction, the proposed algorithm can achieve the same maximum desired peak value with a lower degree of distortion in the original signal. This allows us to reduce the signal's peaks further compared to the standard PAPR reduction method.***

In order to realize this difference in peak detection, the standard PAPR reduction algorithm and, consequently, the broadcast architecture for IBOC must be somewhat modified. Illustration 10 highlights the differences in red to the standard PAPR reduction method shown previously.

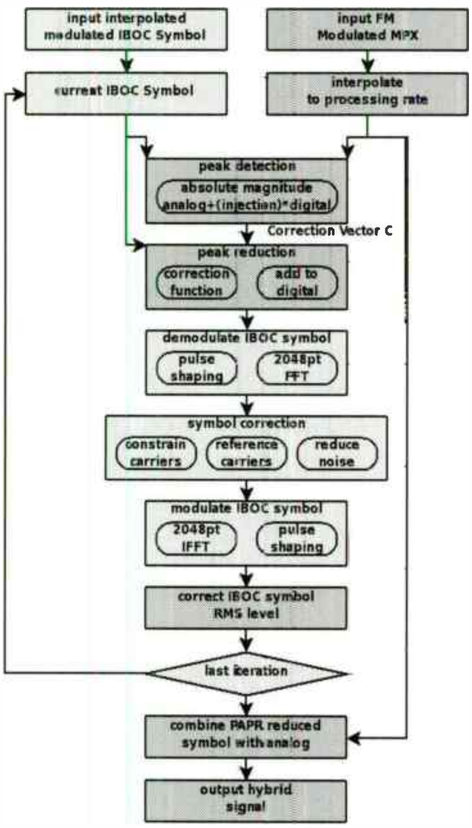


Illustration 10: Proposed PAPR Reduction

Fundamentally, the biggest difference is the fact that the digital IBOC modulator must now know about the FM modulated MPX signal, while the standard PAPR method does not require this input. The FM signal, as well as, the non-PAPR reduced symbol are both interpolated to a higher sample rate. ***While digital sampling theory faithfully preserves a signal's frequency content, there is no guarantee that the signal's peaks fall on discrete sample points. An interpolation process allows us to captures peaks more reliably.*** Performing PAPR reduction at the standard IBOC sample rate of 744 kSa may miss actual signal peaks by 30-40%. By interpolating by a factor of 2, this error is reduced to around 5% for the IBOC signal, but it significantly increases computational requirements.

Peak Reduction

With the standard PAPR reduction algorithm it was found that hard clipping provides an efficient and effective means of peak reduction. In our discussion thus far, this simply means that the correction vector C is added to the digital vector D on a sample-per-sample basis. A similar approach can be employed for our case. ***By not simply clipping the hybrid signal, but keeping the correction vector C separate and only applying it to the digital component allows us to use the established correction techniques described with the standard PAPR reduction.*** It also uses the FM signal only during the clipping decision process and, thereby, faithfully maintains the FM portion of this signal until it is finally added to the digital component to form the hybrid signal stream. ***Therefore, the FM transmission is not impacted by the proposed PAPR reduction technique.***

Applying the correction vector directly to the digital symbol, essentially clips peaks. However, based on the correction vector we could also create an error signal as follows:

$$E[n] = \sum_{k=0}^{length-1} correctionFunction[n-k]C[k]$$

This allows us to shape the spectral impact of the reduction via the correction function to concentrate the introduced noise in more convenient frequency bins rather than the wide impact of the delta function introduced through clipping. Tone or pulse injection techniques may be applicable here. Depending on the choice of the correction function, the error signal can be small enough that one can even safely bypass the constellation and spectrum correction step altogether.

Constellation and Spectrum Correction

Just as for standard PAPR reduction, the modification of the digital signal in the time domain can negatively impact the signal constellation, as well as, increase out-of-band noise requiring correction. Unlike the standard PAPR reduction, different correction functions in the proposed method will have varying impacts on the constellation and injected noise level.

While other implementations of this step are conceivable, the standard implementation can work very well on our modified signal at this point. However, key parameters, such as the number of iterations and correction thresholds can be adjusted to achieve various levels of PAPR reduction performance.

PROPOSED PAPR REDUCTION PERFORMANCE

While intuitively the proposed PAPR reduction method should provide superior results in comparison, the theory must be put to the test, first using simulations and second using real hardware. Nautel is presently assembling an IBOC modulator proof-of-concept prototype system able to perform rigorous hardware tests. In the mean-time, this paper reports on the simulation results obtained thus far.

Since -10dBc carrier levels are of interest to the broadcast industry at this time, our first simulation case is aimed at obtaining a comparable IBOC constellation to the standard PAPR reduction with comparable noise performance. Basic clipping and no other advanced options, such as using the extended carrier spectrum are used to compile these results in order to provide a fair head-to-head comparison of the two reduction methods.

For comparative purposes, a standard PAPR reduced symbol stream is captured from the exgine modulator and the PAPR reduction is removed by moving all constellation points back to their ideal location. A symbol stream comprising a particular bad power spike is selected in order to ensure the proposed PAPR reduction can effectively deal with a worst case scenario.

The resultant proposed PAPR reduced symbol stream is subjected to a noise performance test to ensure similar noise performance to the standard PAPR reduced symbol. The standard and proposed symbols are compared to ensure no bit errors are introduced in the PAPR reduction process.

Illustration 11 graphically reports the results of this test. The blue plots pertain to the standard PAPR reduced IBOC symbol, the red plots pertain to the proposed PAPR reduced IBOC symbol.

The spectrum plot reveals that carriers are indeed increased by 10dB with respect to the more stringent IBOC emission mask and both methods maintain this level throughout. Noise performance tests reveal that the standard PAPR reduced symbol performs 12% below the ideal IBOC symbol, while the proposed PAPR reduced symbol performs 14.8% below the ideal IBOC symbol well within comparable levels.

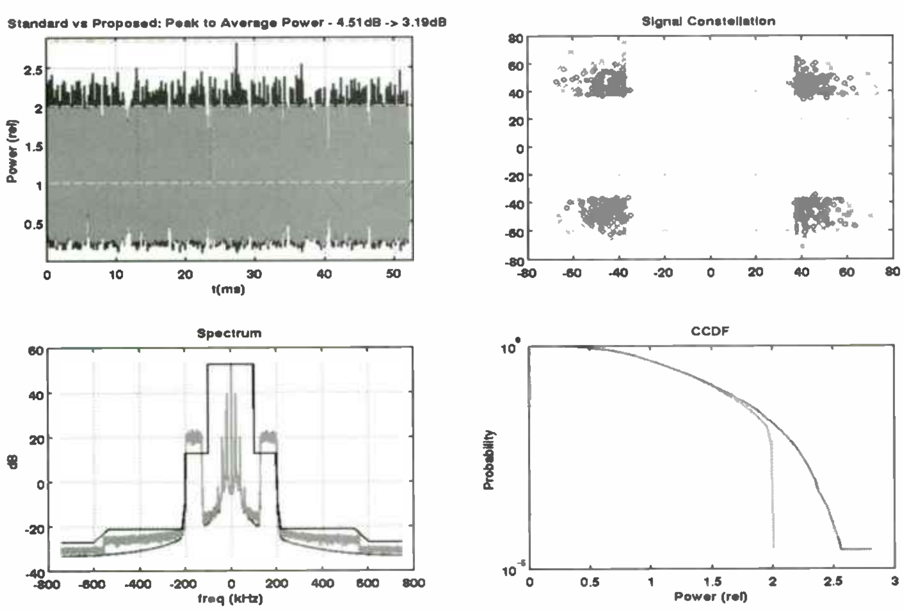


Illustration 11: Performance Comparison of Standard (blue) vs Proposed (red) PAPR Reduction

While this does not represent the best maximum gain possible with the proposed PAPR reduction, it does show a very substantial reduction in the PAPR from 4.51 dB down to 3.19 dB. Graphically the presented time domain plot shows how the required transmitter overhead is reduced. The CCDF, however, presents a better picture of the situation. The discontinuity in the blue curve is explained by the fact that we have specifically selected the input signal based on a maximum peak, but this should not be disregarded, as the frequency of this peak in an actual symbol stream is still significant. However, the proposed PAPR reduction has been able to effectively remove this singular peak. One caveat of the proposed PAPR reduction method is the fact that the proposed reduction has a much sharper drop off in the CCDF compared to the standard PAPR reduced symbol. This means that this method won't be able to be driven into amplifier compression as much, but at -10dBc carriers the amplifier simply may not be driven into amplifier compression by any significant amount and maintain spectral compliance.

EXISTING HYBRID INSTALLATIONS

A regulatory move to allow the transmission of higher IBOC carriers, may place low level combined broadcasters that have already converted to IBOC at a disadvantage. Therefore, it is an interesting exercise, to see how the gains achieved using our proposed PAPR reduction can be applied to existing low level combined stations. At this point in the discussion, it should be clear that power levels can be very easily be increased, if no attention is paid to the underlying signal constellation. So we must consider the impact of constellation degradation.

Our objective is to increase the signal's noise performance, not the signal's output power. A 10 dB power increase may not translate into a 10 dB noise performance improvement, if the underlying signal constellation is modified.

While our proposed reduction provides significant gains, a 10 dB carrier increase is too large to be absorbed by these gains without seriously deteriorating the signal constellation.

However, ***a 6 dB carrier increase is possible with some impact on noise performance using back-off figures for current low-level combined transmitters***. Illustration 12 provides the results for this case. The resultant PAPR of 1.52 is very close to the compressed PAPR used to spec current low level combined transmitters. For this case, various advanced techniques had to be employed, such as using the extended carrier space, pulse injection, and scaling individual symbols to maintain constant symbol-to-symbol peaks.

The noise performance impact in this case is significant and requires 67% additional power to achieve the same BER as an ideal IBOC symbol. Considering that a

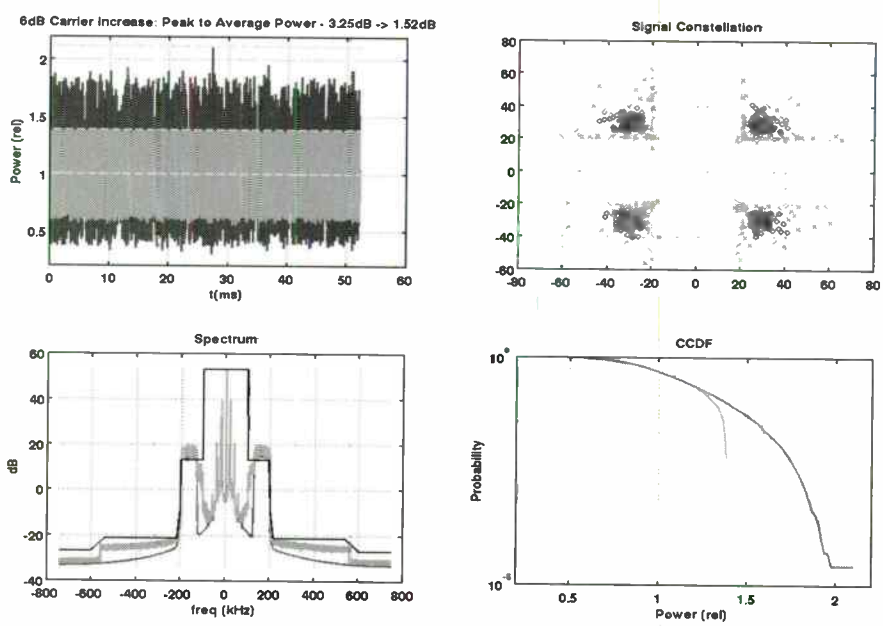


Illustration 12: Aggressive PAPR Reduction Utilizing Extended Frequency Partitions

PAPR Reduction Method	Injection Level	Modulator PAPR	Compressed PAPR	Noise Performance	Signal Improvement	Peak Power for 8kW
Standard PAPR Reduction	**-20dBc**	**1.85**	**1.50**	**0.49 dB**	**0.0 dB**	**11.3 kW**
Standard PAPR Reduction	-14dBc	3.25	3.25*	0.49 dB	6.0 dB	16.9 kW
Aggressive Reduction using Extended Partitions	-14dBc	1.52	1.52*	2.22 dB	4.3 dB	11.4 kW
Standard PAPR Reduction	**-10dBc**	**4.51**	**4.51***	**0.49 dB**	**10.0 dB**	**22.6 kW**
Proposed Reduction	-10dBc	3.19	3.19*	0.60 dB	9.9 dB	16.7 kW
Proposed Reduction using Extended Partitions	-10dBc	2.53	2.53*	1.34 dB	9.2 dB	14.3 kW
Aggressive Reduction using Extended Partitions	-10dBc	2.19	2.19*	2.14 dB	8.3 dB	13.2 kW

Table 2: Comparative PAPR Reduction Options

standard IBOC symbol requires 12% additional power, we have effectively improved our noise performance by 6 dB – 2.22 dB (degradation from ideal) + 0.49dB (with reference to standard constellation) = 4.3 dB at a BER of 10^{-2}. Not quite the 6 dB corresponding to the power increase, but considering it requires no additional transmitter hardware, this represents a significant improvement in IBOC transmission.

If the station has some initial head room available, we don't have to compress quite as heavily, this will allow us to first improve the signal constellation and recover some losses, we can then either free up the extended carrier space, or further increase carrier power. The optimal operating point will have to be determined on a station by station basis depending on the available headroom, transmitter type, and station preference.

Table 1 presents a number of simulation cases at varying injection levels. These results should be taken for reference only and don't represent any official transmitter performance specifications. The compressed PAPR for most of these cases has yet to be determined and are marked with an asterisk*. Assuming a TPO of 8 kW, the table also lists the comparative required FM transmitter size required to handle the signal's peak power. It is apparent that a wide variety of choice exists and broadcasters will likely have to make choice such as is it worth installing additional 2.4kW of transmitter power in order to marginally improve our signal and free up our extended carrier space.

These choices should neither be dictated by manufacturers, nor iBiquity Digital Radio, but should truly be a broadcaster's choice.

CONCLUSION

This paper has demonstrated the potential gains to be obtained using this novel PAPR reduction approach.

It is understood that this paper only scratches the surface on a very extensive topic. Simulation results must be verified in real hardware, both in the laboratory and through field trials. The purpose of this paper is to detail Nautel's novel PAPR reduction approach with the hope of setting the framework for more extensive testing and experimentation involving the broadcast community at large.

Providing a stronger HD Radio™ signal and increased coverage area for a lower additional investment may prompt additional broadcasters to adopt HD Radio™ delivering more HD content choice to the listener.

APPLICABLE REFERENCES

[1] iBiquity Digital Corporation; 2005
HD Radio Air Interface Design Description – Layer 1 FM, Rev. E
Doc. No. SY_IDD_1011s

[2] Maxson, David *The IBOC Handbook*
Oxford: Focal Press, 2007

[3] Kroeger, Brian William; 2005
Peak-to-average power reduction for FM OFDM transmission
United States Patent 20050169411

Field Tests for Service Area and Handover Service in T-DMB

Sang-Hun Kim, Gyu-Young Kim, and Man-Sik Kim
Korean Broadcasting System
Republic of Korea

Abstract

This paper presents the field test results of T-DMB (Terrestrial Digital Multimedia Broadcasting) services in Korea. In December 2005, the commercial broadcasting service of T-DMB was firstly launched in Seoul metropolitan area using VHF band. The T-DMB system, based on the Eureka-147 DAB (Digital Audio Broadcasting) system, provides streaming video, CD-quality audio and data services at speeds of up to 300km/h using H.264 video codec along with additional channel coding and interleaving schemes. In order to provide seamless high-quality services on various radio channels, we measured CIR (Channel Impulse Response), TII (Transmitter Identification Information), electric field strength, video quality, audio quality and data quality using specially-designed mobile measurement system. From the field tests, we obtained the service coverage for error-free reception and have been installing new transmission sites to extend its coverage. Furthermore, we performed the field tests to verify and adopt the handover service in T-DMB networks.

1. Introduction

T-DMB (Terrestrial Digital Multimedia Broadcasting) can provide high-quality video, CD-quality audio and data services to fast-moving vehicles as well as pedestrians even in highly deteriorated channel environments. Streaming video service is realized by applying MPEG-4 AVC (Advanced Video Coding) or H.264, RS (Reed Solomon) coding, and convolutional interleaving on the existing DAB (Digital Audio Broadcasting) system [1]-[9]. The T-DMB system also supports the simultaneous reception of the same information originated from more than one transmitter by using COFDM (Coded Orthogonal Frequency Division Multiplexing) modulation [1], [9]. In SFN (Single Frequency Network), all signals, direct or reflected, are added positively if they can reach the receiver within the given guard interval [1], [9]. This characteristic is normally called to network gain based on spatial diversity [10], [11]. It also allows all the transmitters in a network to use the same frequency - thus providing excellent spectrum efficiency.

In the T-DMB standards proposed in Korea, the underlying standards for audio, multiplexing, channel coding and modulation follow the conventional standards of Eureka-147 DAB with the addition of the character codes for the Korean language, the MPEG-4 multimedia service systems, and various data services [1]. The T-DMB standards have been approved as ETSI standards in June and July 2005, and they have been also approved as ITU standards in December 2007 [3]-[5].

In December 2005, the commercial broadcasting service was launched in Seoul metropolitan area using VHF band. In August 2007, the nationwide service was launched. The number of T-DMB users has been rapidly increased up to 8 million as of November 2007. At present, one or two video services, several audio services and data services are in operation at each of 6 ensembles, using VHF TV channels. One 6MHz bandwidth TV channel is divided into 3 T-DMB ensembles with 1.536MHz bandwidth, respectively. The nationwide SFN is currently impossible because there is no available common frequency until analog switch off. The wide-area SFN and nationwide MFN (Multi-Frequency Network) are implemented for T-DMB in VHF band.

Since 2004, the T-DMB system has been tested in Korea to provide seamless high-quality services to the mobile subscribers under various deteriorated reception environments. Before December 2005, the objective of the field tests is to measure and analyze the quality of T-DMB system in order to derive optimal transmission and reception parameters prior to launching commercial broadcasting services [12]. Currently, the objective of field tests focuses on estimating and extending the service coverage for error-free reception by adding new transmission sites. In addition, the objective is also to measure the feasibility of T-DMB handover and its performance between the neighboring MFN.

In this paper, we mainly describe the field test results especially for service area and handover service in T-DMB networks. The service area field test results and the handover field test results are presented in Sections 2 and 3, respectively. Finally, some conclusions will be given in Section 4.

2. Service Area Field Test Results

2.1 Transmission Network

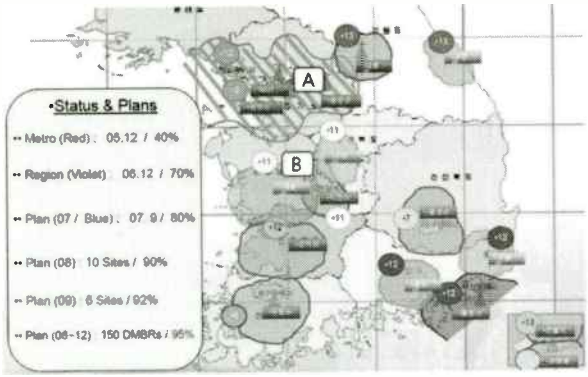


Fig. 1. Nationwide frequency allocation

In T-DMB of Korea, the different frequency channels in VHF band are allocated in each region to provide the nationwide services. The allocated frequency and the current state of networks are illustrated in Fig. 1. By the end of 2007, 80% of the whole country was covered although there are still some uncovered areas depending on geographical conditions.

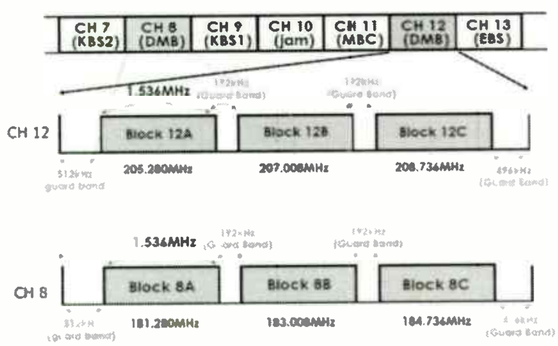


Fig. 2. VHF band structure in Seoul metropolitan area

At present, two VHF TV channels, 6 ensembles, are allocated for T-DMB in Seoul metropolitan area. The configuration of each ensemble in VHF band is depicted in Fig. 2.

The streaming video, CD-quality audio and data services are multiplexed into each ensemble, which is transmitted via 1.536MHz bandwidth. Although the total bit rate of T-DMB in one ensemble is 2.304Mbps, the available bit rate is reduced into about 1.152Mbps excluding the redundancy data of 1/2 code rate for error correction in the receiving side. 8 video services, 12 audio services and 8 data services are in operation in Seoul metropolitan, which is illustrated in Fig. 3.

Service / Broadcaster	Video (8)	Audio (12)	Data (8)
U1 Media (8A)	U1 544		Data: 64
	U KBS Heart 496		Null: 48
YTN DMB (8B)	mYTN 512	TBN 160	Data 320
		Satio+ : 160	
Korea DMB (8C)	1to1 496	CBS Audio 128	Data 32
	MBC net 496		
MBC (12A)	My MBC 544	MBC FM 128	Data 224
		MBN 128	
		Arirang 128	
KBS (12B)	U KBS Star 544	U KBS Music : 160	U KBS Clover 192 (BWS 96, TPEG 96)
		U1-Radio 128	
		OZIC-Mnet 128	
SBS (12C)	SBS U 544	SBS U Radio . 128	Hangyeorye 96
		TBS 128	LGT 96
		Gyeonggi 128	SBS 32

Fig. 3. Ensemble structure in Seoul metropolitan area

TABLEL I

Transmission parameters for service area field tests

Site	Center Freq. (MHz)	Transmit Power (kW)	ERP (kW)	Ant.	Ant. Height (m)
Mt. Gwanak	207.008	2	16.3	24P	642
Mt. Nam	207.008	1	9.9	16P	325
Mt. Yongmun	207.008	1	17.9	4P	1177

Three transmission sites consist of SFN to cover Seoul and neighboring Gyeonggi Province. Each antenna is composed of omni-directional two dipoles with vertical polarization and 24, 16 and 4 panels each, and mounted at Mt. Gwanak (615meters above sea level), Mt. Nam (230meters above sea level) and Mt. Yongmun (1157meters above sea level) in Seoul. The RF transmit powers are 2kW, 1kW, 1kW and 0.1kW, respectively and complied with the RF signal spectrum mask of Eureka-147 DAB [1]. Each RF output signal is transmitted via block B of TV channel 12 which center frequency is 207.008MHz and bandwidth is 1.536MHz. The applied RF transmission parameters of T-DMB are summarized in Table I.

2.2 Measurement System

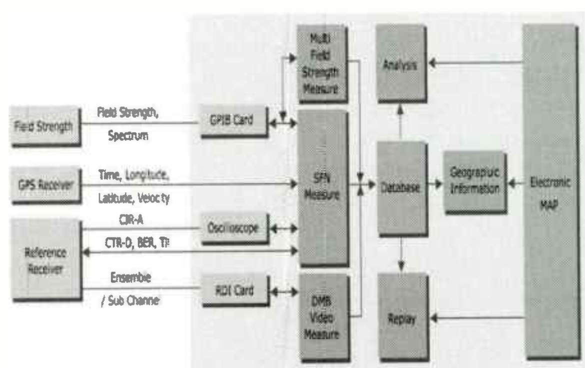


Fig. 4. Integrated measurement and analysis system

To implement SFN and estimate service area, we developed the integrated measurement and analysis system for T-DMB. Fig. 4 illustrates the integrated system used to measure SFN quality, electric field strength, video quality, audio quality and location coordinates acquired using GPS (Global Positioning System) receiver. Our measurement system consists of a SFN measurement system, a video quality measurement system, an electric field strength multi-measurement system, a replay system, a geographic information analysis system, database, electronic map and the related hardware equipment including spectrum analyzer, GPS receiver, reference receiver, oscilloscope, and interface cards. Two whip antennas with 0dB gain are mounted at the rooftop of test vehicle for receiving T-DMB signals. Using this system, SFN quality, electric field strength, video quality, audio quality at specific locations can be obtained. Every measured value is recorded and saved into database with the corresponding location coordinates. The test vehicle is depicted in Fig. 5.

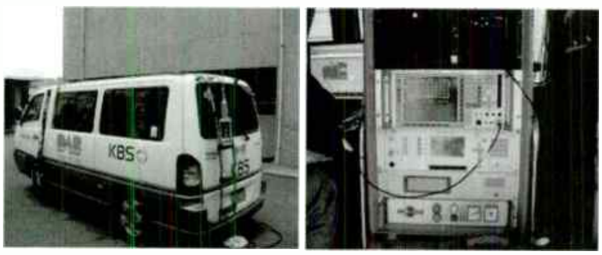


Fig. 5. Test Vehicle (Outside and Inside)

2.3 Field Test Results

The field tests were performed on the routes in Seoul and neighboring Gyeonggi Province area using test vehicle which install the measurement system. The routes include the inner city, major highways and intercity roads toward all directions to reach marginal areas of service coverage. The prior field tests before the commercial broadcasting service mainly focused on measuring electric field strength and BER (Bit Error Rate) performance according to electric field strength. We have already investigated the appropriate electric field strength level for error free reception in T-DMB services with the aid of the prior field tests. Therefore, in this paper, we estimate the service area depending on the electric field strength level.

The T-DMB system, based on the Eureka-147 DAB system, enables to construct SFN [1], [9]. The time synchronization with neighboring transmit site is the most important factor to decide SFN quality. Badly synchronized signals degrade S/N (Signal to Noise) ratio and cause inter-symbol interference. Normally, SFN quality is measured by CIR (Channel Impulse Response), and all of the peaks from transmitters should be carefully controlled and located within the guard interval. In addition, the distance between the first and last peak should be equal to the distance difference from the receiving side to the corresponding transmit sites. Fig. 6 illustrates SFN quality in the receiving side. The dotted box on the right side represents CIR, and the two peaks are located within the guard interval which is shown as a bright area in the center. The dotted box on the left side represents TII (Transmitter Identification Information), and it gives the information where the received signals come from. The TII shows that the received signals are constructed from the two transmit sites, Mt. Gwanak (second peak in CIR) and Mt. Nam (first peak in CIR) with nonzero strength. The distances from the vehicle to two transmit sites are 10.850km and 7.030km, respectively. The difference is 3.82km, which is equal to the distance calculated between two peaks in CIR. Therefore, we can decide the time synchronization of two transmit sites is satisfactory enough to construct SFN.

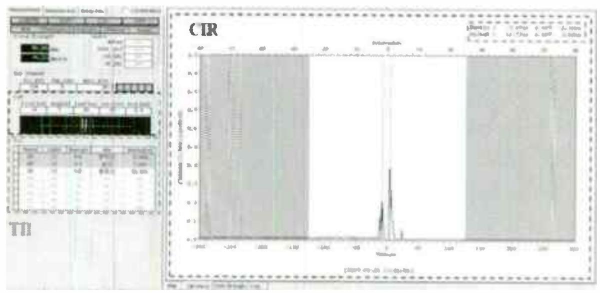


Fig. 6. SFN quality in the receiving side

The service area corresponding to the electric field strength level along the inner city roads of Seoul is depicted in Fig. 7. The electric field strength of 37dBμV/m means that commercial receivers of the best sensitivity can provide an error-free reception by rooftop mounted antenna. 45dBμV/m means that commercial receivers of the normal sensitivity can provide an error-free reception by the rooftop mounted antenna (That is the level defined in domestic T-DMB standard). 50dBμV/m means that commercial receivers of the worst sensitivity can provide an error-free reception by the rooftop mounted antenna. 65dBμV/m means that the error-free reception is possible in the inside of the vehicle or in-building by T-DMB mobile phone. It includes the penetration loss (about 8~14dB) and the unfavorable antenna gain of mobile phone. 75dBμV/m means the average electric field strength measured along the downtown roads. These designated values are determined by several times of experiments. The percentage of satisfying the designated electric field strength with respect to the total measured electric field strength is summarized in Table II. An error-free reception in the vehicle by the rooftop mounted antenna is possible more than 99% of the locations. In addition, an error-free reception by the mobile phones inside the vehicle is possible in about 75% of the locations.

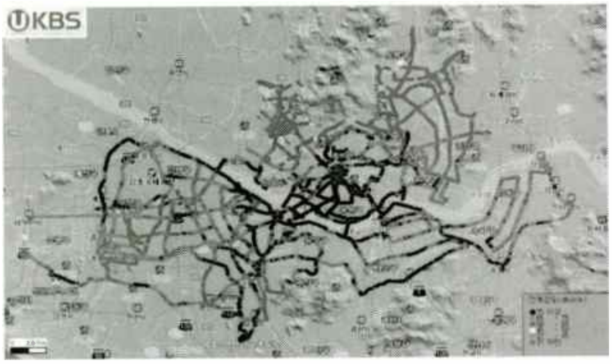


Fig. 7. Electric field strength along the inner city roads of Seoul

TABLE II

Electric field strength at inner city roads of Seoul

	Total distance (km)	Designated distance (km)	Ratio (%)
37dBuV/m	837	837	100
45dBuV/m		836	99
50dBuV/m		828	99
65dBuV/m		629	75
75dBuV/m		321	38

Fig. 8 and Table III show the service area along the inner city, major highways and intercity roads toward all directions to reach marginal areas of service coverage in Seoul and neighboring Gyeonggi Province. With no obstacles between the transmitter and the receiving vehicle, an error-free reception is possible despite the much longer distance in some locations. The average electric field strength level is considerably decreased at the outside of Seoul, and it is mainly due to geographical factors.

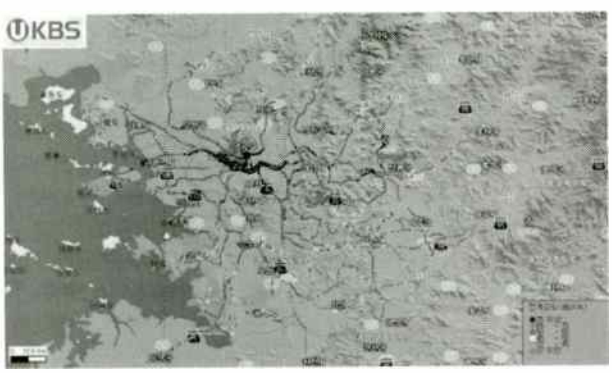


Fig. 8. Electric field strength in Seoul and neighboring Gyeonggi Province

TABLE III

Electric field strength at Seoul metropolitan area

	Total distance (km)	Designated distance (km)	Ratio (%)
37dBuV/m	3052	3040	99
45dBuV/m		2592	85
50dBuV/m		2139	70
65dBuV/m		933	31
75dBuV/m		430	14

The current service area of T-DMB in Korea is illustrated in Fig. 9. The blue, green, yellow and red colors represent the areas of more than electric field strength of 65dBμV/m, 50dBμV/m, 45dBμV/m, and lower than 45dBμV/m, respectively. The shadow area is being reduced as the new transmit sites and the relay stations are gradually installed to increase the electric field strength up to higher than the required level for error-free reception.

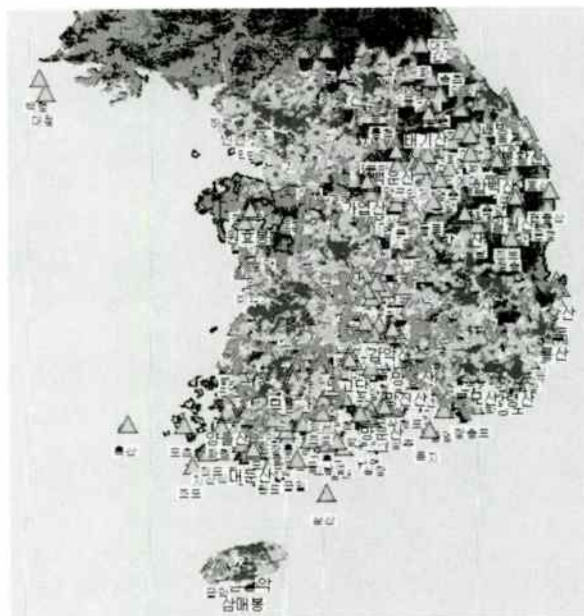


Fig. 9. T-DMB Service area in Korea

3. Handover Field Test Results

3.1 Transmission Network

The advantage of T-DMB handover is to provide users with the convenience of not manually following the tuned service when they move beyond the current regional broadcast network. It also gives the effect of reducing the uncovered areas by mutually compensating for the neighboring networks. To confirm the performance of handover of T-DMB system, handover parameters need to be measured in field tests.

TABLEL IV

Transmission parameters for handover field tests

Region		Site	Center Freq. (MHz)	Transmit Power (kW)	ERP (kW)	Ant.	Ant. Height (m)
Seoul	A	Mt. Gwanak	207.008	2	16.3	24P	642
		Mt. Nam		1	9.9	16P	325
		Mt. Yongmun		1	17.9	4P	1177
Chungcheong	B	Mt. Gyeryong	201.008	2	9.8	16P	814
		Mt. Sikjang		2	9.3	16P	580

The objective of the handover field test is to measure and analyze the quality of T-DMB handover in order to derive optimal handover parameters and to estimate an effect of handover on the change of the service area prior to launching commercial handover service. The field tests were performed on the route in Seoul metropolitan area and neighboring Chungcheong Province. The applied RF transmission parameters for the field test are summarized in Table IV.

3.2 Handover algorithm

In general, the handover algorithm can be defined as Fig. 10. In order to increase the handover performance, deciding the appropriate moment to trigger handover process and reducing the handover processing time need to be more carefully considered. The parameters applied in this field test are given in Table V. Note that those parameters are decided by several times of local tests, and they are proved to be efficiently applied for T-DMB handover although other parameters can be also selected.

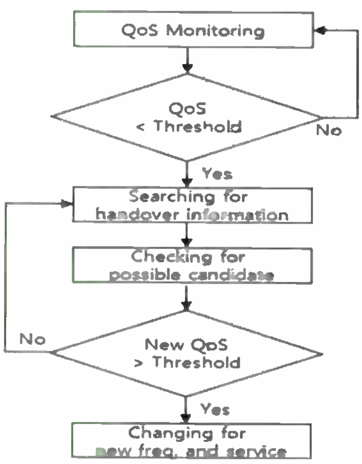


Fig. 10. Handover algorithm

In our field tests, FIB CRC (Fast Information Block Cyclic Redundancy Check) was selected as the QoS (Quality of Service) parameter. In the T-DMB system, the transmission frame is made up of FIC (Fast Information Channel) which is normally used to signal the carried services and MSC (Main Service Channel) which is used to carry video, audio and data services [1], [7]. FIB is transmitted in FIC, and the receiver should firstly handle this information to access each service. In addition, the time interleaving (360ms in T-DMB) is not applied to FIC, so there is no time delay to confirm the validity of FIB, which is contrary to MSC. Therefore, FIB CRC can rapidly reflect the changes of the receiving environments while moving. Since it is also obtained after channel decoding process in the receiver, it can become the reference value to represent error-free reception similar to BER. In case of the electric field strength and S/N, there are frequent changes in their level depending on the geographic locations or the vehicle speed, and it is impractical to discern the interference caused by analog TV channels which are possibly located in the same frequency band with T-DMB only by electric field strength.

TABLEL V

Handover algorithm parameters for field test

QoS Parameter	FIB CRC
QoS Threshold	Number of FIB CRC error = 10
Handover initiation	Time window length = 3.6 second
Handover completion (channel change)	To find the alternative service satisfying QoS threshold
Handover completion (no channel change)	To find the current service satisfying QoS threshold again
Timeout	Not to complete the handover process within the designated time of 10 second

The handover process is initiated whenever the QoS parameter dissatisfies the designated level during some period of time (about 3.6s in our algorithm). The time window length should be carefully designed by considering the trade-off with the handover processing time. The longer time window increases the ability to manage short-term fading, which can be recovered rapidly without any measures as the vehicle moves away the location, but it can increase the time delay for finding the proper alternative.

3.3 Measurement System

To measure the feasibility and the performance of T-DMB handover service, we developed the integrated measurement and analysis system. The structure of the integrated system is similar to that of illustrated in Fig. 4 except that two reference receivers are installed to measure the current service and the alternative service simultaneously. Fig. 11 shows the handover field test system to measure FIB CRC error, electric field strength, video quality, SFN quality, handover quality and location coordinates. Every measured value is recorded and saved into database with the corresponding location coordinates.

The parameters of handover quality measured in the field test system are coverage discontinuity, coverage overlap, handover location, timeout, channel change, current channel, and best QoS channel. The coverage discontinuity can be defined when there are no channels to satisfy the required QoS at some locations. On the contrary, the coverage overlap can be defined when more than two channels satisfy the required QoS.

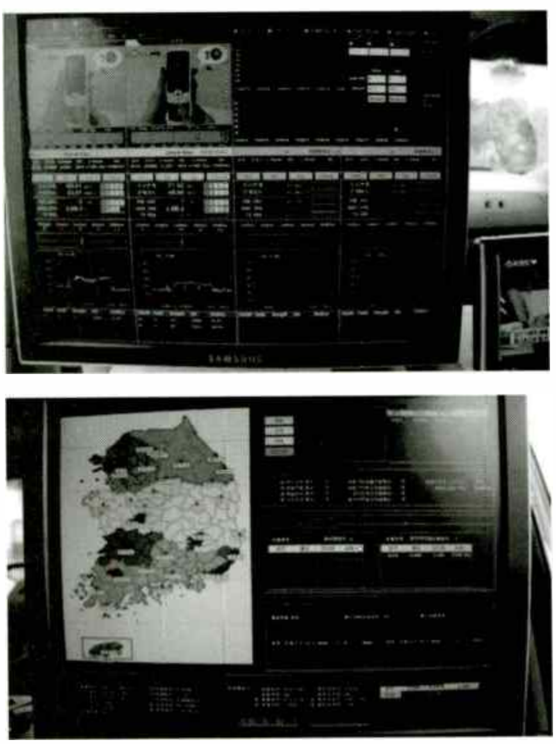


Fig. 11. Handover field test system

3.4 Field Test Results

The handover field test results are given in Table VI. The routes include the major highways and intercity roads between the two neighboring T-DMB networks which are given as A and B in Fig. 1, and the distance is 103.1km. Although the number of handover trials is 50, there are no more than 16 times of the channel changes to the alternative service, and the other cases still remain in the tuned channel. It implies that several times of short-term fading in the current channel exist along the routes, but its QoS can be recovered again as the vehicle moves away. The integrated system can decide to stay in the current channel or to change into the other channel during 11.9km, but it is more favorable not to change its frequency because the frequent change can cause an interruption or confusion to users. The discontinuity distance of 4.1km accounts for 8 times of timeout.

Table VI
Handover field test results

Item	Value
Total distance	103.1km
Overlap distance	11.9km
Discontinuity distance	4.1km
Current channel distance (207.008MHz)	82.6km
Current channel distance (201.008MHz)	20.5km
Number of handover	50
Number of channel change	16
Number of no channel change	34
Number of timeout	8

Fig. 12. Handover trials

Fig. 12 shows the handover trial locations along the route. In detail, the quality of the currently tuned service is degraded at the locations shown in Fig. 12, and the field test system attempts to change its current service into the alternative. The current channel at each location is illustrated in Fig. 13. The red color represents 207.008MHz, and the blue color represents 201.008MHz. In some of the handover trial locations in Fig. 12, the current channel is not changed into its alternative because the quality of the current service can be recovered again before the receiver finds the appropriate alternative.

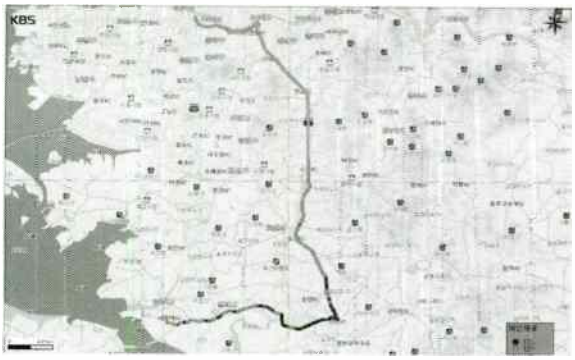


Fig. 13. Current channel

Fig. 14. Number of FIB CRC error of 207.008MHz

Fig. 15. Number of FIB CRC error of 201.008MHz

Figs. 14 and 15 show the distribution of FIB CRC error in 207.008MHz and 201.008MHz, respectively. The higher FIB CRC error leads to a decrease in the possibility of error-free reception, and the handover process begins when the number of occurred error surpasses the threshold level. Note that some of the badly received locations in one channel are covered by the other channel. It means that the effect of spatial diversity can be obtained between the neighboring networks and the error-free reception with no additional transmission sites can be achieved at more locations only by adopting the handover service.

4. Conclusion

The objective of our field tests is to measure and estimate service area of error-free reception, and it is also aimed to measure and control the quality of the constructed SFN. With no obstacles between the transmitter and the receiving vehicle, an error-free reception is possible even in the considerably long distance. Although the average electric field strength level is higher than the required value for mobile reception in downtown area, it is decreased at the outside of Seoul, and it is mainly due to geographical factors. The time synchronization for high quality SFN is satisfactory within the guard interval.

Furthermore, our field tests were performed to verify the feasibility and effect of the handover service in T-DMB network. In this test, we found the T-DMB network performance can be upgraded by introducing the handover service because the neighboring networks compensate for each other's shadow areas.

REFERENCES

[1] Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to Mobile, Portable and Fixed Receivers, ETSI EN 300 401 v1.4.1, June 2006.

[2] W. Hoeg and T. Lauterbach, Digital Audio Broadcasting, Principles and Applications. Hoboken, NJ: John Wiley & Sons, 2001

[3] Digital Audio Broadcasting (DAB), DMB video service; User Application Specification, ETSI TS 102 428 v1.1.1, June 2005.

[4] Digital Audio Broadcasting (DAB), Data Broadcasting-MPEG-2 TS streaming, ETSI TS 102 427 v1.1.1, July 2005.

[5] Broadcasting of multimedia and data applications for mobile reception by handheld receivers, ITU-R BT.1833, Dec. 2007.

[6] Advanced Video Coding for Generic Audiovisual Services, ITU-T Rec.H.264-ISO/IEC 14496-10 AVC, Mar. 2003.

[7] Digital Audio Broadcasting (DAB); Guidelines and rules for implementation and operation, Part 1 System outline, ETSI TR 101 496-1 v1.1.1, Nov. 2000.

[8] Digital Audio Broadcasting (DAB); Guidelines and rules for implementation and operation, Part 2 System features, ETSI TR 101 496-2 v1.1.2, May 2001.

[9] Digital Audio Broadcasting (DAB); Guidelines and rules for implementation and operation, Part 3 Broadcast network, ETSI TR 101 496-3 v1.1.2, May 2001.

[10] M.C.D. Maddocks, I.R. Pullen, and J.A. Green, "Field Trials with a high-power VHF single frequency network for DAB," EBU Technical Review, pp. 36-55, Autumn. 1994

[11] A. Lau, M. Pausch, and W. Wutschner, "First results of field tests with the DAB single frequency

network in Bavaria," EBU Technical Review, pp. 4-27, Autumn. 1994

[12] YoungJin Lee, SangWoon Lee, YongHan Kim, SooIn Lee, ZungKon Yim, ByungHo Choi, SangJin Kim, and JongSoo Seo, "Field Trials for Terrestrial Digital Multimedia Broadcasting System," IEEE Trans. Broadcasting, pp. 425-433, Mar. 2007

AN IMPROVED COVERAGE PREDICTION METHOD FOR HD RADIO

John Kean
NPR Labs, National Public Radio
Washington, DC, USA
jkean@npr.org

ABSTRACT

As part of a comprehensive study of broadcast signal coverage, NPR Labs developed a model to predict the signal coverage of In-Band On-Channel DAB ("HD Radio®"). In the course of this project several new techniques for testing of HD Radio receivers in the laboratory, and broadcast signals by vehicle. This paper presents these measurement techniques, the development of the model and the results in predicting actual station coverage.

INTRODUCTION

With the introduction of HD Radio, a need has arisen for an accurate method to predict the signal coverage of stations broadcasting this new digital audio broadcast system. The transmission characteristics and compatibility of HD Radio was tested by the National Radio Systems Committee before adoption by the FCC.[1] Seen as an elective addition to existing FM stations, however, less attention was given to determining the independent coverage of HD Radio, particularly in the presence of signal interference from neighboring stations that are 1st- or 2nd-adjacent channel. As no model was available to NPR Labs to prepare accurate coverage maps of U.S. public radio stations a development program was undertaken.

The project began in late 2006 and is near completion in early 2008. All of the measurement programs are completed, the data is compiled, and the prediction model is developed and tested. The model was implemented in software to produce coverage maps of approximately 850 public radio stations.

Prior Measurements by NPR Labs

The coverage potential of HD Radio has been of interest to NPR Labs since the initial member station rollouts in 2004. In our first program, we constructed a portable measurement system capable of recording the locations at which an HD Radio receiver operated successfully and the analog host FM signal strength.[2] This study reported on reception of 26 public radio stations using vehicle-based measurements.

All the drive-test measurement data of each station was analyzed statistically to derive a distribution curve of HD Radio reception availability versus field strength. Figure 1, below, shows the results for three different FM stations, along with the combined mean of all three. It is apparent that the required field strengths vary widely for a given reception availability, such as 95%.

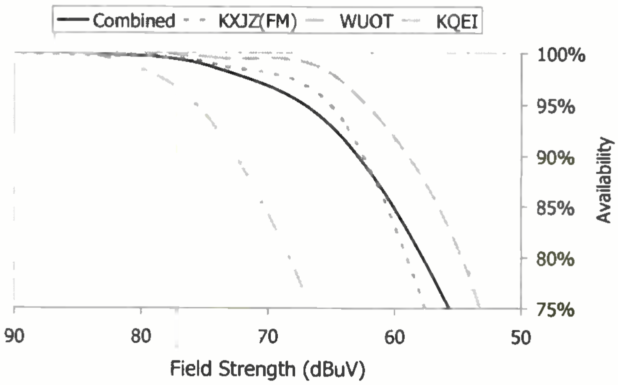


Figure 1. Example of HD Radio reception availability versus field strength measured for three stations.

While the study derived an average for all 26 stations, the standard deviation of the results was wide: a variety of factors in the methodology was affecting the individual station results. The wide spread in field strengths for the same availability meant that we would not be able to confidently use field strength alone to predict the locations where HD Radio reception was available.

Several factors contributing to the variations were discussed in the report, including differences in gain of the magnetic-mount antenna used on the roof of different vehicles, differences in electrical noise generated by the measurement vehicles, environmental noise and multipath around different stations, and signal interference from other stations. During this follow-on study, considerable effort was taken to identify the factors affecting reception the most and minimize extraneous factors. Recent laboratory testing has determined that, given average conditions for vehicular reception, the most important factors affecting reception, after desired channel signal strength, are the signal ratios of stations on first-adjacent channels and the same channel.

HD RADIO RECEIVER MEASUREMENTS

To determine the interference ratios affecting HD Radio reception, as well as other potential impairments, an RF Test bed was built to measure consumer receivers. Using threshold digital receive failure as the criteria, the measurement objectives of the receivers were:

- Unimpaired sensitivity
- Sensitivity with Additive White Gaussian Noise
- Co-channel interference ratio
- 1st-adjacent channel interference ratio (single and dual)
- 2nd-adjacent channel interference ratio (single and dual)
- 3rd-adjacent channel interference ratio
- Tests of the above with analog FM interfering signals as well as hybrid interfering signals
- Tests of the above with Raleigh fading on either the desired channel, interfering channel, or both

For the hybrid signal generation systems a Harris Dexstar HD Radio exciter was combined with a Hewlett Packard 8647A FM generator. The Dexstar and the FM generator operate at a constant +3 dBm output; with a 20 dB attenuator connected to the Dexstar, the combined output level of the system was approximately 0 dBm. Three of these systems were used; one for the Desired Channel and two for Undesired Channels #1 and #2. All of the HD Radio exciters were connected to a central GPS antenna on the roof of NPR Labs. A 10 MHz GPS-derived signal from one Dexstar was used to synchronize the other RF test instruments.

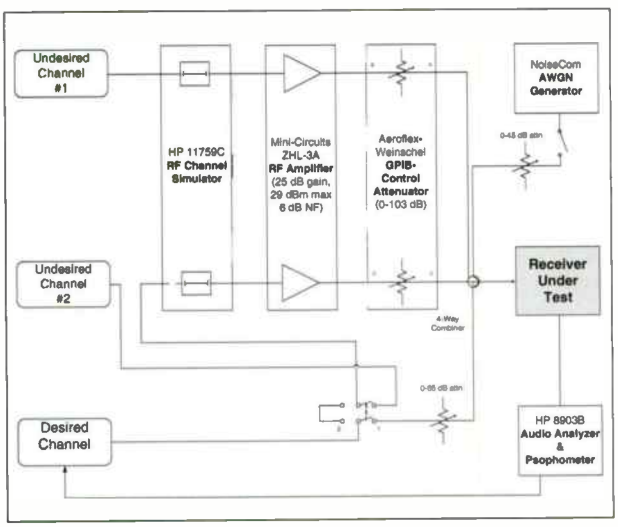


Figure 2- Receiver test bed for analog and IBOC DAB receiver measurements

As shown in the simplified diagram of Figure 2, the signal generator systems were connected to a Hewlett Packard 11759C Channel Simulator, which was used to introduce Rayleigh fading. A two-channel RF attenuator system by Aeroflex-Weischel provided control of the RF levels over a 0 to 103 dB range in 1 dB steps.

System levels were made up by two high-performance RF amplifiers made by Mini-Circuits so that the maximum input level to the Receiver Under Test was approximately -7 dBm on any FM carrier. This is equivalent to a dipole field strength of 109 dBuV. Higher input powers to the receiver were found to be necessary to cause 2nd- and 3rd-adjacent reception failure with some receivers, but the dynamic range of the system was a good tradeoff between noise floor and high-level receive conditions.

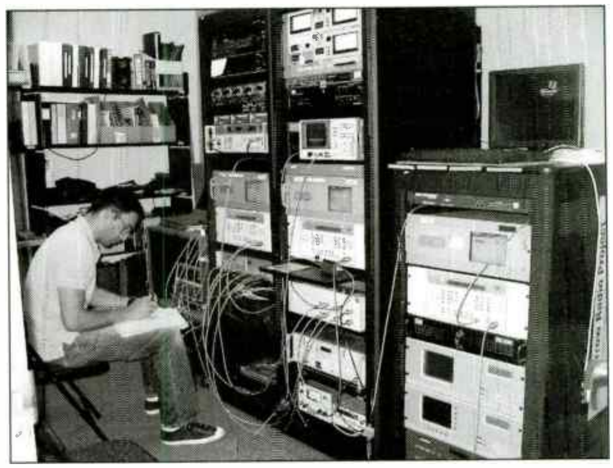


Figure 3- The RF Test bed in use

A photo of the RF Test bed is shown in Figure 3. Not visible is the computer workstation that was networked to the instrumentation. This computer runs MATLAB software to automate the measurements. (It was decided early in the project that the number of tests would be cumbersome to collect manually. As we learned later, as the number of measurement points climbed, automation was a fortunate decision.) An Ethernet switch is connected to the Dexstar exciters, to control emission and frequency, and to an Audemat-Aztec FMX480 stereo generator on the Desired Channel generator system. A GPIB network was set up to communicate to the other instrumentation. A Hewlett Packard 8903B audio analyzer was included in the test bed for audio SNR testing and generation of audio test signals under MATLAB control.

To the left of the tall racks, and just visible in the photo past our lab intern, is a Ramsey Electronics STE5000 shielded test enclosure. Used with double-shielded coaxial cables, this cabinet allows us to test unshielded consumer receivers without signal leakage from broadcast stations and avoids the need for a large screen room.

An Additive White Gaussian Noise Generator using the NoiseCom NC6110 noise source was available for noise impairment tests. Noise levels of 30,000°K and 300,000°K (degrees Kelvin) were chosen for sensitivity impairment tests. These levels were proposed by iBiquity Digital, which commissioned two consulting firms to determine the noise level caused by co-channel analog interference. The noise level was approximately 30,000°K, which is equivalent to a field strength of 15 dBuV [3]. While iBiquity intended this noise to be added for analog compatibility testing of IBOC DAB, AWGN is a good representation for outdoor environments in which to test HD Radio receivers. For additional comparison, a level 10 dB higher (300,000°K) was included in our testing.

Sensitivity Measurements and Link Budgets

To date, a total of 17 HD Radio receivers have been measured on the Test bed. The types represented are: after-market car radios, hi-fi tuners, professional tuner/monitors, and table radios. As mentioned earlier, sensitivity was performed for all receivers automatically by the test bed with two levels of Gaussian noise as well as with no added noise.

Table 1- Block Error Rate Results with AWGN and Rayleigh Fading Profiles

Test	Cd/N	Fading	Block Error
Gaussian Noise	54.1	None	0.16
	54.5		0.032
	55.1		0.0029
9-Ray Fading	55.4	Urban Fast	0.8
	56.4		0.056
	57.3		0.012
	59.3	Urban Slow	0.106
	60.4		0.054
	61.4		0.0202
	55.9	Rural Fast	0.6
	56.9		0.087
	57.8		0.007

Additionally, several receivers were tested manually to verify the results with Gaussian noise and Rayleigh fading that were supplied by iBiquity to the NRSC and ITU.[4] The iBiquity results are listed in Table 1 using block error ratios. The threshold of muting was used to estimate performance (consumer receivers do not provide a means of measurement BER), however, good agreement was found between the production receivers and the published results at the highlighted BER values.

The receiver faded performance threshold (FPT) was applied to the link budgets in Table 2, Table 3 and Table 4. The basic formula for the link budgets is:

FPT=Vi − Nr + Cd/No + IBACr + Kd + C - G + L + Lf

where the coefficients are as listed in the left column.

Using kTB with an equivalent noise power bandwidth of 140 kHz (both 70 kHz carrier groups combined) the thermal noise bandwidth Vi, of the receiver is -152.5 dBW. The noise figure of the receiver is estimated from performance tests and literature published by the manufacturers of tuner modules.

The carrier-to-noise value is taken from *Cd/N* in Table 1. Since the carrier power is expressed in dBm the thermal noise of the receiver is converted to the same units. A normalization is made to *Cd/N*, where *N* is across one Hz, to the receiver noise bandwidth *B*, which is combined with the noise figure *Nr* for the required input power in dBm. It is NPR Labs' practice to express RF signal levels for testing and mapping in terms of the field strength of the analog host. Thus, 20 dB is added to the required input power for the required analog host FM power.

The next section converts the signal power into field strength by first calculating the dipole factor at 90 MHz (the middle of the FM Reserved Band, which most public radio stations operate), to which the 50-ohm dBm-to-dBu conversion factor *C* is added. The antenna gain relative to a dipole is based on NPR data and discussions with automotive receiver manufacturers.

Table 2- Link budget for vehicular reception

k	Boltzmann's constant	1.38E-23		W/K/Hz
T	reference noise temperature	290		degrees K
B	noise equivalent bandwidth of input of both carrier groups	140,000		Hz
Vi	thermal noise of receiver bandwidth	-152.5		dBW
			-122.5	dBm
Nr	noise figure of receiver input		6	dB
Cd/N	minimum CNR for acceptable service (9-ray terrain-obstructed for TOA) (1.0% BER, urban fast fading, ref. Cd (dBm) to No (dBm/Hz)	57.3		dB-Hz
	normalization of Cd/N to 1 Hz		5.9	dB
	required input power	-110.6		dBm
IBACr	1% IBAC ratio adjustment		20	dB
	required analog host FM power	-90.6		dBm
f	frequency of operation		90	MHz
Kd	dipole factor [$20\cdot\log(9.73/(\lambda\sqrt{G})$], where G=1.64		7.2	dB
C	dBm (50Ω) to dBuV conversion factor		107.0	dB
	antenna gain relative to dipole (NPR data)		-5	dB
L	transmission line loss		0	dB
FPT	incident field at 1.5m rcv. height		29	dBuV/m
Ne	Environmental Noise-adjusted Faded Performance Threshold per ITU-R P-372 model, noise above *kTB*: F = *c* - 27.7·Log(*f*), where *c* = 76.8, (business)	20		dB
	Normalization per TIA Report TSB-88.2-C, 10*log(1+Ne/Nr)		15	dB
FPTadj	FPTadj =FPT + 10*log(1+Ne/Nr)		43	dBuV/m

The result is the incident field strength at the vehicle in the absence of other RF noise. It should be noted that this field strength is not comparable to the FCC's F(50,50) curve predictions due to differences in reference height (9.1 meters vs. 1.5 meters). Also, the field strength is reduced due to the effects of signal scattering and absorption at low antenna heights.

To provide the minimum field strength suitable for vehicular reception, the last portion of the link budget adjusts the Faded Performance Threshold for environmental noise *Ne*. This may come from utility power lines, traffic and even the vehicle electrical systems. ITU-R Recommendation P-372 is used to estimate the adjustment relative to the thermal noise power *Vi*. Finally, *Ne* is normalized to avoid adding it to the receiver noise figure, per TIA Report TSB-88.2-C.[5] The resulting field strength, 43 dBuV/m, compares closely with NPR Labs' calibrated field strength drive-test measurements, discussed later herein, and our point-to-point pathloss prediction model.

Table 3- Link budget for indoor reception

k	Boltzmann's constant	1.38E-23		W/K/Hz
T	reference noise temperature	290		degrees K
B	noise equivalent bandwidth of input of both carrier groups	140,000		Hz
Vi	thermal noise of receiver bandwidth	-152.5		dBW
			-122.5	dBm
Nr	noise figure of receiver input, estimated		7	dB
Cd/N	minimum CNR for acceptable service (9-ray terrain-obstructed for TOA) (1.0% BER, urban fast fading, ref. Cd (dBm) to No (dBm/Hz)	55.1		dB-Hz
	normalization of Cd/N from 1 Hz to *B*		3.6	dB
	required input power	-111.9		dBm
IBACr	1% IBAC ratio adjustment		20	dB
	required analog host FM power	-91.9		dBm
f	frequency of operation		90	MHz
Kd	dipole factor [$20\cdot\log(9.73/(\lambda\sqrt{G})$], where G=1.64		7.2	dB
C	dBm (50Ω) to dBuV conversion factor		107.0	dB
	antenna gain relative to dipole, (500mm whip, BBC 1990)		-15	dB
L	transmission line loss		0	dB
LB	building loss factor, 50th percentile, (single-story, 90 MHz, Skomal & Smith)		9	dB
Lf	location variability factor (20% likely to exceed building loss)		8.4	dB
FPT	incident field at 1.5m rcv. height		55	dBuV/m

Table 4- Link budget for portable reception

k	Boltzmann's constant	1.38E-23		W/K/Hz
T	reference noise temperature	290		degrees K
B	noise equivalent bandwidth of input of both carrier groups	140,000		Hz
Vi	thermal noise of receiver bandwidth	-152.5		dBW
			-122.5	dBm
Nr	noise figure of receiver input		8	dB
Cd/N	minimum CNR for acceptable service (9-ray terrain-obstructed for TOA) (1.0% BER, urban fast fading, ref. Cd (dBm) to No (dBm/Hz)	61.4		dB-Hz
	normalization of Cd/N to 1 Hz		9.9	dB
	required input power	-104.6		dBm
IBACr	1% IBAC ratio adjustment		20	dB
	required analog host FM power	-84.6		dBm
f	frequency of operation		90	MHz
Kd	dipole factor [$20\cdot\log(9.73/(\lambda\sqrt{G})$], where G=1.64		7.2	dB
C	dBm (50Ω) to dBuV conversion factor		107.0	dB
	antenna gain relative to dipole (BBC 1990, NPR data)		-20	dB
L	transmission line loss		0	dB
LB	building loss factor, 50th percentile, (single-story, 90 MHz, Skomal & Smith)		9	dB
Lf	location variability factor (20% likely to exceed building loss)		8.4	dB
FPT	incident field at 1.5m rcv. height		67	dBuV/m

The link budgets of Table 3 and

Table 4 cover the parameters for indoor and portable HD Radio reception, respectively. These tables include a median building loss at 90 MHz and a location variability factor to adjust to 80 percent of the best locations (i.e., 20 percent of locations are expected to exceed this loss factor).[6] These table indicate the incident field required at the exterior of a building, which accommodates prediction with NPR Labs' pathloss model mapping.

The indoor and portable link budgets follow the same layout as for vehicular reception, except without the adjustment for environmental noise. Since these tables include additional losses for building penetration and location variability, we find that the required field strengths exceed the (outdoor) environmental noise predicted by ITU-R P-372. Although local sources of indoor noise may elevate the RF noise level substantially and require higher minimum field strengths, these are potentially-correctable by the listener, at least in the case of fixed indoor reception.

These field strengths may again appear too low to the reader, when considered against familiar FCC field strength levels. NPR Labs has conducted indoor measurements with a special portable measurement system at numerous locations in the Washington, DC, area, including homes, office buildings and shopping centers. We find the results of HD Radio reception availability correlate with the portable receiver link budget of

Table 4.

HD Radio Receiver Lab Measurements Leading to Interference Model

In the following section we present some lab measurement results of the Kenwood KTC-HR100MC after-market car radio, one of the best-performing receivers and the basis of the NPR HD Radio Logger measurements (discussed later herein). The results were obtained with hybrid (HD Radio) interferers operating service mode MP3 (including two additional Extended Hybrid Partitions).

Figure 4 plots the ratio of cochannel desired and undesired signals at which digital receive failure occurs. It is apparent that for unfaded (steady) signals the D/U ratio remains relatively close to 4 dB over a wide range of desired signal powers. With a "Trimmed Urban Fast" Rayleigh fading profile (60 km/hr) used by the NRSC, the cochannel measurements required and increase of approximately 3 dB in the D/U protection ratio, increasing gradually at lower signal powers [7]. This change is due to 30,000°K noise in the Test bed filling in larger portions of the faded signal and increasing the data error rate. (For comparison, FCC allocation rules require a minimum 20 dB protection ratio between desired and undesired analog FM signals.)

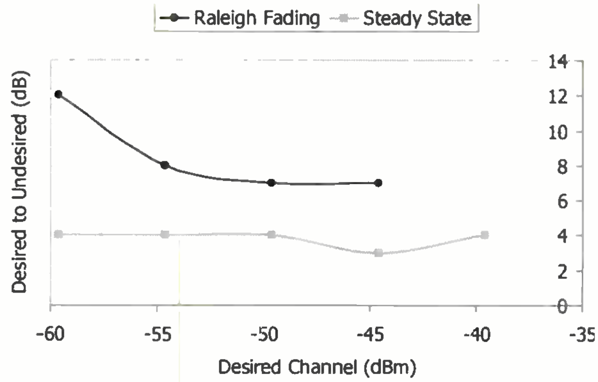


Figure 4- Co-channel interference ratio for steady and Raleigh fading signal

Figure 5 plots the ratios of 1st-adjacent interference necessary to cause digital receive failure with the Kenwood tuner. The upper (89.9 MHz) and lower (89.5 MHz) adjacent channels are shown for a Desired Channel frequency was 89.7 MHz with steady signals. The D/U ratios remain relatively constant, around -14 dB ±2 dB, over a very wide signal range. This receiver also shows good symmetry in upper and lower ratios, which is an important factor in the reliability of the following measurements with dual (upper and lower) interferers. (Note that the FCC's allocation rules require a minimum D/U of 6 dB between 1st-adjacent stations at the 60 dBu contour of the desired channel. This is approximately 20 dB less tolerant of interference.)

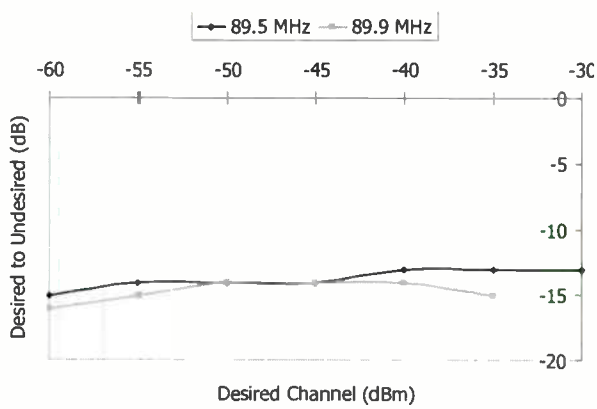


Figure 5- Single 1st-adjacent channel interference ratios for upper and lower channels

The reader should be aware that HD Radio transmits data simultaneously in upper and lower data carrier groups, which extend from approximately 129 to 199 kHz on each side of the FM carrier frequency. Interference to one carrier group by an adjacent channel interferer has a minor effect, compared to interference that may occur to both carrier groups.

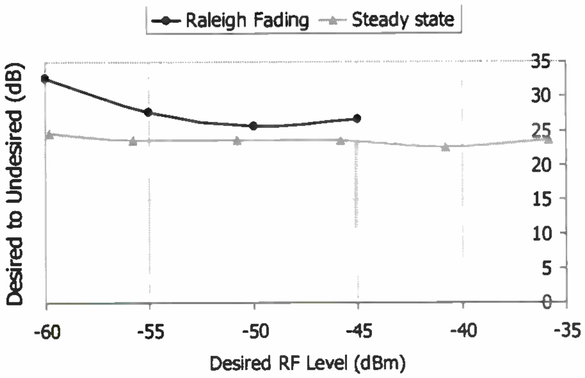


Figure 6- Ratio of dual 1st-adjacent channel interference at equal level, relative to a Desired signal

Figure 6 plots the results of interference from dual adjacent channel interferers (one above, one below the desired channel) and of equal level. The Desired Channel signal was fixed. The graph shares some similarity to Figure 4, except that the D/U scale is shifted much higher, that is, requiring a much higher level of protection than co-channel interference and very much higher than with a single adjacent interferer. This test is more to demonstrate a point of dual-adjacent interference, as the likelihood of identical levels is remote.

A more probable condition for interference is illustrated in Figure 7, in which threshold interference ratios are plotted as a function of both adjacent interferers. Considering both on a two-dimensional plot requires a unique presentation: the graph lines show the signal ratio of the Desired Channel to the *stronger* adjacent channel interferer on the x-axis *versus* the ratio of the *weaker* to stronger adjacent channel interferer on the y-axis. Tests were performed at Desired Channel signal levels of -60 dBm and -50 dBm and show good agreement.

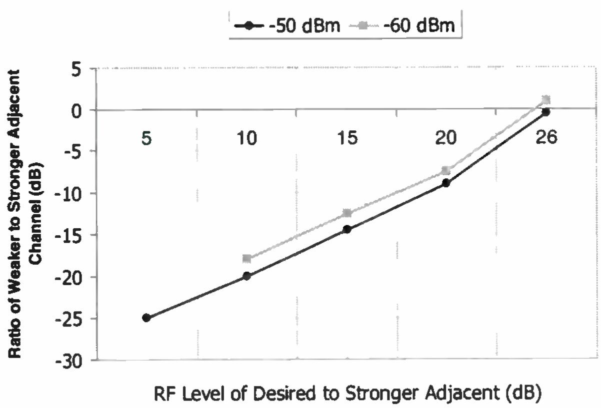


Figure 7- Dual 1st-adjacent channel interference ratios with desired channel at -50 dBm and -60 dBm

The preceding graphs are fundamental to the prediction of HD Radio interference-limited coverage. Using the observed behavior and measurement data a coverage and interference prediction model was developed, using the signal strengths of the desired station and upper- and lower-first adjacent and second-adjacent channels. In the next section we discuss the methods for measuring and verifying these laboratory measurements.

FIELD TESTS TO VERFIY THE MODEL

One of the concerns about the initial measurements of HD Radio coverage was the accuracy of the magnetic-mount antenna affixed to the top of vehicle for drive-testing. Not only did the antenna have an uncertain gain, relative to a dipole antenna, but it was also expected to have pattern distortion in the horizontal plane that would affect its gain depending of the orientation of the vehicle relative to the station being measured. Consequently a custom ground plane antenna was designed to help isolate the antenna from the vehicle's irregular roof shape. It was built by Kintronic Laboratories of Bristol, Tennessee, of expanded aluminum mesh in pie-shaped sections forming a disc of approximately five feet radius. A stainless steel whip is mounted to a metal plate at the center of the disc for connection to the receiver instrumentation.

The antenna was tested on a vehicle at the Table Mountain Test Range in Boulder, Colorado, operated by the Institute for Telecommunication Sciences, part of the U.S. NTIA. The test range has a large turntable that rotates the entire vehicle for pattern measurements. RMS gain of the antenna was estimated to be within 2 dB of a dipole at four test frequencies in the FM Band. A typical car installation is shown in Figure 8.

Figure 8- Ground plane antenna atop a test vehicle

For the drive test measurements, NPR Labs reconfigured its original HD Radio Logger system to permit simultaneous recording of up to four signals: HD Radio reception status, field strength of the Desired Channel station (carrying the HD Radio signal being tested) and field strength on two assignable frequencies, such as upper and lower 1st-adjacent channels. Figure 9 shows the basic components of the system, including an 88-108 MHz bandpass filter, RF amplifier and splitter feeding the four receivers. A micro-computerized data logger in one unit collected the receiver data along with GPS location and time stamps.

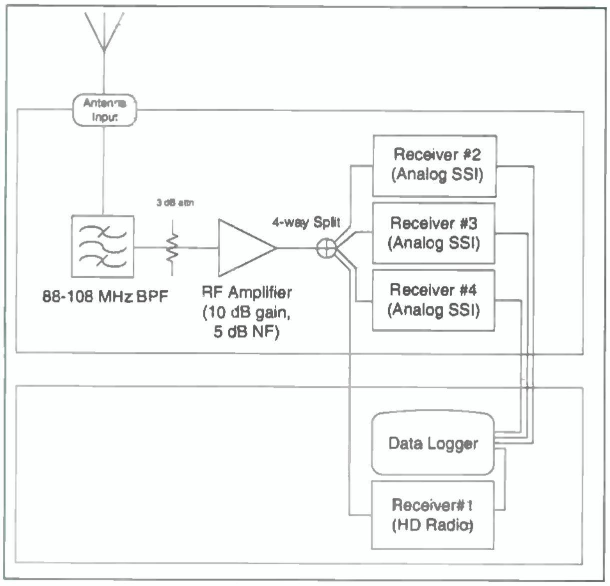


Figure 9- Simplified diagram of 4-receiver data logger system

The system was housed in two instrumentation cases that were powered by the vehicle's 12-volt system. All of the receivers were Kenwood KTC-HR-100MC "black box" tuners. The tuners provide a digital status of HD Radio reception along with a wide-dynamic range measurement of received signal power. The three tuners used for signal strength measurement were modified by replacing the ceramic filters used in the analog FM side chain with filters of greater selectivity, resulting in an adjacent channel response (±200 kHz) approximately 30 dB down. This permitted the tuners to collect more accurate field strength data in proximity of adjacent channel signals.

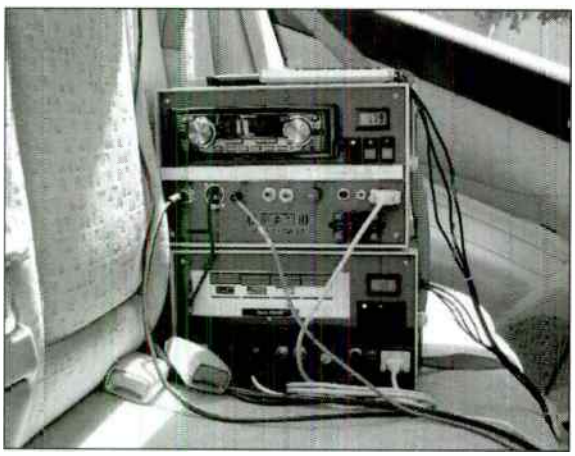


Figure 10. The HD Radio Logger system in service

The intention of the drive-test measurements was to drive around the Desired Channel station, probing the boundaries of the HD Radio coverage area. This yields information about the field strengths necessary to provide digital reception, as well as signal strength conditions on potentially-interfering channels when digital reception failed. No special attention was made to driving radially as the optimized pattern of the antenna helped minimize the effects of vehicle orientation.

The logger system is programmed to collect data from all sources (HD Radio and analog receivers, and GPS) four times per second. Active filtering in the loggers before A/D sampling remove signal fast fading effects. The completed data is post-processed to remove portions at speeds below about 20 miles per hour where log normal fading may be confused with Rayleigh fading. The data also is converted to received signal power and then into field intensity (in dBuV) after taking the antenna factor into account.

Using the logger calibration measurements shown in Figure 4 through Figure 7, a computer program was developed to model the reception expected of the HD Radio receiver. This model considered:

- The Desired Channel signal strength, and
- The ratios of any 1st-adjacent channel signals.

In a complete model the ratio of co-channel, 2nd-adjacent channel and 3rd-adjacent channel signals would be included, however, the logger system cannot measure cochannel signals and lacks additional channels to record other potentially-interfering channels. Laboratory measurements indicated that with the Kenwood receiver, interference from 2nd- and 3rd-adjacent channel signals is small enough to be neglected for the basic model. However, coefficients for these other channels, based on laboratory test data, are included in the final model.

Table 5- Public radio stations measured

Call	Channel	City, State	Interference
WWFM	Ch. 206A	Trenton, NJ	30%
KMPO	Ch. 204B	Modesto, CA	28%
WMUB	Ch. 203B	Oxford, OH	26%
KQEI-FM	Ch. 207A	North Highlands	23%
KVPR	Ch. 207B	Fresno, CA	14%
WNCU	Ch. 214C2	Durham, NC	14%
KXPR	Ch. 205B	Sacramento, CA	8%
WFYI-FM	Ch. 211B	Indianapolis, IN	4%
WHRV	Ch. 208B	Norfolk, VA	3%
WBEZ	Ch. 218B	Chicago, IL	0%

A four-hour drive test for a station broadcasting in HD Radio is illustrated in Figure 11. The station, WJFK-FM, Channel 294B, Manassas, Virginia, serves the Washington, DC metro area. It has two close-spaced adjacent channel neighbors to the north: WWMX, Channel 293B, Baltimore, and WWEG, Channel 295B, Hagerstown, both Maryland. The drive test ranged through parts of southern Maryland, the District of Columbia, and northern Virginia. The received signal powers of the three stations are plotted along the lower half of the graph. At the top of the graph the status of the HD Radio reception ("Measurement") is shown with the model's result ("Prediction"). Examination of these tracking lines show a high degree of correlation between measured and predicted performance. In the case of this drive-test, the model correlated 94% of the time, with 2.5% false positives and 3.6% false negatives.

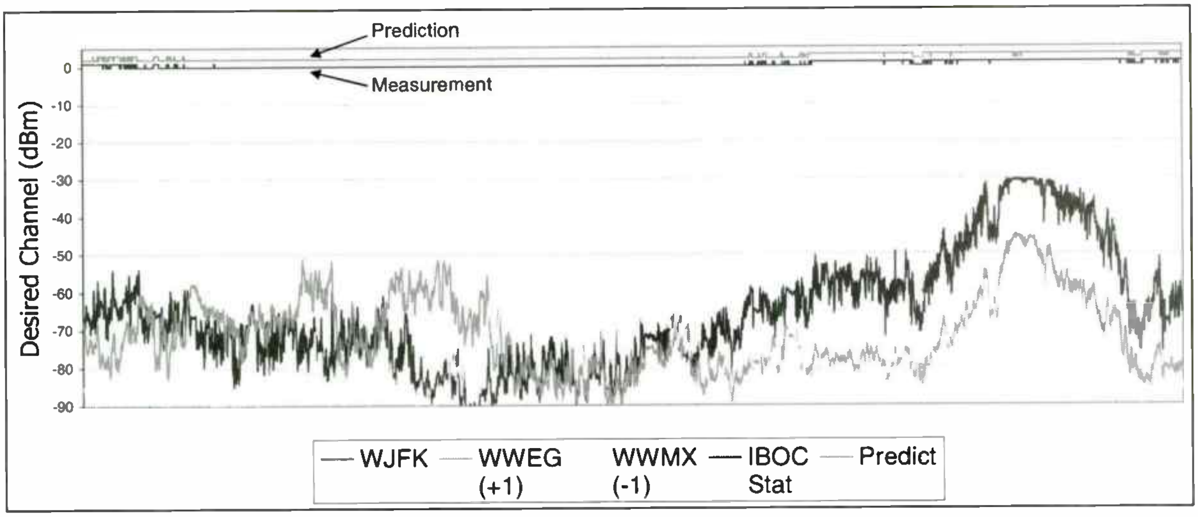


Figure 11- Four-hour timeline of drive-test measurements for WJFK-FM and two first-adjacent channel stations showing the RF levels in dBm; top compares measured and predicted HD Radio reception status (high=received, low=not received)

Ten more stations across the country have been measured with the logger system and analyzed against the model. The stations listed in Table 5 were selected to represent a cross-section of facility size, topography and estimated HD Radio coverage loss due to interference. The same equipment and measurement techniques as WJFK were used. After data processing, the model correlated with actual HD Radio reception between 85% and 95% for all station drive-test routes. The reliability of the algorithm in prediction interference-limited service and its simplicity encouraged NPR Labs to seek a patent for the model.[8]

Final Results

The final steps for the current project are the preparation of coverage maps for approximately 850 CPB-qualified public radio stations in the U.S., using the HD Radio coverage model with field strength predictions with the TIREM® point-to-point model.[9]

Included as Figure 12 and Figure 13 are maps showing the HD Radio coverage prediction for two of the ten stations in the drive-test. The presence of digital reception is indicated by "1" along the routes, while "0" indicates no reception occurred. The red-shaded areas show where NPR Labs' coverage model predicted interference to HD Radio service, that is, areas where reception would be possible in the absence of interference. The blue-shaded areas show interference-free reception of HD Radio. Comparison of the blue areas with the drive-test results indicates that the model performs well in predicting the availability of HD Radio Service. It is apparent that interference-limited digital coverage varies widely with these stations.

Figure 12- Predicted IBOC coverage of WNCU, Durham NC

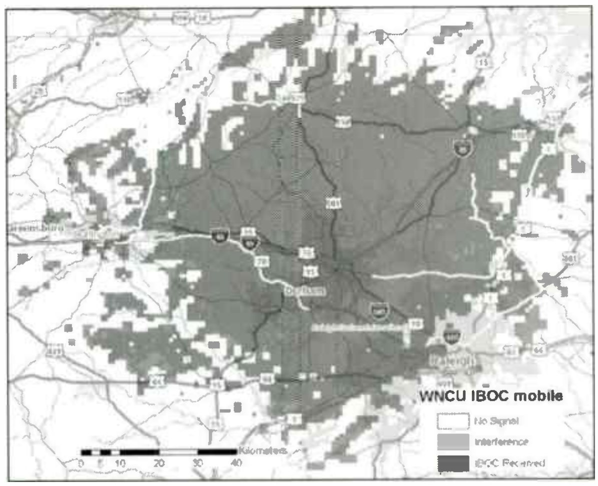
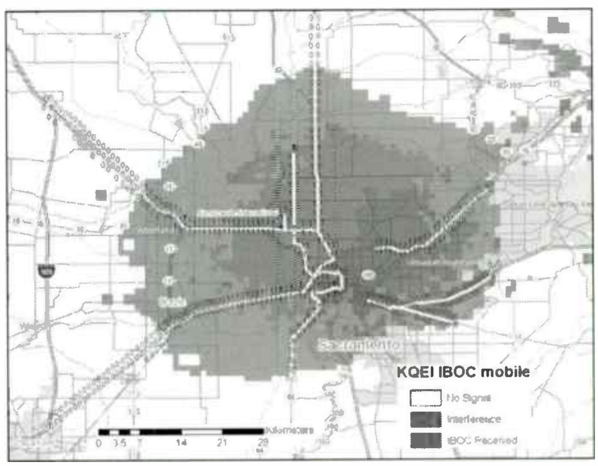


Figure 13- Predicted IBOC coverage of KQEI-FM, Sacramento CA



ACKNOWLEDGEMENTS

Thanks to our interns, Sam Goldman for developing the original MATLAB software to automate the Test Bed, and Babak Monajemi for optimizing software and conducting many of the receiver measurements. Kyle Evans, of the NPR Labs staff, deserves credit for generating and tracking the thousands of station maps required. Our appreciation to Harris Broadcast for loan of some test instrumentation, and our thanks to the Corporation for Public Broadcasting for the vision to fund the *Digital Radio Coverage and Interference Assessment project.*

REFERENCES

---

[1] National Radio Systems Committee, *NRSC-5-A, In-band/on-channel Digital Radio Broadcasting Standard*, September, 2005, and associated Reference Documents supplied by iBiquity Digital Corporation, Columbia MD.

[2] Kean, J., *HD Radio Coverage Measurement And Prediction*, presented at the 2006 NAB Engineering Conference, available on the NPR Labs web site at http://www.nprlabs.org/publications/reports/20061116_HDRadio_CoverageMeasurement&Prediction_JCK.pdf

[3] iBiquity Digital Corporation, *The FM Broadcast Band: Service Area Noise Floors in the US*, (presented to NRSC TPWG 11/09/2000).

[4] ITU-R Question 107/10, Document 6/203(Rev.1), *Systems for terrestrial digital sound broadcasting to vehicular, portable and fixed receivers in the frequency range 30-3 000 MHz*, 8 March 2002.

[5] Telecommunication Industry Association, *Wireless Communications Systems – Performance In Noise And Interference - Limited Situations*, Telecommunications Systems Bulletin TSB-88-B, Ballot Working Document PN3-4744, 2004.

[6] Skomal, Edward and Smith, A. A.: *Measuring the Radio Frequency Environment*, Van Nostrand, Reinhold Company Inc, 1985.

[7] Advanced Television Technology Center, *IBOC Laboratory Test Procedures – FM Band, Document No. 01-03*, Alexandria, Virginia, 2001.

[8] *A Method for Determining the Usable Signal Coverage of In-Band On-Channel Digital Audio Broadcast Transmissions*, USPTO # 60984421, Nov. 1, 2007.

[9] Alion Science & Technology Corp. (licensor), *Terrain Integrated Rough Earth Model (TIREM™)*, described at http://www.alionscience.com.

ADVANCES IN DIGITAL MEASUREMENT TECHNIQUES FOR FM BROADCAST

Written By Frédéric Allard - Vincent Philippe - Tony Peterle
Audemat - Bordeaux, France

ABSTRACT

Advances in Digital Signal Processor (DSP), Receive Signal Processor (RSP) and accurate high-speed analog-digital converter (ADC) technology have made possible new concepts and techniques in measuring and analyzing signals used in FM Broadcast facilities. We will explore the advantages and limitations of this technology as it applies to FM Broadcast measurement and analysis. We will begin by exploring the differences between the architecture of traditional analog measurement equipment and the new digital model, and expand on the new theories and mathematics employed in making accurate measurements within the more open structure of digital equipment. The digital designs allow us to approach theoretical ideals in terms of filtering and accuracy, and this presentation will attempt to explain in solid terms the relevance and expectations of this capability in real-world applications.

One of the novel concepts to be explored is undersampling. We will explore the Shannon theorem, and present practical and mathematical evidence that the use of undersampling for signal analysis can produce excellent, accurate, results. The presentation will then provide a brief explanation of digital measurement techniques as they apply to each part of a broadcast FM signal – audio, composite/MPX, and final complex RF emissions.

DEMODULATOR VS. RECEIVER

Let us first define the major difference between a receiver and a demodulator in terms of FM performance measurements and applications.

Receiver

In general terms, a receiver has a high dynamic range, and accurate performance can be obtained at fairly low RF levels. Selectivity is high, but performance and accuracy are somewhat limited. A receiver accepts the entire FM band at its RF input.

Demodulator

A demodulator has a low dynamic range (around 30dB), and a high RF level is needed to guarantee accurate performance (>-25dBm). It is a wide-band device with excellent accuracy. When used for signal analysis, only one RF carrier is allowed at the RF input.

BENEFITS OF COMPLETELY DIGITAL EQUIPMENT

In order to understand why using a demodulator for FM signal analysis is innovative, and how it improves performance and accuracy, we have to first examine the current and most commonly used equipment architecture, and how digital techniques can improve not only the performance, but also add capabilities in terms of the number of instantaneous measurements.

Analog receiver / measurement equipment

Up until today, FM demodulation and analysis has most commonly been done with analog components; the task is to demodulate a frequency somewhere between 87.9MHz to 108MHz, with channels spaced every 200kHz.

In the case of analog measurement products, the critical areas of the receiver are the RF mixer, RF filters and the demodulation component. We have to ensure certain things:

- That the tuner is able to tune every frequency in the FM band, without creating problems of tuning-frequency images
- That the tuner in the equipment is selective enough to extract only one channel (and not any portion of the adjacent channels)
- That the filter has a ±150kHz bandwidth around the selected frequency, with the flattest in-band frequency response possible

Long ago, the architecture for this type of receiver became standardized: analog mixers are used to blend the incoming frequencies against a tunable local oscillator, a technique known as heterodyning. The oscillator is tuned so that desired reception frequency is output on an Intermediate Frequency, or IF, of 10.7MHz. Ceramic filters are then used to separate the desired 300kHz-banwidth (or >400kHz-banwidth in the case of HD Radio tuners). But the 10.7MHz-filter is often a compromise between out-of-channel rejection, and flatness of the in-channel frequency response. It is very difficult to reduce in-channel ripple to less than 0.1dB when using analog components.

This inherent non-linearity of analog FM reception and demodulation inevitably adds a certain amount of noise and distortion to the output signal that is to be measured and analyzed. The tolerances of the analog components can also induce disparity in the performance from one measurement to the next, or the measurements made on different pieces of equipment. Regular calibration and adjustment of each stage of the equipment is also a typical need for these types of analog devices.

Today, there are many examples of measurement devices that digitize the signal for display on a graphic interface. In most cases, these devices digitize and convert the analog signal once it is completely demodulated (conversion of MPX signal or conversion of audio signals) and/or decoded. So, even if they are described as 'digital' equipment, they digitize the signal solely for the display of the levels. Measurements can be digitally displayed, but the underlying performance is limited by the analog demodulator, and analog stereo decoder.
As the analog to digital converter hardware can be very sensitive, most equipment has only one digitization chain: that single chain is used to convert only one signal at a time (or two, for audio). So to display deviation, pilot and RDS injection simultaneously, you would need 3 pieces of equipment stacked on your desk.

In the area between the older, completely analog devices and the newest, completely digital demodulators, we can find several other hybrid architectures, before we arrive at the best one in terms of performance and functions.

Narrow band receiver

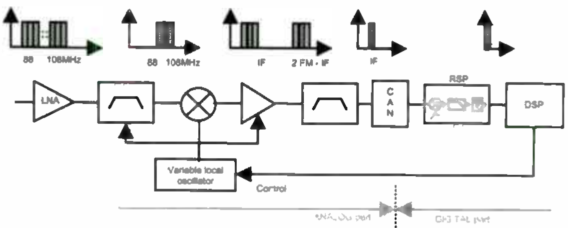


Figure 1 : Narrow Band Receiver

In this architecture, the first digital stage of the demodulation chain is a demodulation of the IF signal. In this example the analog tuner stage, from RF input to IF filter, is retained, along with the benefits and problems inherent to that design. Only one channel, translated from the desired frequency to the IF frequency (commonly 10.7MHz), is digitalized. RSPs and DSPs complete the FM-demodulation process and also stereo decoding. Measurements are extracted – except for the RF level which can not be measured because of Automatic Gain Control (AGC) used in the tuner stage. A digital filter can be incorporated into the RSP to increase out-of-band rejection.

This method allows simultaneous display of many different measurements. Digitalization in the IF stage eliminates analog FM demodulation non-linearity, noise and distortion, and demodulation does not depend on component tolerance.

The primary problem of this type of receiver is the analog RF stage which can add some noise and non-linearity. Some AM ripple will be created due to the IF-filter response, as its in-band frequency response can vary by several dB over the 400Khz bandwidth. Performance and accuracy will decrease, and will be less than optimal. As analog circuitry is still prevalent in the RF stage, it will require adjustment and regular calibration.

Wide band receiver

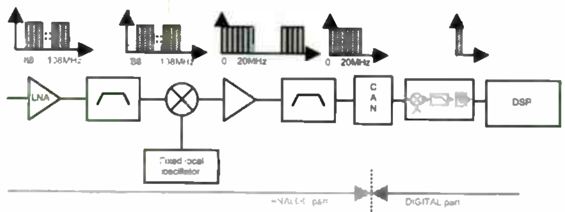


Figure 2 : Wide Band Receiver

With the new generation of accurate high-speed converters, we can now sample a signal over a large bandwidth with flat in-band response. This method is based on the sampling of the entire FM band, which is translated to 0-20.5MHz with a fixed frequency local oscillator (ex: 87.5MHz). An RSP, with its digital mixer, is used as the tuner and also as a digital filter to select the desired channel.

This method is more accurate than the previous one described. The use of an Intermediate Frequency and especially the IF filter are no longer needed, and the in-band filter response is flat (less than ±0.05 dB of ripple over ±150kHz), with high rejection (>90dB at ±200kHz).

In this method, almost every analog limit has disappeared: adjustments on RF stage are no longer needed, and FM demodulation noise, distortion and linearity are also perfect.

The weakness of this method is the need for a local oscillator, and the RF signal level which needs to be low enough to not saturate the analog to digital converter.

Innovative architecture of demodulator

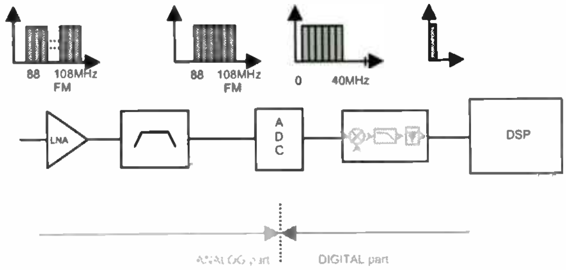


Figure 3 : Innovative architecture of demodulator

Some high speed AD converters are able to digitalize a bandwidth of up to 200MHz. The entire FM band can be converted without any mixing stage. Analog components are reduced to a single wide band RF filter (to reject signals outside the FM band), and overall hardware is simplified. The demodulator is reduced from 6 stages to 3, improving performance. The mixer (for tuning), channel filter, demodulation, and measurements: every thing is done in the digital realm.

To have the best accuracy, it is important to not saturate the A to D converter, and also to ensure that the level of your carrier is high enough. Calibration consists only of checking levels in the band and the exact frequency of the oscillator.

This architecture enables highly accurate RF, MPX, Pilot SCA, and Audio measurements to be displayed simultaneously and with perfect synchronization. Displays are not limited to a few bar graphs, or a single value on the front panel, but can now integrate multiple advanced RF measurements (Frequency, level, RF channel spectrum).

'IDEAL' VS. 'REAL' EQUIPMENT

Simulation of the perfect performance of a completely digital system is easy to do within the confines of advanced mathematic software. But some limitations unfortunately exist in the real world.

Hardware limitation

In an analog receiver, the tuning accuracy is defined by the local oscillator used in the analog mixer. A slight drift of the oscillator will induce a measurable mistuning. Even if the signal is mistuned, the FM demodulation will still occur, but a DC signal will appear on the MPX signal, with a level proportional to the mistuning of the oscillator.

In a completely digital system, the heart is also the oscillator. To avoid synchronization troubles, you have to monitor the performance of all of your clocks. A drift of the oscillators could have some other dramatic consequence (glitches in demodulated signals ...).

Thus the choice and performance of the oscillator has an important impact on some measurements (RF frequency ...).

Mathematics and DSP

Some mathematical operations are not native to DSP processors. Sine, cosine, tangent^{-1}, and square root calculations are approximations, with precision defined by the designer.

When implementing the algorithms, you have to have exact knowledge of what the required precision is; for example, you have to know this precision to ensure that your IIR[1] Filter will not diverge, because of rounding of calculations.

By using floating-point operations in a DSP, you avoid the scaling problems of a fixed-point DSP. Your filters are easy to import from your simulation, and the result is very close to that of your simulation.

RSPS AND DSPS IN BROADCAST MEASUREMENT

RSP – Receive Signal Processors

Commonly used in Wireless Infrastructure, this component includes a frequency translator (or digital mixer), and cascaded programmable digital filters.

Even if we could put frequency translator and digital filter into a DSP, an RSP is a device uniquely dedicated and better adapted to such computations, at a very high speed (80-100MHz). The use of a RSP in conjunction with a DSP, relieves the DSP of these basic mathematical manipulations.

The resulting performance, when these methods are used in digital measurement equipment is excellent: you can select any frequency you want in the FM band, within 0.1 Hz (theoretical value: it depends on your oscillator). The digital filters reject the out-of-band signal down to -90 dB, with less than 0.1 dB of in-band ripple...

DSP performances

DSPs are ubiquitous in signal processing today. Wider availability of high-speed floating-point DSP chips allows fast implementation of complex algorithms, and incredible calculation power. A 1 Giga 32 bit Floating point operation (1 GFLOPS) DSP is common today, and allows nearly unlimited possibilities in broadcast measurement.

[1] Infinite Impulse Response

DSP possibilities

Imagine you want to analyze an entire high performance FM signal, while performing 4 FFT computations (RF, MPX, Audio Left, Audio Right), to check noise.

You also would like to simultaneously analyze or check every MPX signal parameter (Pilot, RDS, SCA, Audio), apply some specific filters (de-emphasis, weighted, un-weighted) and detectors (Peak-Peak/2, RMS, QuasiPeak), and matrix the signals while avoiding objectionable high-frequency levels in your headphone by applying a de-emphasis filter only for headphone output..

While you are checking this, your boss would also like to be sure that RF level and RF frequency are OK, and check that RDS data is being correctly set and sent (TMC, ...), from the same equipment, without any interruption in your measurements.

That is what DSP can do - compute all of those measurements simultaneously, and make them all available. Then system just needs to be able do send the information over an IP network connection.

Some broadcast measurement equipment offers such simultaneous measurements, but you then need to have a powerful PC connected, and sometimes basic information such as RF level is not available.

DSP Flexibility

You want to deploy new technologies (FMeXtra for example), and measure the performance? Your existing analog equipment does not have the correct filters, or it suffers interference from this new modulation? You need to buy new dedicated equipment, or retrofit new filters and hardware.

With DSP technology and a thorough knowledge of algorithms, just create the new filters, adjust the others if needed, and upload them into the equipment: That's all! Technology upgrades have never been so easy.

DIGITAL MEASUREMENT TECHNIQUES

Now let us explore the mathematics that must be applied to this innovative architecture in order to generate useful data.

Filtering

Just like in analog filtering, we can define two kinds of digital filters: filters without feedback, so-called Finite Impulse Response filters (FIRs) and the Infinite Impulse Response filters (IIRs) whose final output depends to a degree on feedback from previous output of the filter

An Nth-order FIR digital filter can be performed in the time domain using the following equation:

$$y[n]=\sum_{i=0}^{N} x[n-i]b[i]$$

Equation 1: FIR filter computation

- x : digital input signal derived by sampling the analog signal
- y : digital output
- b : the filter's impulse response.

By definition, there are N+1 terms to compute and add. The main advantage of using this kind of filtering is stability but if the value of N is too large (i.e. if the order of the filter is too high), computation time becomes an issue, since it is a convolution computation and demands a lot of processor speed and power.

An Nth-order IIR digital filter can be defined by a recursive time-domain equation as follows:

$$y[n]=\sum_{i=0}^{N} x[n-i]b[i]-\sum_{i=1}^{N} y[n-i]a[i]$$

Equation 2: IIR filter computation

Though the IIR filter can be regarded as an infinite order FIR, it has the advantage of having a very short recursive equation that is much easier to carry out. The main drawback of this approach is its potential instability. To ensure the stability of the equation and the accuracy of the resultant filter, one has to make sure that all of the poles are located within the unit circle. In digital filtering, we define the poles as simply the roots of the polynomial defined by vector a.

If any one of the poles approaches the unit circle, the resultant filtering could become inaccurate.

These types of filters are very convenient when it comes to common applications such as low-pass, high-pass or band-pass filters. Very high order filtering can be done without placing too much demand on the processors, and accuracy can be maintained by ensuring that the poles remain well within the unit circle.

The filter order is set depending on the desired steepness (sharpness) of the filtering.

Example:

For instance, let's filter a signal sampled at frequency of 50kHz. The analog signal is the combination of 2 sine waves at frequencies of 2.5kHz and 10kHz.

We want to remove the 10 kHz component from that signal. A low-pass FIR filter of order 62 is chosen.

Here is the b vector graph of this filter:

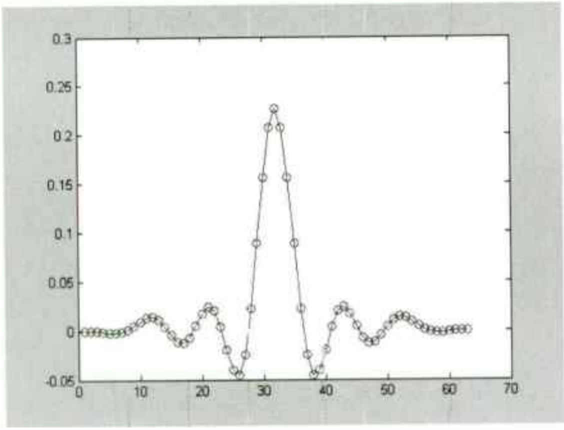


Figure 4 : Time-realm filter impulse response

Using a Discrete Fourier Transform of vector b one can easily compute its frequency response:

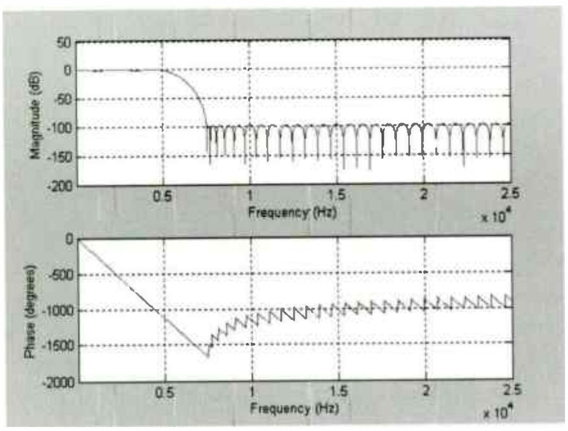


Figure 5 : FIR Filter frequency and phase response

It has a linear phase response in its bandwidth, which is an interesting characteristic of FIR filters.

The output (red) is derived by convoluting the filter function b[n] with the digitized input (blue). The result is a single sine wave at 2.5kHz, which lags the 2.5kHz component of the input signal slightly (0.62ms) due to the delay inherent to an FIR filter.

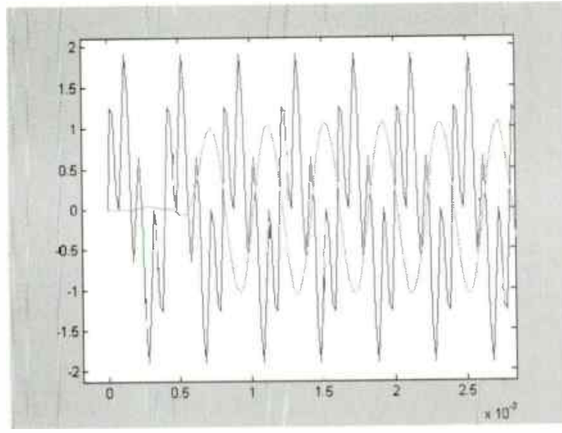


Figure 6 : Result of digital filtering

One can get a similar result using an IIR filter. Here are the filter's b and a vectors:

```
b = [   1.0679e-004
        -1.2500e-004
        2.7007e-004
        -1.3871e-004
        2.6516e-004
        -1.3871e-004
        2.7007e-004
        -1.2500e-004
        1.0679e-004      ]

a=[     1.0000e+000
        -6.6277e+000
        1.9942e+001
        -3.5470e+001
        4.0714e+001
        -3.0848e+001
        1.5060e+001
        -4.3312e+000
        5.6220e-001      ]
```

Now let's check the stability, by examining the positions of the poles in the unit circle.

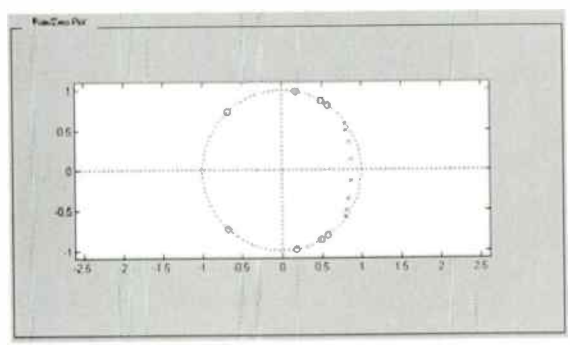


Figure 7 : Poles (x) and zeroes (o)

The filter in this example has 8 poles and zeroes which are inside the unity circle and not too close to it, so this filter is stable and the filtering will be accurate.

Z transforms, applied to the IIR's recursive equation, enable us to analyze its frequency response. A presentation of Z-transforms is beyond the scope of this paper, but the frequency response of the example filter can be graphed as below.

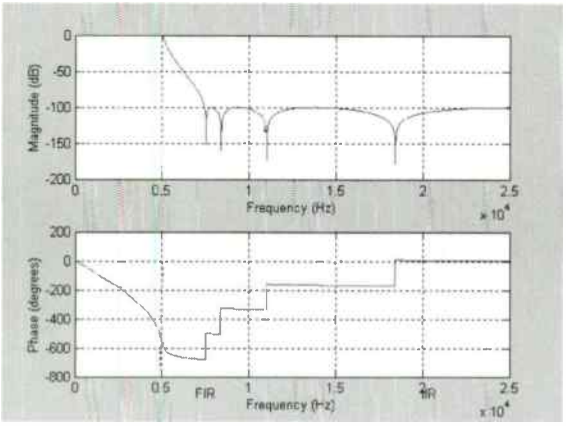


Figure 8 : Result of IIR digital filtering

This filter is as sharp as the previous FIR example, but is only an order of 8 as opposed to the FIR filter which had an order of 62.

Now let us compare the computation power required in both cases:

In this case an IIR solution is clearly more computationally efficient: only 17 multiply-add operations (MAC) are performed per sample to compute the output, whereas the FIR filter required over 60 such operations.

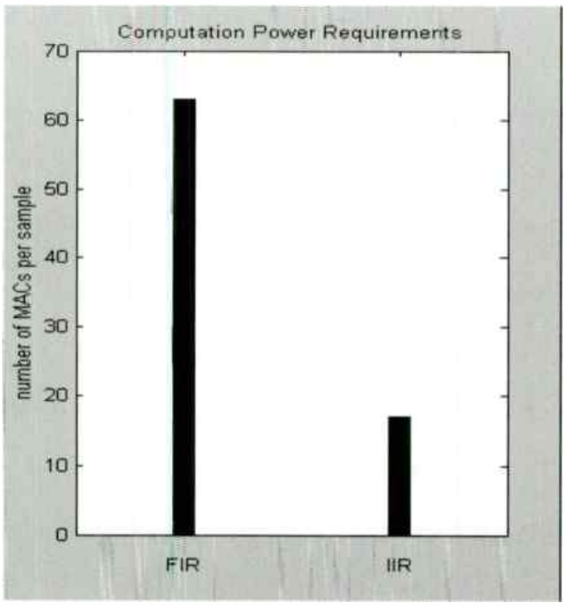


Figure 9 : Computation power requirements

Advantages of Digital filtering:

Using FIR digital filtering, it is possible to approach theoretical ideals, provided that one uses an FIR order great enough. But this approach also requires more computation power. It is also possible to achieve excellent performance using IIR filtering while requiring far fewer calculations, but the accuracy of the filter must be tested and confirmed.

With digital filtering it is possible to design accurate reproducible filters. As a result, 2 channels can be processed, (for instance left and right audio), in an absolutely identical manner, unlike the case with analog filters. The response of the filters does not change with variations of temperature or over time. And of course no calibration is ever required.

The FFT convolution approach to digital filtering:

As an alternative to traditional FIR filtering, it is possible to employ FFTs to carry out the filtering in the frequency domain. The filtering equation then becomes a simple product of FFTs in that domain.

Two FFT operations, an FFT and an inverse FFT, are necessary to enter and leave the frequency domain. This approach is more computationally efficient than FIR filtering if the order (N) is greater than approximately 60, depending on the hardware used. So this technique is strongly recommended in cases of high-order FIR filtering.

Measuring the Instantaneous Derivative

Thanks to digital mixing, complex samples can be derived from a real-time analog signal which allows the calculation of the exact current instantaneous phase. In FM modulation, this phase equals the antiderivative of the instantaneous frequency.

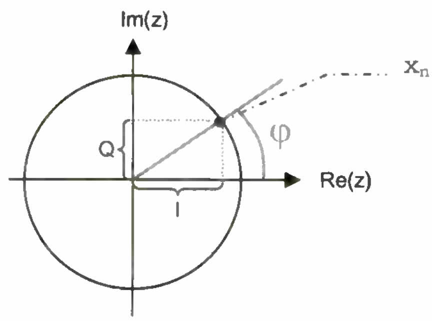
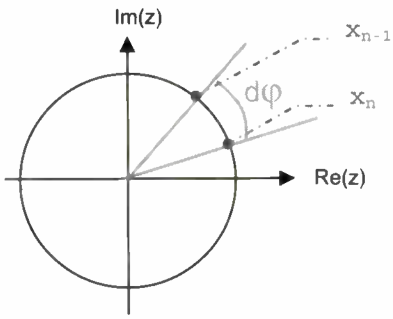


Figure 10 : Complex sample representation

By analyzing the variation of the phase between two samples, one can get an approximation of the instantaneous frequency value. If the sampling frequency is high enough, it is possible to derive the modulating signal.



$$U(n)=e^{j.\varphi n}.e^{-j.\varphi n-1} = e^{\underbrace{j.(\varphi n-\varphi n-1)}_{\text{Instantaneous phase derivative}}}$$

Figure 11 : Complex sample representation

Thus we can perform digital FM demodulation.

Undersampling and Measurement Accuracy: the Shannon theorem

According to the Shannon theorem, in order to retrieve all the information of an analog signal, it must be sampled at a frequency at least twice that of its maximum frequency component.

For instance, let us consider an analog audio signal with a bandwidth of 15kHz. The sampling frequency should be greater than 30kHz, otherwise spectrum aliasing will occur. However, if one considers a signal of limited bandwidth, using a sampling frequency greater than 2*f_max is inefficient. By using a lower sampling frequency, it is possible to recover the entire FM band at a lower frequency, thanks to the fact that a digital signal has an Fs periodical spectrum. By using a sampling frequency of 80MHz, it is possible to recover the FM band and translate it to the 8 to 28MHz range.

When considering using undersampling, one must respect the following rules:

The occupied bandwidth of a narrow band signal must be less than half the sampling frequency and the entire bandwidth to be retrieved needs to fall between any multiple of the sampling frequency and that multiple plus one half the sampling frequency so that the signal is not affected by the sampling operation.

This is a generally accepted extension of the Shannon theorem.

Given that the spectrum of a digital signal is Fs-periodical, the relevant spectrum can be recovered between 0Hz and half the sampling frequency.

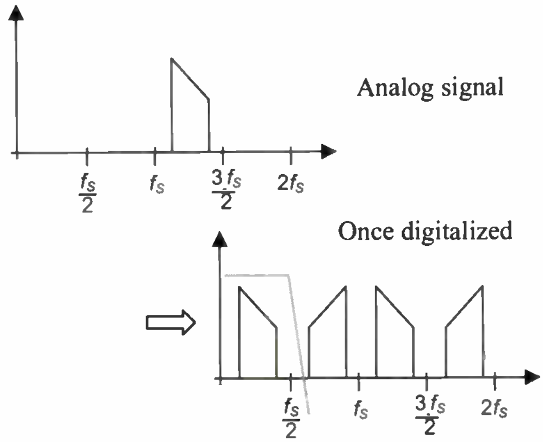


Figure 12 : Complex sample representation

Thus undersampling enables the processing of the signal at a lower frequency, without any additional demands or requiring any additional processing power.

In the case of the Navigator Modulation Analyzer, we chose an Analog to Digital Converter operating at 80MHz. Applying the technique of undersampling enables us to instantaneously recover the entire FM band and transpose it to 8 to 28MHz for analysis. In this particular case one can observe that this undersampling also acts as a frequency translator.

Using this method one can sample the entire FM band at once thanks to an 80MHz ADC provided that the incoming signal is limited to the FM band.

The technique of undersampling can introduce unwanted frequency components which may occur as a result of aliasing. It is a simple enough matter to eliminate those unwanted frequencies using an analog FM pass band filter on the input.

SPECIFIC BROADCAST MEASUREMENTS AND EXPECTATIONS

RF analysis

When checking the RF section of a transmitter, technicians must be able to test some critical parameters.

RF Frequency stability

When a transmitter emerges from the manufacturing process, technicians have to adjust the transmitter frequency. This adjustment requires a precise RF frequency meter to calibrate or validate an existing calibration. Required accuracy is within 100Hz.

After the transmitter is installed on site, we need to periodically check transmitter frequency. Even if a slight drift will not cause any perceptible problems with reception, it could be a sign of the aging of the transmitter. A large amount of drift is not acceptable, as it could disturb an adjacent channel, and must be avoided.

RF power meter

To maintain the coverage area of the transmitter, the RF power level has to be checked. Aging of some stages of the transmitter can introduce AM modulation, or cause a slow continuous decrease of transmitted RF power.

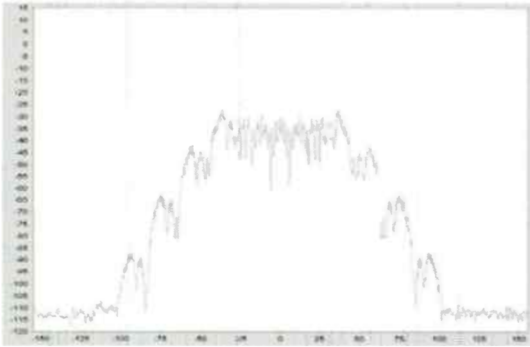


Figure 13 : RF spectrum

Measuring AM modulation on an FM signal requires a high speed measurement of the RF power, in order to detect the rapid fluctuations. Usually only specific equipment could be used – a traditional RF power meter which averages the signal will not be able to correctly measure the AM modulation.

Spectrum analyzer to check out-of-band noise, and total deviation

When a new transmitter is tested, technicians have to check that undesired noise is not being generated by the transmitter. Noise could consist of harmonics of the carrier, or other anomalies in the system. Out-of-band emissions are completely forbidden as it could disturb an adjacent channel, but it is also necessary to check for emissions outside the FM band.

A spectrum analyzer can also be used to precisely adjust your total composite deviation. The carrier cancellation method is the most precise way to adjust your modulation. Generate a pure sine wave at a specific frequency (13587Hz if you want to configure 100% total deviation), bypass your stereo-encoder, and adjust the level of the sine wave as necessary to make the carrier disappear or be reduced as low as possible. You are now sure that your transmitter is correctly configured.

Composite signal analysis

The composite signal is a multiplex of the combined audio channels, a pilot signal (used for stereo decoding), data stream (RDS), and other SCA.

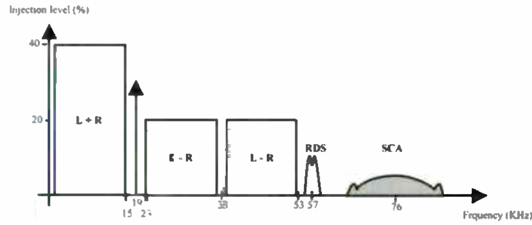


Figure 14 : Multiplex description

Modulation

As total deviation is limited by regulation authorities, broadcasters have to respect the 75-kHz total deviation limit of their Multiplex signal (MPX).

In order to keep audio as loud as possible within these limits, the RDS and other modulation need to be set correctly: high enough to correctly decode RDS, and achieve lock on the Pilot, but not too high as to reduce the available modulation for the audio.

Adjustments of all of these signals are generally simplified by using a graphical display. Since changing one level changes the total deviation, simultaneously displaying every level with bar graphs, allows quick and accurate sound processor settings.

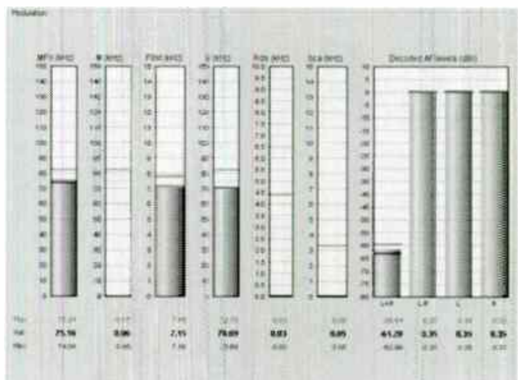


Figure 15 : Bar graph display

Spectrum

To analyze certain problems, such as an unstable pilot or noisy audio output, spectrum analysis can be the best tool to gather information. What could be easier to see than a moving noise floor, or a Pilot signal buried in audio?

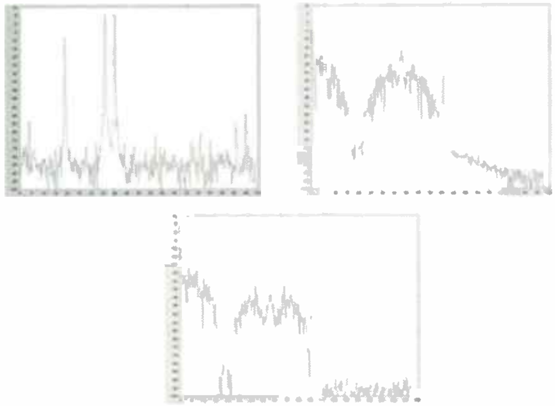


Figure 16 : MPX spectrums

Unstable levels or a bad noise floor can indicate a sound processor problem (second figure), such as bad audio filters or incorrect sound processor settings. High-frequencies in the audio signal can disturb the Pilot or RDS. You can easily check for this on the MPX spectrum display. Pilot and RDS have to be correctly isolated (correct in last figure).

RDS decoder & analyzer

Once the RDS injection level is correctly set, it is useful to check that the RDS data is being sent correctly.

Figure 17 : RDS decoder

You need to use a RDS decoder and RDS analyzer to display data being received. An efficient RDS analyzer should display the basic RDS information (PI code, PS, AF) and specific application data (TMC, paging).

Audio signal analysis

The main audio measurements of the transmission chain consist of checking the frequency response, distortion, and also separation. With new digital studio and transmitter equipment, these parameters have never been so good, and the performance of the old analog test equipment is sometimes worse than the broadcast chain being checked; the performance of the measurement equipment needs to be perfect. What could be better than measuring digital equipment with digital equipment?

Digital techniques can help us to have the best separation, distortion and accuracy in station performance and measurements.

Filter response

The (left + right) and (left – right) audio signals are extracted from the composite signal using FFT-convolution pass-band/low-pass filters as seen before.

This allows for steep, accurate and phase-limited filtering, with a perfect control of the cut-off frequency. As a result the audio channels are not altered by this retrieval process, and the ensuing audio measurements will only measure the transmitter's defects.

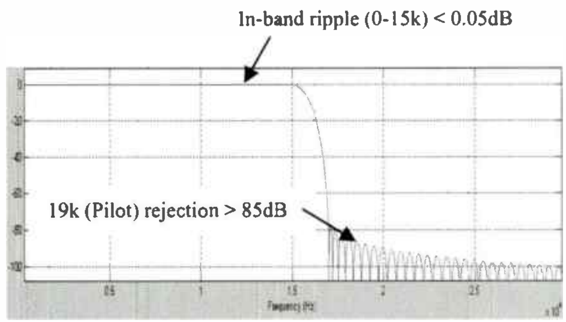


Figure 18 : Audio digital filter

Distortion

The aim of the distortion (THD) measurement is to check for non linearities in the chain: saturation, clipping, compression…

There are two kinds of THD measurements: THD and THD+N on both audio channels.

Basically a THD measurement consists in comparing the level of the fundamental signal to the level of the remaining harmonic frequency components.

$$THD = \sqrt{\frac{H_2^2 + H_3^2 + \ldots + H_N^2}{H_1^2}}$$

THD+N compares the fundamental to the entire spectrum minus the fundamental, thus it also takes into account noise.

$$THD + n = \sqrt{\frac{H_2^2 + H_3^2 + \ldots + H_N^2 + n^2}{H_1^2}}$$

A THD compares the fundamental to all the other harmonics of superior order.

An easy way to assess distortion is to use an FFT computation, however it lacks accuracy, especially in the case of low-frequency fundamentals if no under sampling is carried out.

An accurate distortion measurement can be performed in two steps:

First, one computes the frequency of the fundamental signal, using for instance an array of Discrete Fourier Transforms (DFT) around the fundamentals' assessed frequency. Using a dichotomic algorithm one can fine-tune this frequency measurement.

Second, one computes the harmonics' levels using DFTs at frequencies that are multiples of the fundamentals'.

A THD+N measurement also requires measurement of the power in the audio channel; notch-filtered to remove the fundamental. A notch filter can be a mere 2nd-order IIR digital filter.

Thanks to this all-digital approach, we only measure the distortion caused by the transmitter. And it is measured very accurately.

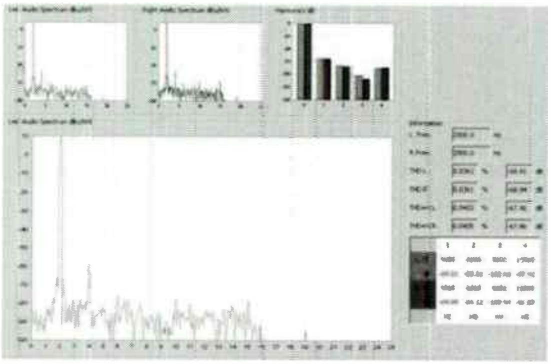


Figure 19 : Distortion display

Audio separation

The last critical point is separation. When transmitting right and left channels, one expects than right channel will arrive in right loudspeaker without being disturbed by left channel.

In transmission chain, cross-talk can be created by bad synchronization of Pilot and 38kHz (center frequency of L-R channel), or poor filters. If this phase relationship is bad, the receiver can not correctly frequency-translate the (Left – Right) signal, and then recombination of Multiplex signals will not correctly produce the discrete Left and Right audio channels.

The performance of the demodulator depends on two major parameters: flatness of the audio filters and accuracy of Pilot recovery. As described before, the flatness of digital filters is more than sufficient to ensure no noise is added. Digital Pilot recovery is based on a Digital Phase Lock Loop (DPLL). The 38kHz center of the (Left – Right) signal is easily computed with an arithmetic formula, and the frequency-translation of (Left - Right) becomes easy.

Once again an all-digital technique proves to be the best approach: it does not add any significant cross-talk, so what is measured is only the performance of transmitter; and it is measured precisely.

CONCLUSION

In conclusion, we can safely say that Moore's law, about the capabilities of integrated circuits improving at an exponential rate, has as much impact on the equipment that we use for broadcast measurements and analysis as it does for cellular phones and laptop computers.

The capabilities of these new circuits, along with innovative and advanced application of advanced mathematics, give us the ability to observe the signal components of an FM broadcast chain in new and better ways, and with higher and higher levels of accuracy. Where this will lead us in the future we can only imagine, but it will certainly advance the quality and reliability of the connection between the broadcaster and the public, and that will undeniably be good for all of us.

Communicating with Management

Monday, April 14, 2008
10:30 AM – 12:00 PM

Moderator: Chriss Scherer
Radio magazine, Overland Park, KS

*John Bisset Presentation

John Bisset, Broadcast Electronics, Manchester, NH

*Engineering Communicating with Management Panel

Don Kelley, Greater Media, Inc., Boston, MA
Gary Kline, Cumulus Media, Atlanta, GA
Paul Tinkle, Thunderbolt Broadcasting Company, Martin, TN
David Isreal, WFYV-FM/WMXQ - FM, Jacksonville, FL

*Paper not available at the time of publication

DTV Reception Issues

Monday, April 14, 2008
10:30 AM – 12:00 PM

Chairperson: Al Grossniklaus
WTHR NBC 13, Indianapolis, IN

***DTV Reception in an Urban Environment**
William Meintel, Meintel, Sgrignoli & Wallace, LLC, Warrenton, VA

New Neighbors: Can Wireless Microphones and Consumer Devices Coexist In the White Spaces?
Christopher Lyons, Shure Incorporated, Niles, IL

***Measurement Results of Consumer Indoor Antennas**
Dennis Wallace, Meintel, Sgrignoli & Wallace, Waldorf, MD

*Paper not available at the time of publication

NEW NEIGHBORS: CAN WIRELESS MICROPHONES AND CONSUMER DEVICES COEXIST IN THE WHITE SPACES?

Christopher Lyons
Shure Incorporated
Niles, Illinois USA

INTRODUCTION

For decades, the broadcast industry has worked relentlessly to improve the television viewing experience for the American public. Advances in video technology like in-car cameras, computer animation, and high-definition video have been matched by upgrades in audio technology like wireless microphones and multi-channel surround sound. Together, improvements in audio and video tools and techniques have raised what was once derisively termed "lights and wires in a box" to the point that it is worthy of the name "home theater."

But the spectrum that has been home to television stations and wireless audio devices for decades may soon be invaded by a horde of new consumer wireless devices.

Recent developments at the FCC are threatening the viewing public's ability to receive over-the-air television broadcasts and the ability of content producers to use the wireless production tools on which they depend. If these critical tools can no longer operate with the reliability that producers have depended on for years, their ability to deliver the exciting and compelling programming that viewers (and advertisers) demand will be hampered.

BACKGROUND

For decades, several types of wireless production tools have been permitted to operate in the VHF and UHF television bands under a classification known as "Low Power Auxiliary Stations". These include:

- Wireless microphones used by reporters, news anchors, talk show hosts, and actors
- Wireless intercom systems used by production technicians and crew
- Wireless personal monitor systems used by musical performers
- Wireless interuptible foldback (or "IFB") systems used to cue talent
- Wireless video assist devices used to transmit the image from a film camera's viewfinder to a remote video monitor

The FCC grants these devices secondary status, which means that they may not cause interference to broadcast stations, but may receive interference from them. Low Power Auxiliary Stations (detailed in Title 47, Part 74, Subpart H of the Code of Federal Regulations) face strict limitations on power output and other operating parameters. They may not operate on channels that are occupied by a broadcast station within a 70 mile radius. Because the TV channels that are unoccupied in any given location vary, wireless microphone users must investigate the active TV channels in their location and choose operating frequencies accordingly.

Desired-to-Undesired Ratio Is Key

Under current FCC regulations, wireless audio equipment is not permitted to use spread spectrum or frequency-hopping transmission such as is used in mobile phones, Bluetooth, Wi-Fi, and other devices. This means that it is critical that the ratio of the Desired Signal to the Undesired Signal is quite robust at the receiver. Improving this ratio can be achieved in two ways: increase the level of the Desired Signal, or reduce the level of the Undesired Signal. There are some ways of increasing the amount of Desired Signal reaching the receiver, such as using directional receiving antennas and locating the receiving antennas closer to the transmitter. Most professional wireless microphone users (including content producers) have already taken advantage of these techniques to increase the reliability of their systems to meet the requirements of televised events. Increasing the power output of the transmitter is not too feasible, because it reduces battery life and reduces the number of systems that can be operated in close proximity to each other. So, that makes it important for the strength of Undesired Signals at the receiver to be minimized – and that requires that there are no transmitters operating on the same frequency nearby.

HOW THE SPECTRUM IS CHANGING

The transition from analog to digital television broadcasting in the U.S. has a significant technical side benefit: TV stations can occupy adjacent channels without interfering with each other. This means that the TV stations can be accommodated in a smaller piece of spectrum, which allows the remaining "recovered" spectrum to be allocated to other uses.

The plan for spectrum reorganization involves both licensed and unlicensed use, and consists of multiple elements occurring on different timetables and involving different industries. This has created widespread confusion and misinformation among both incumbent users of the spectrum as well as the TV-watching public. To clarify, there are four major elements of the reorganization of the TV broadcast spectrum in progress.

DTV

Over-the-air digital television stations have been assigned to VHF channels 2 to 6 (54 to 88 MHz) and 7 to 13 (174 to 216 MHz) and UHF channels 14 to 51 (470 to 698 MHz). This spectrum has been termed the "Core TV Band." In February 2009, all full power analog TV stations will cease operation, leaving only LPTV, Class A, and translator stations operating in analog. Full power stations must switch to DTV transmission. All TV stations will all operate in channels 2 to 51.

Public Safety

Two sections of UHF spectrum have been reallocated for Public Safety radio communications. These include TV channels 63 and 64 (764 to 776 MHz) and channels 68 and 69 (794 to 806 MHz). This spectrum will be used in addition to the spectrum in TV channels 14 to 20 that is used for communications by public safety and emergency response agencies in 13 U.S. cities.

Advanced Wireless Services

Television channels 52 to 62 (698 to 764 MHz) and 65 to 67 (776 to 794 MHz) have been reallocated for use by Advanced Wireless Services. This spectrum has been divided into blocks which are being auctioned to licensees who wish to offer new services. Most of the licenses cover only a small geographic area; a few cover a large portion of the continental United States. To provide a nationwide service, an operator would need to acquire either six large area licenses or 176 small area licenses, depending on the particular block of spectrum in question. Some of these services are already up and running. Qualcomm's MediaFlo service that delivers television clips to mobile devices on the Sprint and Verizon networks uses TV channel 55, and is currently available in 40 U.S. cities.

White Spaces

The FCC also plans to use the unoccupied TV channels in the Core TV Band, which have become known as "White Spaces." These channels would be used by unlicensed consumer wireless equipment to access the internet and provide other services. While the channels that are unoccupied in any particular location vary, the combined amount of spectrum available is significant.

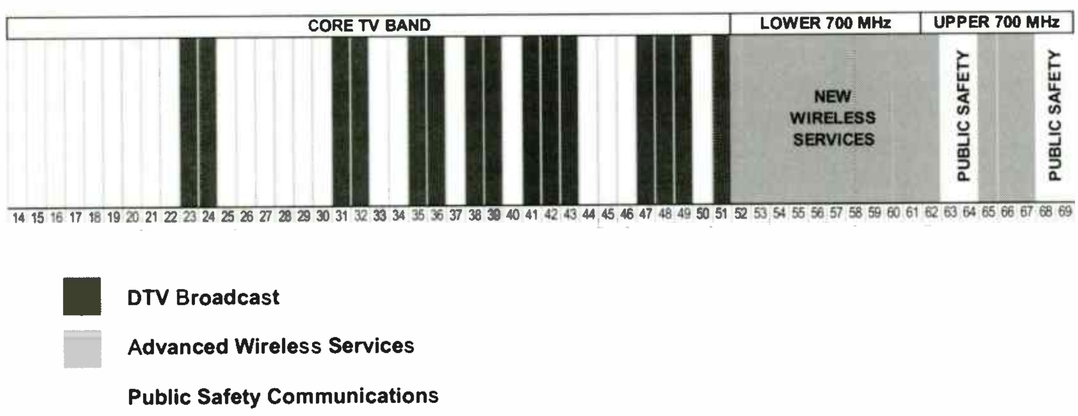


Illustration 1: TV channel usage in Hollywood, California after the DTV transition in February 2009

WHITE SPACE DEVICES

Two distinct types of devices would be permitted to use the White Spaces. So-called *Fixed/Access* devices include stationary transmitters and rooftop receiving antennas that would provide wireless broadband internet access as a "third pipeline" to compete with cable TV or landline telephone service providers. *Personal/Portable* devices could include mobile phones, personal digital assistants (PDA's), portable music/video players, and other portable consumer products that might transmit or receive content through a wireless internet connection. They might also operate on a peer-to-peer basis; e.g. to provide wireless interconnection services between multimedia equipment.

The proposal to allow portable consumer wireless devices to operate in the White Spaces has caused concern among a variety of incumbent users of the Core TV Band. Television stations, broadcast networks, and even DTV receiver manufacturers are concerned that the White Space Devices would interfere with DTV signal reception. Cable television service providers are concerned that the devices would interfere with set-top cable boxes. Content producers who use wireless audio devices – ranging from TV networks to sports leagues to equipment rental agencies to individual audio engineers – are worried that their wireless microphones and other equipment will be interfered with during a production or broadcast. With the often huge costs of producing films and television programming, any audio signal dropout at all is considered unacceptable.

Is Concern Justified?

In a word, yes. In 2006, Shure Incorporated conducted a test to determine how a typical wireless microphone might be affected by a White Spaces Device. To approximate the performance of a White Space Device, Shure was granted an experimental license by the FCC to modify a commonly available wireless router to operate within the TV band. A Shure UHF-R series wireless microphone was turned on indoors and a walk-around test conducted to establish a baseline operating range. The system performed reliably at more than 100 feet and around corners as would be expected in similar interior office environments.

When the modified wireless router was turned on, the walk-around test was repeated. This time, the wireless microphone signal began to drop out when the transmitter was only about ten feet from the receiver. Beyond thirty feet, the signal was badly broken up to the point that it would be completely unacceptable for a typical wireless microphone user, whether on a TV production or in a business meeting.

PROPOSED INTERFERENCE PREVENTION MEASURES

A number of techniques have been proposed to prevent White Space Devices from interfering with DTV reception and wireless microphone signals.

Reserved Channels

One approach to preventing interference is to provide a few TV channels in each market that would be restricted from consumer wireless device operation. The broadcast industry has strongly recommended that the channels immediately adjacent to each occupied DTV channel be designated as off-limits to consumer devices, in order to protect DTV sets from interference. These adjacent channels would also provide a safe operating environment for wireless audio equipment. In a major metropolitan area, the adjacent channels would likely allow enough wireless audio equipment to meet the needs of a small or even medium-sized production. Large productions that require many wireless devices – such as sports events, award shows, and music or theatrical productions – would still require additional spectrum to meet their wireless needs however.

Listen Before Talk

Another approach to interference mitigation is to employ "Listen Before Talk" technology. With this scheme, the device would scan the core TV band before transmitting, detect and identify DTV broadcasts and wireless microphone signals, and avoid operating on those frequencies. While this is simple in theory, the differences in technical and operating characteristics between DTV broadcasts and wireless microphone signals could make it quite difficult.

Television broadcast signals are like skyscrapers: they're big, they exist all the time, and they stay in one place. Their power output ranges from tens of thousands to hundreds of thousands of watts, and they occupy a full six megahertz chunk of spectrum – an RF 'city block'. In most metropolitan areas of the U.S., TV stations broadcast twenty-four hours a day, so the signal that was there yesterday is there today and will be there tomorrow. Their transmitting antennas are also fixed – located high up in the air, on towers carefully (and expensively) placed on top of the tallest building or mountain in the area, so as to deliver a predictable and consistent signal strength throughout the entire viewing area. Altogether, this makes it easy to detect and identify a DTV broadcast signal.

Wireless microphones are more like fireflies. Their power output is typically from ten to fifty milliwatts – less than one twentieth of a watt. (FCC regulations permit power output as high as 250 milliwatts for UHF wireless microphones, but such high power is seldom used in practice because it results in decreased battery life and allows fewer systems to operate in one location.) Limitations in antenna efficiency and factors like body absorption of the signal can reduce the effective output to as little as one milliwatt. In addition, wireless microphone signals are restricted to a maximum bandwidth of 200 kHz (.2 MHz).

Compared to TV stations, wireless microphones operate intermittently. An entire day of production might only involve a few hours of actual transmitting time, and many users turn the transmitter off whenever possible to extend battery life. A wireless microphone signal can be there for a minute (or an hour, or a day) and gone the next. This means that the consumer device would need to re-scan the spectrum frequently to see if a new wireless microphone had begun transmitting since the last scan was made.

Finally, wireless microphones are often used indoors. The signal from the transmitter antenna emanates outward in all directions and is reflected off of metallic objects in the room, causing the received signal strength to vary wildly from one point to another as the performer moves. In some cases the receiving antenna may see virtually no signal at all, while a full-strength signal could be had just inches away. Wireless microphone systems address this through the use of diversity antennas – two antennas spaced apart, so that when the signal briefly vanishes at one antenna, it can still be received at the other antenna. If a consumer device was equipped with only one antenna, it is possible that the wireless microphone's signal might not be detected from some points.

	Power Output	Bandwidth	Transmitting Characteristics
DTV Broadcast	10 – 100 kW or more	6 MHz	Always on; fixed location
Wireless microphone	10 – 50 mW	200 kHz	Short-term or intermittent; varying locations

Table 1: Comparison of DTV and wireless microphone operating characteristics.

Detecting Wireless Microphones Will Be More Difficult Than DTV Broadcasts

While technology will no doubt be developed that can reliably detect a DTV broadcast signal, the challenge of detecting wireless microphone signals is made even more complicated by the differences in transmitting distance between wireless microphones and the proposed White Space Devices. The power output proposed by the FCC for the new Personal/Portable devices is 100 milliwatts conducted or 400 milliwatts EIRP (although some proponents have suggested even higher levels). As stated earlier, wireless microphones typically have a power output in the 10 to 50 milliwatt range, although actual radiated power can be far less. This means that a consumer device operating in the white spaces could potentially interfere with every wireless microphone that is within a radius of more than 1,000 feet. But, the signals from those microphones could be far below the RF noise level by the time they reach the device.

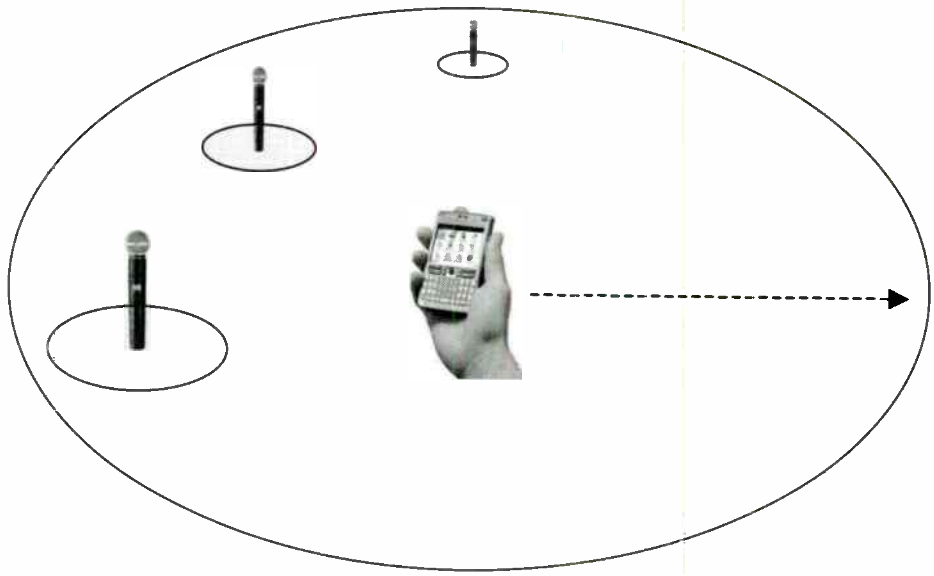


Illustration 2: Relative transmitting distances for wireless microphones and proposed White Space Devices

It is no surprise that the signal from a wireless microphone transmitter loses strength as the distance from the transmitter increases. To make matters worse, indoors – where most production and performance takes place – the signal strength is further degraded by reflected signals that arrive out of phase with each other. When these 'multipath' effects are taken into account, the actual signal level at any given point in the room could easily be 40 dB lower than calculated simply by distance.

Within the transmitting radius of the consumer device, there could be multiple wireless audio devices. At a major sports event or production location, there could be dozens or even hundreds of wireless microphones, intercoms, personal monitors, and IFB systems in operation. The consumer device will need to detect and avoid all of them.

TESTING

The FCC tested two prototype White Space Devices (one supplied by Microsoft and one supplied by Philips) in 2007. The results of that testing indicated that neither device satisfactorily detected and avoided DTV broadcasts and wireless microphone signals. The results of this test were later deemed unreliable, however, because it was discovered that one of the prototype devices was not functioning properly.

As of this writing, a second round of testing is expected to commence in February 2008. Both laboratory and field tests will be conducted, and unlike the first round, the tests will supposedly be open to interested parties.

So far, it has not been demonstrated that new devices can occupy this band without causing harmful interference to incumbent users of the band, including television broadcasters and users of wireless microphones and other wireless audio equipment. Interested parties of all types have been working diligently to arm the FCC's Office of Engineering and Technology with all relevant information. Manufacturers of White Space Devices as well as wireless audio equipment have provided sample units and performance data for their products. At the same time, representatives from major content producers have provided insight into their production needs, technical standards, and operating techniques.

CONCLUSION

The viewing public, whether consuming news, sports, music, theater, or feature films, has made it clear that they prefer – and in fact demand – programs that dazzle them in every way. Content producers have invested heavily in technology to meet this demand, with the result that content is richer, more informative, more exciting, and more available than ever before. Wireless audio equipment has been and will continue to be a key enabler that gives content producers vastly greater creative freedom.

For this reason, the introduction of new types of devices into the Core TV Band must be managed carefully.

The benefits of this new use of the spectrum must not come at the expense of existing users. Hopefully, with collaboration between content producers, equipment manufacturers, and the FCC, a new wireless landscape that accommodates both new and existing users can be created.

Alternative STL Technologies

Monday, April 14, 2008
1:00 PM – 2:30 PM

Chairperson: Paul Shulins
Greater Media, Boston, MA

The HD Radio STL: Issues, Options and Technologies

Bob Band, Harris Corporation, Mason, OH,

Best of Synchronous with the Best of IP?

Guy Gampell, APT, Belfast, UK

Robust HD Radio™ Exporter to Exgine Architecture

Tim Anderson, Harris Corporation, Mason, OH

The HD Radio STL: Issues, Options, and Technologies

Bob Band
Harris Corporation
Mason, Ohio

INTRODUCTION

This paper addresses the challenges involved in getting the HD radio signal from the studio to the transmitter. Unlike a traditional FM STL, HD brings up a new set of questions, such as: Should the Importer and Exporter be at the studio or the transmitter site? What are the bandwidth requirements for each of these options? How does adding HD2 and HD3 affect the requirements?

What are the pluses and minuses involved in using 950 MHz, T1, IP, or license-exempt radios for the link? What else can be done on the STL - additional audio programs, return audio, off premise extensions, LAN bridging, control data?

We'll look at these issues, and more, in the quest to find the ideal STL for your particular situation.

THE EVOLUTION OF THE STL

Not so many years ago, an STL meant one of two things: an analog RF link at 950 MHz, or an analog conditioned line leased from the telco. While many analog 950s are still in use, conditioned leased lines have almost completely disappeared as the telco world has moved to digital networks, and the old-school technicians needed to maintain them have retired and been replaced by IT techs more familiar with protocols than impedances.

Well before the advent of HD Radio, STLs began moving from analog to digital, taking advantage of the higher fidelity and greater flexibility the digital world offers. First to move were land lines, as the conditioned analog circuits were replaced by T1 digital links. As far back as the late 1980s digital pioneers were installing Intraplex T1 multiplexers, capable of combining multiple services such as STL and TSL audio, telephone circuits, and control data over a single digital service.

A decade later, 950 MHz radios also moved into the digital realm, beginning with the Harris CD Link in 1997, followed by several others including Moseley and TFT.

The move from analog to digital opened up a host of new possibilities. Anything that could be converted into ones and zeroes could be carried across these new digital media, subject to the total capacity of the link.

For example: T1 operates at a fixed rate of 1.544 Mbps. Converting a 15 kHz FM stereo broadcast signal to digital using linear PCM encoding gives us a data rate of 1.024 Mbps.

While this fits comfortably within the confines of a T1, the desire to fit more channels onto one link led to the development of a number of audio compression algorithms, including apt, Dolby, the various flavors of MPEG (up to and including AAC), and ITU-T standards such as J.41 and G.722. While each of these has its pluses and minuses, taken together they enable us to carry anywhere from two to a dozen or more stereo channels across a single T1 circuit.

In the 950 realm, digital carrying capacity is limited by a combination of factors, including how much RF bandwidth is available (typically 300 kHz, but sometimes more or less), and what kind of modulation you can perform on the signal, which is itself influenced by the distance, terrain, and other elements affecting RF transmission. In an ideal case, a well-designed, robust system can manage to get up to about 2 Mbps of throughput across a 950 MHz link.

One big difference between the two technologies is that 950 MHz radios operate one-way only, while T1 is inherently duplex, with the same amount of bandwidth on the TSL path as on the STL. Thus, the return path can be used for monitor audio and for backhauling downlinked satellite programs, as well as for two-way traffic such as intercoms, off-premise telephone extensions, and LAN interconnectivity.

Today, T1 has begun to give ground to IP, as digital networks move from TDM (time division multiplexed) links with fixed bandwidth to the greater efficiency of packet-based services. But packet networks, which were not originally designed for transport of real-time services such as broadcast radio, present their own challenges for use as STLs.

In the wireless world, as 950 MHz licensed frequencies have become ever more difficult to obtain, particularly in more crowded urban areas, license-exempt radios, particularly in the 2.4 GHz and 5.8 GHz bands, have begun popping up with ever-greater frequency. Here the leased-line and wireless formats begin to merge, as these new types of radios typically have telco-standard interfaces such as T1, E1, and Ethernet ports, allowing the same equipment that is used on digital land lines to be used in wireless installations.

THE CHALLENGE OF HD

The HD radio system is designed to carry multiple programs and services, including:

- Main Program Service (MPS), which by current FCC regulations must be the same in content and quality as the FM analog broadcast signal.
- Main Program Service Data (MPSD), information about the artist, song title, etc that is displayed on the HD radio receiver.
- Supplemental Program Services (SPS), the secondary audio programs. Currently the HD Radio protocol supports up to two SPS programs, SPS1 and SPS2.
- Supplemental Program Service Data (SPSD) for each SPS, similar to the MPSD.
- Advanced Application Services (AAS), data services that may include such things as stock market, traffic or weather updates, e-commerce, and many other possibilities.

Physically, from the broadcasters' perspective the HD Radio system consists of three main components:

- The Importer, which encodes and formats secondary audio programs and their associated data, as well as any AAS services being offered.
- The Exporter, which takes in and encodes the main audio program and its associated data, as well as the output of the Importer, to create the multiplexed HD signal.
- The Exgine Exciter, which receives the formatted output of the Exporter and creates the HD Radio signal.

In HD terminology, the data connection between the Importer and Exporter is referred to as the I2E signal, and that between the Exporter and Exciter is the E2X signal. The I2E signal is a TCP data stream, while the E2X signal can be either UDP or TCP (Figure 1).

The STL bandwidth requirement of the HD Radio signal depends on two basic factors: where the physical components of the system are located, and which Service Mode is in use.

Scenario 1: For those who do not plan to add any secondary programs to their HD broadcast, no Importer is required and it can make more sense to place the Exporter at the transmitter site, thus eliminating the need to carry either the I2E or the E2X signal across the STL at all. In this configuration, we need only transport the main program audio in AES/EBU format at a 32 kHz or 44.1 kHz sample rate, plus the associated MPSD data channel as a low-bit-rate (< 400 bps) UDP channel with no duplex connectivity required (Figure 2).

Scenario 2: Placing the Importer at the Studio and the Exporter at the transmitter site means we need to add I2E transport to the above. The I2E is a TCP data signal that can run as high as 156 kbps – not a huge amount of additional data, but being TCP it requires duplex connectivity (Figure 3).

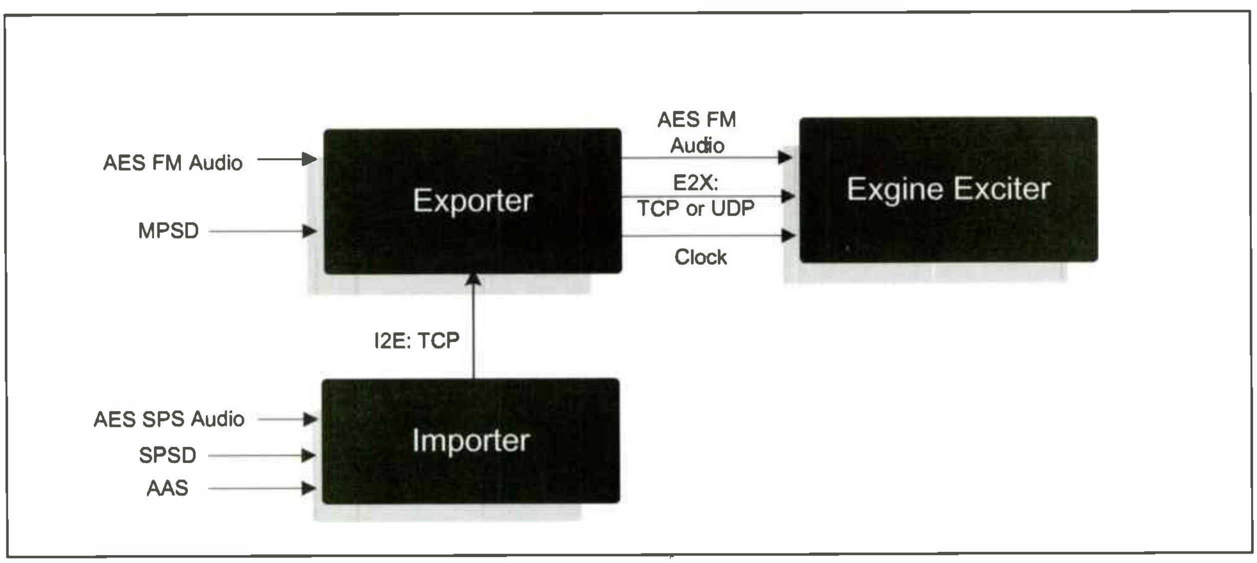


Figure 1: HD Radio Components and Signals

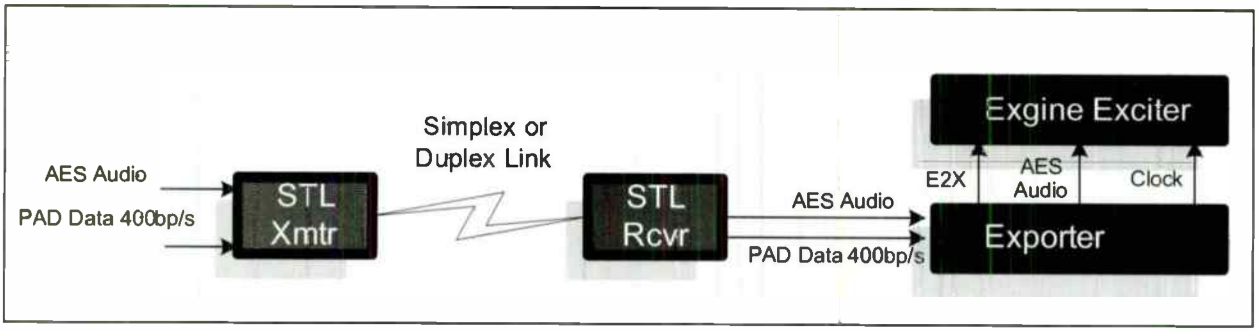


Figure 2: No Importer, Exporter at Transmitter (Scenario 1)

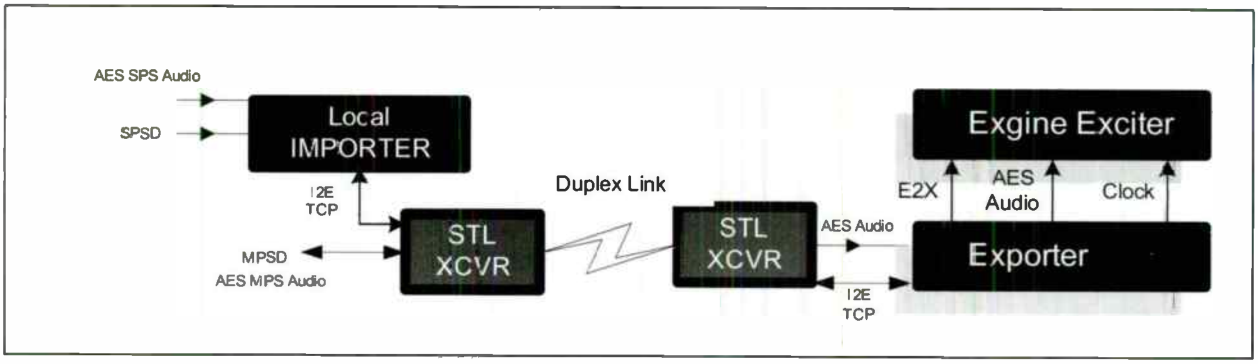


Figure 3: Importer at Studio, Exporter at Transmitter (Scenario 2)

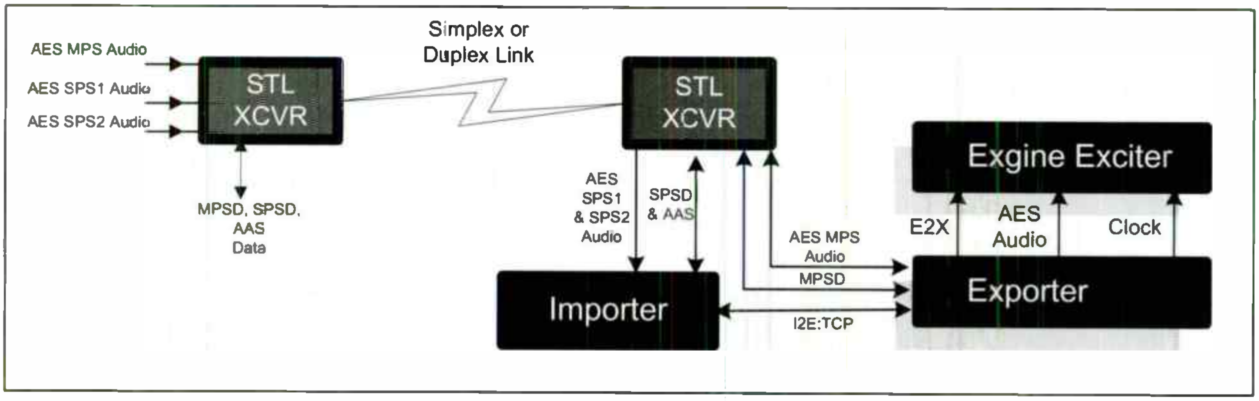


Figure 4: Importer and Exporter at Transmitter (Scenario 3)

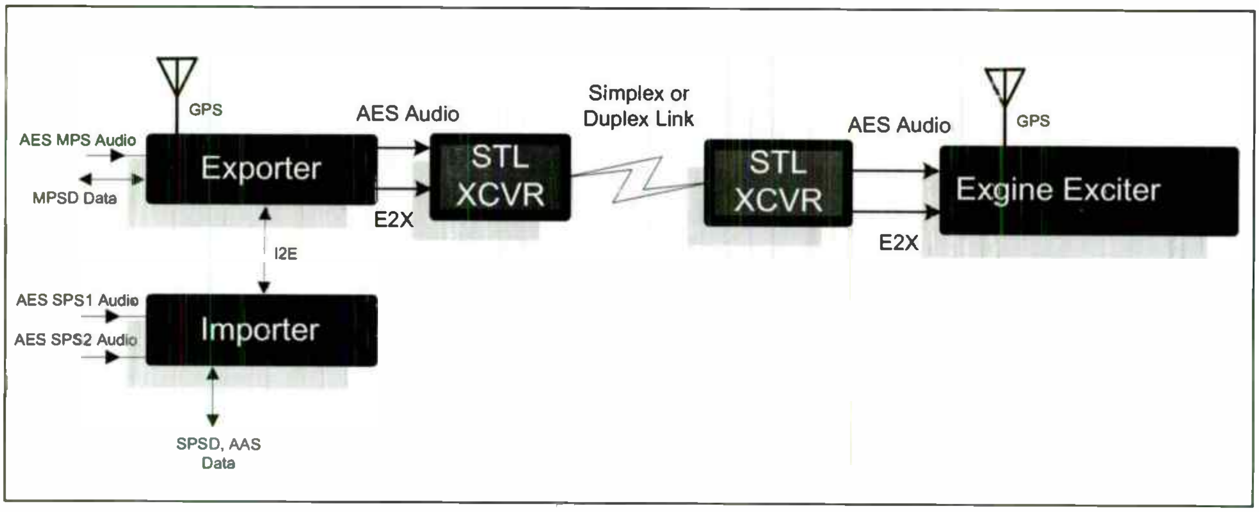


Figure 5: Importer and Exporter at Studio (Scenario 4)

Courtesy of iBiquity Digital

Scenario 3: Positioning both the Importer and Exporter at the transmitter site places heavy demands on the STL (unless, as previously mentioned, only the Main Program is being broadcast, in which case the Importer is not required) as we would then need to transport all of the audio programs across it, plus all their associated data channels, plus any additional data channels we're using for AAS. Of the available STL options, only wideband IP has the capacity to support this without using audio compression (Figure 4).

Scenario 4: In most cases, the most effective and bandwidth-efficient solution is to place both the Importer and the Exporter at the studio and feed the E2X over the STL to the transmitter site, and in fact we find that iBiquity Digital recommends this configuration for most HD installations (Figure 5). Yet it is precisely this transport requirement that presents the greatest challenge in the design of the STL. Current implementations of the E2X use UDP, a one-way protocol which does not allow for the resending of lost or errored packets, and iBiquity Digital reports from their lab testing experience that IP packet loss must be less than 10^{-5} for successful performance.

The latest version of the iBiquity Digital standard for HD Radio allows the use of TCP, which increases robustness by allowing missing packets to be retransmitted. With TCP in place, the signal remains usable under severely impaired conditions with bit error rates up to 0.03%. However, TCP requires two-way communications to operate, and thus presents a problem for traditional 950 MHz STLs. T1 and IP links, whether as land lines or over microwave radios, are inherently duplex and thus well suited to handling TCP traffic.

An additional consideration when transporting the E2X signal is the fact that both the Exporter and Exciter must be kept in close synchronization in order to prevent overflows and underflows in the Exgine buffer, which can cause audio dropouts. The simplest method of achieving this is to use GPS clocks at the studio and transmitter which provide 10 MHz reference clock outputs.

The other major factor in determining the bandwidth required on the STL is the Service Mode. In the hybrid broadcast of an FM analog signal along with HD Radio, the HD Radio signal travels in the side lobes of the analog FM within the limits of the FCC mask. The Service Mode determines how many bits are used and where in the side lobes they are placed. That subject is worthy of a paper all to itself, but suffice it to say that there are currently three modes available, referred to by the acronyms MP1, MP2, and MP3. Within each Service Mode, the available bandwidth can be broken up in different ways depending on the type and number of services being offered (Figure 6).

MP1 mode provides 96 kbps of bandwidth that can be divided into multiple channels. The minimum requirement for the Main Program Service is 32 kbps, which leaves 64 kbps for the other program and data services. The minimum bandwidth for an SPS audio service is 10 kbps; however, any audio service running at less than 32kb is going to be mono and anything less than 24kb is not likely to sound terribly good, though there is a lot of work being done with very low-bit-rate codecs for speech.

Typically, stations use 48 kbps for the main program service, and the remaining 48 kbps for SPS services. This whole 48 kbps can be dedicated to just SPS1, or it can be further subdivided to a 24 kbps SPS1 and a 24 kbps SPS2. Alternatively, the original 96 kbps in MP1 mode could be divided up 32/32/32 or 64/32 or 64/16/16, or even further subdivided if AAS data services are in use. We expect that next year SPS3 will be introduced, giving the broadcaster even more options from which to choose.

MP1 requires ~232 kbps of TCP E2X transport bandwidth or 160 kbps UDP no matter how it is

Data Rates and Provisioning Required for Modes and Services					
HD Protocol	**Direction**	**IP Protocol**	**Service Mode**	**Avg. Bandwidth Kb/s**	**Provision Kb/s**
I2E	Duplex	UDP	MP1 SPS 12Kb	13.0	17.3
			MP1 SPS 32Kb	34.9	46.6
			MP1 SPS 48Kb	43.5	58.0
			MP2 SPS1 12Kb	21.5	28.7
			MP2 SPS1 32Kb SPS2 12Kb	57.0	76.0
			MP2 SPS1 48Kb SPS2 12Kb	65.2	87.0
			MP3 SPS1 24Kb	36.5	48.6
			MP3 SPS1 32Kb SPS2 24Kb	69.4	92.5
			MP3 SPS1 48Kb SPS2 24Kb	77.7	103.6
		TCP	MP1 SPS 12 Kb	16.3	27.2
			MP1 SPS 32 Kb	37.6	62.7
			MP1 SPS 48Kb	53.8	89.6
			MP2 SPS1 12Kb	29.8	49.7
			MP2 SPS1 32Kb SPS2 12Kb	65.2	108.7
			MP2 SPS1 48Kb SPS2 12Kb	80.6	134.2
			MP3 SPS1 24Kb	42.3	70.4
			MP3 SPS1 32Kb SPS2 24Kb	78.2	130.3
			MP3 SPS1 48Kb SPS2 24Kb	93.3	155.5
E2X	Simplex	UDP	MP1	119.7	159.5
			MP2	132.1	176.1
			MP3	149.3	199.0
	Duplex	TCP	MP1	139.3	232.0
			MP2	155.6	259.2
			MP3	167.8	279.5

Figure 6: Service Modes

(courtesy of iBiquity Digital)

divided up.

MP2 is an extended mode which provides another 12 kbps for services. Unfortunately this additional bandwidth cannot be combined with the 96 kbps available in MP1, so it's not too useful except for data services. The E2X transport bandwidth requirement for MP2 is ~260 kbps TCP or 176 kbps UDP.

MP3 is a further expanded mode which provides 24 kbps of bandwidth on top of the MP1 allotment. As with MP2, this extra 24 kbps can't be combined with the original 96 kbps but it can be used as a single extra 24 kbps channel, and thus could support a mono audio program. Stations might use this mode to run two audio programs at 48 kbps and one at 24. MP3 requires ~280 kbps of TCP transport bandwidth for its E2X signal, or 200 kbps in UDP mode.

This is the maximum that we have to be concerned with today, but we expect that the Fall 2008 release of the HD Radio standard will bring MP11, which will provide yet another 24 kbps data block in addition to the 96 in MP1 and 24 in MP3. Estimated transport requirements for MP11 are ~330 kbps in TCP mode and 240 kbps UDP.

THE COMPLETE STL: FM AND HD COMBINED

T1

The T1 bandwidth of 1.544 Mbps is subdivided into 24 time slots (sometimes referred to as channels or DS0s) of 64 kbps each, plus 8 kbps of overhead. The digitized but uncompressed FM stereo signal occupies 17 time slots, leaving seven for other purposes.

Of the four scenarios described above, in Scenarios 1 and 2 (Main Program only with no Importer, or Importer at the studio and Exporter at the transmitter site) only a relatively low-speed digital signal (MPSD or I2E) needs to be carried alongside the FM signal, using at most three time slots.

Scenario 3, placing both the Importer and Exporter at the transmitter, requires us to carry multiple audio programs and multiple data signals. This is achievable by using compression on the audio signals. With a high-quality audio compression algorithm such as Enhanced-apt, all three audio programs (Main, SPS1, and SPS2) can easily co-exist on a T1 line with their associated data.

For Scenario 4 (Importer and Exporter at the studio), we need to carry the E2X signal beside the analog FM. As we see from the numbers in the previous section, depending on the Service Mode (MP1, MP2, MP3, or MP11) and transport format (UDP or TCP), the E2X signal can range from 160 kbps to just over 330 kbps, which translates in T1 terms to three to six time slots, meaning it can fit comfortably alongside the uncompressed FM signal, as long as the hardware in use allows that portion of the T1 to be configured as an Ethernet bridge.

However, as T1 systems are often used for additional purposes beyond the pure STL audio, in many cases it may be desirable to use compression to free up extra bandwidth, which can allow transport of multiple audio programs, provide telephone connectivity, and enable remote control over the same T1 link. Many broadcasters find it expedient to establish two distinct Ethernet bridges across the T1, one to carry the E2X signal and one to carry all other Ethernet traffic such as remote control data and LAN bridging, to ensure that nothing can interfere with the critical E2X packets.

All in all, T1 provides an excellent, versatile, and highly reliable medium for an STL.

950 MHz

As with T1, digitization of the 950 MHz STL brings the ability to multiplex different types of programs together, and the capacity of the digital 950 STL is thus in many ways comparable to that of a T1 circuit. The major difference is the lack of a return path, which means that the 950 is unable, for example, to support Scenario 2, which requires transport of the I2E signal in TCP format. It also makes the 950 unsuitable for STL use where intercoms, telephones, or other two-way services are desired.

Unlike T1 with its fixed 1.544 Mbps bandwidth, the total bandwidth available on the 950 is in part determined by geographic and environmental conditions; in some cases it may be capable of transporting more data than a T1; in other cases, less.

IP

Internet Protocol (IP) is a packet network protocol originally designed for data transport, and as such presents particular challenges in carrying real-time, mission-critical signals such as STLs. Some of these issues, such as network jitter, are eliminated by using dedicated point-to-point services such as microwave radios (covered in detail in the next section).

But when the IP transport is provided over a shared network from a service provider, reliability becomes a serious problem. In general, only networks that support the Multi-Protocol Label Switching (MPLS) protocol are suitable for use as STLs. MPLS allows network service providers to guarantee users a fixed amount of bandwidth at an assured quality of service level. Such services can come close to the reliability of T1 at a lower cost, and as such have become increasingly popular, where they are available. Techniques such as Forward Error Correction (FEC) can further enhance reliability, albeit at the expense of increased bandwidth.

On an MPLS network, you can get as much duplex bandwidth as you are willing to pay for, so it can easily support any of the component placement scenarios. But even here, the one time, up-front cost of audio compression may be preferable to the higher recurring costs of greater network bandwidth.

License-exempt radios

License-exempt radios (often referred to as spread spectrum radios, though this is only one of the technologies such radios employ) are available for several frequency bands, most commonly 900 MHz, 2.4 GHz, and 5.8 GHz. While all are freely available for use by the general public, the lower two, 900 MHz and 2.4 GHz, are part of the ISM (Industrial, Scientific, and Medical) bands, originally intended for use by devices that emit electromagnetic radiation for purposes other than telecommunications, while 5.8 GHz is a communications band used by, among other things, the emerging WiMax standards.

All other things being equal, the lower the frequency, the greater the distance over which these radios can operate; however, both the 900 MHz and 2.4 GHz bands are more likely to encounter interference from other devices, including cordless telephones and microwave ovens, so most broadcasters these days are opting for 5.8 GHz.

These radios offer standard network interfaces on the user side, most commonly Ethernet (IP), and in many cases T1 and its higher-bandwidth cousin, E1[1].

On a well-designed microwave link, using either the T1/E1 or IP ports provides the same capabilities and carries basically the same limitations as for land-line T1 and IP described above. On the IP side, MPLS is not required, as there is no network through which the IP packets need to get routed. However, although the E2X signal is a packet stream that can be routed through the Ethernet port directly, many broadcasters prefer to use the T1/E1 ports to carry their audio, with an Ethernet bridge within the T1/E1 for the E2X signal, to keep the actual broadcast signals separate from LAN and other traffic that they may be transporting using the radio's Ethernet port.

One thing that differentiates license-exempt radios is whether they use Frequency Division Duplexing (FDD) or Time Division Duplexing (TDD). FDD radios use a pair of frequencies with a guard band between them, one for transmit and one for receive, while TDD radios use the same frequency half-duplex at alternating 0.5 mS intervals. TDD radios thus make better use of the available bandwidth, and have more flexibility when seeking out a "clean" frequency on which to operate. FDD radios also tend to be frequency-specific in their hardware, while more modern TDD radios tend to be software-tunable.

Another characteristic to look for in these radios is whether the modulation scheme they use is Single Carrier (SC) or Orthogonal Frequency-Division Multiplexing (OFDM).

SC systems use a single carrier frequency to carry all the data sequentially, while OFDM uses multiple carriers to transport the data as interleaved orthogonal signals. OFDM modulation provides greatly enhanced resistance to multipath interference, and even allows the radios to operate under Near Line of Sight (NLOS) conditions, where more traditional radios require clear line of sight. OFDM also incorporates FEC for enhanced reliability.

The HD Radio broadcast signal itself makes use of OFDM to help ensure interference-free reception.

WHAT'S NEXT?

In the near term, a new protocol called HD Protocol or HDP will be introduced later this year. HDP will add FEC and other techniques to further enhance the reliability of E2X transport. This will also add to the bandwidth requirements of the STL link by an as-yet-undetermined amount.

Out there on the horizon in the HD Radio world is something called Host Audio Transport (HAT). In HD terminology, the Host is the legacy analog FM program on whose signal the HD Radio content piggybacks. HAT will allow the FM signal to be encoded and multiplexed together with all the rest of the HD programs into the E2X signal, so that only one IP stream needs to be sent across the STL.

---

[1] T1 was originally developed as a means of digitizing and transporting telephone conversations. Each telephone connection became one 64 kbps time slot, so a T1 enabled 24 conversations to be carried at 1.544 Mbps on a single set of wires. It is still the standard used by telcos in the US, Canada and a few other places. E1 is a similar protocol that allows the transport of 30 telephone connections at 2.048 Mbps, and is the standard throughout most of the world outside North America. When connecting to radios that do not interface to the public switched telephone network (PSTN), either standard may be used, and broadcasters often prefer E1 because they can get 1/3 more bandwidth at the same price as T1.

This is still in the development stage but will probably be a 256k kbps low latency channel that will require ~500 kbps TCP and 400 kbps UDP of additional transport bandwidth. Expect to see this sometime in 2009. It will, of course, not be a requirement, as many broadcasters will not be willing to give up the fidelity of uncompressed audio (or the ability to choose their preferred method of compression, if compression is to be used) for their primary program signal.

Altogether, these additions to the HD Radio standard mean that a couple of years from now, the total bandwidth for the HD Radio signal on an STL may well exceed 1 Mbps.

On the transport side, while T1 is an established standard not subject to any significant change, digital 950 MHz radios are only ten years old and have yet to see the kind of improvements (like OFDM) that have made digital license-exempt radios so robust in recent years. Future developments here may significantly improve their ability to reliably carry greater amounts of traffic.

As the radio STL bands grow more crowded, we expect to see more attention paid to recently-introduced radios in the 5.3 & 5.4 GHz bands. Although these license-exempt bands have strict power limitations that preclude their use for long-distance shots, they may prove quite useful in crowded but short-distance urban environments.

IN SUMMARY

There are a number of ways of establishing an STL that can support both legacy analog FM and HD Radio. While T1 provides the most reliable system to support all the possible equipment placement scenarios, a digital 950 MHz link can handle most of them as well, and the newer IP and license-exempt radios open up yet more options. Which system is right for your station depends on three things:

1. Technical considerations such as which HD services you choose to put on the air and where you decide to place the Importer and Exporter.

2. External constraints such as line of sight, availability of 950 MHz licenses and T1 access to the transmitter site.

3. Commercial factors such as the recurring costs of land lines vs. the up-front cost of wireless.

The choice is yours.

References:

Tim Anderson *et al, HD Radio™ Data Network Requirements*, iBiquity Digital Corporation White Paper

BEST OF SYNCHRONOUS AND BEST OF IP?

Guy Gampell

APT, Belfast, UK

INTRODUCTION

In the past broadcasters have utilized a number of different telecommunication mediums to move audio content between locations. Balanced analog copper circuits, microwave systems both licensed and unlicensed and synchronous based systems such as V.35 or T1 have been at the heart of broadcast networks for many years. Now, however we are at the beginning of a sea change in the way broadcasters move content for applications such as Studio Transmitter Links, Transmitter Studio Links, Inter-Studio Networking and Remotes. We are at the beginning of the IP age.

This paper seeks to clarify the pitfalls and challenges faced by those looking to create an IP Broadcast Chain. In doing so, I have divided these challenges and pitfalls into two areas, Network related parameters and Equipment Related Parameters. We will also examine the advantages to implementing or migrating to an IP transport mechanism. Finally we will look at all the issues covered and answer the question posed...Audio Over IP Can It Really Work?

NETWORK RELATED PARAMETERS

Network Selection

It is of course of paramount importance that the network be fit for purpose. It must be of sufficient quality to support the uninterrupted flow of packets from point A to point B and it must be configurable with regards to the parameters which I will describe in detail later in this paper. In short, Audio Over IP for broadcast quality applications cannot be achieved on the Open Internet, contentious ADSL links or contentious WANs or LANs. If you cannot control the IP Network, then you cannot control the quality of the audio emanating from that network.

Protocol Selection

There is a common mistaken assumption when broadcasters first broach the subject of audio over IP that TCP will be the most appropriate protocol. TCP is a connection-oriented protocol. This means that a connection is created between sender and receiver and, as packets flow over this connection in the order they were sent, every packet must be acknowledged as received by the receiver to the sender. For the purposes of real time audio delivery, the fundamental flaw with this protocol comes when a packet is dropped. Having sent a packet over TCP/IP, the sender will repeatedly ask the receiver for an acknowledgment that the packet was received and, if the packet is dropped, this acknowledgement can never come. This produces the ugly side effect of data peaks on the IP link as bandwidth is swallowed up in a fruitless request for a receipt acknowledgement. The repeated acknowledgement request packets and the compressed audio payload are now in competition for the available bandwidth and the audio will begin to glitch.

UDP as a connectionless protocol employs a "send and forget" strategy with no acknowledgement required making it much more suited to real time audio applications over reasonable data bandwidths. RTP is now also being deployed along with UDP to aid the time stamping and sequencing of packets. Strategies for dealing with packet loss are discussed in further detail later in this paper.

Specifying Network Bandwidth and Packet Size

The choices made with regard to audio settings will define the data bandwidth you need to transport encoded audio over the IP network. For synchronous connections this actually equates to the bandwidth required to transport the compressed audio.

However, in IP we must add an overhead that is required to packetize the audio data. Encapsulation into an IP packet adds header bytes containing information related to the routing of the packet along the IP Network. The packet information is examined by the routers and switches which constitute the routing intelligence on the network and, based on the packet header information, decisions are made on how, when and where to route a packet. Included in the IP Packet Header are:

- 8 Bits - Containing the Type of Service or Quality of Service
- 16 Bits – Containing the length of the packet

- 16 Bits – Containing an identification tag to assist in reconstruction of fragmented packets
- 8 Bits that contain the protocol (TCP, UDP, RTP)
- 32 Bits that contain the source IP Address
- 32 Bits that contain the destination IP Address

A packet header containing this data (and more) must be included in every packet that originates from an IP device on an IP network. There is a correlation then to be found between packet size and bandwidth requirement. The table below details the IP Data Rate for audio transportation using the Enhanced apt-X algorithm as an example.

Table 1: IP Data Rates using Enhanced apt-X

Audio Algorithm	Frequency Response	Audio Data Rate (kbps)	Packet Size Bytes	IP Packet Size Bytes	IP Pkts/Sec	Packetization Delay mS	IP Data Rate (kbps)
Eapt-X 16-Bit	3.5kHz Stereo	64	128	194	62.5	16	97
Eapt-X 16-Bit	7kHz Mono	64	256	322	31.25	32	80.5
		64	512	578	15.625	64	72.3
		64	1280	1346	6.25	160	67.3
Eapt-X 16-Bit	7kHz Stereo	128	128	194	125	8	194
Eapt-X 16-Bit	15kHz Mono	128	256	322	62.5	16	161
		128	512	578	31.25	32	144.5
		128	1280	1346	12.5	80	134.6
Eapt-X 16-Bit	15kHz Stereo	256	128	194	250	4	388
Eapt-X 24-Bit	20kHz Mono	256	256	322	125	8	322
		256	512	578	62.5	16	289
		256	1280	1346	25	40	269.2
Eapt-X 16-Bit	20kHz Stereo	384	128	194	375	2.7	582
		384	256	322	187.5	5.3	483
		384	512	578	93.75	10.7	433.5
		384	1280	1346	37.5	26.7	403.8
Eapt-X 24-Bit	20kHz Stereo	576	128	194	562.5	1.8	873
		576	256	322	281.25	3.6	724.5
		576	512	578	140.625	7.1	650.3
		576	1280	1346	56.25	17.8	605.7

It is clear that choosing a larger packet size will reduce the bandwidth as the header bytes will represent a smaller proportion of the total bytes i.e. more payload, less overhead. However, with larger packets there is a downside. Firstly, latency will increase as the packetization delay is greater. Secondly, if a large packet is dropped there will be a correspondingly larger amount of payload dropped - in the real world this equates to dropped audio.

Choosing the optimum packet size will always be a balance between bandwidth efficiency, network performance (in terms of dropped packets) and audio quality.

Packet Loss – Correction, Concealment, SLAs & Performance Monitoring, Synchronous Backup

All packet-based systems are susceptible to dropped packets, resulting in dropped audio. As discussed, we can negate the problem by choosing smaller packet sizes but in so doing we incur greater bandwidth requirements. The other options for dealing with packet loss are concealment, correction or temporarily abandoning the packetized network in favor of an automated backup to a synchronous network.

Correction

No forward error correction is currently used in the UDP packet. The disadvantage of this is that, as the integrity of the received data is not verified once received, packets are susceptible to undetected bit errors. On the flip side, bypassing the additional correction stage increases the delivery speed of packets. Forward Error Correction techniques have been implemented successfully in audio streaming applications, however the implications for real time audio delivery due to the processing and data overheads have yet to be reconciled. Other issues that limit the effectiveness of Forward Error Correction in audio broadcast applications over IP are the dynamic nature of the packet losses on a network. For example, will the Forward Error Correction implementation be able to handle bursty packet losses or a percentage of lost packets if the packets lost are non consecutive? Also, to maintain compatibility between IP codec vendors, the FEC Information should be sent on a different port to ensure that the decoding codec does not become confused if it cannot handle the sent implementation of an FEC scheme. Finally, even very powerful multi dimensional FEC processes only recover a proportion of the lost packets, usually around 50% meaning that severe overheads have been incurred without even resolving the issue to 100% satisfaction.

Concealment

With correction not a viable option for dealing with dropped packets the next viable option is to look at concealing those dropped packets.

There are a number of methods of concealment ranging from simple repetition of the last good packet received to silence/noise injection or interpolation and retransmission. All have an impact on the reproduced audio.

In listening tests injection of silence predictably produced unacceptable breaks in the audio that led to a degree of incoherence. The injection of white noise improved intelligibility of the reproduced audio but was again noticeable. The use of repetition of the last known good frame produced favorable results. The use of interpolation/pattern matching/waveform substitution to conceal the loss of packets is also possible and again produced favorable results. However, when complexity is measured against the benefit to the listener, these last two techniques are delivering a diminishing return. Also the results of all these techniques will be governed by subjective improvements in audio quality and by the amount of packets being concealed.

SLAs & Performance Monitoring

Although not a short term fix to packet loss the use of SLAs (Service Level Agreements) and Performance Monitoring can provide long term stability on an IP Broadcast link in guarding against packet loss. Any Telco or provider will issue a SLA against an IP link typically guaranteeing uptime in percentage terms. This percentage can be reconciled to lost packets and actual down time on the link. However, to truly measure that SLAs are being adhered to, the terminating codec on a link must run non-intrusive performance monitoring on that link.

Figure 1 Packet Loss Concealment

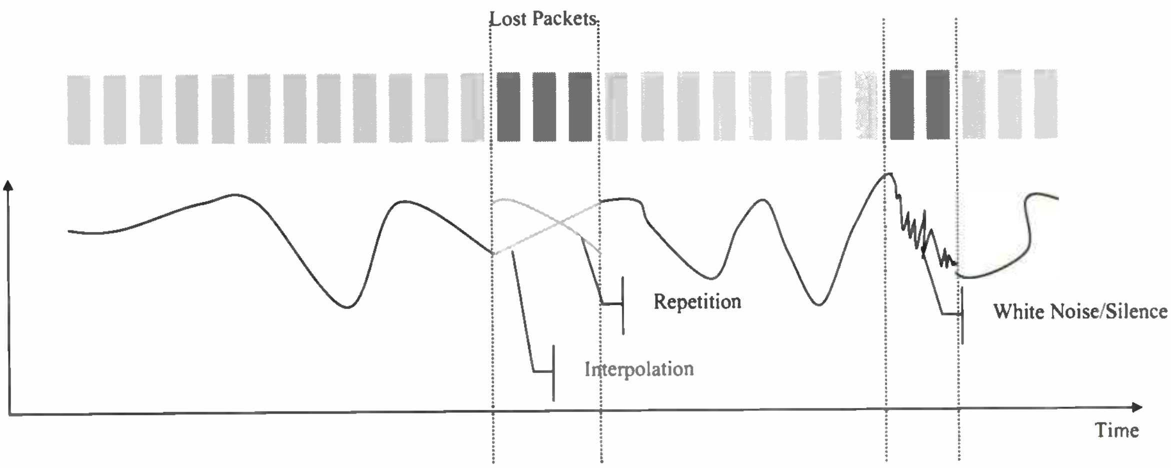


Figure 2 Perceived audio quality

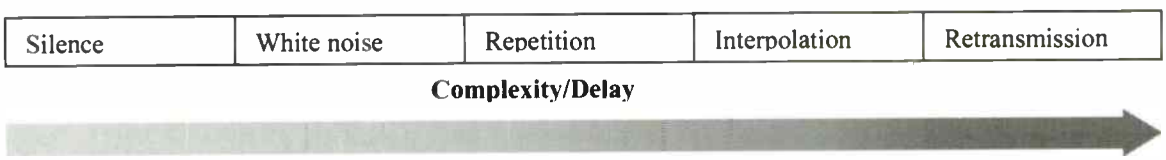
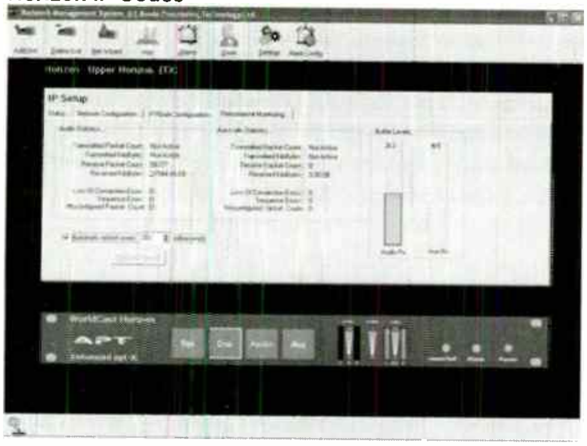
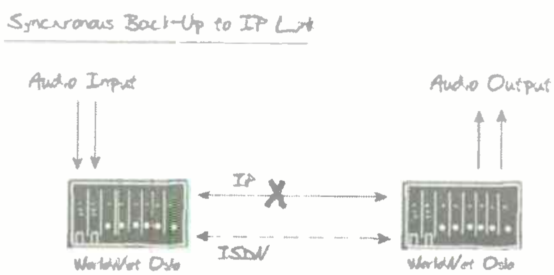


Figure3: Performance Monitoring on APT's WorldCast Horizon IP Codec



Synchronous Backup

Even with all necessary due diligence applied in the selection of the IP Network and Service Provider the network is likely to suffer a major outage at some stage. This can equate to consecutive dropped packets over a sustained period of time but is more likely to be large consecutive bursts of dropped packets. In this instance, the broadcaster is effectively off-air unless they have a backup. With a synchronous link over IP or T1 a point-to-point backup can be created. A many to one backup is also possible, however this backup will only be effective if the Primary IP links are diversely routed.



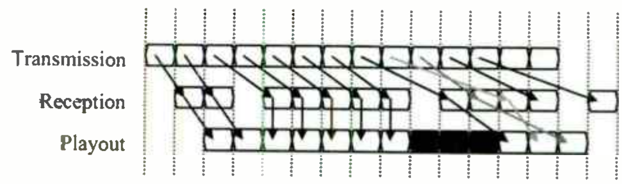
The backup from the primary IP link to the secondary synchronous link can be triggered by a number of different criteria on the APT WorldNet Oslo system. The trigger to switch to backup could be silence on the audio output of a specific audio module or a defined threshold in the Performance Monitoring log. Similarly, the automated restore back to the primary IP link could be defined in the Performance Monitoring log i.e. number of consecutive packets received without a single drop would equate to a restoration of the primary link.

EQUIPMENT RELATED PARAMETERS

Jitter

Jitter can be defined as packets that are received either side of the predicted arrival time. Given that any packet can take any route from transmit to receive, jitter is an inherent characteristic of packet switched networks. Packet jitter can be stipulated and defined under an SLA with the service provider, establishing an upper limit of the permitted jitter value. However it will always be present in an IP based network and so codec selection is vital in dealing with this. The buffer depth is critical in allowing the codec to provide enough time for the packet to be received and decoded before its play-out time. Reducing the buffer size reduces the jitter time mask available and increases the possibility of being forced to drop packets that have arrived beyond their play-out time.

Figure 3 Network Jitter Effects:



This diagram shows the effect of network jitter on the reception of audio and its subsequent play-out through an audio system. The buffer depth is usually set in milliseconds; in this case it is set to a two-packet buffer. For the purpose of this example this allows for up to a two-packet jitter delay in the system. Provided the network jitter is low the system is unaffected and plays out the packets received in sequence. However, should the jitter increase, there is the distinct possibility that the packets will arrive after the determined play-out time. In this example the packets have to be dropped, which results in the audio being corrupted.

Latency

All networks have transport latency due to the natural laws of physics. Transporting an electronic signal through whatever medium takes a finite amount of time that cannot be removed. In a switched network there is both the standard transmission delay and also the packetizing delay to contend with. This packetizing delay ensures that the delay on an IP network is always greater than on a synchronous network. Therefore special attention must be paid to latency if real-time audio delivery is to be achieved.

By definition a packet must be assembled and consists of a header plus payload. The size of that payload can be varied but ultimately it consists of an audio sample.

Take for example the Enhanced apt-X algorithm, a system that uses a four to one compression algorithm and has a packet size of 128 bytes. That's 512 audio samples, equal to 666 μsec in Mono and 333 μsec in stereo. Then take the time it takes to propagate through the UDP stack after being assembled into a Real-time Transport Protocol packet. In real, live-unit tests this is approximately 20-30 ms using the most optimal codec equipment available on the market. It is important to realize that this natural latency increases as the sample frequency decreases. With the inherent latency in the protocol stack, the need for efficient management of the UDP stack and the packetizing process by the hardware codec is crucial for the maintenance of low latency broadcast circuits. Also crucial to countering latency is the implementation of QOS within the packet header. QOS can be set through the IP communicating equipment, for example an IP codec, and determines that the packet is given priority over other packets on the network that do not have a QOS value. The additional delay of the audio coding is also critical in real time audio applications and is my final consideration in equipment selection.

Algorithm Selection

Having thoroughly investigated the intricacies of IP as a method of moving program content from Point A to Points B, C, D and through to Point X, the next step is to look at the best method of layering in audio on top of the transport stream. In essence there are two options – PCM / linear or using compression to reduce bit rates. Within compressed there are two sub-options - perceptual or ADPCM.

PCM or linear audio is well defined in terms of the audio - what you get in should be what you get out, assuming there are no problems relating to analog-to-digital conversions, signal-to-noise ratios or quantization issues. The compelling reason not to choose linear is directly related to the data bandwidths required.

A stereo signal sampled at 44.1 kHz, with a word depth of 16 bit, will require a data rate of 1.411Mbit/s (plus 10 – 15% overhead and additional for FEC and synchronization algorithms). This data rate bandwidth will cause stress on the IT network passing the data. If the broadcaster adds in additional channels (5.1 or more Stereo signals), deepens the word depth to 24 bit and increases the sampling frequency to 96kHz (or even 192kHz for the small furry animals that happen to be listening), it soon becomes apparent that what was a benign solution has now turned into a network nightmare.

Making the decision to use compression opens up an interesting argument. Two options are available:

The perceptual based algorithms using psycho-acoustic based principles that can generally be described as "Lossy". Some examples are MPEG Layer II, MPEG Layer III (MP3) and AAC (including the myriad of derivatives). These algorithms are heavily processor hungry and remove content that is perceived to be irrelevant. As such, they result in content that vaguely resembles the original (especially after several passes) and has a long latency i.e. 50+ milliseconds.

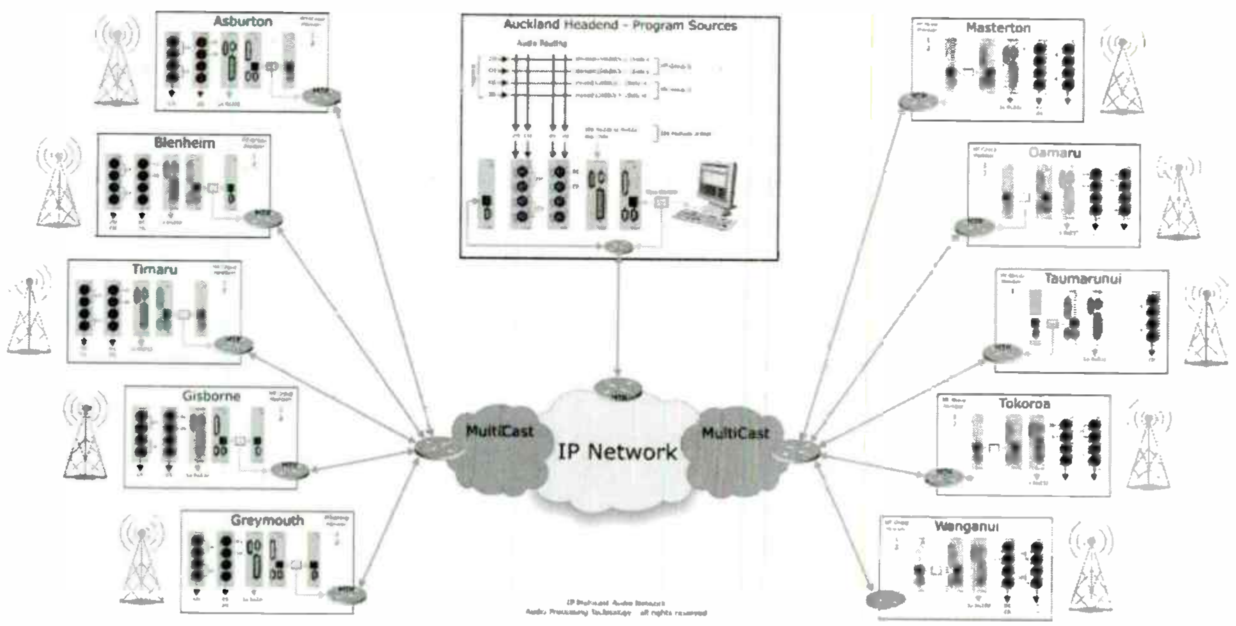
The other option is to use the relatively non-destructive Enhanced apt-X algorithm, which is based on ADPCM principles. This algorithm offers a low delay of less than 2 milliseconds and has exceptional acoustic properties. These acoustic claims have been confirmed by independent listening tests. For example, one listening test undertaken was with a group (approximately 20) of Chief Engineers from GCAP group. This was a double blind listening test with 10 audio samples (different genres (Classical, Pop), a cappella, spoken voice (Male / Female)). We tested Enhanced apt-X, MPEG and J.41. Enhanced apt-X was shown to be indistinguishable from the original PCM. The Enhanced apt-X algorithm can also offer word depths of 16, 20 & 24 bit, thus significantly improving the dynamic range to greater than 110dB.

Working on the assumption that the IP transport stream will naturally introduce a minimum delay of 20+ milliseconds, reducing the latency of the compression algorithm becomes an imperative when considering the design of a broadcast network. In essence, using a perceptual coder will render the solution unusable for any level of live event that requires off-air monitoring, whereas using Enhanced apt-X will offer broadcasters a viable alternative. Any overhead using apt-X relates to the RTP/UDP protocol used for the transmission of the compressed audio data and will be approximately 10-30% depending on packet size.

Along with the well-documented features of low latency and audio performance, Enhanced apt-X also has an embedded word pattern to aid connection and synchronization. AutoSync aids the ability to quickly synchronize i.e. 3 milliseconds on start up or in the event of a drop out. In addition the predictive nature of Enhanced apt-X allows for the masking of lost packets. As such, both features allied together act as a form of FEC.

On a more subjective issue, using multiple passes of a perceptual codec (for example, consider the final emission for HD Radio or DAB) will result in content heavy with artifacts. Ultimately these will cause "listener fatigue," swiftly followed by users tuning to another station that sounds better because it uses less destructive coding algorithms.

MAJOR BENEFIT OF IP



Many broadcasters emphasise the benefits of migration to IP backhaul as being the reduction of operating costs. But indeed there are significant other benefits too.

1. Convergence of distribution on to a single platform.
2. Control given to the broadcaster as opposed to the Telco
3. Flexibilty of a multi casting Network,

By way of illustration lets consider the above example, of a proposed multicasting implementation in New Zealand

In this example a head end is origination four different audio services, split in to two multi casts. Ten different regions then are on the same multicast network and can select either or both services simply to specifying to receive one multi cast or another. This concept is easy to extend to many more "groups" of services and many more destinations. Indeed the WorldNet Oslo illustrated here can establish 24 different routes from a single frame, and single frame, each route being able to be a multi cast or a unicast transmit, or the reception of a unicast or multi case stream . Since the routing is performed at a level of control GUI for the codec rack, the broadcaster has complete "authority" of what goes where as opposed to relying on a telcos ability to route and distribution T1 timeslots on a point to point basis.

CONCLUSION

It is patently clear that today's broadcasters are convinced not only that Audio Over IP can work but are determined to ensure that it does. The benefits of packetized networks over existing synchronous networks for audio transport have proved too persuasive for them to ignore:

- ❑ Synchronous networks are not easily scalable in comparison to IP
- ❑ Synchronous networks can be inefficient i.e. only supplied as T1 (1.5Mbits) or E1 (2Mbits)
- ❑ Synchronous links tend to be more expensive
- ❑ Not easy to achieve flexible multipoint configurations over Synchronous
- ❑ Widespread viability of IP Networks
- ❑ Consolidation of Engineering skills

However, despite these overwhelming arguments in favour of IP audio transport, broadcasters must also ensure that the quality and reliability of their existing networks are not sacrificed in the quest for greater efficiency and cost-savings. Therefore, the following checkboxes must be ticked

- ❑ An IP Network which is fit for purpose running a Protocol suitable for real-time delivery of Audio
- ❑ A robust hardware codec solution with redundant options and powerful performance monitoring to help enforce SLAs
- ❑ A hardware codec solution that allows parameters to be adjusted to take account of network conditions, parameters such as Jitter, Packet Size and Latency
- ❑ A low delay algorithm such an Enhanced apt-X to counter the inherent delay associated with IP Networks.

ROBUST HD RADIO™ EXPORTER TO EXGINE ARCHITECTURE

Timothy B. Anderson

Harris Corporation
Mason, Ohio

INTRODUCTION

The HD Radio system presents opportunities as well as significant challenges from the perspectives of both traditional broadcast systems engineering and IT infrastructure engineering. The various signals involved in the creation and distribution of HD Radio are digitally based, with a mixture of AES audio, IP based streaming, timing and control. The ability to support main program and multicast audio services with minimal delay over a moderately low bandwidth unidirectional link makes UDP well suited to the real-time streaming nature of the Exporter to Exciter data transported over STL and satellite. However, the lack of native Forward Error Correction and the Engine's source synchronous dependence, create a virtually insurmountable challenge. The result is a less-than robust transport resulting in dropped audio, unstable diversity delay and link failures. These issues have often been a source of frustration for listeners and station engineers alike.

This paper explores the realities and challenges of the Generation 3 HD Radio system architecture, network communications and synchronization requirements. It discusses the latest system enhancements that have been incorporated as well as improvements that are in development to make the E2X transport more robust and reliable.

HD RADIO BROADCAST SYSTEM ARCHITECTURE

System Overview

Figure 1 shows the physical architecture of the Generation 3 HD Radio system. The two primary components required to provide the HD Radio Main Program Service (MPS) and Programs Service Data (PSD) are the Exporter and the Exgine/Exciter. Additionally, an Importer may be implemented to support Supplemental Program Services (SPS), SPS PSD and other Advanced Application Services (AAS). The Importer provides compression and multiplexing of these services and sends them to the Exporter via TCP/IP for inclusion into the HD radio stream.

The Exporter accepts the Main Program Audio and Program Service Data (PSD) as well as the Advanced Applications Services and SPS streams from the Importer, compresses the MPS audio stream, and combines all of these services into a single data stream for transport to the Exgine/Exciter via the Exciter Link interface. The Engine subsystem of the Exciter accepts the Exciter Link data from the Exporter and performs the Orthogonal Frequency Division Multiplexing (OFDM) modulation for the digital portion of the HD Radio waveform. In FM+HD modes, the waveform is then scaled and summed with the output of the FM modulator, and passed to the Exciter's digital up-converter to create the Hybrid FM signal. In non-hybrid modes, the digital and analog waveforms are passed to separate up-converters and output for external summation.

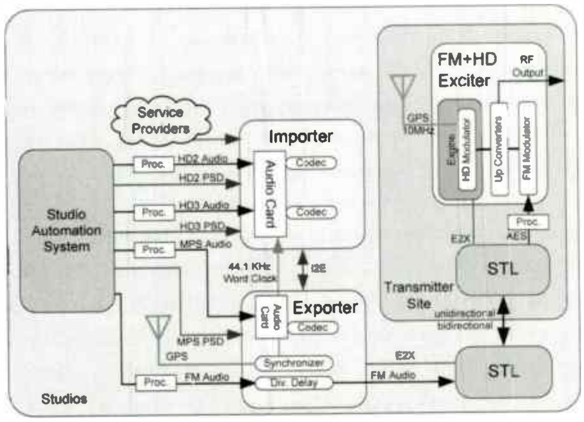


Figure 1 - Generation 3 HD Radio Architecture

E2X Message Structure

The Data Link stream between the Exporter and Exgine is the "Exporter to Exgine" or E2X protocol. The E2X protocol is defined as part of the standard HD Radio System Broadcast Architecture and follows the OSI networking model. The E2X protocol works within the application, presentation and data link layers. Internet Protocol (IP) is utilized at the network layer with Ethernet operating at the physical and data link layers.

There is no request/response exchange between the Exporter and Exgine in the E2X protocol. It instead implements source synchronous clocking. Messages are sent from the Exporter to the Exgine at a defined rate established by the Exporter's reference clock. The Engine/Exciter receiving the data must be synchronized with the Exporter.

The Layer 1 (Physical Layer, contains the modulation, framing, and signal processing) messages are structured into three distinct Program Data Units (PDUs) based on the rate at which they are delivered. A complete HD Radio Layer 1 frame is sent every 1.486 seconds and is divided into 8 block pairs each containing 2 blocks for a total of 16 Blocks per Frame.

Frame Rate PDU is sent every 1.486 seconds and contains the entire P1 message, a P3 and a P4 message (if present), a PIDS and a Clock message.

Block Pair Rate PDU period is 185.75 milliseconds and contain a P3 and a P4 logical channel message (if present), a PIDS and a Clock message.

Block Rate PDU period is 92.875 milliseconds and contain a PIDS and a Clock message.

P1 Logical channel messages are sent in the E2X stream at the Frame rate of one every 1.48 Seconds. These are relatively large packets at 19,416 bytes, which is larger than the 1518-byte IP limit, and so are broken down into 13 IP packets by the transport layer for transmission.

P3 and P4 Logical channel messages are sent in the E2X stream at the Block-Pair rate of one every 185 milliseconds. There are 8 each, P3 and P4 messages in each frame if configured for MP2/3 (P3) or MP11 (P4). The individual P3 and P4 message packets are relatively small at around 488 bytes in MP2 and 688 bytes in MP3 and MP4.

PIDS (Primary IBOC Data Service) messages contain information required for fast acquisition of IBOC operational information by receivers. PIDS are sent at the Block rate in the E2X protocol. These are very small packets of 100 bytes each.

Clock There are 16 clock messages in each frame, sent at the Block rate of 92.875 milliseconds. Clock messages are used to maintain a constant throughput delay from the Exgine buffers to the Digital UpConverter (DUC). When the Exgine has counted 16 clock messages, a line is toggled to the DUC indicating a complete frame has been received and the frame buffer contents are transferred to the DUC for processing. Clock messages may also be used to recover the Exporter's processing rate from across the E2X link providing 10MHz phase and frequency process synchronization to the Exgine. These are very small packets of 64 bytes. The clock message is sent as a synchronization counter in each block.

Within each E2X data frame are the messages bound for the active channels. Messages for channels P1 MPS (and SIG if present), PIDS message, and the clock message will always be present. P3 messages will also be sent as part of the frame if the Exporter is configured for Extended Hybrid mode. MPS and SIG services are encoded into the P1 logical channel. The supplemental services may be encoded within P1 or P3 depending on the service mode and configuration of the Exporter.

Figure 2 represents an MP3 E2X data frame and shows the timing relationship of the PDUs in the E2X stream as sent from the Exporter. *Note that the PDU sizes are not to scale in this illustration.*

At the starting boundary of each frame, in Block Zero

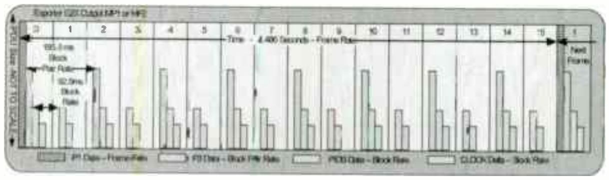


Figure 2 - E2X MP3 Data Frame

(0), the Exporter sends the frame rate PDU, the block rate PIDS and clock message PDUs, along with the block pair rate P3 message PDUs. The PIDS and clock messages are sent in each of the 15 subsequent blocks while the P3 messages are sent in each of the subsequent block pairs.

The bandwidth required for each mode depends on the logical channels configured and is shown in Figure 3.

E2X Bandwidth Requirements				
IP Protocol	Service Mode	Frame Size in Bytes	Avg. Bandwidth Kb/s	Provision Kb/s
UDP	MP1	22,232	119.7	**159.5**
	MP2	24,536	132.1	**176.1**
	MP3	27,736	149.3	**199.0**
TCP	MP1	25,870	139.3	**232.0**
	MP2	28,894	155.6	**259.2**
	MP3	31,166	167.8	**279.5**

Figure 3 - Minimum Transport Bandwidth Requirements

The following formula may be used to determine the link bandwidth required over a given time period:

$$Link\ Bandwidth, bps = \left(\frac{8 * Frame\ Bytes}{1.486\ (FramePeriod)} \right) * X$$

$$Where\ X = 1.67\ for\ TCP\ or\ 1.33\ for\ UDP$$

This gives an average data rate that may be used to calculate minimum STL bandwidth requirements. An MP3 UDP frame contains 27,733 bytes, and so would require an STL bandwidth of 199 Kb/s.

In order for an HD Radio stream to function properly, great care must be taken to avoid dropped IP packets in the STL/TSL or WAN transport. A single dropped IP packet will result in a lost frame and the loss of at least 1.5 seconds of program. The average rate of the UDP E2X stream should not exceed 75% of the link's available bandwidth. This is necessary to avoid any potential congestion, resulting in packet loss caused by peak/burst traffic. Using TCP, for data stream to function properly even under adverse conditions, the average bandwidth should not exceed 60% of the available link bandwidth. This is necessary to accommodate the higher data rate demand that occurs during retransmission of dropped packets.

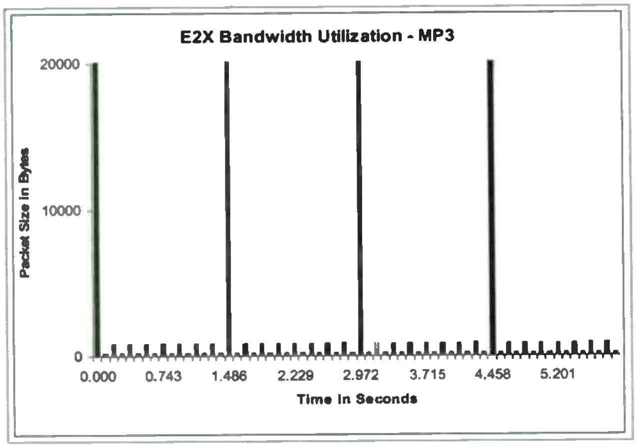


Figure 5 - Distribution of Messages in the E2X Data Stream

The E2X payload is not evenly distributed across the frame as shown in Figure 5. The largest packet is the Frame Rate PDU at 19,416 bytes for MP3. The Clock packet is a block rate PDU with a period of 92.9 milliseconds. By applying the same formula as above, using the frame rate size and the block rate period of the clock, a dedicated bandwidth of 2.3 Mb/s is required in order to avoid any interference to the timing of the clock packets. Buffering in the STL smoothes these peaks for transport and the receive buffer in the Exgine absorbs the imbalance but E2X based timing between the Exporter and Exgine will be degraded. Figure 4 show the resultant data smoothing caused by buffering through a bandwidth limited STL. Over a 320Kb/s link, the initial clock packets in each frame will be delayed by as much as 500ms

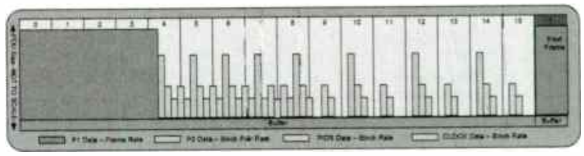


Figure 4 - Effects of Bandwidth Limited Transport

System Synchronization

Clock synchronization between the Exporter and Exciter is critical to the proper operation of the HD Radio system. The transfer of HD Radio audio frames between the Exporter and Exciter is performed synchronously. PDUs are sent from the Exporter to the Exgine at predetermined rate of 0.673 Hz or one frame every 1.486 seconds. In turn, the Exciter expects a complete frame every 1.486 seconds to form the next HD radio message to be passed to the modulator. The reception rate of the E2X stream directly affects the digital symbol rate. The OFDM symbol-clock frequency and carrier frequency must be maintained within 2ppm for proper operation (Level II compliance) and .01ppm for best performance (Level I compliance).

In order to ensure precise time synchronization of OFDM clock-symbol rate, accuracy of the carrier frequency and to allow for rapid station acquisition the entire system must be GPS synchronized. Systems not locked to GPS will not benefit from fast tuning since they cannot be synchronized with other stations.

Synchronization between the Exporter and Exgine clocks is required to maintain lock-step between the processes. Without synchronization, variations between the clock's frequencies and phases will result in analog-to-digital blending artifacts and buffer underflow or overflow in the Exgine buffers. As buffer levels change, so to will the apparent diversity delay between the analog and digital signals. When the buffers are depleted the E2X stream will stop to recover, resulting in a loss of the HD stream and a receiver dropout of to 3-4 seconds.

It should also be mentioned that the Importer's and Exporter's AES audio cards must be synchronized to a 44.1 KHz word clock that is derived from the same reference as the Exporter's 10MHz clock. This is accomplished internally on the Exporter but requires external word clock from the Importer be provided to either the incoming AES audio or to the Importer audio card. Without the Exporter referenced audio word-clock synchronization, the audio buffers will under or overrun resulting in SPS service dropouts. Figure 6 illustrates the necessary timing distribution for the Importer, Exporter and Exciter.

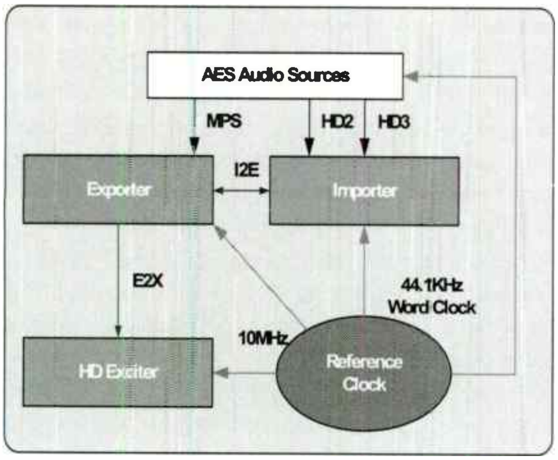


Figure 6 - Timing and Synchronization Distribution

There are three methods that may be used to assure proper synchronization between the Exporter and Exgine:

Local Direct Synchronization

GPS Referenced Remote Synchronization

E2X Remote Clock Synchronization

The simplest and most reliable E2X synchronization method is to connect data through a network or crossover cable, to provide a direct connection between the Exporter's 10MHz output and the Exciter's 10MHz external reference input. This method of course, is only practical when the Exporter and Exciter are co-located.

The most reliable method of remote synchronization is to provide a GPS receiver feeding the Exporter at the studio and a separate GPS receiver located at the transmitter site feeding a locked 10MHz reference signal to the Exciter. When using a limited bandwidth or marginal performance STL system, this is highly recommended as the preferred method of remote synchronization.

E2X Remote Clock Synchronization has been recently implemented as part of the Generation 3 system architecture. As illustrated in Figure 7. the Remote Synchronizer receives a pulse each time the Exgine receives an E2X Modem Frame Boundary Clock Message from the Exporter and continuously outputs a 10 MHz clock to the Digital Up-Converter that is now frequency and phase-locked to the Exporter's 10 MHz clock. In order to reject high frequency jitter, the PLL circuit is very slow to respond to changes and attempts to average the frequency of the incoming pulses over several hours to train the 10MHz oscillator.

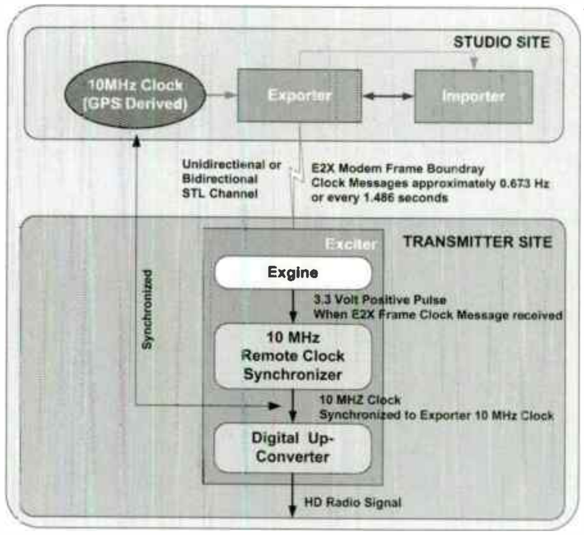


Figure 7 - E2X Remote Clock Synchronization

This method may be deployed successfully when GPS satellite synchronization is not possible at the transmitter site but it must be understood that because the synchronization is now reliant on the precision cadence of the clock packets to indicate a modem frame boundary, this method has significant limitations that must be understood.

Sufficient E2X link bandwidth must be allocated to accommodate the high peak data load presented by the frame rate PDU so as to minimize the impact to the block rate Clock packets. Ideally then, the STL would have a bandwidth of around 2.3 Mb/s in order to not interfere with the timing of the 93ms period clock packets, as discussed in the previous section. This is not necessary to maintain the E2X Remote synchronization so long as sufficient STL bandwidth and Exgine receive buffering exists to allow the current frame's 16 clock packets to arrive before the next frame boundary message.

Any significant continuous jitter in the STL system will result in a skewing of the 10MHz reference frequency. The E2X Remote Clock Synchronizer requires a relatively low jitter network/STL combination to work reliably. End-to-end E2X clock message jitter must be below ±10 milliseconds (or preferably below ±5 milliseconds) for reliable operation. Because of the very long settling time inherent in the PLL circuit, if the jitter is frequent and excessive, the PLL's 10MHz output will continue to wander resulting in continuously varying apparent diversity delay. A variation of +/- thousands of samples is not uncommon.

Deployment

Transporting the Exporter to Exgine or E2X stream is the most bandwidth efficient method of deploying HD Radio. The STL link must now be able to carry Ethernet/IP based traffic along with traditional audio streams. This requires a digital STL with sufficient band width to accommodate the traditional audio stream and HD Radio IP data stream or a separate digital STL specifically for IP based HD radio services with sufficient bandwidth overhead to accommodate any.

Traditionally, RF based STL systems have been unidirectional and many systems now provide a unidirectional Ethernet path that can be used for the E2X stream. While the E2X data stream protocol does not require a bidirectional circuit, TCP is the recommended E2X transport protocol and should be used if a bidirectional link is possible. It is important that future developments continue to support unidirectional systems such as microwave and satellite, but be able to take advantage of bidirectional circuit capabilities for network functions, such as Address Resolution Protocol and Transport Control Protocol (TCP) and Automatic Retransmission Request (ARQ).

All HD Radio devices – Importer, Exporter, and Exciter – should use statically assigned IP addresses within their own subnet. It is imperative that broadcast, multicast, and other extraneous traffic be kept off of the network path to the transmitter site and that the number of switches in the network path is minimized. The implementation of VLANs or connection of devices through a dedicated physical network will substantially reduce packet loss and data collisions.

RECENT SYSTEM ENHANCEMENTS

TCP/IP Support for E2X

A recent addition to the Exporter to Exgine communications is the ability to use TCP/IP to provide "guaranteed delivery" of data between the Exporter and Exgine/Exciter. The TCP implementation of E2X protocol is nearly identical to UDP except that each E2X IP packet will contain an additional 20 byte TCP header and there is a 64-byte acknowledgement message sent back from the Exgine to the Exporter for each IP packet received. If there is a discrepancy, between what was sent and what was received, the missing data will be resent by the Exporter. This requires a bidirectional link with at least 60% overhead to accommodate retransmission of any lost packets. The use of TCP adds an additional 7% to 12% overhead to the data stream's average bandwidth over UDP depending on the configuration and service mode.

If a bidirectional data link is available, TCP protocol should be used for guaranteed packet delivery. The legacy FM host audio must still be transported separately over a traditional audio or composite STL.

With TCP and 2 frames of buffering in the Exgine, so long as the average data stream rate occupies less than 60% of the link's bandwidth, the HD Radio data stream can tolerate up to 1% packet loss, 80 milliseconds of latency, and up to 15ms of jitter although this would be considered an unhealthy network.

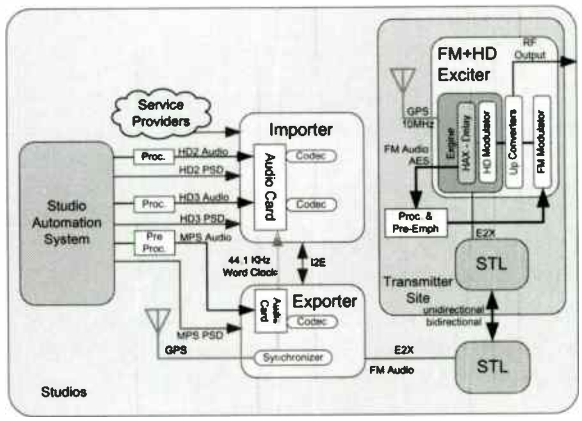


Figure 8- Embedded Host Audio Extraction System

Embedded Host Audio Extraction (HAX)

A new technique has been developed whereby the encoded MPS from the Exporter is decoded back to audio at the Exgine for use as the legacy analog program audio, thus eliminating the need for a separate STL audio stream dedicated to the analog FM signal. All program content for both the HD and analog are provided over the E2X link. Figure 8 illustrates the HAX system configuration.

The audio for the host FM signal is extracted from the E2X stream by the Exgine through the HAX daughter board and presented as AES-3 audio on the rear XLR connector of the Exgine. The FM audio may now be processed externally using traditional FM audio processors and must have the standard 50 or 75ms pre-emphasis added before being sent to the FM analog exciter.

As the system delay is now identical for both the analog and HD audio streams the only additional diversity delay required is 4-5 seconds to compensate for the receiver's additional digital processing latency. This delay is adjustable and incorporated into the HAX adapter.

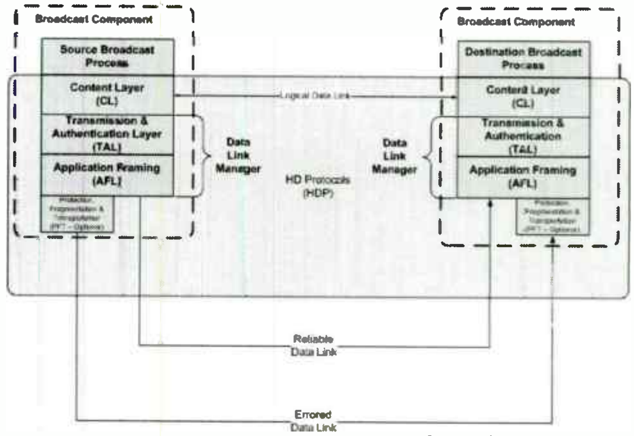
Mode Control Message Retransmission

The FM mode initialization message that is sent from the Exporter to the Exgine/Exciter, which configures the Exciter's operating mode and bandwidth had only been sent at startup of the Exporter. The mode control packet is now retransmitted periodically by the Exporter to the exciter. This assures the correct mode configuration when starting an exciter without the need to restart the Exporter.

HDP Protocol

HD Protocol (HDP) began as an effort by iBiquity to support the unification of MPS and SPS PSD. It has since been expanded by iBiquity and included in the 4.0 release to enhance and unify communications between all of the various HD Radio components support creation, distribution, command and control of the entire HD Radio system and its content from local, centralized and/or remote locations. HDP is HDP is an extensible general purpose protocol that is suitable for both unidirectional and bidirectional links, and is now the standard protocol used by the HD Radio Broadcast System throughout the industry. HDP provides options to improve robustness to network errors by providing fragmentation and error correction and to improve security by enabling the ability to digitally sign messages being received from other broadcast components and systems.

HDP is very similar to the DCP (Distribution and Communications Protocol)[1] used by Digital Radio Mondiale (DRM). In order to make use of the many features of the DCP standard, add security features and make it more suitable for use in the HD Radio broadcast architecture, HDP has been defined as the DCP standard with additional information at the AF Layer and a redefinition of the TAG Layer. The HDP stack has been created to bring the various broadcast system components logically closer together by defining a common interface to all communications between these components. The HD content is carried from the source to the destination through a



Courtesy iBiquity Digital Corp.

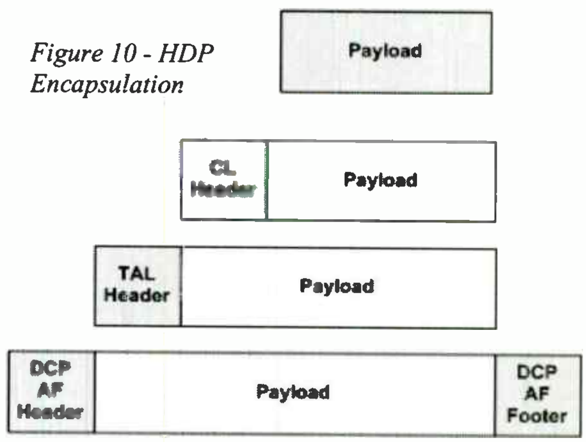
Figure 9 - HDP Protocol Stack

number of layers as shown in Figure 9.

HDP Stack Layers and Functions

The data at each layer is encapsulated in a series of packets as shown in Figure 10.

Figure 10 - HDP Encapsulation



Courtesy iBiquity Digital Corp.

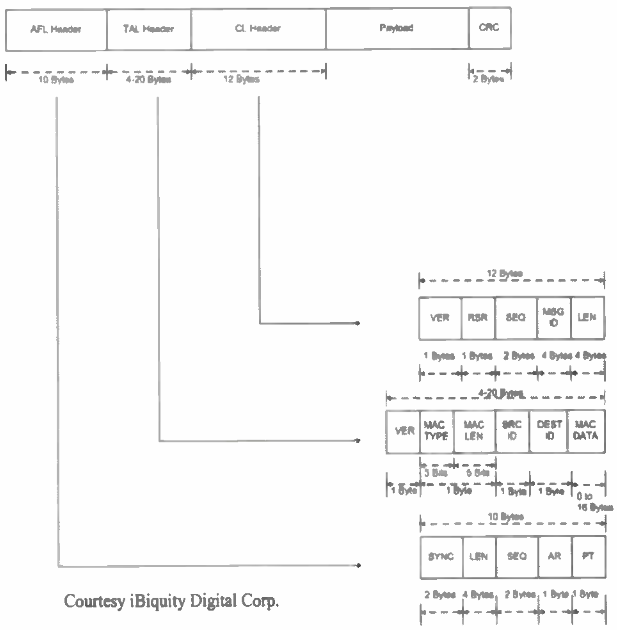
The Content Layer (CL) is specific to the destination process but typically consists of information about the payload needed by the destination process such as message identifier, sequence number, or any special processing required. **The Transmission & Authentication Layer (TAL)** authenticates data received from the AFL and does routing to different processes in the same broadcast component.
The AF Layer Header (AFL) moves packet from one broadcast system to another and combines the

[1] ETSI TS 102 821 Digital Radio Mondiale (DRM); Distribution and Communications Protocol (DCP)

elementary data into a cohesive block of related data. **The AF footer** includes a CRC check that allows the detection of transmission errors at the destination. **An optional Protection, Fragmentation and Transportation (PFT)** layer allows fragmentation of the potentially large AFL packets, and adds the possibility of having addressing and FEC. The AF Packets or the PFT fragments can then be transported by any one of a number of physical links.

The Data Link Manager indicated in Figure 9 is the process that exists on all broadcast components and is responsible for processing the TAL and AFL layers.

Any information or content transmitted by HDP is called application data. The entire HDP stack requires an additional 24 to 44 bytes and is summarized in Figure 11.



Courtesy iBiquity Digital Corp.

Figure 11 Detailed HDP Stack

FUTURE ENHANCEMENTS

Encoded Host Audio Transport (HAT)

Encoded Host Audio Transport is an expansion of the E2X protocol, whereby an additional high-bandwidth, low-latency audio stream is encoded into the existing HD streams for transporting the main legacy analog program over the E2X/IP stream. At the Exgine, the Host program audio will be extracted much the same way Decoded Host Audio is extracted. It is expected that the HAT will provide up to a 256Kb, HDC encoded, stereo audio channel and will require approximately 500Kb of additional transport bandwidth.

This could potentially result in an overall HD+FM system bandwidth reduction of 600kb for transporting.

Forward Error Correction (FEC) and Quality of Service (QoS)

The reliability of the HD Radio E2X transport over unidirectional UDP/IP has been mediocre due to the potential for packet loss using this "best-effort" delivery protocol. While the addition of TCP to the E2X protocol suite is extremely important, it does not answer all of the issues. It is important that we continue to support and develop improvements for the unidirectional transport systems such as microwave and satellite.

The requirements of real-time delivery and the lack of a return path limit the possibilities for Automatic Repeat Request (ARQ) retransmission. It is for these reasons that various forms of forward error correction (FEC) are being explored to compensate for packet loss over unidirectional links. Simple schemes such as "send the message 3 times and use a best 2 out of 3" voting scheme are inefficient error-correction methods and cannot guarantee that a block of data can be communicated free of error. Advanced techniques such as Reed–Solomon codes and, more recently, turbo codes come much closer to reaching the theoretical Shannon[2] limit, however, this comes at the cost of higher complexity and increased bandwidth requirements. In any FEC implementation, the E2X clock packets must not be re-sent, as a late clock packet will do more harm to system synchronization than a missing one.

These issues have already been addressed in currently available STL systems where packet filtering and these FEC techniques have been implemented. In these systems, it is possible to assign Quality of Service (QoS) levels to individual packets, giving clock packets the highest priority, the audio data packets the next lower priority, with all other data assigned the lowest priority. It is also possible to

[2] The Shannon theorem states that given a noisy channel with channel capacity C and information transmitted at a rate R, then if $R < C$ there exist codes that allow the probability of error at the receiver to be made arbitrarily small. This means that theoretically, it is possible to transmit information nearly without error at any rate below a limiting rate, C.

provide additional GPS timing signals to both the transmit, and receive terminals of the link to assure correct clock synchronization alignment regardless of link latency and jitter conditions.

By embedding low-density parity-check (LDPC) codes or turbo codes into the HDP PFT layer along with packet filtering and QoS mechanisms, it is possible to greatly improve the performance of the unidirectional UDP E2X link as part of the Exporter/Exgine Architecture to reduce component count, system complexity and cost.

CONCLUSIONS

The real-time streaming nature of the technology presents some very difficult challenges to implementing a fully robust HD Radio architecture. The E2X protocol itself is sound and works as intended if implemented properly. It requires a very high degree of accuracy and reliability in order for the HD Radio system to function properly, while relying purely on the physical link layer to provide that reliability.

E2X protocol as it stands today may be used reliably, if best practices in network engineering are followed.

- Providing sufficient link bandwidth overhead with low jitter and latency is critical.
- Providing for proper system synchronization is also critical. GPS synchronization of the Exporter and Exciter should be considered as the first, best choice.
- Restricting extraneous traffic through network segmentation and proper switch configuration is essential. The HD Radio system must be on its own subnet.
- Use TCP/IP as the transport protocol of choice for the E2X link if this is option is available.

Many improvements have been made, particularly with the implementations of TCP, and HDP. Development will continue in the areas of synchronization and data loss mitigation.

HD Radio™ is a trademark of iBiquity Digital Corporation, Columbia, Maryland

The HD Radio™ transport design and description are copyright iBiquity Digital Corporation and used with permission.

TV News and Live Production

Monday, April 14, 2008
1:00 PM – 5:30 PM

Chairperson: Wayne Kube
Belo Corporation, Dallas, TX

***News ML-G2: Metadata for News Exchange**
Jean-Pierre Evain, European Broadcasting Union (EBU), Grand Saconnex, Switzerland

Advanced Video Image Technologies for Sports TV Productions
Kimihiro Tomiyama, NHK Science & Technical Research Laboratories, Tokyo, Japan

Managing Multiformat Images for the Broadcast News Environment
Karl Paulsen, AZCAR Technologies, Canonsburg, PA and Markham, Ontario, Canada

Best Practices: Using IP in Broadcast TV
Joel Wilhite, Harmonic Inc, Sunnyvale, CA

Live Integrated Production Systems Streamline Live Workflow
Ken Swanton, Broadcast Pix, Billerica, MA

***Making Field Applications Bandwidth Efficient**
Mick Gardina, iDirect Technologies, Herndon, IL

High Definition Electronic News Gathering (HD-ENG) Field Test Report
Walter Sidas, CBS, New York, NY

NBC Universal's New IPTV Distribution System
Robert Goldfarb, NBC-Universal, New York, NY

***Understanding and Implementing an Ultra-fast Time-to-Air Workflow by Integrating Metadata**
Ed Casaccia, Thomson, Beaverton, OR

*Paper not available at the time of publication

Advanced video image technologies for sports TV productions

KIMIHIRO TOMIYAMA†, KENSUKE HISATOMI†, MIWA KATAYAMA†, YUICHI IWADATE†
TETSUYA HAYASHIDA† HIROTAKA MARUYAMA†, T. GOJI ETOH‡, and TETSUO YOSHIDA*
†Science & Technical Research Labratories NHK (Japan Broadcasting Corporation)
Kinuta, Setagaya-ku, Tokyo, Japan
‡Kinki University, * Hitachi Kokusai Electric co. Ltd.

ABSTRACT

There is strong demand for more dynamic and impressive images for live sports TV productions. In response to this demand, NHK (Japan Broadcasting Corporation) is seeking to establish a new visual production process using a multi-view HDTV system and an ultrahigh-speed, high-sensitivity color TV camera.

The Multi-view HDTV system can generate multi-view HDTV images of a sports scene by matching the shooting directions and shooting angles of the individual camera images by using projective transformations. We describe an example of practical use of this system in a live broadcasting of a gymnastics competition.

The other system, the ultrahigh-speed, high-sensitivity color TV camera, has been used in broadcasting of sports events, such as baseball night games and golf tournaments. The camera is able to take clear slow motion images of the moment of impact of the bat or club with the ball. It is a powerful tool for new forms of video expression.

Multi-view HDTV system

Background

Productions using multi-view image systems such as the "Eye Vision" system[1] have been a much-discussed subject recently. These systems are designed to produce three-dimensional and dynamic views during playbacks of the action by using multiple cameras set around the playing fields, stadiums, and athletes.

To use these systems in a live sports broadcast, the following three points are important. First is simplicity of system installation. It is necessary to set up the system quickly in order for it to be used in various sports programs. Second is immediate output. In a live sports program, it is necessary to broadcast the multi-view images immediately after the event. The third is a high-resolution image.

In consideration of these points, NHK has developed the Multi-view HDTV system. The features of our system are as follows.

1) The system is simple; it can be set up in about a day.
2) It can immediately output multi-view HDTV images during a live sports broadcast. It takes approximately 15 seconds to deliver the images.
3) Smoothly switched images can be displayed by performing image processing matching on the shooting directions and shooting angles of the camera images.

This system supports commentary in live sports broadcasts. In the following, we explain a Multi-view HDTV system that uses 12 HDTV cameras.

CONFIGURATION

This system consists of 12 HD cameras and frame memories (Fig. 1). The uncompressed video streams (1080/30p /YUV:422) in HD-SDI format from these cameras are simultaneously recorded onto frame memories, and are then composed on a workstation. The shutter timings of the cameras are synchronized with the genlock signal. Note that recording and playback are controlled by a master workstation.

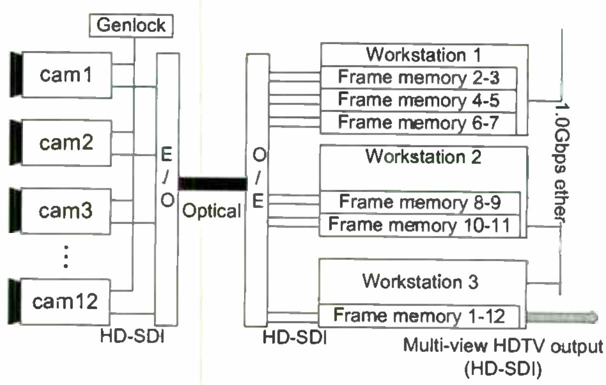


Fig. 1 Multi-view HDTV system

COMPENSATION OF THE IMAGE BY THE PROJECTIVE TRANSFORMATION

Video images of the cameras are switched along with the camera allocation. To switch camera images smoothly, we need to converge the shooting directions of all cameras at a target point on the subject and need to match the shooting angles of all cameras. There are difficulties in doing this in a live sports broadcast because there's very little time to adjust the cameras. Hence, our system adjusts them by using image processing to create a smoothly switched sequence of images.
The procedure of image processing is shown as follows.

1) Camera calibration
The camera calibration must be carried out to obtain camera parameters such as three-dimensional position, orientation, and focal length of each camera. A pattern (Fig. 2) is set up in the target area, and the camera calibration is automatically done on the image of this pattern with software.

2) Select target point
The target point is indicated by the operator. The operator clicks the target point on two of the camera images. The three-dimensional coordinates of the target point is calculated by stereo matching using the camera parameters.

3) Converge shooting directions of all camera images
To converge the shooting directions of all cameras at the target point, we pan and/or tilt each camera image through a projective transformation that uses the three-dimensional coordinates of the target point and the camera parameters.

4) Match sizes of all camera images
The three-dimensional distance from each camera to the target point is calculated using the camera parameters. To make the size of the subject consistent in all camera images, each camera image is enlarged or reduced in proportion to the camera's distance to the target point.

The images of each camera are converted in semi-real time by using the graphics board of the workstation. It takes about five seconds to process the sequence of 2 or 3 seconds.
As a result, the target point and the image size are automatically compensated to synchronize the camera images. The result of this processing is multi-view HDTV in which the shots smoothly switche.

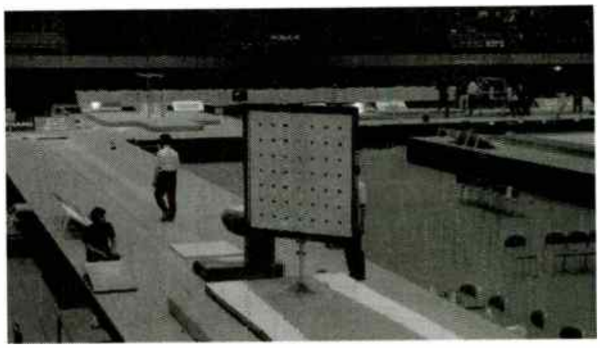


Fig. 2 Calibration pattern

PRACTICAL EXAMPLE IN LIVE BROADCASTING OF A GYMNASTICS GAME

System setting

We describe a practical example of a live gymnastics broadcast. The 12 cameras were set up the front row of audience seats, on the 2nd floor of the arena, and they lined the side and front of the runway and vaulting horse (Fig. 3). The workstations equipped with the frame memories were set up on an audience seat. Four staff took only half a day to complete the setup.

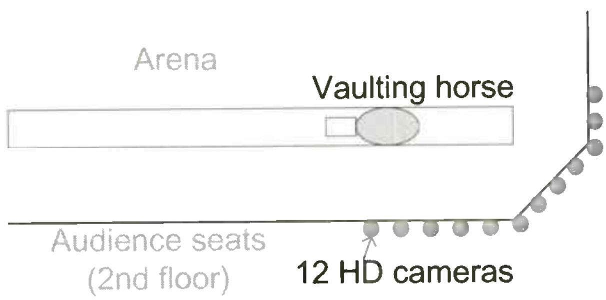


Fig. 3 Camera setting (for the vaulting horse)

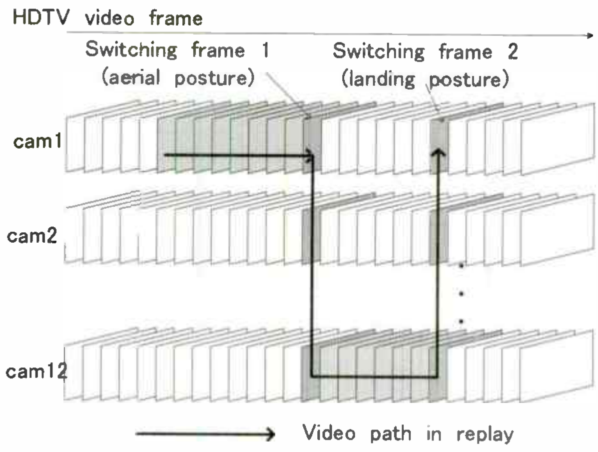


Fig. 4 Video path of multi-view HDTV

Operation

One operator and one communication staff member conducted all the operations of the live broadcast.
The recording of the video streams started when the gymnast started to sprint down the runway. As soon as the gymnast landed and the performance ended, images which had been captured in the aerial posture and landing posture were chosen as the switching frames to create the multi-view HDTV images (Fig. 4).
It took approximate 15 seconds to deliver the multi-view HDTV images (10 seconds for choosing the switching frame, 10 seconds to read the images from the frame memories and to edit it). These images were used during commentaries after each performance.
Fig 5 shows a broadcast sequence of multi-view HDTV. In the vaulting horse performance, the operator first shows the moment that the gymnast touches the vaulting horse. The operator then pauses the images for a second, and shows the aerial and the landing postures by switching between the 12 camera images. This sequence made it easy for viewers to understand the motions and the positions of the gymnast in the sequence (Fig. 5).

In addition, we developed an image expression method for multi-view HDTV and multi-motion images (Fig. 6).

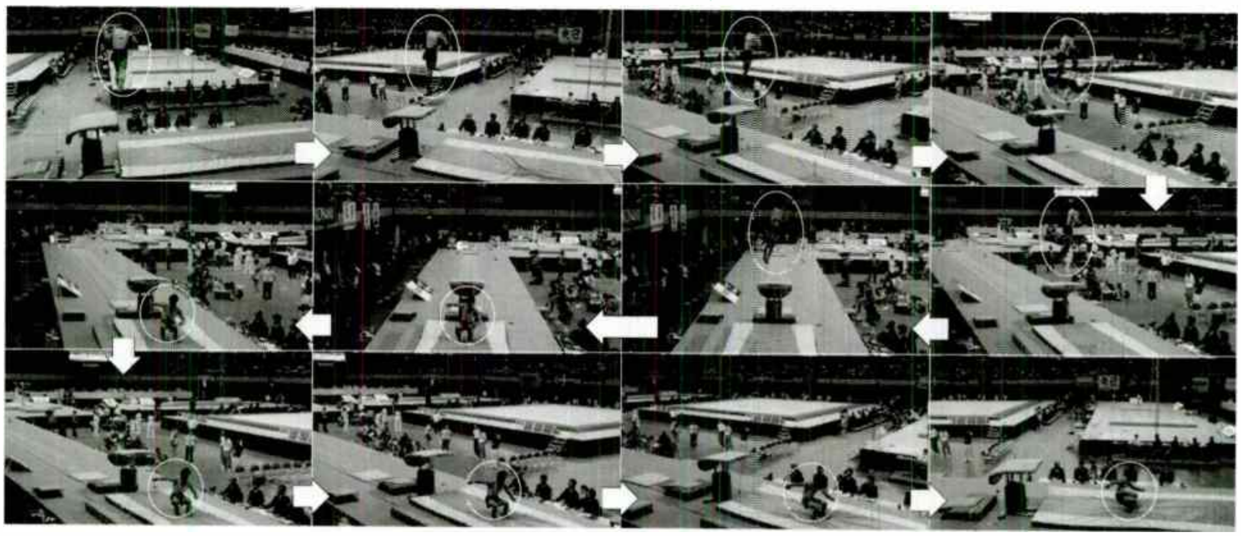


Fig. 5 Multi-view HDTV (vaulting horse)

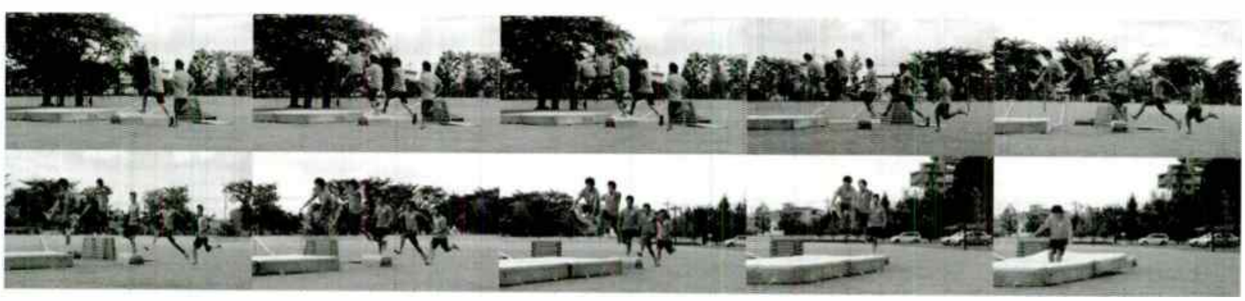


Fig. 6 Multi-view multi-motion shot

Ultrahigh-speed Camera

Background

Recent advances in cameras' ability to capture fast-moving phenomena that cannot be perceived clearly with the naked eye, and to represent these in slow-motion video, are drawing interest not only for broadcast applications but also for industrial applications[2) 3)].

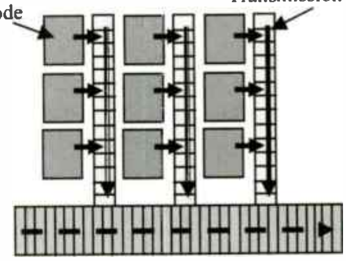
Many high-speed cameras incorporate a CMOS imaging device[4)], which can read out signal charges at high speed by using an X-Y matrix switching scheme. However, there are problems with CMOS imaging devices; they require especially intense lighting to obtain video with a good signal-to-noise (S/N) ratio during high-speed shooting with a short exposure time because they are affected by noise and have inadequate sensitivity. This requirement has made it difficult to shoot high-quality high-speed images for relay broadcasts of nighttime sports events at facilities with inadequate on-site lighting.

Our goal has thus been to develop a high-speed camera suitable for broadcasting programming. To this end, we have developed a 300,000-pixel ultrahigh-speed, high-sensitivity color camera that employs a special CCD that provides both high-speed operation and excellent sensitivity.

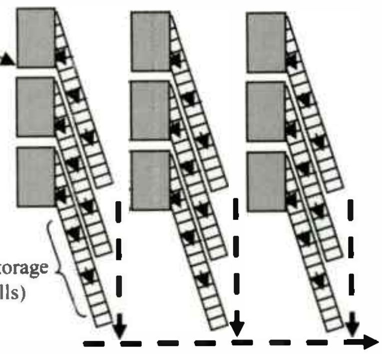
We have used this camera in relay broadcasts of professional baseball games and golf tournaments. In the broadcasts, the camera proved to be a powerful tool for shooting new forms of video expression; for instance, it captured extremely vivid footage of a ball's impact with the bat during a baseball game.

PRINCIPLE OF ULTRAHIGH-SPEED HIGH-SENSITIVITY CCD

Fig. 7 illustrates the principle of the ultrahigh-speed CCD in comparison with that of an ordinary CCD. In ordinary CCDs, signal charges, which are generated in photodiodes by light, must be read outside the devices over long transmission paths (1,000 or more transfer steps), making it difficult to attain higher frame rates. In ultrahigh-speed CCDs, each CCD sensor pixel has its own in-situ storage area for image signals, which reduces the number of transfer steps involved to one per frame, making ultrahigh-speed imaging at up to 1,000,000 frames/sec possible. This operation is called ISIS (In-situ storage image-sensor) [5)] mode. High-sensitivity was obtained by designing a CCD with large photodiodes and with characteristic low noise. The photodiode area of the ultrahigh-speed CCD is about ten times larger than that of conventional CMOS imaging devices.



(a) Ordinary CCD



(b) Ultrahigh-speed CCD

Fig. 7 Principle of ultrahigh-speed CCD and that of an ordinary CCD

FABRICATION OF 300,000-PIXEL ULTRAHIGH-SPEED CCD

Each of the 144-step memory cells for each of the 300,000 pixels in this CCD must be handled similarly to a regular CCD pixel. Hence the manufacturing technology required is equivalent to that for a 300,000 x 144 = 40 megapixel CCD. Therefore, fabricating this CCD is very difficult.

This led us to adopt a buttable[1] structure in which two 150,000-pixel CCDs are joined to create a 300,000-pixel CCD. In this butting technique, one side of the CCD is cut in such a way as to support the layout, and the photodiodes are positioned to the side of the pixels. To perfect this butting technique, we developed new technologies for precise cutting and alignment of the CCD chips. We also investigated a technology to cut silicon wafers precisely and developed an precise alignment system using a piezoelectric stage.

Fig. 8 shows a photograph of the ultrahigh-speed CCD. The CCD was equipped with an on-chip color filter for the single-chip portable color camera.

[1] Method of seamlessly joining multiple CCDs to build a single CCD with a greater number of pixels.

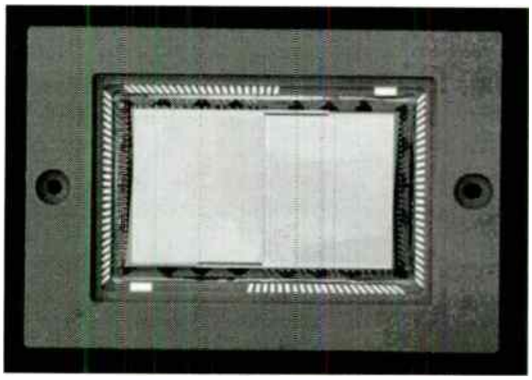


Fig. 8 300,000-pixel ultrahigh-speed CCD

ULTRAHIGH-SPEED, HIGH-SENSITIVITY COLOR CAMERA

Fig. 9 shows a camera mounted with the single-chip CCD with a color filter. The camera weights 7 kg and is small enough to be hand-held. The lens-mounting interface of this camera is an F-mount, so general-purpose lenses can be used with it.

Table 1 shows the camera's major specifications. Sensitivity is F (F-number) 22½ at a light intensity of 2000 lx and a recording rate of 30 frames/sec. The S/N ratio is 54 dB. This camera has two shooting modes, ISIS mode and external memory mode. ISIS mode's frame rate is up to 1,000,000 frames/sec. The number of consecutive images is 144 frames at any frame rate. Image signals are recorded in the image sensor's internal storage. The external memory mode's frame rate is up to 1,000 frames/sec. The number of consecutive images is 5,000 frames (2-GB memory) at any frame rate. The image signals are recorded in SDRAM.

Image sensing device	Ultrahigh-speed high-sensitivity color CCD
Pixels	720 x 410 pixels
Aspect ratio	15.9 : 9
Frame rate	30-1,000,000 fps (ISIS mode) 30-1,000 fps (External memory mode)
Number of stored images	144 frames (ISIS mode) 5,000 frames(External memory mode, 2GB)
Sensitivity	2,000 lx F22½ (30 fps, 5,600 K)
S/N	54 dB
Video output	HD-SDI : 1 output
External trigger	TTL (5 V: positive or negative) Switch closure (normally open)
Lens mount	F-mount
Camera head weight	7 kg
Power consumption	75 VA
Dimension	160(W) x277(H) x 485 (D) mm

Table 1 Major specifications of the 300,000-pixel ultrahigh-speed, high-sensitivity color camera

Fig. 9 300,000-pixel ultrahigh-speed, high-sensitivity portable color camera

PRACTICAL EXAMPLES IN LIVE BROADCASTING

Live professional baseball

We have used this ultrahigh-speed camera in TV programs of live professional baseball games. Owing to the high sensitivity of this camera, we could shoot pitching and hitting scenes with a telelphoto lens clearly in spite of the low-light conditions of the night games. Fig. 10 shows an example image of a pitching scene. The details of the pitcher's grip and ball release are clearly visible.

Fig. 10 Pitching shot

Live professional golf

Fig. 11 shows example images taken at a professional golf tournament. The lighting at the outdoor location of the subject varied greatly from about 60,000 lux under clear skies to about 10,000 lux under cloudy skies. Even under cloudy conditions, this camera was able to clearly capture the impact of the club on the ball at the rate of 1000 frames/sec. The portability afforded by the light weight of the camera enables us to take the lay-up shots anywhere on the golf course.

Fig. 11 Sequence of images showing a golf club impact

CONCLUSIONS

We reported on our new Multi-view HDTV system and our ultrahigh-speed, high-sensitivity color TV camera.

The Multi-view HDTV system is able to display images of sports player's movements from various aspects by switching between HDTV camera images. In addition, this system can generate muti-view HDTV images that enlarge an arbitrary point on the images. Therefore, the system can display the multi-view HDTV images centered on the objects such as balls that move at high speed.
The system uses a graphic board to enable broadcast of smoothly switched multi-view HDTV images within about 15 seconds. Because images from 12 HDTV cameras are recorded with three workstations and the system configuration is simple, it is possible to apply it with various sports programs. So far, we have used the Multi-view HDTV system in live programs on gymnastics and sumo wrestling. In the future, we would like to move on to adapt this system to more general and practical uses for sports TV programing including golf, baseball, soccer, etc.

The 300,000-pixel ultrahigh-speed CCD is for a camera with a very high maximum frame rate (1,000,000 fps), high sensitivity, and high functionality for broadcast applications. A 150,000-pixel CCD was designed and fabricated to ensure an adequate fabrication yield, and techniques to cut and position two 150,000-pixel CCDs with a high level of accuracy were developed to join the two CCDs into one unit with 300,000 pixels.
Applying an on-chip color filter to the 300,000-pixel CCD resulted in a single-chip, ultrahigh-speed portable color camera. The camera has been used to shoot various high-speed phenomena, and its performance in high-speed image capture is excellent. To complement its extremely high speed, this color camera is portable and can be hand-held, hence it is suitable for use in broadcasting live footage in various TV programs.

REFERENCES

1.http://www.ri.cmu.edu/events/sb35/tksuperbowl.html
2.T. Ogasawara et al., 2005. Development of an HDTV high speed camera using three 2.2M-pixel CMOS image devices. Proceedings of the international Broadcasting Convention, September, 2005. pp. 209 to 217.
3. H. Maruyama et al., 2004. Color Video Camera Capable of 1,000,000 fps with Triple Ultrahigh-speed Image Sensors, SPIE 26th International Congress on High-Speed Photography and Photonics, 2004. pp.244 to 249.
4. A. Krymski et al., 2001. A High Speed,240 Frames/s Mega-pixel CMOS Sensor, Proceedings of 2001 IEEE Workshop on CCD and AIS, 2001.
5. T.G. Etoh et al., 2002. A CCD Image Sensor of 1M frames/s for Continuous Image Capturing of 103 Frames, Proceedings 2002 IEEE International Solid State Circuits Conference, 2002. pp. 46 to 47.

MANAGING MULTIFORMAT IMAGES FOR THE BROADCAST NEWS ENVIRONMENT

KARL PAULSEN
AZCAR Technologies, Inc.
Canonsburg, PA, USA - Markham, Ontario, CANADA

ABSTRACT

Developing approaches for handling 4:3 and wide screen standard definition to produce high definition content for broadcast news requires different workflows be integrated into the acquisition, transport, editing, storage-and-archive, playout, and interchange of media. Addressing the challenges faced by the industry in managing wide screen formats, while also retaining compatibility with others is the topic of this paper which give examples on the practical design of sub-systems that allows for flexibility and adaptability going forward.

Using implementation and actual cases where the users wished to maintain one portion of an existing workflow while migrating to full high definition operations for live broadcast, these practical applications for conversion hardware and software will be described by example. How hardware solutions can be applied to real workflows, including automated systems that prevent negative results by keeping humans' hands off, will be cited.

THE TRANSITION AND MIGRATION

Those that recall the migration from film to video may have a sort of Déjà-Vu feeling when they start the conversion from full screen 480i to wide screen and/or high definition. Workflows, processes, time to air, and even content collection all underwent a tremendous transition in those years – one that the users might now say, in relief, 'I'm glad that era's over!'

Fast forward from the late 1960s to the present and we find that over the previous decade news organizations have been steadily migrated from a purely analog, videotape based acquisition and release format, through the standard definition digital era and are now squarely in the midst of a wide screen, high definition capture, edit and release to air domain. Beyond the basic direct-to-air requirements has emerged a multitude of ancillary delivery platforms ranging from Internet streaming to cell phone and on to the mobile handheld world. On the horizon are hopes for additional services that will be added in concert with the conclusion of analog over-the-air terrestrial broadcast – services that will harmonize the desires for ease of mobility in the digital over-the-air world.

Between these analog and digital extremes are literally dozens of variations, extensions and renditions in image and media formats. Within this mix are a growing set of image formats that challenge the artistic composition and production capabilities; the makeup of media and storage technologies; video, audio and data compression methodologies; and of course the emergence of new delivery platforms affecting transmission, emission and signal carriage.

For the local community, television news remains one of the principle means of getting information, and sustains the rationale for the continuation of free over the air broadcast. The modern news organization faces continual pressure to deliver program content that is one step ahead of its neighbor - for market competition yields ratings which results in revenue. As the half century plus of the analog television broadcasting era heads toward its sunset, the emphasis for television is shifting more to high definition broadcast – or perhaps the 'perception' of a high definition broadcast - which drives the local broadcaster as well as the network delivery mechanisms to change hardware, software and production workflow in order to meet these demands.

COMPROMISING FOR THE PAST

Mixing the methodologies of production and image footprints of the past with the digital presence of the future involves not only technical challenges, but the transition necessitates that certain philosophical, aesthetic and financial choices be made. The path to these choices are varied, with some choices having to be made much earlier than desired, and in turn forcing the broadcaster into dealing with varying impacts on how the organization may evolve going forward. Decision makers may find they must compromise desires for a single format/single platform system; in turn seeking alternatives that address the issues of mixed aspect ratios, multiple delivery platforms, retention of legacy media components, and image quality for workflow functionality.

Future implementations and systems design must therefore consider flexibility and extensibility while preventing (or limiting) obsolescence. In planning for growth, it may be that the facility must move in baby steps rather than a massive leap to the next generation in digital high definition.

It's further unlikely that newsroom operations can simply ignore the past three decades of the 4:3 standard definition, 480i world they grew up in; regardless of the inevitable desire to compete in the high-def market space. The inevitable discussion then becomes how to handle the mix of material and options given the qualifications and quantifications of the existing infrastructure. Systems designers should consider all the elements of both the past with the future, placing levels of importance for the continuation of some sub-systems while simultaneously weighing cost, functionality, extensibility, practicality and the importance of the station image against revenue.

ADDRESSING THE SYSTEMS

A broad look at the systems which should be considered when determining the changes that must be made include:

(A) Image Capture Systems
- Field recorders
  - Linear Tape
  - Solid State Media
  - Optical Media
- Field cameras
  - 4:3 only (480i)
  - 16:9 (480i anamorphic)
  - SD (480i) on an HD imager
  - Native HD
  - Professional vs. pro-sumer/consumer
  - Cell phone, other mobile imagers
- Remote Platforms (fixed)
  - Tower cameras
  - Point of View (POV)
  - Secondary contribution (highway cameras)

(B) Transmission and Contribution Systems
- ENG microwave systems
  - Analog 480i only
  - Digital 480i only
  - Digital high definition
- Airborne platforms
  - Analog 480i only
  - Digital relays for 480i and/or high definition
- Satellite platforms
  - DSNG/SNG trucks
  - Digital transmission technologies, including compression
  - Satellite receivers and transmitters
- Alternative Delivery
  - Fiber local loops, private carriers
  - Internet

(C) Ingest, Editing and Graphics Systems
- Media
  - Linear Tape
  - Solid State
  - Optical
- Platforms
  - Non-linear Server Based
  - Desktop Platforms
  - Purpose Built Platforms

(D) Play to Air Systems
- Delivery Media
  - Servers
  - Tape (digital, compressed, analog)
- Live Studio systems
- Streaming Media Platforms
- On Demand
- Mobile/Handheld

(Note: only systems A, B, and D will be discussed in this paper)

AN INVESTMENT IN HD? FROM THE IMAGE FORMAT PERSPECTIVE

For those news departments that recently made significant investments in an SD non-linear solution for editing, and are still shooting in the field on videotape; the upgrade to a full HD-implementation may not be on their roadmap. Certainly, if that investment was in the previous 2-3 years, the consideration for high definition news may only have been a blip on their radar screen. For these broadcasters, their systems may be forced into retention of an all 480i SDI-based 4:3 infrastructure.

In this scenario, the balance of the station's live-to-air production system probably remains 480i SDI, with only the master control functioning as a typical HD/DTV pass through (see Figure 1); that is, the output of the 480i master control chain is simply upconverted to high-def. Network delivered HD programs (e.g., prime time or sports) are A/B-switched between the HD-network feed and the upconverted SD-master control. Effectively the station produces an SD signal rendered to a live HD format, usually in a 16:9 frame with a 4:3 image accompanied by black side panels.

A logical 'next step' for live production and local program delivery is wide screen 480i, as in *standard definition-wide screen* (SD-WS). This model gives the impression of high-def, but without the need for building an entirely new technical infrastructure. In this mode of operations (Figure 2), camera imagers are shifted into SD-WS scanning mode whereby 16:9 images are captured into a native 4:3 frame, but are horizontally squeezed to fit into the 4:3 aspect ratio footprint. Should the SD-studio cameras also have the wide screen optical and electrical switching option, and the video production switcher can be configured for a wide screen mode; then the station has a relatively straight forward path to wide screen on air news broadcasting – at least for the in studio portions of the newscast.

Further in the broadcast chain, should master control air operations be either HD-based or a hybrid-SD with downstream network-HD switching, the output of the live studio production switcher might be input to an upconverter whose HD output is passed into a separate HD sub-switcher just head of the processing and encoding chain. The result, even if the editorial and field capture remained SD-only, is a full time wide screen broadcast portrayed in a pseudo-HD image format giving the station the appearance of an HD-news broadcaster.

If the production or master control rooms are equipped only with smaller CRT monitors, the image just lost a noticeable amount of perceived detail based upon the

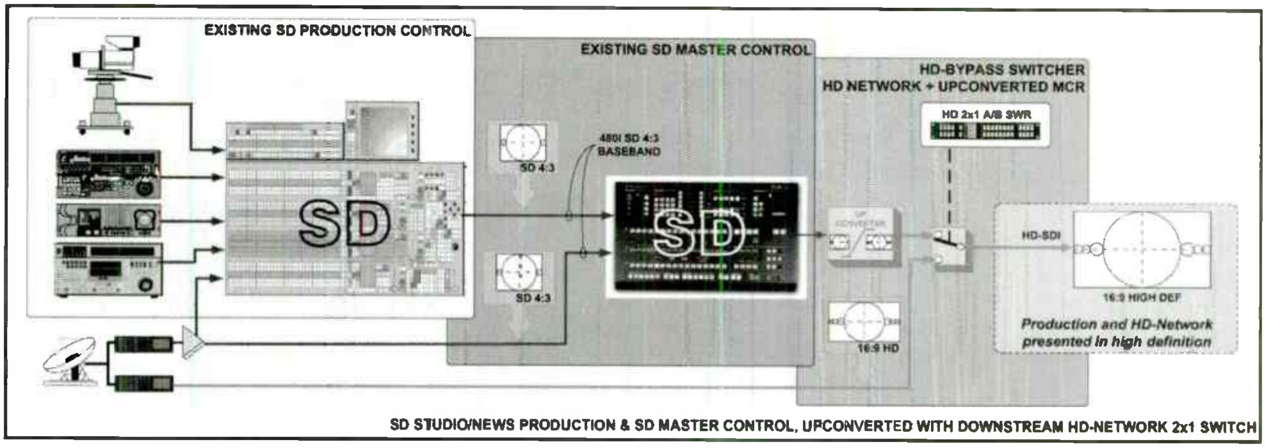


Figure 1

FIELD ACQUISITION AND PROCESSING

Extending the previous scenario to the field camcorders: if the station already uses field cameras that are wide-screen (16:9)/full-screen (4:3) switchable, they can begin capturing stories in the wide-screen 480i format. The editorial process, using tape-to-non-linear editors (or even tape-to-tape editors), can then be configured for editing in a wide-screen mode. At this juncture, the production infrastructure now begins a solid shift to a 16:9 aspect ratio, even though the content is stored as 480i WS-SD images (i.e., in a 4:3-SD frame).

distance the operators are from the monitor wall. This problem is significantly reduced if the monitor wall is a multi-image generator shown onto a set of flat panel displays that can be configured to stretch the WS 4:3 frame into a 16:9 aspect ratio, making the images on the display themselves larger.

GRAPHICS

Graphics generation must also take on a new dimension depending upon where in the signal path graphics are added and how keyers or CG devices can be configured

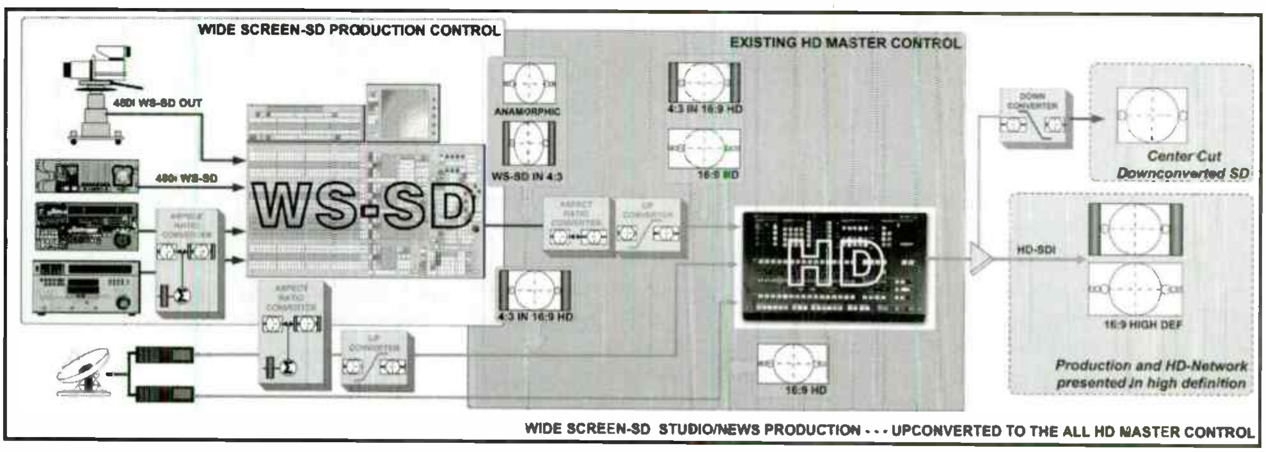


Figure 2

However, this is where several things have to change. It now becomes paramount that consistency throughout the production chain is maintained. Monitors in the control rooms that are displaying WS-SD must be capable of producing a wide screen image on a 4:3 raster. WS-images will now appear in a letter boxed format and you'll want to see these wide screen images even if the monitors are only 4:3 (CRTs). This does present somewhat of a visual acuity issue as the apparent size of the image is now reduced vertically due to letter boxing.

for wide screen anamorphic images. There are no hard fast rules for dealing with a 4:3 graphic in a wide-screen or anamorphic domain, and it may require that the point where graphics signals are inserted in the video chain must also change. This may indeed be one sub-system where replacement of the graphics generator with a device capable of producing either an anamorphic version of a character string, a dual mode SD/HD output, or even an HD-generator set up for an interim use in an SD mode is necessary.

When placing a wide screen graphic into the 4:3 WS-raster, artists should generally protect the 4:3 image space (see Figure 3), as the air signal when seen on an NTSC channel will most likely remain a 4:3 frame that is derived from a center-cut/down-conversion of the high-def release to the transmitter – and truncating the front and back of the string is not advised. One should probably not assume this problem disappears after the analog shutoff in 2009, as for the foreseeable future, a 4:3 image is most likely to remain throughout the viewing world.

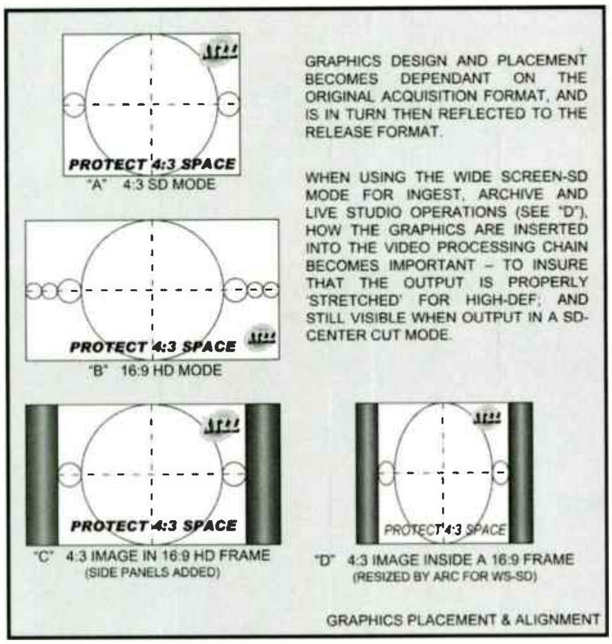


Figure 3

INTERMIXING LEGACY AND LIVE 4:3 CONTENT

For those legacy tapes originally captured in a 'full aperture' 4:3 format, a further bit of care must be exercised. Once a station has taken the 480i WS-SD path, the options become more prescribed for the handling of legacy or live material when it is created in a full aperture, non-anamorphic format. Of concern are those signals from the helicopter equipped with only a 4:3 imager and lens, or the State's traffic camera network, or even the weather graphics equipment which has no wide screen mode. Outside the facility, the interchange with other news services that only receive or deliver 480i SD (non-wide screen) content may present issues that will require a firm definition of how to structure or modify incoming content for consistency, or how to redistribute your content to others for use in other mediums.

As cited previously, in those facilities that undertake a WS-SD or hybrid-SD/HD approach, the graphics departments must design for a workflow process that maintains compatibility for the both legacy and foreign (out of house) content. At the same time they must prepare their own material for future integration once an all HD-environment is enabled. PC-workstations with commercial off the shelf (COTS) software generally don't address the issues of developing and creating files for a mixed format environment; and only recently have dedicated hardware platforms been able to properly deal with the WS-anamorphic representations of graphics as well – leaving these functions to external devices, such as aspect ratio converters or production switchers.

IMMEDIATE CONVERSION TO WIDESCREEN

Another option in the process is stretching all native 4:3 content into a wide screen format, and outputting that image back into an anamorphic 4:3 raster. And some broadcasters do not want to see side panels on their content – so they use a device that performs a non-linear stretch at the vertical edges of the 4:3 image which retains the center portion of the image in a 1:1 ratio but as you approach the edges of the 4:3 region, the image is the stretched in an almost logarithmic fashion that drags only the last 10% or 15% of each side of the image out to the edge of a 16:9 'frame'. Depending upon the treatment of the 'stretch', this can be annoying to the viewer and can alter known shapes into oblong images that change aspect ratio as they enter or leave the edges of the frame.

Regardless of how you pre-process the 4:3 images, once the facility moves to an SD-WD production chain, any '16:9-frame' must then be linearly 'squished' back to a 4:3 raster using an SD-SDI based aspect ratio converter (ARC), a process easily handled by today's modular terminal equipment. Provided the CRT displays and/or multi-image viewers can stretch the image back into a 16:9 viewing frame, the entire process of ARC-ing (i.e., converting the signals from the storage media) becomes transparent.

In the long run, material should be archived, edited and/or played back in a matching format that mitigates the sometimes unpleasing switching between wide screen and full aperture with side panels often seen on HD-broadcasts. How this is accomplished is set by the capabilities of the processing equipment and how the station engineers configure the systems – so be certain the features needed for your workflow can be met by the modular terminal equipment or the utility conversion devices selected, before making the purchase.

SIDE PANEL ALTERNATIVES FOR WS-SD

For aesthetic reasons, some news organizations may not wish to distort the 4:3 image, preferring to retain side panels for native (legacy) 4:3 images when upconverting and presenting in a wide screen 16:9 format. At least two base line options are available: one places the 4:3 image into a 16:9 frame then adds black side panels, but then horizontally aspect ratio compresses the image back into a full frame 4:3 raster. The second option incorporates the same functions, but places either a static curtain image into the side panels or an animated graphic generated either externally, by a continuously changing graphics panel incorporated into the modular conversion terminal gear, or by a static image also loaded externally into the AR-converter.

There are pros and cons to these side panel alternatives (most evident in both the editorial process and in the playout to air), which is why news and graphics departments need to establish a policy on conversion during the system design process. When the first option is used (i.e., adding only the black side panels onto all the 4:3 images), color or animated curtains can then be added using a vertical curtain wipe insert over the black side panels during the live to air playout. This can easily be accomplished either externally (i.e., a curtain generator inserted ahead of the input to the switcher) or internally (i.e., the production switcher mix effects system adds a current of choice live). Automated news production systems or EMEM registers in manual production switchers can add substantial ease to this live process.

NON-LINEAR SD EDITORS

Having established a baseline structure for handling 4:3 legacy material *and* native 480i widescreen field images – most editing systems can are the capabilities to configure their edit displays to handle the stretching of the images (in their edit windows) through their preferences set up control panels. The effect of the 'squished' wide screen image is unnoticeable, and treated in much the same fashion as the multi-image viewers for control room flat panel displays. The station's investment in the SD-non-linear systems is not lost, and is further positioned to move to HD-widescreen as budgets or market completion dictates.

As long as the content is treated consistently, including new and legacy material; and that the finished segments or completed stories are archived so they can be recovered properly, the continued use of an SD-NLE should not be an issue. That being said, the policies and operational procedures of the news department should establish their own guidelines, and carry them through the entire production process – from capture through ingest, on to editing and playout, and finally to archive.

TAPE TO NON-LINEAR INGEST

For the ingest and tape transport portions, this 480i WS-SD model requires that the output of videotape transports be outfitted with an aspect ratio converter (operating in SD-mode) such that all content ingest to the non-linear editing system's storage is consistent. All material is then prepared as though it were going to be a wide screen image. This procedure readies the station up for the future, at a time when either a partial conversion to high-def occurs, or for the eventuality of a full conversion to HD. This planning further prevents discrepancies in image framing formats, whereby only the image *resolution* is affected. In other words, all 480i content is already configured for wide screen, and all future 720p/1080i content is ready for intercut with 480i WS-SD through the upconversion process.

From the design side, some stations have added utility frame synchronizer and aspect ratio converters that also incorporate upconverters, then added them on the house router for ease of use throughout the facility. This may not be advised for news organizations with multiple edit stations and an intense set of live shots, field stories or production elements that occur throughout the day; as this time-sharing of utility converters proves to be a bottleneck for news and the entire station. Ultimately the station ends up dedicating devices to each edit system, another set to the live remote inbound feeds, and another set to the production switchers just to keep flexibility at a maximum and minimize the congestion when all signals must pass through a single few converters.

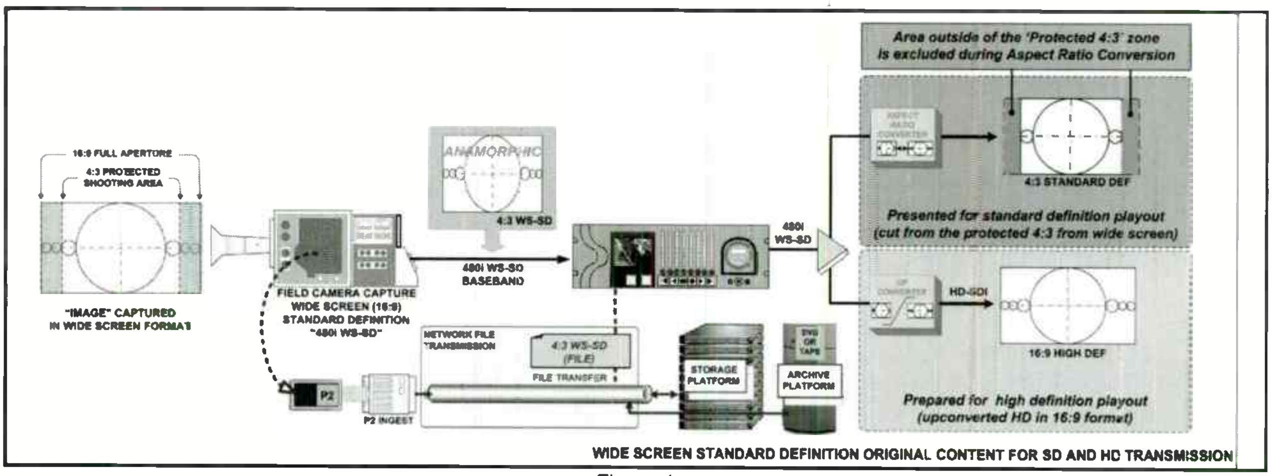


Figure 4

TECHNIQUES FOR EMPLOYING ASPECT RATIO CONVERSION

Finding a solution that meets all the needs of the organization can be daunting if one fails to think through all the options and really concentrate on the workflow before deciding which converters to use and how to implement them. Fortunately broadcast equipment manufacturers now have a large tool set of devices, utilizing sophisticated software based set-up menus and configurable control panels that can achieve most of the needed processes without over complicating the design or risking an operator making a bad choice resulting in permanent image format incompatibility.

Examples of the components and system configurations for meeting these example operations are shown in the next set of diagrams. Figure 4 depicts a wide screen, standard definition work flow with the image frame 'appearance' depicted at various stages in the signal path.

integration into a live program. However, this example facility also has a standard definition sub-channel that features a continuous news program with its program content derived from the same (wide screen) edited stories. To accommodate this form of playout, the material must first be restored to a proper aspect ratio, and then the areas outside the 4:3 space (the side panels) must be removed.

The ARC devices employed for this type of application both stretches the 4:3-anamorphic back to normal shape and then creates a 4:3 'center-cut' version for the sub-channel program stream – a feature that can be programmed and altered by GPI control if desired. If the captured content were a full wide screen image, the viewer would see the same wide screen captured material, but on the sub-channel any non-informative material that falls outside the 4:3 protected shooting space would be removed, i.e., provided the videographer protected that area during shooting; or if the material had

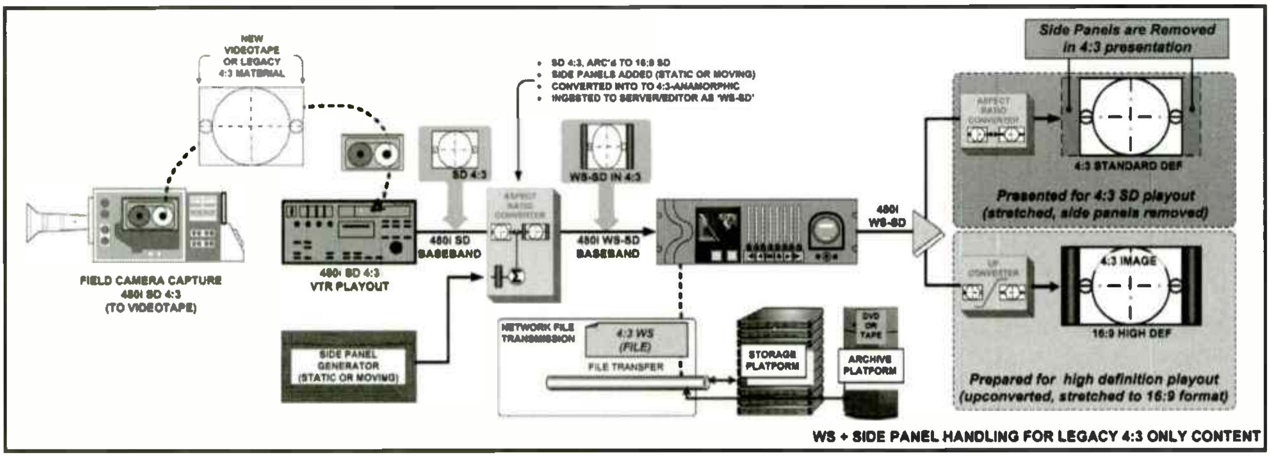


Figure 5

In this diagram, field content is ingested – either direct or as a file – into a server, editor, a storage platform or playout device. All the content captured in this mode is then preserved in its intended format, i.e., WS-SD, and playback may occur according to the needed or intended transmission mode (as SD or HD). This concept further allows an archive to remain in a full wide screen aperture which can be utilized for SD or HD purposes in the future.

Figure 5 shows how native 4:3 (legacy) SD-content, as videotape from the field camera or from the station's library, might be handled. All material is pre-processed and side panels are added, then the content is aspect ratio converted so that it may be edited in a wide screen format, but as SD.

For this example, the playout for the main digital broadcast channel is in high definition, so the playout from the standard definition file (as a WS-SD format ready for HD), must be stretched to 16:9 and then upconverted to high definition. A full raster HD image can then be presented to an HD-production switcher for

been ingested with the side panels, those would be gone.

INTEGRATING WITH VIDEOSERVERS

To meet the needs of this type of workflow, today's broadcast servers now may employ these conversions directly on their hardware platforms, thus eliminating the need for external upconversion and/or ARCs. For older servers and some 'standard definition only' newsroom editing systems, i.e., those legacy devices from two or more years ago, one can achieve the same results, but must utilize external 'glue' devices for the stretch or high def conversions.

The next condition depicts the handling of legacy videotape in a hybrid standard definition and high definition environment. As discussed earlier, this legacy material may come from archived videotape, stringer video, feeds from microwave links or other affiliates that do not use a wide screen format. Again, Figure 5 depicts the workflow that starts with legacy tape transferred to a server or storage platform through converters that format the material for playout on high definition channels (using side panels); or playout on standard definition sub-channels.

Each path maintains the same picture format continuity. Legacy material may be mixed with wide screen or native content during playout, or added to the archive with side panels, maintaining those images in a 4:3 anamorphic (WS-SD) format, and promoting the continued editing process on standard definition platforms.

As previously described, side panels generated in this ARC device may be color or black static panels, an animation delivered from an external player or another graphics file loaded into the ARC. These diagrams may be interpreted for any of these or other alternatives, but each application should be fully validated with the equipment vendor as to needs and requirements before purchase.

The primary importance to remember when applying an of these various methodologies is the maintenance of consistency in the content's appearance, for both aspect ratio converted and full aperture images; for SD-native, for WS-SD native, and for high definition images - allowing for flexibility, regardless of the source material.

AUTOMATING THE PROCESS

For organizations that are retrofitting their facilities with new capabilities and hardware, or for those building a new studio or complete broadcast center – the ability to enable these various requirements from a hands free or minimal human intervention is very important. Many of the products available to the station have the capability of pre-setting configurations that can be recalled by automated news systems, or by remote control panels with menu driven soft button displays. Help is on the way in getting these systems to intelligently process the signals based upon some parameters being inserted into the SDI signal at various touch points in the broadcast systems chains.

The SMPTE standards committees recently completed its two years of work on a process that articulates a means to handle these various image formats and representations – from content creation through the entire processing and air chain. The extension of their work flows nicely into the ATSC A/53-Part 4:2007 work and allows the carriage of messages to the viewing displays through the transmission, reception and decoding processes.

The recent SMPTE 2016-1 standard creates an Active Format Descriptor (AFD) that can be embedded into the SDI transport as metadata, and in turn can signal processing devices to reconfigure there outputs to match the intentions of the image format and aspect ratio accordingly. AFD signaling and handling has already been implemented in most of the major broadcast vendors' modular terminal equipment – especially in ARCs and up/down converters. This set of parameters will soon be implemented in many products such as cameras, distribution fanouts, A-to-D converters, transmission chains, editing systems, and production switchers.

An Engineering Guideline (EG) is currently in the works that will suggest or establish the proper implementation of the rich set of AFD parameters that include how to handle their subsets called 'Pan and Scan', and 'Bar Data'. Further extensions of AFD are also under development that will extend the signaling into the file domain through KLV (key-length-value) coding, which in turn moves these concepts directly into the MXF (Material eXchange Format) domain.

IN THE FUTURE

For the most part this perspective has focused mainly on the transition and implementation of a hybrid SD (480i WS-SD) and HD world at a baseband digital video level. Work continues in the field of migrating these capabilities and workflows into the file-based world – thus potentially allowing for conversion of files to the necessary formats without utilizing secondary or external terminal gear components. While time frames and performance targets aren't yet set (as of this writing), it is inevitable that these capabilities will be extended into an environment that lets the users and creative talents focus on their craft – and not have to worry about image placement, editorial technical process or whether their work will be unknowingly cropped, stretched or distorted – in the production process or in the home.

In conclusion, we can all hope for the day when wide people on flat panel displays disappear forever!

ACKNOWLEDGMENTS:

SMPTE Engineering Committees on Television Systems Technology (S22) and Television Recording & Reproduction Technology (V16)

Manufacturer data, concepts and technology input from Miranda, Evertz, Thomson, Harris/Leitch and others.

TV Technology magazine, various installments of Media Server Technologies, Karl Paulsen, feature columnist.

Best Practices: Using IP in Broadcast TV

JOEL WILHITE
Harmonic Inc.
Sunnyvale, California

ABSTRACT

Broadcasters and their viewers want their television pictures to be perfect. To deliver them using an efficient and flexible architecture, providers look to find and refine the best practices to deliver video services on their IP network. There are several techniques to deploy and deliver transport streams, using newer IP-based technologies. With new technology come the struggles and problems engineers encounter during integration. Since moving transport streams over IP, much has been learned from early implementations. This paper will explore some requirements for integration, describe some problems and offer solutions based on experience.

INTRODUCTION

In the last two years, many of us in the industry have witnessed the shift to more systems being supported on IP network infrastructure. As more and more systems and their interfaces show up on the IP network doorstep looking for a place to live, the IP network infrastructure community is pushed or in some cases pulled into working on or integrating a new protocol to allow a service to survive. At the same time, new revenue and business opportunities arise seemingly out of nowhere that add to the challenge of maintaining the level of service and picture quality while adding new capabilities. Implementation of new business systems while maintaining the current architecture is a challenge, and can be more entertaining than watching a three-ring circus. What was once operated over a twisted pair or coax cable will all of a sudden be moved onto a CAT 6 cable tomorrow.

TELEVISION SYSTEM OVERVIEW

Systems in general are designed to support an application or a set of applications. When the application requires a collection of equipment that uses one or more communication protocols, the system design must consider all of the protocol requirements to create a reliable product. If an application doesn't work, the troubleshooting process involves looking through the layers of the design for clues using a set of tools to look at the different layers. A set of system drawings with a protocol overlay is a simple way to trace out a signal route. When the drawings aren't accurate, the fun begins.

The following are the three drawings one usually needs to find message routes in a system:

Overall System Diagram – This is the 10,000 foot view of everything with very little annotation other than the names of the major blocks, clouds, and links.

Network Topology – This drawing shows all of the network building blocks, demarcation points with some of the routing and VLAN topology. I like to include a simple network protocol stack that shows which of the links are supporting which of the interface standards and protocols.

Subsystem Topology – These views organize a small collection of products connected in a way to provide a signal flow. A typical station will have dozens if not 100 or more pages in the topology to lay out all of the subsystems.

Floor plans and rack elevations are also developed by system integrators so that installers can proceed but aren't usually needed to diagnose a system failure unless you don't know where a piece of equipment is installed. Beyond looking at a set of drawings, other visual aids are available to look at system designs.

When using a cluster of Cisco network switches, a tool I have found particularly useful is the Cisco Network Assistant or CNA for short.

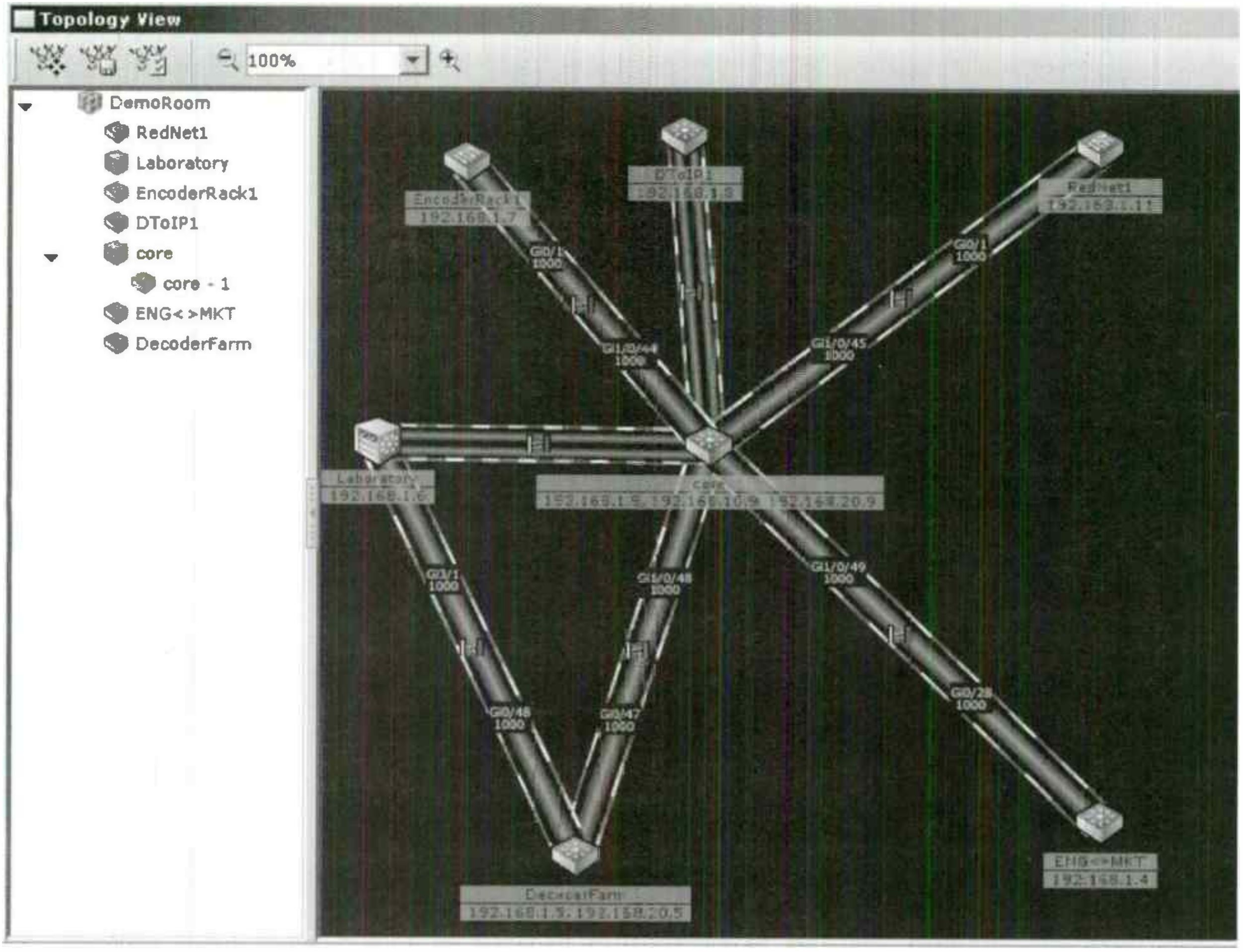


Figure - 1

Though the free version is limited to a small number of routers that can be monitored, the network topology view allows the operator a quick way to see the network health and configuration in a snapshot using their browser like interface. The network diagram drawing should closely match the topology view.

The subsystem topology drawing should give the view of how the signal flow is to be managed into and out of a system. A similar view is also available from Harmonic's NMX Digital Service Manager™ control system.

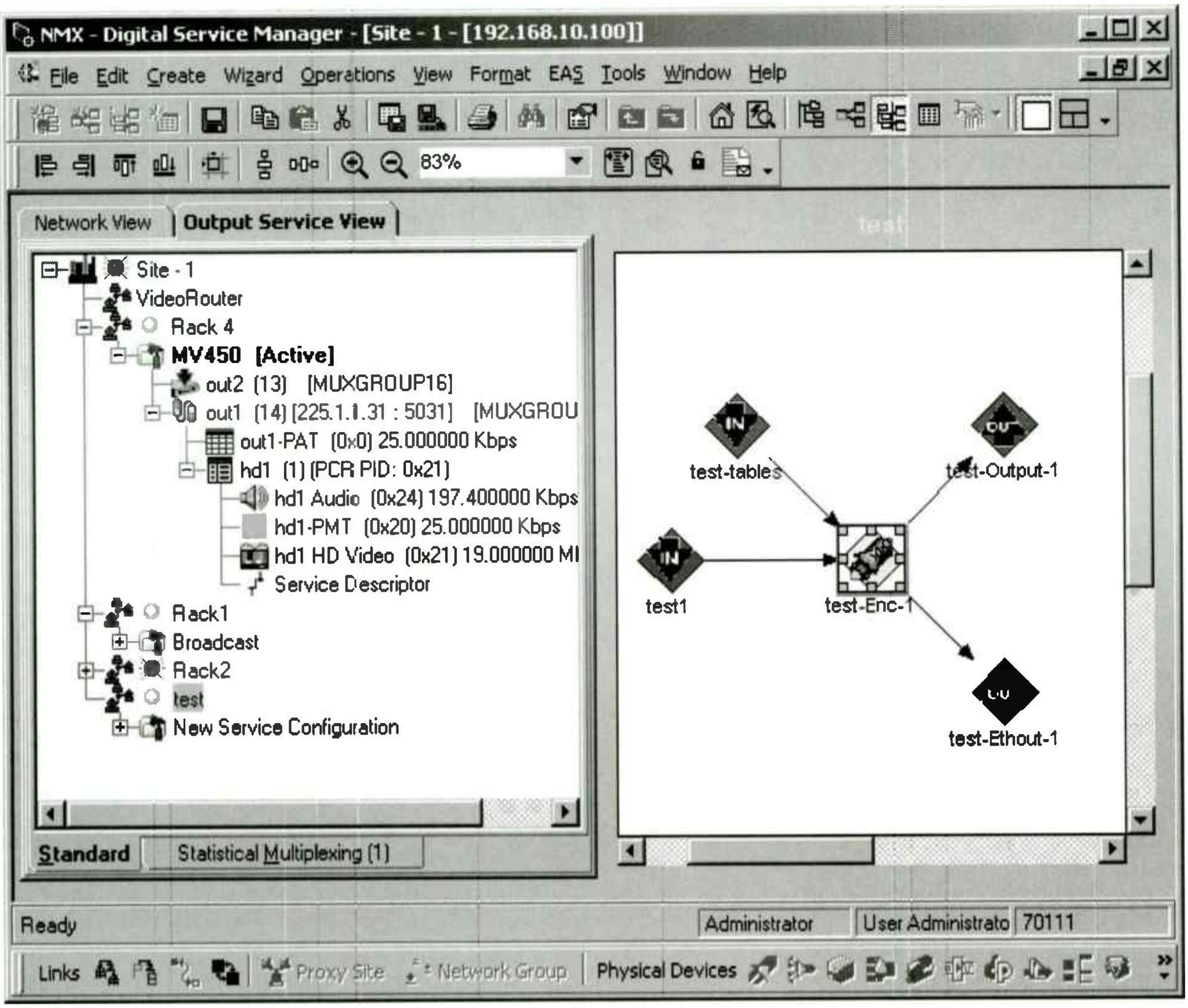


Figure - 2

Much like CNA, the NMX application gives the user a quick view of the compression system's health.

There are a number of other system monitoring tools available to the operator for system overview and monitoring. Knowing how these tools work will provide an efficient overview of the system and allow both engineers and operators to monitor the systems for any problems that may occur.

ETHERNET NETWORKS

Many of the products being managed today over Ethernet use a combination of protocols to start, load code and provision into a particular state. Once the equipment is running, there are yet a few other protocols used to send traffic back and forth to monitor the systems. In almost all cases the equipment has a management interface or console port. Console ports, or serial ports are the back door and support a simple terminal interface. (Though handy in a pinch and necessary for installation, typing exercises for setting up a switch environment while standing in front of a bank of whistling fans is not my idea of fun. Once a Telnet session can be opened at another, more comfortable location, life is much better. One nice thing about console ports, they are there no matter what and as long as the device is behind a locked door, you can feel reasonably assured no unauthorized access will take place.) Can we simplify this a bit?

Ultimately, a majority of the work is performed behind the scenes via Telnet sessions. Telnet has been around since the early dawn of the creation of the Ethernet protocol and is still used today for most day to day operation. Most networks don't require constant attention but when the network spans huge geographical areas, having a good system design for using this tool is a must. The only pitfall is if, during a Telnet session, an operation is sent or a parameter is set in a device that could sever the connection. Once lost, the only way to recover is to provide human intervention.

Beyond Telnet are the more highly evolved interfaces, built into the software of a device which delivers a graphical interface through a web browser. Now ubiquitous, web browsers are gaining more popularity with devices that stand alone, home networking equipment being a good example.

When a system is put together, the control interfaces may include a combination of choices that may include terminal ports, browser sessions and client-server software specially designed to manage a more complex design or configuration.

MANAGEMENT LAN SYSTEMS

Most equipment out of the box comes with a minimum set of instructions to get the device up and running on Ethernet so that communication can begin. Though an Ethernet hub will work in some instances, most of those devices are falling out of favor in networks as they are repeaters.

The definition of a repeater means that data that shows up on any one port is copied to all other ports in the same frame and creates a single collision domain. This concept is not a problem with just a few devices. But when the system size is greater, say a dozen devices, the polling mechanisms used to verify the system's health will be slower and response time when commanding changes will be impacted since the Ethernet protocol will fall back to retries when collisions start to occur. This is only one of several drawbacks when using an Ethernet hub. By contrast, Ethernet switches dominate the market and operate very differently from a hub. Switches allow the packets to be buffered in a number of schemes to avoid collisions altogether. Switches do cost more and add some latency but their advantages far outweigh any of their disadvantages.

Management systems are usually offered on one of a number of computer/server platforms. These systems vary in size, capacity and density. Some reach to the deepest ends of the new computer racks nearly 32" in length including the power cord protrusions and require rather significant rack rail supports. Many come with computer rack slide rails that are a huge time saver if you have the correct rack. But if the rail doesn't fit your rack then another system will need to be devised. The biggest mechanical problem is finding the rails that fit your rack that can support the weight and allow other equipment to be mounted below or above.

Once powered, the units should come up with Ethernet ports enabled. Computer/server models are designed to support two independent ports. Most system designs employ one port to manage the collection of equipment, never to route a message through a router or gateway. The other network interface would be used for remote access. Inside the OS, the user has the ability to allow the server to route between the two ports. With that feature disabled, the Internet will never route to or from the management LAN.

From this point, the network controlling the system requires very little attention beyond the obvious connectivity problems like a bad cable. To support most equipment, devices will require at a minimum an address for a server to receive software. Once the information is put into the network management system, the boot up procedure should proceed. The first pitfall occurs right about here.

NETWORK DETAILS

In the Cisco network gear, we need to set the management network switch ports to allow the equipment to establish links very quickly. In the television business, it's not how often the power goes out, it's how fast can you recover from a power outage. In this case, the default port settings in Cisco gear allow Spanning Tree Protocol (802.1D) and Cisco Discovery Protocol to go out on the port first to find other Cisco gear and network loops. Since we know what is connected on the other end of the wire, we can tell the Cisco switch to let that port be enabled with "portfast". A warning message will tell you that you can break the network if you rewire that port to another switch.

The problem occurs when the switch port is looking for responses to STP messages and the encoder is booting up. During the boot cycle, the encoder performs Power On Self Test (POST) and then sends a BOOTP broadcast packet. After the power is restored, the switch port will still be in STP mode, and the BOOTP packet will never make it to the server. (How is this resolved?)

Another reason hubs are not recommended is because the packets are transferred at the physical layer which silences all other transmitters or else they collide. This is the definition of half duplex operation. In a switched environment, each switch port decides where a packet should be sent as all packets are handled on Layer 2, the data link layer. Layer 2 switching allows full duplex communication.

During a start up or recovery scenario, if a number of devices need to download in a hub environment, each additional unit would cause the server to increase the back off timers for packets going to different IP addresses to avoid packet collisions. Another speed up technique to use on some switches is to set the switch port speed and duplex settings in advance to force a state so that the ports don't need to go through the handshake process at startup.

In a switch environment the addresses of each device will need to be loaded into ARP cache to allow the switch to know where the packets are to be routed. After a power recovery all the ARP tables will be empty, so the first one or two boot messages won't go through because the tables need the first packet going into the port to learn the address of the device connected to the port. The switch can be set to learn these addresses forcefully but it comes with a cost. Each time the switch needs to learn an IP address, it adds traffic to the network.

Another reason for using "portfast" is for faster startup when using Dynamic Host Control Protocol (DHCP). This protocol allows a machine to broadcast a message to request an IP address assignment from a DHCP server mounted on the same network segment. This protocol is used for devices that don't stay connected to a switch port on a permanent basis. But why have DHCP enabled on a network with a range of devices using static assigned addresses? There are times when troubleshooting a subnet when a technician will need an address assigned to a sniffer. When using a laptop to connect to a network, if the laptop didn't run in DHCP mode, the potential for an IP address collision is higher. Having two devices connected on the same network with the same IP address will make for some strange network reliability issues. DHCP helps avoid duplicate IP addresses since the DHCP server maintains a list of active IP addresses.

NETWORK TIME

After all of the initial start up sequences have completed, the next phase of equipment operation includes a seemingly mundane house-keeping chore—adjusting the time clocks using Network Time Protocol (NTP). It is a minor detail but one worth noting. In one case, when the daylight savings time date was adjusted ahead of schedule, it made for some strange behavior. Alarm log time stamps jumped during an outage; a dish pointing program that was being used to track moon position made the dish swing out of position; backup servers that were programmed to send off reports at 1 AM jumped over this time. You can relive Y2K failures when the clocks don't get aligned automatically. It is recommended to set up a machine or two in the network to serve out NTP and have them synchronize to machines outside of the network.

OBJECT MANAGMENT

The next sequence of events in the "turn up" process is the provisioning messages used to verify the state of all the devices after they wake up. During new installation this time is used to tell the device what state to be in. As alluded to earlier, some products use more than one protocol to provide status and control because of the complexity of a control system and the requirement to live on a network with many other devices and variables. Two of the most common protocols are SNMP and XML.

Simple Network Management Protocol (SNMP) is an application layer message protocol designed to allow multiple clients to control or monitor any number of manageable objects on an Ethernet network. SNMP is extensible since the protocol allows the messages and their value ranges to be defined by the manufacturer. The messages and variables are defined in the code when the runtime OS is compiled. All of the operation variables of a device are configured and stored in a Management Information Base (MIB), a database by any other name. The extensibility of SNMP provides for some fun and games when two machines with different MIBs try to talk with each other.

In a scenario where a piece of equipment is operating in a suspicious manner, the MIB in the device being controlled can be read from using a commonly available MIB browser. The browser tool provides a crude back door to all of the variables currently loaded on a device. Looking in the opposite direction, the tool for catching what an NMS system is sending is an IP network sniffer. Any difference between the captured messages and the values actually present in the equipment will be a clue something is wrong.

Each object in SNMP is given an ID that is called an "OID". Each OID will have a string of data behind it, sometimes it is all nulls.

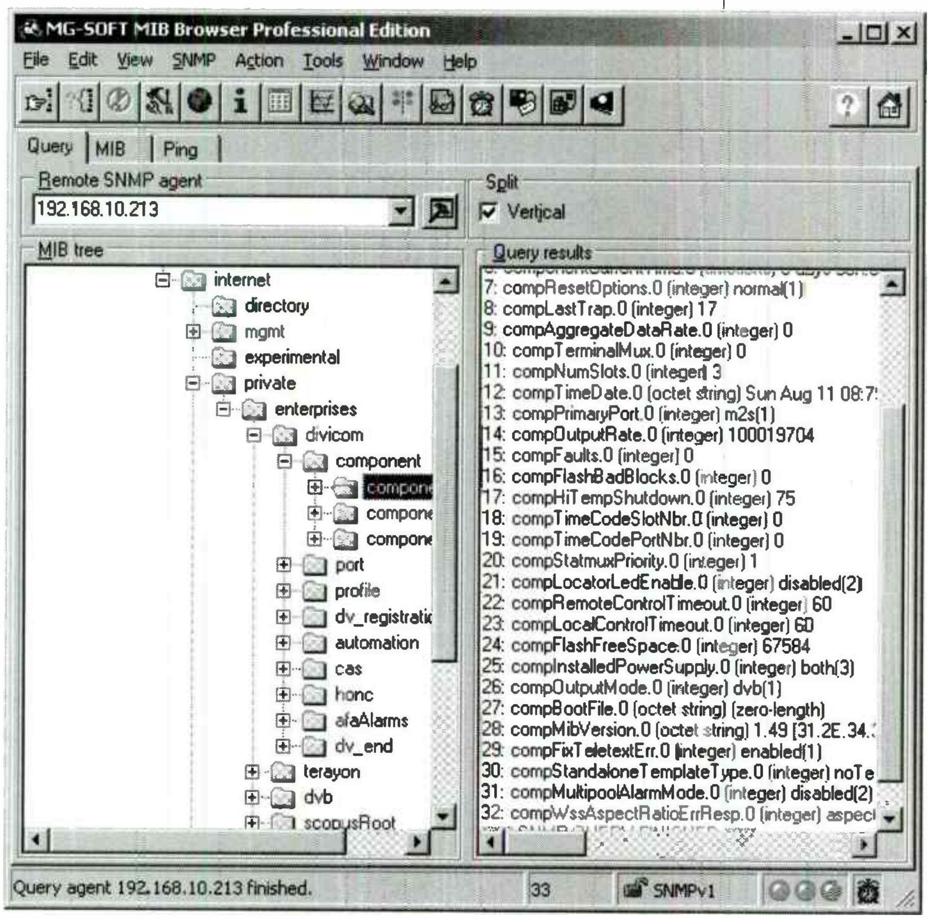


Figure - 3

The test method for testing SNMP reliability is to SET a series of OIDs with some set of variables and then "walk" the MIB with a MIB browser to verify the settings were received by issuing a GET for the same OIDs. As simple as that may sound, SNMP SET and GET messages operate at the application layer (layer 4). Some systems send SNMP over TCP and set up and tear down sockets to open up a pipe, otherwise (and most commonly) they send their messages over UDP. The few trade-offs between socket and socket-less based (TCP vs. UDP) communication are easy to understand but beyond the scope of this paper. Diagnosing SNMP over UDP communication errors can get tedious without some training using IP sniffer tools.

The technique for catching a specific message from an NMS can be a rather daunting task when the system may comprise dozens of manageable objects using multiple protocols. The capture tool of choice for monitoring IP network traffic is WireShark which is freeware. Inside WireShark, the application allows the user to limit or filter traffic to exclude packets from all addresses except the two ports or addresses being analyzed. Using filters will eliminate scanning numerous lines of traffic capture for the two messages you are looking for.

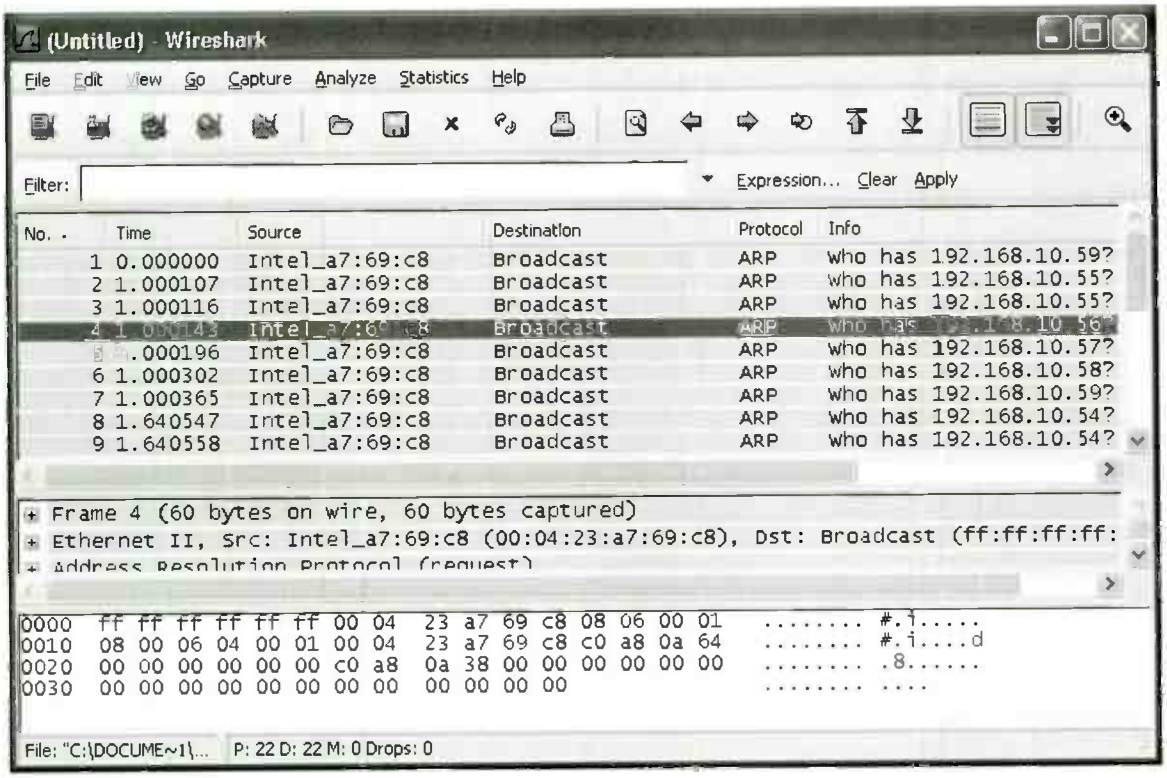


Figure - 4

In typical captures, lines of messages are time stamped along with other relevant data related to the protocols and their structures. Since most control systems are transactional by nature, a series of commands is usually followed by some type of response. The lack of response or the message reply may hold the clues to why an application or a system maybe failing.

Beyond diagnosing control message protocols, network sniffers fill a much more important role in television station designs as more stations are outfitted with the latest in compression technology. Digital television over IP is here and here to stay. This is not the formal announcement of the end of ASI but rather, in addition to ASI, IP will play a more significant role in supporting delivery of the main signal. The consequence of this is the need to be better prepared to understand the mechanisms and how to diagnose an issue if or when they occur.

MEDIA NETWORKS

Video over IP has been around for years. The most common technique is to take seven MPEG transport stream packets and encapsulate them into a single UDP packet and wrap it into an IP packet. UDP is used over TCP because TCP is implicitly unicast only. There are numerous other reasons why TCP isn't used but it is also not impossible. The diagram shows the four layers by color.

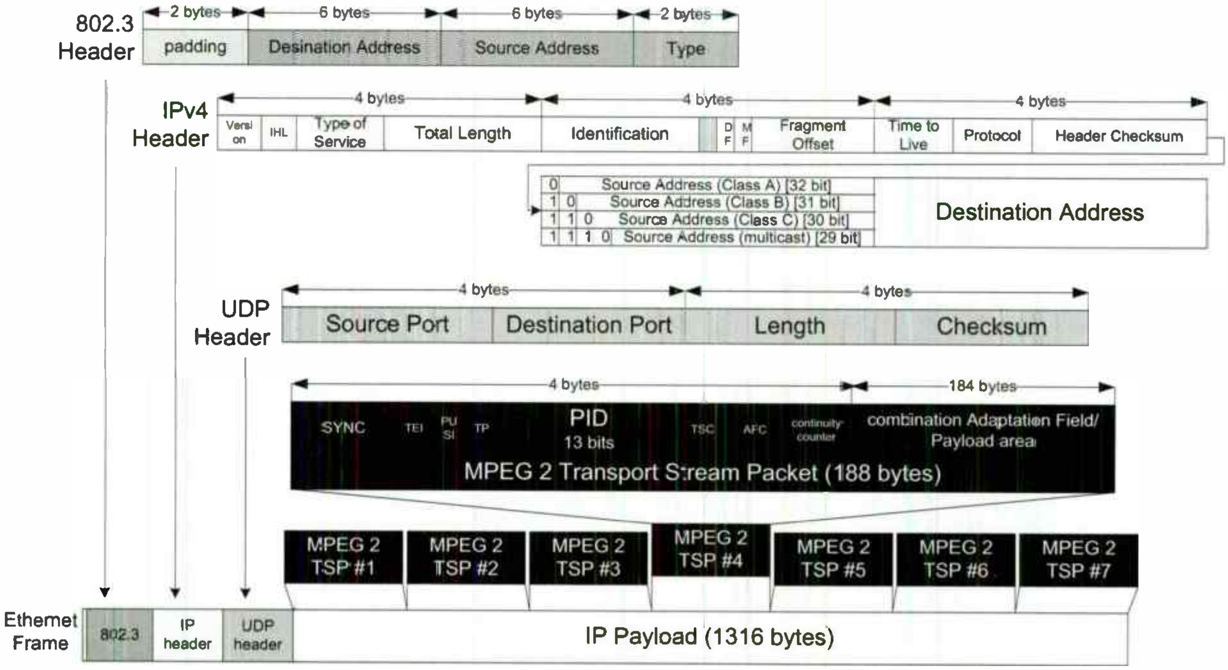


Figure - 5

The UDP layer contains the designation of the source and destination ports. The IP address assignment carries the designation for a unicast or multicast address. In a unicast environment, both source and destination addresses are needed to route packets through a network, and this is how most control networks are organized. When using a multicast address scheme, the destination address is the relevant information to tell a switch or router to know how to replicate a packet. The most common protocol used to start a multicast route is the Internet Group Management Protocol (IGMP) protocol. IGMP is used by an IP host to signal a switch or router the multicast address they wish to 'join". Version 2 of IGMP incorporates the ability to "leave" a session. Another streaming protocol in the same family as IGMP is Real Time Protocol (RTP).

RTP offers additional services like payload identification, sequence numbering, time stamping and delivery monitoring. The stream can be transported more reliably, but only if the far end receiver application knows how to process the RTP signal. Even when RTP is used, networks can still reorder the delivery of a packet. It is up to the downstream device to make any corrections like de-jitter or packet reorder based on the RTP header information.

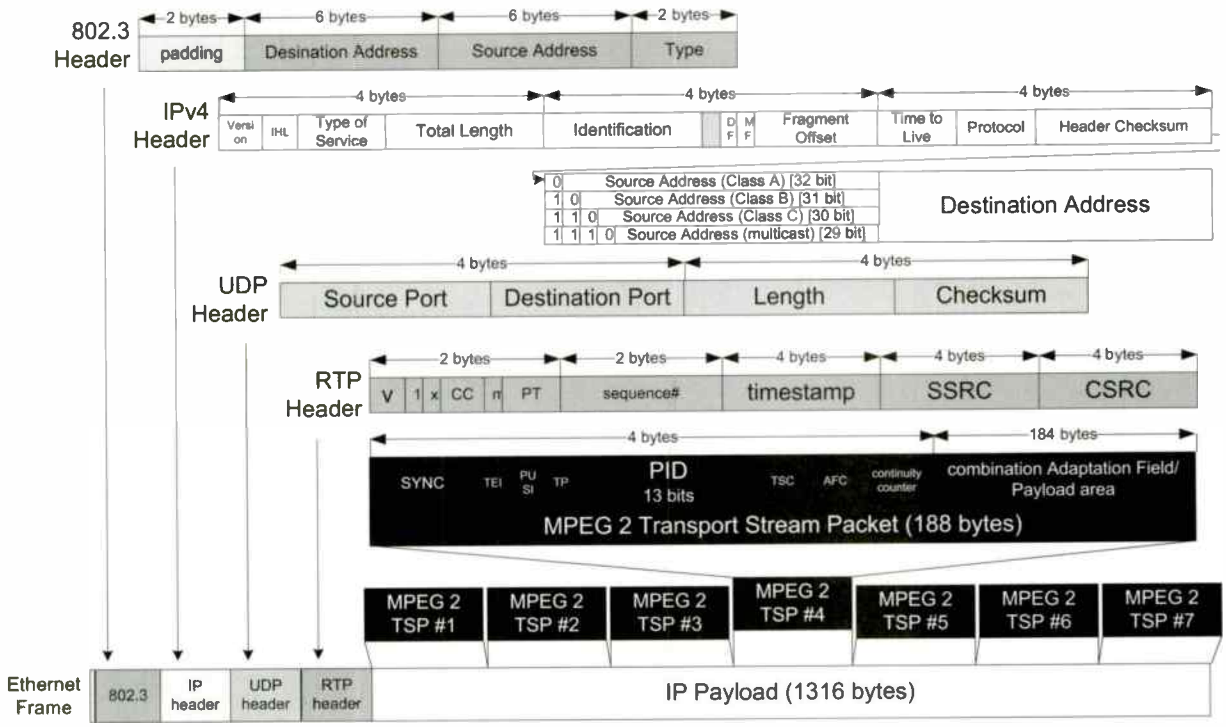


Figure - 6

RTP adds a layer of protection to a steam of compressed data by leveraging the headers. As IP network switching equipment becomes more sophisticated, even those devices will help correct or maintain more accurate delivery of the traffic.

Multicast packets are synonymous with broadcasting, but in Ethernet networks, broadcast packets are defined specifically to route to every port in a collision domain or subnet. For this reason, broadcast packets are used sparingly since every device in a subnet has to listen. The differentiation between broadcast packets and multicast packets are the value of the destination address, a broadcast packet destination address is 255.255.255.255. Multicast packets use Class D destination addresses which were defined by IANA to include a clever technique to allow any Ethernet receiver the ability to see multicast addresses like a broadcast. The same technique helps speed a packet through a network can also flood Ethernet switches if they aren't smart enough to filter packets correctly.

STATISTICAL MULTIPLEXING OVER IP

In a typical closed loop statmux system, a multiplexer or a computer acts like a calculator to divide the bandwidth allowed from a pool of bits to be divided between a number of encoders based on the complexity message being reported by the encoder. Periodically each encoder sends a message corresponding to the picture complexity based on a number of parameters the encoder monitors to maintain picture quality and compliance. At the same time, the encoder in turn receives a bit rate allocation and follows the messages from the multiplexer, thereby closing the loop. This two way communication is easily distributable over an IP connection using fairly common off-the-shelf Ethernet switches. But not all Ethernet switches are created equal.

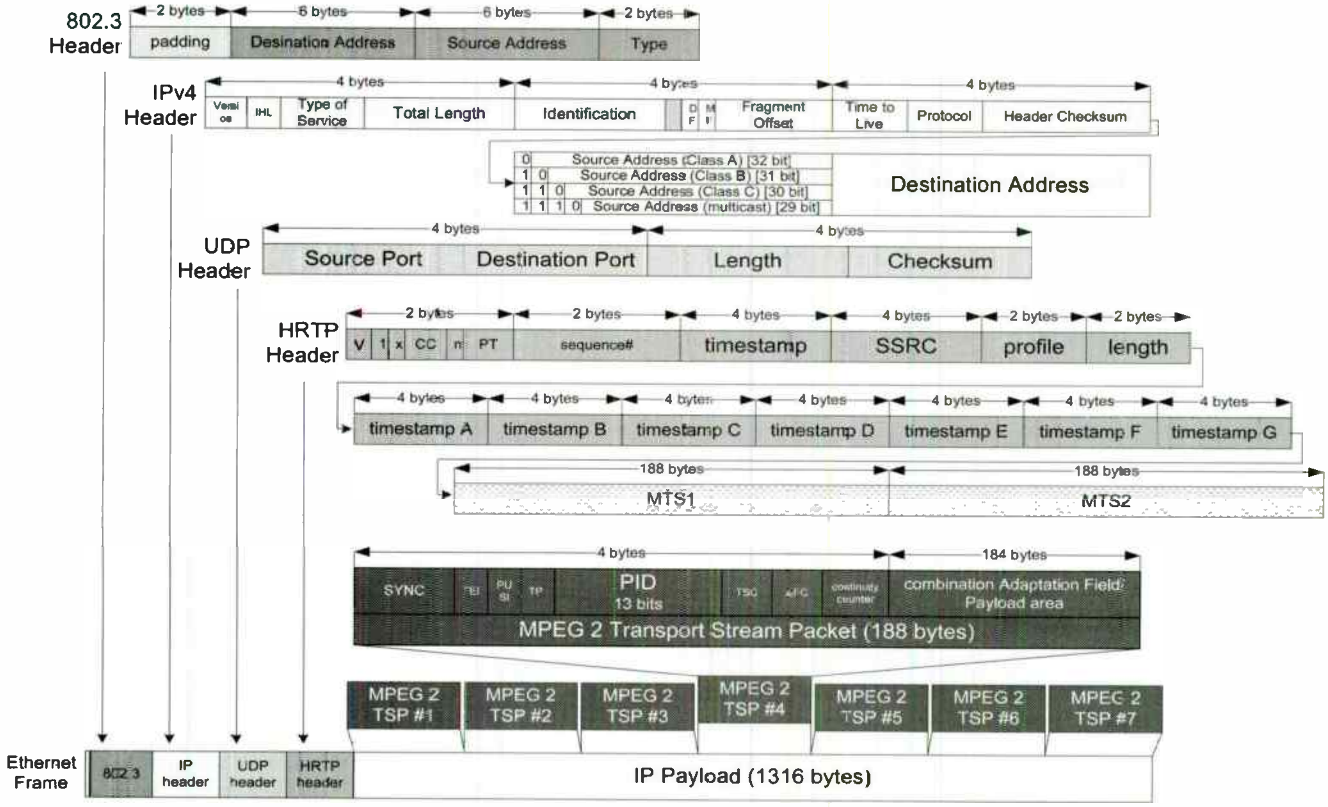


Figure - 7

Over the years we have found that some of the earlier switches had packet loss problems severe enough to corrupt the accurate delivery of the messages going back and forth between the encoders and the remultiplexer. And, considering the amount of traffic generated on the VLAN supporting Harmonic's DiviTrackIP™ statistical multiplexing, we always separate the encoder to remux traffic into a separate VLAN with "IGMP snooping" enabled. This specific switch command tells the switch to watch the IGMP messages as they go by to learn the MAC address of the sending device to build the ARP table it needs to build the routing table. This technique prevents the multicast traffic from flooding the LAN by restricting the packets from a particular location to be added only to certain ports and conserves switching bandwidth across the switch.

IP TRUNK OVER A VLAN

A nailed up trunk allows multiple VLANs to be tunneled through a single connection between two points by tagging the packets. The 802.1q protocol is widely supported across manufacturers and allows the mix of various types of traffic to be comingled in the wire without fear of one type of traffic impacting the quality of service of the others. This is not to say that accidents can't happen since a network loop with spanning tree disabled will bring a network to its knees rather quickly.

FLOODING A NETWORK

Network flooding is a common occurrence and can come in many forms—all of which are generally not pleasant, such as a denial of service attack. When transporting bulk video feeds across a network, the traffic passing through a switch will look exactly like a network flood, so you could say, in some cases network flooding is a good thing. But when things get out of hand, the feeling of helplessness is likely as you frantically look around the network for a culprit. Although spanning Tree Protocol (STP) is used to stop network loops between two or more switches supporting more than one connection, simply avoiding network loops altogether is safer.

CONCLUSION

One sure fire way to avoid influences from an outside network device is to use only trusted devices to communicate on a private LAN. When traversing network segments the routing protocol selection can become an academic research project. Let me be a little more reassuring by providing a couple of specific suggestions:

- Use PIM sparse per VLAN and PIM SSM globally (do not use PIM sparse dense!)
- Look closely at the OSPF routing protocol for routing between manufacturers

No two networks are identical so there will be individual choices to make regarding a specific architecture. At some point, the IT network design will grow to where a network specialist is required to apply their services.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my colleagues at Harmonic Inc., who gave me the inspiration to complete this paper.

Moore Macauley
Principal Engineer
Integrations Group

Bill Kaiser
Sr. Systems Engineer
Integrations Group

Rick Engelbrecht
Staff Technical Support Engineer
CSD Technical Service

Curtis Leo
Sr. Technical Support Engineer
CSD Technical Service

REFERENCES

Some of the material or information presented was provided from the following web sites.

www.cisco.com

www.technet2.microsoft.com

Live Integrated Production Systems Streamline Live Workflow

Ken Swanton
Broadcast Pix
Billerica, Massachusetts

ABSTRACT

Twenty years ago post production and live production were both done with expensive suites of individual components. Since then, computer-based integrated systems have revolutionized the post production industry, but live production is still usually done in control rooms built up from individual components, including a switcher, character generator (CG), clip server, still store, monitoring, etc. Recently, live integrated production systems (LIPS) have begun to be used in many broadcasting, webcasting and other live studios. They provide dramatic improvements in workflow, cost and staffing when compared to conventional control rooms. Moreover, they produce compelling live video that is equal to or better than the more cumbersome traditional approach. This paper describes the contemporary LIPS, its workflow advantages, and three typical installations.

HISTORICAL PERSPECTIVE

First Post Production System Emerged

Live production control rooms have been a part of television from its inception when most shows were live. Twenty years ago, there were other large rooms nearby for post production. Both types of rooms contained a collection of many boxes wired together which were expensive to buy and staff.

Then the post production room began to change radically with the invention of the non-linear editor (NLE). Computer technology dramatically streamlined post production, and editing became much more affordable. Over time a large room filled with equipment was no longer needed, just a workstation.

But the live production control room has continued as a collection of boxes wired together, which has so many control panels, keyboards and mice that a team of professionals is needed to run it. The computerization of live production has lagged behind post production because the application is more demanding. In live, many video and key layers must be combined in real time, and there are no re-dos as the audience is often watching the show as it's created. Until recently, computers have not been powerful or robust enough to transform the live environment.

Live Automation Systems Emerge

Switchers began to reach out and control other devices by adding serial "machine" control to trigger tape decks, CGs, etc. Over time complex software has been developed to do a better job of this, but these automation systems are expensive because each device still needs to be purchased separately. In addition, feedback is limited, most are driven by a mouse rather than a fast action control panel, and the sheer number of devices and brands to control, as well as their various protocols and software versions, often makes these systems difficult to maintain and operate reliably.

All-in-One Systems Emerge

The first integrated production systems to emerge that contained all the devices needed for live video production were low cost all-in-one systems. They also included audio mixing, as well as editing systems to do post production. They introduced live video production to smaller schools, churches and other studios that could not previously afford it.

However, these all-in-one systems had limitations which kept the traditional live control room the preferred choice for most professionals. Several major requirements had yet to be satisfied. First, while the all-in-one systems were reliable enough for simple, non-critical productions, they would occasionally lock-up on-air and go black. Second, they also had a limited number of inputs and key layers. And third, they were mouse driven, whereas live production is so fast paced that most pros prefer a control panel.

LIVE INTEGRATED PRODUCTION SYSTEMS

Five years ago, Broadcast Pix developed a Live Integrated Production System (LIPS) that just did live video production and addressed the limitations of the automation and the all-in-one systems. This LIPS combined traditional live broadcasting technology with computer and networking technology and is now used in over fifty countries for live broadcasting, web-casting, cell-phone casting, cable TV, stadiums, mobile, corporate, religious and government productions.

The LIPS starts by combining the six core components of the live video control room: video switcher, character generator (CG), clip store, still store, monitoring and format conversion. In the conventional control room, each of these is a separate box.

In a LISP, a workstation is the foundation, and its disk drives hold the clips and stills, while its CPU hosts the CG software for creating titles and graphics (Figure 1). In most all-in-one systems, the CPU also hosts the switcher, but switching is a very compute intensive application, because it combines many layers of video in real-time, which can lead to CPU lock-ups. So, in the LIPS, the switching is done on a separate SDI video processing board that plugs into a PCI slot, and then cameras and other video sources plug into the back of this board. Another board handles analog format conversion. The high-speed PCI bus links the switcher to the CPU and disk bays. Since all signals are digitized, a multi-view display is created to show program, preview, sources and much other data in moving windows. The workstation environment also provides file connections with the rest of the studio using high speed networks, DVD, CD and USB.

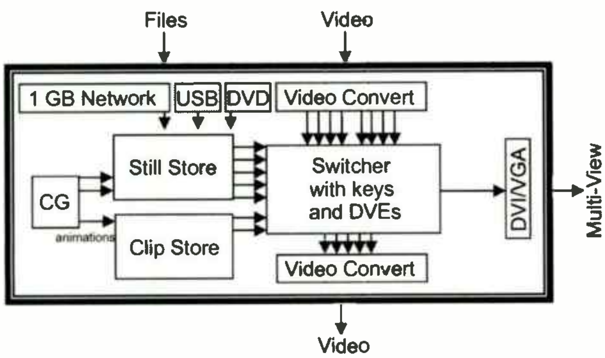


Figure 1 - LIPS workstation

Getting all the studio components into one box is only step one in creating a streamlined LIPS workflow. The other two big ingredients are a file-based architecture, and providing integrated control of all devices to enable the user to keep up with the fast-action inherent in live production.

File-Based Architecture

A file-based architecture has been one of the keys to workflow improvements in post production and throughout the contemporary studio, but it has only recently begun to permeate live production. A file is a graphic, clip or animation that is encapsulated in a computer file, such as a .tga graphic or a .qt QuickTime clip. It includes the content and the attributes of the file. Files can come from a camera or be created in editing systems, and then sent over a network and stored in a studio's servers or directly in a LIPS.

File-Based Streamlines Live Control

Having the file names inside the LIPS enables the names to be visible to the operator so content can be located quickly in the heat of a live production. The file names are shown on displays built into push-buttons on the control panel. When you look at the program row each clip store, still store, CG and logo button displays which file is loaded. In the example in Figure 2, the first channel of clip store has the River clip ready to go, the first channel of CG has the Mens Doubles lower third ready, while the still store has a still called Split on-air. Push-buttons for keyers also display what file is loaded on each key. All these displays change dynamically as different files are loaded, and the names also show on the multi-view monitor. The operator can tell at a glance where everything is.

Figure 2 - Control panel push-buttons show file-names

Integrated Control

With so many video production devices integrated together, the LIPS needs to provide a very fast way to control each device so the operator can capture the spontaneity of a live show. All-in-One and automation systems are usually controlled with a mouse. Mice are great in post production where there is much more time and you can re-do things if needed. But in a live production, where you often never know what the talent is going to do next, a mouse can make it hard to keep up with the action. Many all-in-one systems have retrofitted a panel to meet this need, but these systems were not originally designed for a panel and so their panels control only a portion of the action.

For the best workflow, the LIPS should be designed from inception to be controlled from a panel. In the Broadcast Pix LIPS, the panel looks like a conventional switcher, but it has two portions. One is a conventional switcher panel for control of sources, keys and transitions, while the other portion controls devices. The device control bank can be instantly assigned to any clip store, still store, CG, etc. These device controls make extensive use of buttons that displays file names, so it instantly changes for each device.

Feedback on the Integrated Multi-View

The days of the conventional "monitor wall" are quickly passing. Most new control rooms now install large plasma or LCD monitors on which a multi-view system displays an array of monitors. On a LIPS, rather than being a separate box, this multi-view system is just another output. Both kinds of multi-views show monitors for program, preview and all sources, but because of its better integration, the LIPS multi-view provides far more feedback and data to the operator. The additional information includes the file names and attributes, the contents of each keyer, thumbnail libraries for clips and graphics, key settings, clip counters and clocks. In live production where seconds are precious, more information enables richer productions to be created by fewer people.

For example, as shown in Figure 3, the library for the first channel of clip store shows thumbnails of each clip, which correspond to selector buttons on the panel. The file-name is under each thumbnail and the on-air one is tallied red. The clip attributes (or metadata) are also shown. They indicate that the clip has been set to loop, set to auto-start upon transition to air, as well as set to auto-stop and rewind when it comes off air.

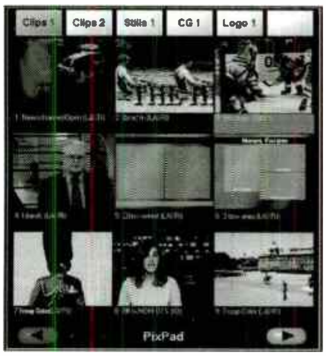


Figure 3 - Clip library portion of multi-view

LIPS WORKFLOW FOR SOLO OPERATION

Conventional control rooms are designed for multiple operators, including operators for the switcher, graphics and clips, plus a director to keep them all together. While the most important live shows are still fully staffed, economics often demand that many of the productions created in these control rooms be done with less staff, and often with only one operator. This results in “dumbing down” or lowering the production values of the resulting video, with fewer layers of graphics, fewer clips, fewer effects, fewer cameras, etc.

With a LIPS, it s not necessary to sacrifice rich production values to create live video with one operator. The LIPS control panel creates such a powerful single point of control that a solo operator can create compelling live video that looks like it came from a team. A separate graphics operator is usually not needed, because whenever the show wanders off its graphics play-list, the solo operator can quickly find the needed graphic. The clip operator is also not needed as clips can be found quickly and can be set to auto-start at the instant they transition to air. The first thing which a new LIPS owner usually notices is that solo operators typically create more compelling video, with more graphics, animations and clips.

Streamline Workflow from Edits Bays to Live

The next thing a new LIPS owner usually notices is how easy it is to get content into it from an editing system. Editing systems have become ubiquitous, and almost every live studio has one or more of them nearby creating clips, graphics and animations that need to find their way into the live production.

In conventional control rooms, the switcher can not accept files, but instead is hooked up to specialized clip servers and still stores which convert the files into streams of video and keys that the switcher can input. These wires consume many inputs in the switcher, and these servers and stores have their own keyboards and mice, which then require more operators. Since the files enter the switcher as video streams, all their metadata has been stripped off, including the name of the file and any attributes, such as when the file was created, mark-in points, settings for auto-start, etc.

With a LIPS, the edit bays simply send files directly into the LIPS’s hard drives, and the files retain their names and attributes. They reside there until they are played out of the LIPS.

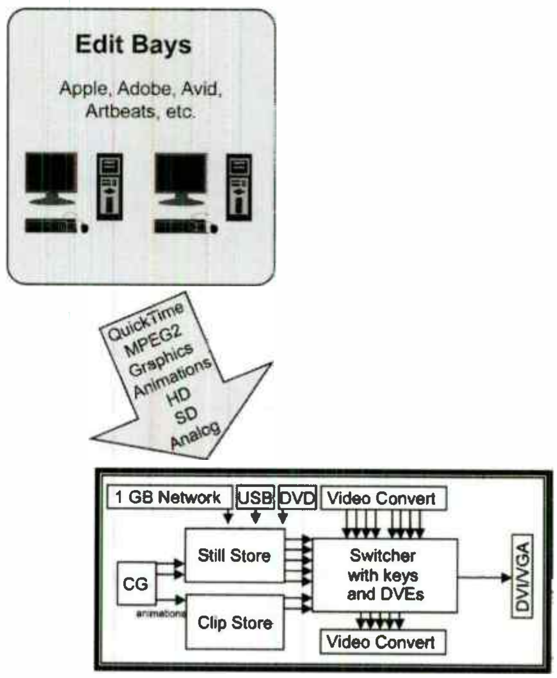


Figure 4 – Workflow from Edit Bays

In the file-based world of the LIPS, it is so much easier and faster to get content into the live show. Files can be created on any editing system, because the LIPS can accept and convert a huge variety of formats (Figure 4). For example, Clips can be QuickTime or MPEG2 files. Graphics can be .tga, .bmp, .gif, .bmp or .png files. Animations can be created on Apple, Adobe, Avid, Motion or most any other system. The files are then sent over the high speed gigabit Ethernet network into a Terabyte of storage inside the LIPS. The file names can automatically appear on the panel's buttons and on the multi-view, along with their thumbnails. Files can even be imported while a show is on air for immediate use.

Extending Solo Control to Audio, Cameras,...

So far, we have discussed how the solo operator can control the devices that reside inside the LIPS, and use files networked into the LIPS from edit bays. But control rooms have other components that often require additional operators, including audio mixers and cameras. Since the LIPS is network and computer savvy, it can reach out and control these with modern network protocols that go far beyond what is possible on the relatively antiquated protocols usually used by conventional switchers, like "machine control".

For example, a LIPS can use MIDI to control an audio mixer. MIDI is ubiquitous on all but the smallest audio mixers and enables a level of control from a switcher that is unprecedented. As a video source comes to air MIDI can trigger commands on the mixer, including volume, fade, pan, effects like reverb, and even an entire mixer scene memory.

Robotic cameras have become very affordable and are a great complement to a LIPS. They can be completely controlled from its panel, including preset positions, tilt, pan, zoom, focus, iris, and CCU functions such as white balance. This eliminates the need for camera operators. It's also so easy that applications which previously used only fixed cameras often add a few robotic ones for specialty shots. In addition to mixers and cameras, LIPS can also control video servers and routers.

File-Based Integrated Memories

One of the milestones in the development of the switcher was the invention of memories of panel snapshots and sequences. Switcher memories have always been heavily used in live production because they enable pretty combinations of layers, squeezed video elements (DVE boxes), and transitions to be carefully prepared before going to air, and then fired off flawlessly in the middle of a live production.

The LIPS makes possible much more powerful memories which recall not just switcher settings but also the exact files to be loaded from the clips stores and still stores, and their attributes, as well as audio settings, camera moves, etc. For example, in a typical news application, a conventional switcher can call up a memory with the news anchor on camera, a CG to hold their name, another CG to hold a crawl below, a logo, and a "DVE box" over their shoulder to hold a squeezed clip. However, the conventional switcher memory cannot tell which graphics to put into the title, crawl and logo or which clip to use. As a result, the operators have to be very careful to get them all loaded into each active source or else the wrong ones will be brought to air when the memory is recalled.

In contrast to this simple conventional switcher memory, the LIPS memory can recall all of the switcher settings, plus the exact file names of the graphics and clips that go with it. The file metadata enables the LIPS to even recall the attributes of each file, including whether the clip and crawl should auto-start when taken to air, any mark-in or mark-out points, the speed of the crawl, and whether the crawl starts on screen or off screen. This deep control of graphics and clips gives the solo operator the freedom to produce high quality content with maximum confidence.

LIPS sequence memories are even more powerful, as they can add in action from all devices. For example, in a four person interview show, shot with two robotic cameras, you can have four sequence memories for each camera, one for each person. If camera 1 has Mary on-air, and you want to bring up John, just push the "John on Camera 2" button, and the following occurs: camera 2 repositions to John on preview, comes to air, John's CG title is added on-air for 5 seconds, and then his title fades off. It's so simple that a novice can run it, and yet it looks very professional.

Switchers are known for doing great transitions. The latest state of the art is "alpha wipes", where a keyed logo is animated between shots, sometimes with a sound effect. This is yet another expensive box in a conventional live studio, while in a LIPS it's just a sequence memory integrating the switcher and an animation file. While some conventional switchers now have some animation capability built-in, these are much limited than in a LIPS as they typically only handle one short animation, take a while to change animations, have no file names or metadata, and no audio.

Streamlining Video Format Conversion

Another control room chore that is streamlined by a LIPS is the sometimes overlooked task of video format conversion. Since a switcher combines many video sources, it must deal with the reality that most studios want to use a variety of video formats. And yet most switchers make you choose one format, and then purchase separate converters to convert all the other formats to that chosen one. For example, you can switch in SD-SDI, but you must buy converters for analog signals. HD switchers make you pick 1080i or 720p, and not use a mixture of HD formats. Most conventional switchers are similarly confining for aspect ratios, as you must pick 16:9 or 4:3, and distort all video that is of the other aspect ratio to the point were people wind up looking too fat or too skinny. Moreover on most conventional switchers, you cannot even attach asynchronous sources, like a Blu Ray HD DVD player, without adding time base correctors.

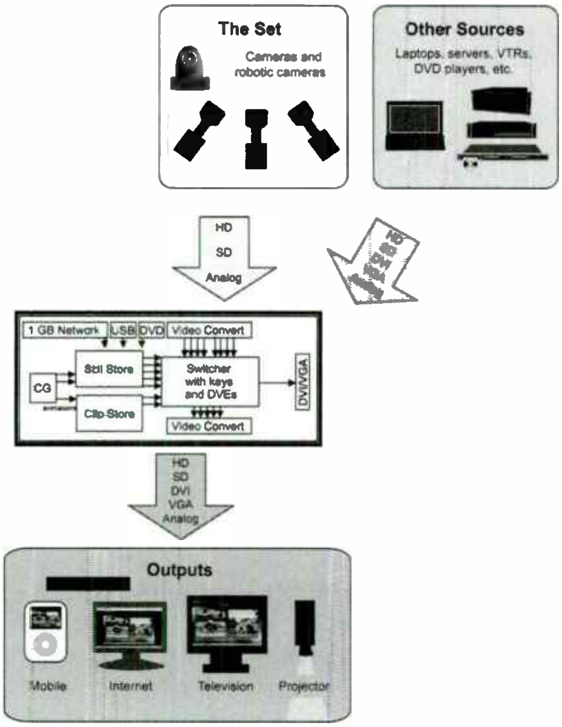


Figure 5 – LIPS has many video formats in and out

The LIPS can solve all of these format conversion challenges, which can shorten installation from weeks to minutes. For example, the Broadcast Pix LIPS can mix HD-SDI inputs in both 1080p and 720p, plus DVI, VGA, or HD analog component, plus SD-SDI, plus analog composite, Y/C and component (Figure 5). The sources can be a mixture of genlocked and asynchronous. Both 16:9 and 4:3 content can be used interchangeably, without distorting the native aspect ratio of each camera, clip and graphic.

The LIPS can also handle the output conversion chores. For example, the LIPS can simultaneously output HD-SDI 720p and 1080p, plus SD-SDI, analog composite, Y/C and component, plus DVI or VGA. And 16:9 and 4:3 shows can be output simultaneously, without making people too fat or thin.

LIPS WORKFLOW FOR TEAM OPEATION

While a LIPS enables a solo operator to create live video with high production values, some productions still require a team of operators because they are so fast paced. The LIPS's network and computer technology enables it to gracefully expand to a team.

For example, in sports the action is so quick, and there can be so many graphics, that a separate graphics operator can be needed. On a LIPS, the graphics operator can be added by simply plugging in a second VGA monitor and handing him or her the keyboard and mouse. The switcher operator works from the panel and uses the first monitor. The graphics operator can use the built-in CG to create new CG pages, crawls, animated titles, etc. When a graphic in the LIPS's libraries needs to be edited, such as creating a lower third for a surprise guest, the graphic's thumbnail on the multi-view can be clicked, and it opens for a quick edit.

For the most graphic-intensive productions, the on-air graphics can be directly linked to a database (Figure 6). Then, whenever the database is updated, the graphic changes. For example, in a sports application, you can have a CG page for a player, which contains a head-shot and various player statistics from the data base. When the database is pointed to a different player's name, these graphic and data fields on the CG page change. It all happens in an instant, so the CG can even be on air when it changes, or when a score is updated. In a conventional studio, this requires an expensive separate box, while in a LIPS it is just more software.

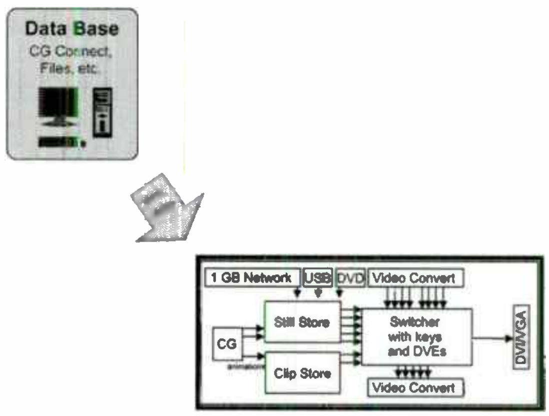


Figure 6– LIPS ties to an external database

Thanks to the networkability of the LIPS, this software does not need to reside on the LIPS, but can be on any computer networked to it. CG connect software is also handy for slower paced applications when there is a large database of names, as it saves many hours of re-keying them all into the CG.

Workflow for larger Teams

While productions in a conventional control room usually require a team of operators to control the many panels, keyboards and mice, the LIPS is usually run by one operator, and sometimes two. However, in some large live productions, like professional sports and large events, a bigger team can be used.

As a network appliance, the LIPS can gracefully scale up to meet the needs of large and even distributed teams. It can be fitted with quad-monitor support for up to four VGA monitors to spread the multi-view out. More significantly, you can add many control panels, either physical ones, or soft panels. Soft panels are displayed on a computer monitor, and look like a panel but are operated with a mouse. Soft panels can open in a web browser, and so can be on any computer, including the LIPS workstation or any computer networked to it. Panels can be locked together or independent, so for example, one operator can control the switcher, another one a channel of the CG, another the clip store, etc.

The network architecture of the LIPS means that the control panels can be anywhere: in the studio, the back room, another building, or even across the continent. For example, a control panel at headquarters can control a remote studio so that an experienced operator does not have to be onsite.

A LIPS can also seamlessly integrate with a large video router, to greatly increase I/O and redundancy. And two LIPS's can share one large router to save money, and greatly simplify wiring and installation.

FAIL-SAFE, NON-STOP, ON-AIR

A problem on most all-in-one systems is not being robust enough for demanding live on-air applications. A LIPS can have many redundancies built-in, so as to never go off-air in the middle of a show.

Workstations have become more reliable each year, and in a dedicated environment like a LIPS, they are even more reliable. However, no matter how rare, you need to protect yourself if a workstation should ever fail in the middle of a live show. It only has to happen once to be a big problem. As mentioned earlier, a LIPS can do its switching on a video processing board inside the workstation. With the "heavy lifting" occurring on this board, the workstation's CPU is not stressed. And in a well designed board, if a failure should still occur, even unplugging the workstation, this board can automatically switch over to a favorite camera, to stay on the air and never go black.

The LIPS can provide still more redundancy by integrating a router. All SDI cameras are connected to the router and then fed from it into the LIPS, and then program and preview video are fed back through the router. Now a separate router control panel can be used to select cameras inside the router, and by-pass the LISP workstation if it should ever fail. On the Broadcast Pix LISP, the control panel has this capability built-in, as the panel has redundant connections to the router, both Ethernet and serial (Figure 7). If a failure is detected in the workstation, it is bypassed, and the panel controls the router via this serial connection, enabling cameras to still be previewed and taken to air. When the workstation comes back on-line the CG, graphics and clips return.

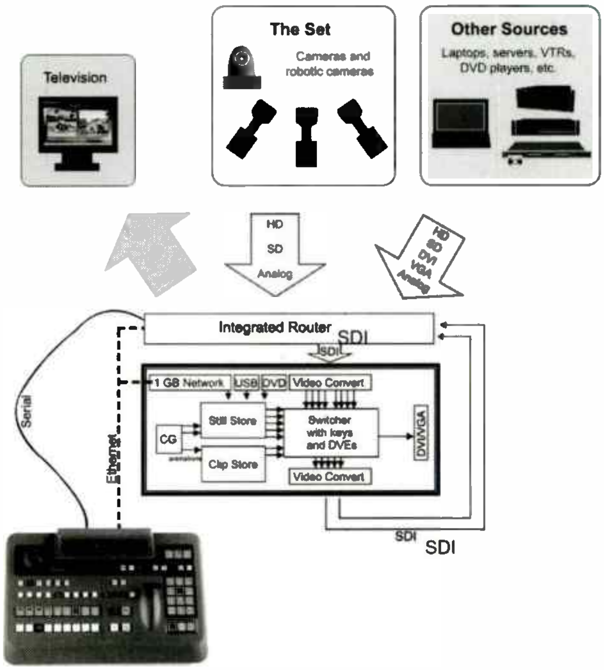


Figure 7 – Integrated router provides not-stop switching

Like conventional switchers, the LIPS can add redundant power for all components, but unlike conventional switchers, it has redundant panels too. With a conventional switcher, if you spill a drink on a panel just before going on air, you are stuck, but with the LIPS you can just open a soft panel, or control it from panel shortcuts on the keyboard.

Figure 10 shows an overview of the LIPS workflow.

EXAMPLE LISP INSTALLATIONS

KHQ

KHQ-DT, the NBC affiliate in Spokane, Washington, uses a Broadcast Pix Slate LIPS for single-operator production of the station's "Weather Plus Updates." The weather report is broadcast daily on its DTV channel 6-2; carried locally by Comcast Cablevision, and available from the station's Web site, www.khq.com. The updates are also fed to two other Central Washington stations owned by KHQ.

The LIPS integrates a production switcher, production control panel, Inscriber CG, transitional effects, chromakeyer, clip store, and multi-view monitoring, among other functions, in a single workstation-based system. The system can be operated with a production switcher-style control panel or by using a keyboard and mouse interface. KHQ found the system's compact footprint ideal for its situation.

Figure 8 – LISP at KHQ

"During the first year that we produced Weather Plus Updates, we had been using our main production control room," said Doug Miles, KHQ's production manager. "We soon found that sharing these resources with the other daily productions that needed to be done from the same rooms made things logistically difficult. Our Weather Plus Updates were tying up the main control room and studio for about five hours a day, and we didn't always have the rooms available to take Weather Plus Updates live in severe weather emergencies. Since Production Control 2 is extremely space-challenged, it was critical to have an integrated production switcher that could be managed entirely by a single operator, as opposed to our main production control room which requires a director, technical director, and audio person."

Another feature that empowers a single operator to manage every aspect of a video production is Scripts, which are sequence memories that automate program execution, including chromakeys, camera takes, dissolves, and graphics and other roll-ins.

"With Scripts, Slate remembers all your button pushes from start to finish as you go through your show–which cameras to take, which lower third supers to display, and rolling in any video and music elements. Then in subsequent production run-throughs, it can automatically recall all of those moves in their precise order," said Miles. "All the operator has to do is push a button and cue the talent. This feature can store the settings for several different show formats, which will be useful should we expand our 'Updates' programming in the future."

Formerly a newsroom edit suite, Production Control 2 is a six by six foot space adjacent to an auxiliary set dedicated to producing Weather Plus Updates, as well as news and elections updates.

Idaho Public Television

Idaho Public Television uses two Broadcast Pix Slate 1000 LIPS to produce its "Legislative Live" service covering Idaho Senate and House proceedings. This "C-Span-style," gavel-to-gavel coverage of the Idaho State Legislature, is distributed as a public service via several different media platforms including over-the-air broadcast, digital cable, the Web, mobile phones, and audio podcasts.

The powerful, compact design of LIPS was a critical consideration since the production control room is located in a 15x9 foot "closet" in the basement of the Idaho State Capitol building. The affordable, HD-ready, LIPS packs the same functionality found in a room full of broadcast equipment—such as character generation (CG); digital video effects (DVE); still store; and a digital production switcher—into an integrated unit that only requires a single person to operate.

"Our live coverage of the State Legislature in action is done with very high production values," said Jeff Tucker, production manager for Idaho Public Television. "We're able to produce this high-quality programming from the Broadcast Pix system and deliver the finished product to a variety of new and conventional media outlets. Our "Legislative Live" service rests entirely on the Broadcast Pix switchers, which have proven their capabilities and reliability."

"The Slate 1000 is designed to enable a single individual to control every aspect of this polished, professional live show, which would otherwise have required a room full of equipment and a three-man crew," said Tucker. "The Slate 1000 enables us to work within our severe space constraints and tight budget without making quality compromises. The control panel offers easy access to everything the operator needs—including the DVE, CG, Wipes, Keyers, chromakeyers, still store, plus camera switching. With integrated VGA monitoring, all of the camera signals can be displayed on a single monitor screen, instead of a bank of video monitors. This is much more ergonomic for us considering that we have essentially two production control rooms operating out of a closet," said Tucker.

"Legislative Live" is broadcast on two of the PBS station's DTV subsystems—digital multicast Channel 3 (Learn) and Channel 4 (Citizen), which carries public affairs programming. Idaho Public Television's SDTV and HDTV programming is broadcast across the large state by five DTV transmitters as well as by CableOne and Time Warner on their digital cable tiers. Anyone with an Internet connection can view the live coverage of the House or Senate proceedings at either 56k or 300k broadband at http://idahoptv.org/idreports/LegislatureLive.cfm).

The First Baptist Church of Texarcana

The Broadcast Pix™ switcher, an integrated live production system (LIPS), is used at the First Baptist Church of Texarkana, TX (FBC) for all its worship services. The unit powers FBC's wide screen displays during services, enhancing the congregation's experience every Sunday. Services are also broadcast locally on Texarkana's KLFI-TV and KAQC-TV as well as KTAL-TV in Shreveport, Louisiana. Services are also seen nationally on religious networks such as The Church Channel and Family Net.

Figure 9 - LISP at First Baptist Church of Texarcana

FBC producer, Jay Budzilowski admitted he had little experience in using switchers before coming in contact with the Slate 1000 and appreciated the ease of use that the clearly marked labeling offers. Displays in the panel's push-buttons that show the exact content on every source and key proved helpful for FBC's worship production. He said he became confident with the unit after reading through its manual and considered most of its functionality to be self-explanatory.

During worship services, the Slate 1000 controls what appears on FBC's two 16x9 wide screens at the front of the church. As many as 2,500 people fill the pews at FBC on Sundays, making the screens a critical part of services. The screens display song lyrics and other video segments that pertain to the pastoral messages. This aspect of production became significantly more streamlined by Slate 1000's clip store functionality.

Prior to installing the Slate 1000, Budzilowski said producing video clips and getting them to run on screen was a complex and labor intensive process. "Edited segments in Final Cut Pro were burned to DVD. Then we had to hook the player to the switcher, cue the scene we wanted, pause it and press play at exactly the right moment," Budzilowski explained. "It became a real headache, especially when there was more than one clip needed for services." Now FBC producers can simply import multiple edited pieces into the Slate 1000 and play them on the fly. With the touch of a button, the right clip plays for the duration needed.

The LISP enhances the congregation's experience of performed music during worship services as well. Budzilowski said FBC previously used a Power Point format to display song lyrics and they had no ability to key the lyrics over moving video. Now they are able to key lyrics over edited video clips or shots of the choir. Church members need not look down at a hymnal to read song lyrics and can sing along with their heads held high.

"With the Slate 1000, we have much more flexibility to employ creative methods and make services more engaging for the church members," Budzilowski explained. "After all, they're here to attend church, not watch TV."

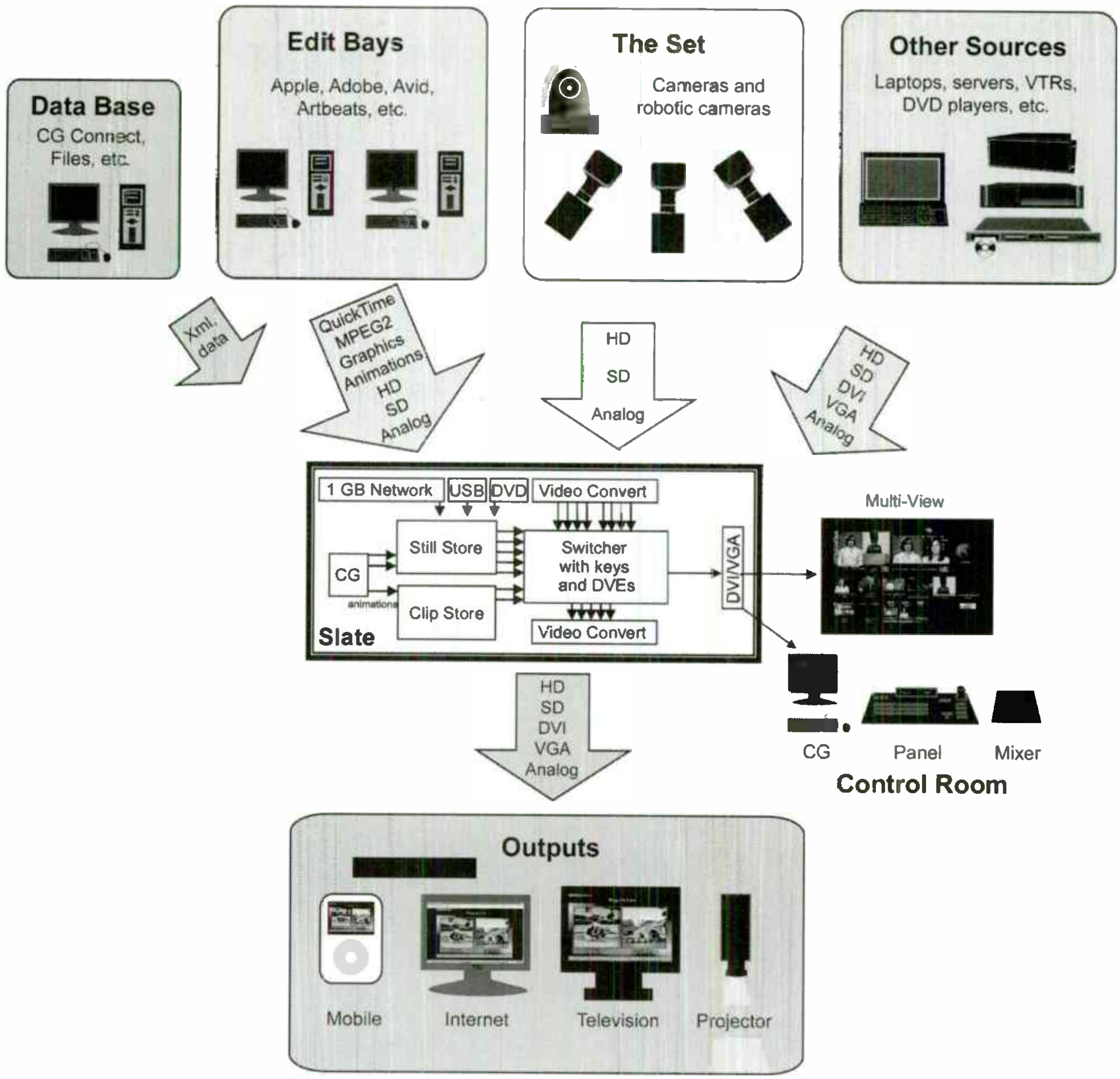


Figure 10 - Live Integrated Production System (LIPS) Workflow

Acknowledgements

Broadcast Pix and Slate PixButtons are trademarks of Broadcast Pix, Inc. All other trademarks are of their respective companies. Patents pending.

Ken Swanton
President and Co-Founder
Broadcast Pix, Inc.
Billerica, Massachusetts, USA
www.broadcastpix.com

HIGH DEFINITION ELECTRONIC NEWS GATHERING (HD-ENG) FIELD TEST REPORT

Walter Sidas, Gregory Coppa, Robert P. Seidel
CBS
New York, NY

ABSTRACT

This paper presents the results of High Definition (HD) Electronic News Gathering (HD-ENG) field tests performed by CBS, in the New York City area. Analysis of the data collected is used to determine the optimal operating parameters and performance of HD-ENG in major US cities. It describes the objectives, methodology, and field tests performed using the new FCC mandated 2 GHz, 12 MHz channel plan, compressed HD signals at bit rates from 18 to 28 Mbps, and reception using Coded Orthogonal Frequency Division Multiplexed (COFDM) modulation. The COFDM signals were transmitted from typical locations used for newsgathering, and received at the Empire State Building. Also tests were performed with Standard Definition (SD) digital (COFDM) signal transmission using the new 2 GHz plan. Adjacent channel performance was verified using 6 and 8 MHz transmission channels ("pedestals") at several modulation and data rates. A 35 Mbps, 16 QAM (non-COFDM), fixed forward error correction (FEC) modulation transmission system was also tested.

SCOPE

This document presents the results of High Definition (HD) Electronic News Gathering (HD-ENG) field tests performed by CBS. It describes the objectives, methodology, and field tests performed to determine the optimal operating parameters of digital microwave transmission and reception in an urban environment using compressed HD signals at bit rates from 18 to 28 Mbps. The microwave signals were transmitted using the new, FCC mandated [1]-[4], 2 GHz, 12 MHz channel plan and employed Coded Orthogonal Frequency Division Multiplexed (COFDM) modulation.

The tests were performed using a typical ENG vehicle, modified to incorporate an HD encoder and the COFDM modulation equipment. Six COFDM configurations were tested to determine the maximum bit rate that could be reliably delivered.

Signals were transmitted from locations frequently used for newsgathering such as the United Nations, City Hall, and the World Trade Center, thus representing "real world" conditions and included near and distant Line of Sight (LOS) and "bounce" transmissions.

The transmitted COFDM signals were received at the Empire State Building using existing RF microwave antennas. COFDM demodulation and HD decoding equipment were added to the receive path to gather received signal quality data.

Analysis of the data collected for the variety of field conditions experienced during the tests is presented. The analysis is used to determine the optimal operating parameters and performance of HD-ENG in New York and that expected in other major US cities.

In addition to the HD-ENG results, this document describes the field tests performed that quantified Standard Definition (SD) digital (COFDM) signal transmission using the new 2 GHz plan. Tests were performed to verify single, adjacent channel and geographically diverse co-channel performance. Interference free adjacent channel performance was verified using 6 and 8 MHz transmission channels ("pedestals") at several modulation and data rates. The result of this work is presented.

Finally, a 35 Mbps, 16 QAM (non-COFDM), fixed forward error correction (FEC) modulation, transmission system was tested. This work was undertaken to determine whether this type of modulation could be used in an urban environment. The result of this work is presented.

The scope of the work herein was not concerned with the compression quality of the audio or video, except when these signals are used as a means to determine that the data has been correctly recovered.

INTRODUCTION

According to a mandate from the FCC the 2 GHz Broadcast Auxiliary Band will transition from the 17 MHz per channel (1990 to 2110 MHz) to a new 12 MHz per channel (2025.5 to 2109.5 MHz) spectrum plan. This transition is expected to increase the use of digital transmission, as the reduced channel bandwidth of the new spectrum plan will degrade the performance of analog SD transmissions below "broadcast quality". In order to have equivalent broadcast quality, digital compression and transmission must be employed.

This microwave spectrum transition also comes at a key point in the transition to High Definition News. An increasing number of News organizations have transitioned their studio operations to HD and are aggressively promoting their HD News broadcasts. The availability of low cost HD News gathering equipment is increasing and, in many cases, dual mode SD/HD camcorders cost less than the SD only equivalent. Thus, the transition to HD News gathering is imminent and the requirement to provide systems with equivalent functionality is paramount.

One such system is ENG microwave. Existing SD microwave systems are a combination of analog and digital equipment. Typical ENG digital microwave systems are based upon the DVB-T specification [5], which uses COFDM modulation and relatively low bit rates. HD-ENG requires higher bit rates to achieve the audio/video quality necessary for further processing for broadcast and yet must be reliably delivered.

In a series of tests representing "real world" conditions, data was gathered and subsequently analyzed. The focus of the data analysis was used to determine the optimal operating parameters for HD-ENG in New York City. Since the urban canyons of New York City represent an extremely difficult RF environment it is expected that these parameters will be valid for other major US cities.

TEST PROGRAM OBJECTIVE

The objectives of the CBS test program were:

- Determine the optimal operational parameters for the highest bit rate possible for reliable HD-ENG delivery.
- Measure transmission paths as exist in the New York City metropolitan area. This includes transmission from "real world" sites frequently used for news events at which ENG coverage is required.
- Determine the operational parameters for adjacent and co-channel SD digital microwave transmission in a 12 MHz channel.
- Provide a uniform series of measurement procedures so that other organizations in different locations can perform similar measurements and compare results.
- Identify the variables in the environment and recommend the minimum set of variables to be measured.
- Collect data useful in creating operational guidelines for digital HD-ENG at 2 GHz in the narrower 12 MHz band plan using 8 MHz pedestals.
- Test and document results of a prototype 35 Mbps non-COFDM transmission system.

DESCRIPTION OF TEST PROGRAM

The test program's primary objective was to determine the optimal operating parameters of digital microwave transmission and reception using the new 2 GHz, 12 MHz channel plan in an urban environment using compressed HD signals at bit rates from 18 to 28 Mbps.

Adjacent and Co-channel SD Tests

A series of tests were performed in June 2005. The objective of the tests was to determine feasibility of adjacent and co-channel SD transmissions using a variety of transmission parameters e.g. power, polarization and signal paths. The tests were conducted using mostly LOS paths. These tests, with up to three ENG microwave vans, utilized 6 and 8 MHz bandwidth channels, in the narrower 12 MHz channel plan at 2 GHz.

HDTV Data Rate Tests

A second series of tests were performed in March 2006. The objective of these tests was to determine operational parameters required to achieve the highest bit rate possible for reliable HDTV transmission. These tests used one ENG van and recorded parameters for quasi-error free HDTV transmission. The tests concentrated on collecting reception data from "difficult" transmit sites, that required a bounce. A variety of LOS locations, considered "easy" transmit sites with increasing distances from the Empire State Building, were measured for all data rates to determine the relationship between data rate and transmit distance.

High Data Rate 35 Mbps Tests

A prototype 35 Mbps, 16-QAM (non-COFDM), transmission system was evaluated as part of the March 2006 tests. The objective of this test was to evaluate the performance of a non-COFDM modulation transmission system in LOS and bounce applications.

Data Collection Procedures

Transmit and Receive Antenna Alignment

The control of the receive antenna was switched to local control, allowing the operator at the Empire State Building to position the receive antenna.

When the microwave van was in position, and transmitting, the van operator contacted the operator at the Empire State Building, who then adjusted the azimuth of the receive antenna until the highest receive signal level (RSL) was measured. It was further refined in coordination with the transmit van, by adjusting azimuth, and tilt of the transmit antenna, and selecting the combination of transmit power, antenna pointing and polarization, that resulted in the best combination of RSL and Link Quality (LQ). The LQ is an index that rates from 10, which is perfect, to 0, which is no reception. The LQ metric is a combination of Carrier to Noise (C/N), RSL and Bit Error Rate (BER).

At this setting, the receive angle of the antenna azimuth, LQ, and RSL were recorded on the data sheet.

A novel approach to antenna alignment for the HDTV data rate tests was developed. A low bit rate, 9.68 Mbps, QPSK, Forward Error Correction (FEC) 7/8 and Guard Interval (GI) 1/8, COFDM configuration was preset (antenna alignment preset) in the Microwave Radio Communications (MRC) transmission equipment. This preset is meant to provide a robust transmission signal that can be used to align the antennas for optimal receive performance, without having to use an analog transmission. Once the antenna position is optimized the operating preset is chosen.

Data Collection

Software provided by MRC displays Signal to Noise (SNR), Pre-FEC, Post-FEC and LQ parameters on a laptop computer. These parameters were recorded on the data sheet for each of the configurations that were tested.

Adjacent and Co-channel Data Collection

Once the link was established and the parameters recorded, the received picture was observed and any anomalies or breakup were noted during a five-minute period and recorded on the data sheet.

The transmission settings were then changed to the preset corresponding to the next data rate and the data was recorded for that combination of parameters. This procedure was repeated for all the data rates and both the adjacent and co-channel cases

HDTV Data Rate and 35 Mbps Data Collection

Using the antenna alignment preset, the operator at the Empire State Building adjusted the azimuth of the receive antenna until the highest RSL was received. The picture was then observed on a high definition picture monitor, and recorded for 5 minutes using a Sony XDCAM-HD disc recorder. Anomalies and/or breakup in the picture were noted. Then the microwave van changed to the next preset corresponding to a higher data rate and the data was recorded for that combination of preset parameters. This was done for Presets 1, 2, 3, 4 and 5. The corresponding parameters and data rates are shown in Table 1.

Upon completion of the COFDM testing, the 35 Mbps non-COFDM system was tested. For the 35 Mbps tests, the RSL, SNR, and Error Vector Magnitude (EVM) were recorded on the data sheet. The picture was then observed on a high definition picture monitor, and recorded for five minutes using a Sony XDCAM-HD disc recorder. Anomalies and/or breakup in the picture were noted and recorded on the data sheet.

Calibration

System Calibration was performed at the start of the tests and repeated when necessary. The transmit van would go to a known LOS site and the link performance and picture quality would be recorded and compared against known good settings. Any anomalies would be corrected before testing would continue.

Bandwidth (MHz)	Modulation System	Code Rate	Data Rate (Mbps)	Guard Interval	Preset
8	QPSK	7/8	9.68	1/8	6
8	16-QAM	5/6	18.43	1/8	1
8	16-QAM	7/8	19.35	1/8	2
8	64-QAM	2/3	22.12	1/8	3
8	64-QAM	3/4	24.88	1/8	4
8	64-QAM	5/6	27.65	1/8	5

Table 1 HDTV Data Rate COFDM Preset Settings

SYSTEM DESCRIPTION

Modulation Characteristics

The DVB-T transmission system utilizes COFDM made up of 2,000 carriers that are spaced a few KHz apart across the operating signal bandwidth of 8, 7 or 6 MHz. Each carrier is modulated with one bit of data at a time. The modulation format of all carriers is operator selectable as QPSK, 16-QAM or 64-QAM.

The DVB-T also defines five selectable forward error correction (FEC) settings or "Code Rates". These are 1/2, 2/3, 3/4, 5/6 and 7/8. The more error correction is used the less data will be transmitted for a given data rate, for example, an FEC of 1/2 means that for every two transmitted bits one is an error correction bit.

Another parameter that is important when using COFDM is the "Guard Interval" (GI). This guard interval improves the multi-path performance. Guard intervals are defined as 1/4, 1/8, 1/16 or 1/32 of the transmitted interval. The guard interval size affects the maximum data rate.

Typical SD digital ENG operations currently use an 8 MHz wide signal, QPSK, 1/2 FEC and GI of 1/4 resulting in a bit rate of 5.53 Mbps. Most SD digital ENG equipment in use today is delivered with these settings yielding very robust transmission however, the resulting video quality in not optimal.

Transmit System

All testing was done using an ENG van equipped with MRC CodeRunner-2 digital microwave transmitters and MPEG-2 audio/video encoders for the SD and HD signals.

The SD adjacent and co-channel tests used up to three similarly equipped ENG vans with MRC digital microwave transmitters and SD encoders.

Transmit Block Diagram

Figure 1 is the ENG van transmit block diagram. This diagram shows both the SD and HD configurations.

In both the SD and HD tests a video camera was used to provide source signals. For the HD signal, a Sony HDW-700A camera was used.

MPEG-2 compression was used to encode both the SD and HD source audio and video signals. The SD encoder used for the adjacent and co-channel tests was a Tandberg model 5750. The HD signals were encoded using a NTT Electronics (NEL) HE-3000 encoder.

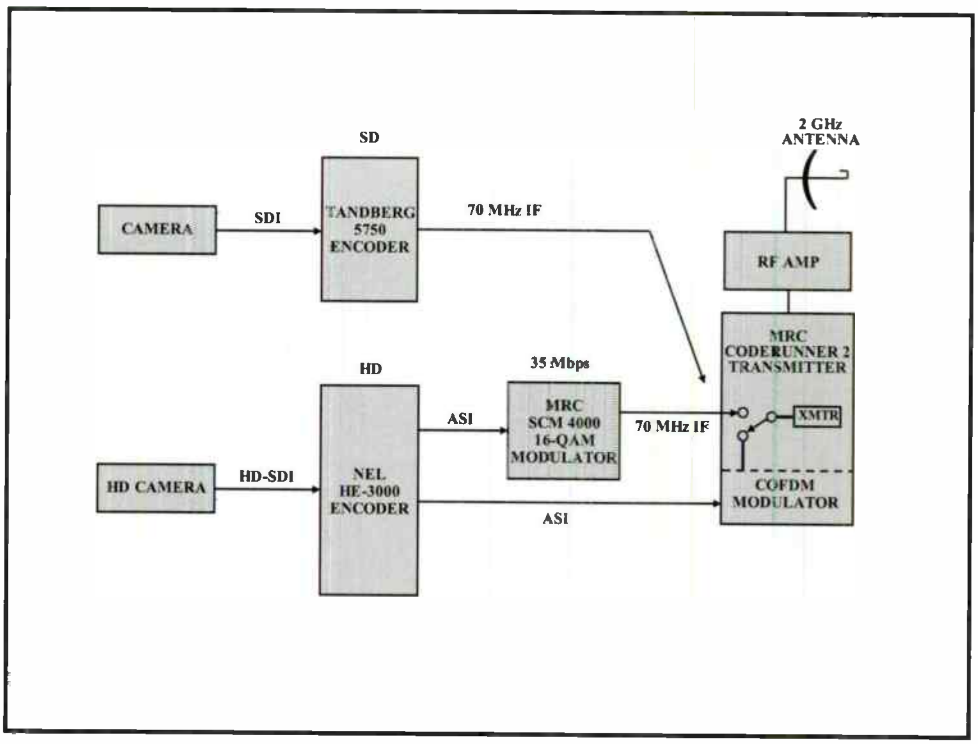


Figure 1 Digital ENG Transmit Block Diagram

The SD encoder's output is a 70 MHz IF signal. The HD encoder's output is an ASI signal that is connected to the input of a MRC CodeRunner-2, COFDM modulator/transmitter. The CodeRunner-2's input is selected between the IF or ASI signal depending upon the test being conducted. The CodeRunner-2's output is a 2 GHz RF signal that is delivered to the mast mounted RF power amplifier.

For the 35 Mbps testing an MRC model SCM-4000 modem that outputs a 70 MHz IF signal was used. The IF signal is connected to the CodeRunner's IF input where it is then delivered to the mast mounted RF power amplifier.

Each van has a telescoping mast, capable of raising the antenna to 30 feet. Controls in the van provide means to adjust the transmitted power level and azimuth, tilt, and polarization of the antenna.

For the HD tests, presets on the NEL encoder and the CodeRunner-2 are used to change operating parameters for the different data rates.

Receive System

The receive block diagram appears in Figure 2. MRC UltraScan DR-2 receive antennas are mounted on the northwest and southeast corners of the 84th floor roof of the Empire State Building. These antennas can be positioned to cover most of the sites required for newsgathering operations. The receive antennas are normally controlled remotely by WCBS-TV Microwave Control however, for the DENG tests the antennas were controlled locally in the WCBS-TV transmitter facility on the 83rd floor of the Empire State Building.

RF signals at 2 GHz are input to the MRC CodeRunner-4 receiver. The receiver down-converts the received COFDM modulated signal to a 70 MHz IF signal.

For the HD-ENG tests, the received IF signal is demodulated using the MRC model MRX-4000, which outputs an ASI signal. The ASI signal is connected to the NEL decoder, which decodes the compressed audio and video signal and outputs a baseband HDSDI signal with embedded audio.

For the 35 Mbps tests, the received IF signal is demodulated using the MRC model SCM-4000, which outputs an ASI signal that contains the 35 Mbps signal. This signal is input to the NEL decoder, which decodes the compressed signals and outputs a baseband HDSDI signal.

For the adjacent and co-channel tests, the received IF signal is distributed to the input of a MRC Strata RXU receiver that demodulated the COFDM signal and decoded the MPEG-2 (4:2:0) signal. The output of the decoder was displayed on a picture monitor. A quad-split video display device was used to display the multiple received SD signals.

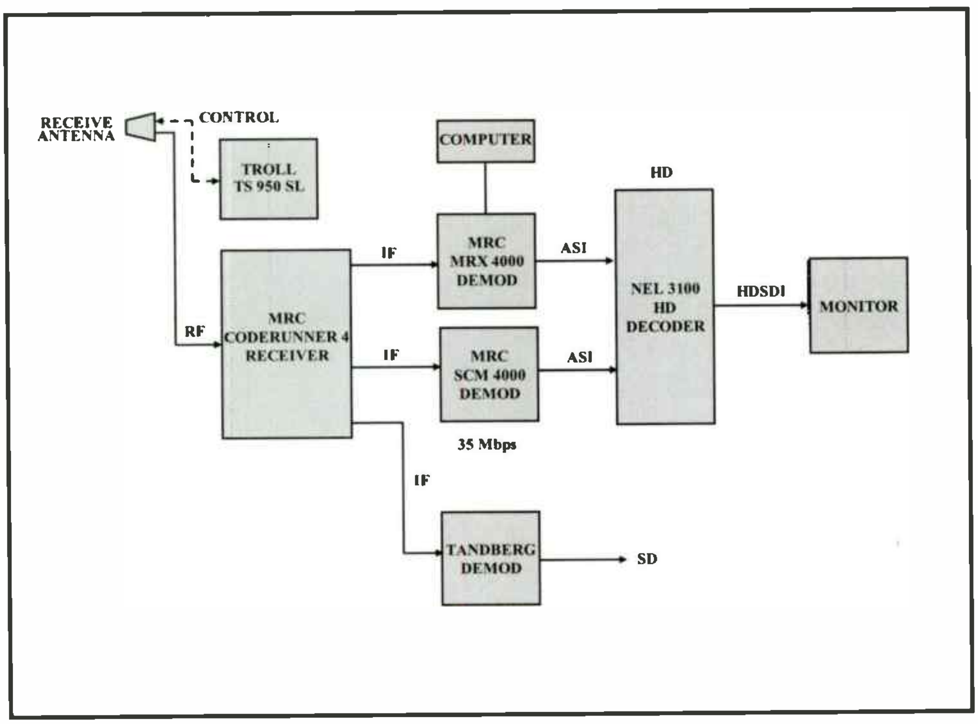


Figure 2 Digital ENG Receive Block Diagram

TRANSMIT SITES

HDTV Data Rate and 35 Mbps Transmit Sites

Transmit sites in the New York City region were selected to provide a variety of transmit paths including sites regularly used by WCBS-TV for electronic News gathering.

Selection criteria included "difficult" sites, that require the transmitted signal to be reflected or "bounced" off of an adjacent building to reach the receive site, and "easy" sites that have LOS to the receive site. For the LOS locations, various distances, near and far were selected to collect data on the performance of the path. as data rate increased.

.

Table A.1 in the Appendix is a list of the transmit sites, site number and distance from the Empire State Building for the data rate tests.

Figure 3 shows the sites within 6 miles of the Empire State Building.

Figure 4 shows the transmit sites located 10 to 25 miles from the Empire State Building.

Figure 5 shows the transmit sites located 25 to 51 miles from the Empire State Building.

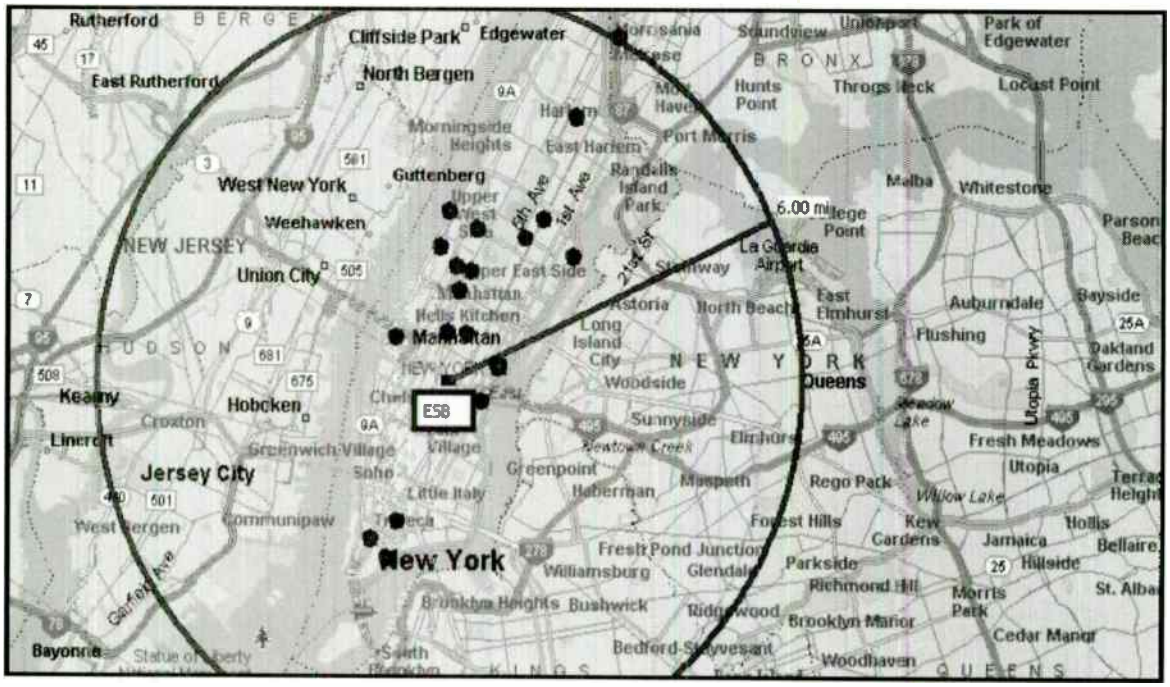


Figure 3 Sites Within 6 Miles of the Empire State Building

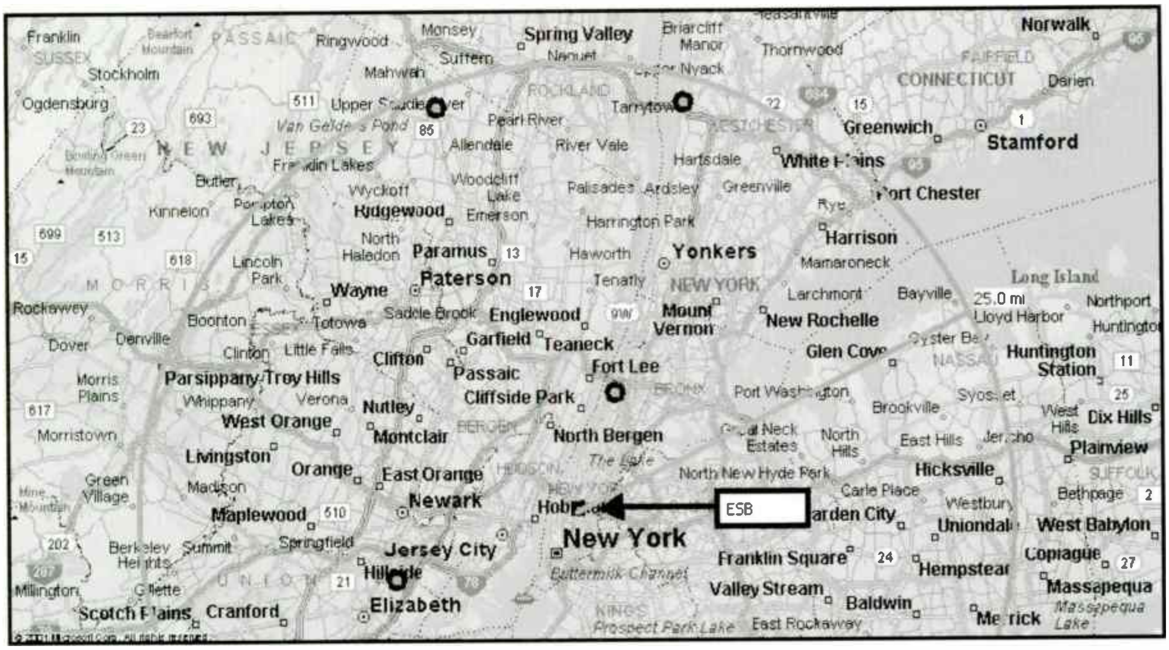


Figure 4 Sites Within 10 to 25 Miles of the Empire State Building

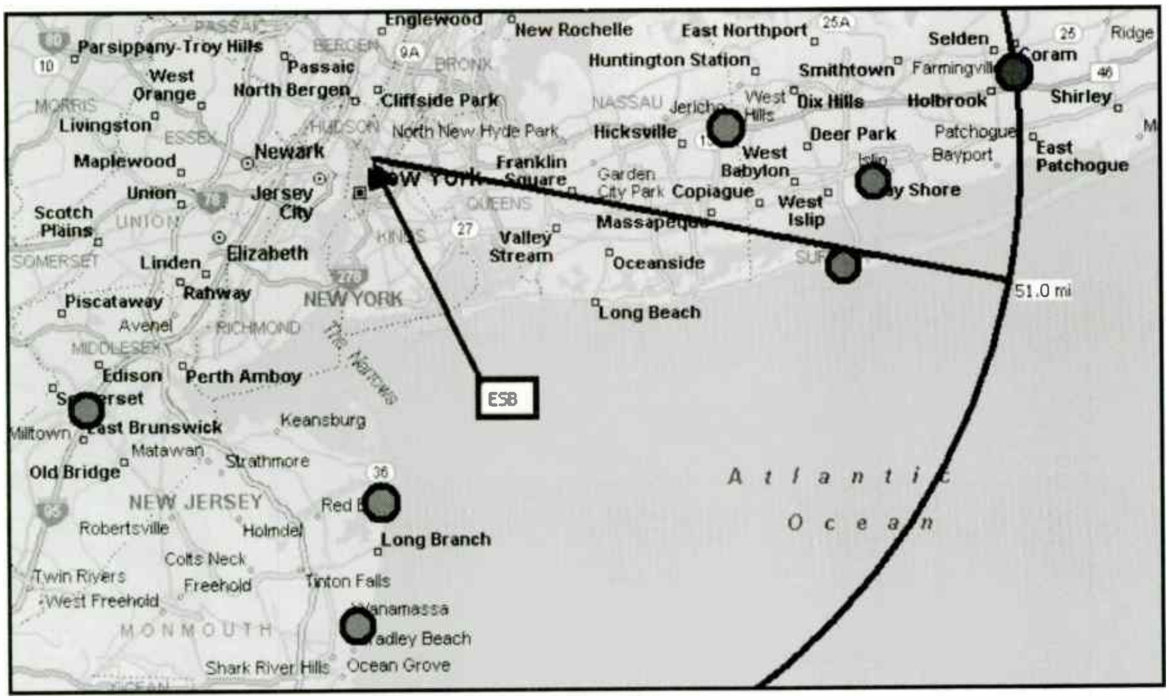


Figure 5 Transmit Sites Within 25 to 51 Miles of the Empire State Building

Adjacent and Co-channel Transmit Sites

Table A.2 in the Appendix is a list of the transmit sites for the adjacent and co-channel tests. Figure 6 is a map of the test sites for the adjacent and co-channel tests. The Javits Center located on the west side of Manhattan is one mile W-NW of the Empire Sate Building and affords a clear LOS path. This site was used to determine baseline transmission parameters, and determine that both the transmitting and receive equipment was operating properly. Most of the sites used for adjacent channel operation were LOS, and up to a distance of 36 miles from the Empire State Building.

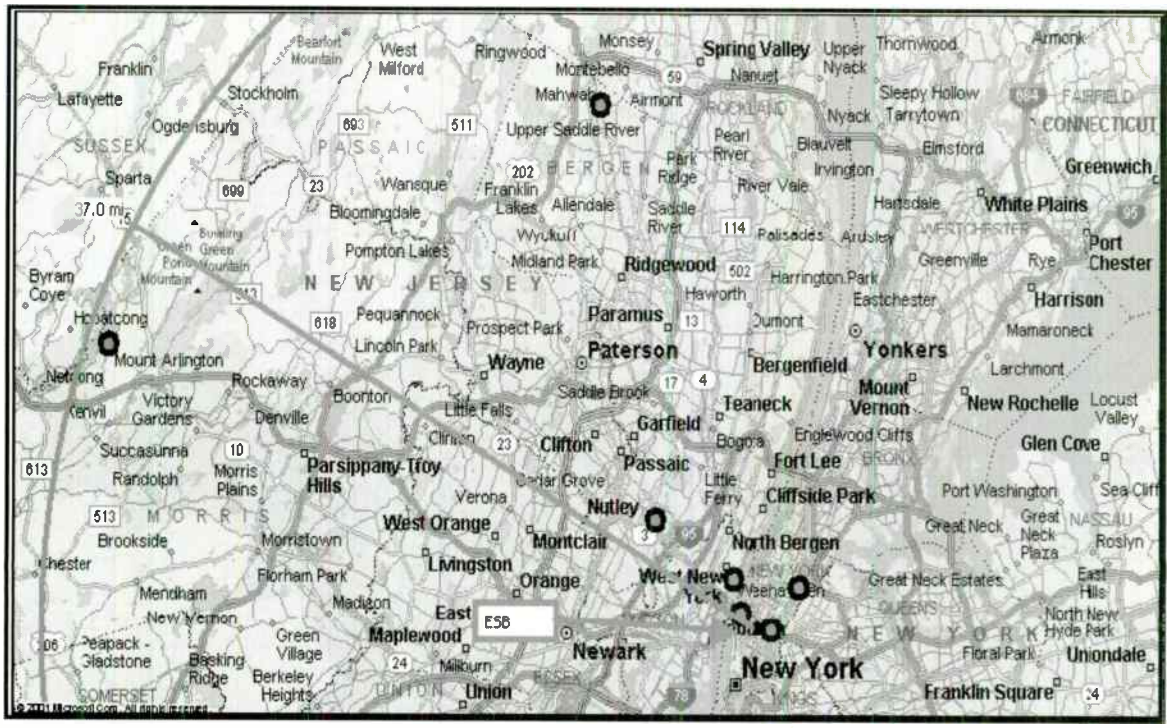


Figure 6 Test Sites For the Adjacent and Co-channel Tests

DATA ANALYSIS

Data analysis for the HDTV data rate, 35 Mbps, and SD adjacent and co-channel tests are presented below.

DVB-T Modulation and Data Rates

When analyzing the data, it is important to understand that the different data rates encompass three different modulation schemes. The different modulation characteristics are detailed in Table 2.

Bandwidth (MHz)	Modulation System	Code Rate	Data Rate (Mbps)	TV Resolution
8	QPSK	1/2	5.53	Manufacturer Default
8	QSPK	7/8	9.68	Antenna Aiming
8	16-QAM	5/6	18.43	HD
8	16-QAM	7/8	19.35	HD
8	64-QAM	2/3	22.12	HD
8	64-QAM	3/4	24.88	HD
8	64-QAM	5/6	27.65	HD
6	QPSK	2/3	5.53	SD
6	QPSK	3/4	6.22	SD
6	QPSK	5/6	6.91	SD
6	QPSK	7/8	7.26	SD
6	64-QAM	3/4	18.66	HD
6	64-QAM	5/6	20.74	HD
6	64-QAM	7/8	21.77	HD

Table 2 DVB-T Modulation Parameters, 1/8 Guard Interval

DVB-T Carrier to Noise Requirements

The different DVB-T modulation systems have different minimum Carrier-to-Noise (C/N) ratios that must be taken into consideration when analyzing the data. The minimum C/N is also different for the type of channel model. A Gaussian channel, where the direct receive signal is only impaired by white noise, will have the lowest C/N requirement. A Ricean channel, where the signal path is impaired by a number of echoes of varying levels and phases, will have a slightly higher C/N for the same modulation. A Rayleigh channel, which has no direct line of sight and has multiple echoes similar to the Ricean channel, has the highest C/N requirement of the channel models and is closest to actual field conditions. For Quasi-Error-Free (QEF) reception, a bit error rate of $2x10^{-4}$, after FEC is applied, is required. Table 3 lists the C/N ratio required to achieve QEF for the Gaussian, Ricean, and Rayleigh channels.

Required C/N for a $2x10^{-4}$ BER to have Quasi-Error-Free Reception after FEC					
Data Rate (Mbps)	Modulation	FEC	Gaussian (dB)	Ricean (dB)	Rayleigh (dB)
9.68	QPSK	7/8	7.7	8.7	16.3
18.43	16-QAM	5/6	13.5	14.4	19.3
19.35	16- QAM	7/8	13.9	15.0	22.8
22.12	64-QAM	2/3	16.5	17.1	19.3
24.88	64-QAM	3/4	18.0	18.6	21.7
27.65	64-QAM	5/6	19.3	20.0	25.3

Table 3 QEF Performance for Gaussian, Ricean, Rayleigh Channels with 1/8 Guard Interval and 8 MHz Channel Bandwidth

HDTV Data Rate Testing

HDTV Data rate testing was undertaken to determine the maximum data rate that could be transmitted using a DVB-T 8 MHz pedestal in a 12 MHz channel. It should be noted that the 9.68 Mbps MPEG-2 data rate was only used for antenna sighting and is insufficient to provide "broadcast quality" HDTV transmissions, since the HD-ENG signal will be concatenated with the ATSC compression process when it is broadcast to the home. Achieving the highest data rate possible with quasi-error free receive performance is thus highly desirable.

Site Performance

A total of 31 test locations were used for the HDTV data rate tests that included 11 LOS and 20 totally obstructed, bounce sites.

BER and RSL were recorded at each site. The audio and video signals were observed for a period of five minutes; during which time the number of audio/video disturbances or "hits" was recorded. A site was deemed to have been a success if there were less than or equal to three audio/video disturbances during a five minute interval.

Locations Within a 35 Mile Radius

Since the vast majority of ENG sites are within a 35-mile radius of the television station, data was collected and analyzed for sites with locations within a 35-mile radius. LOS and bounce sites with this radius are analyzed below.

LOS Locations

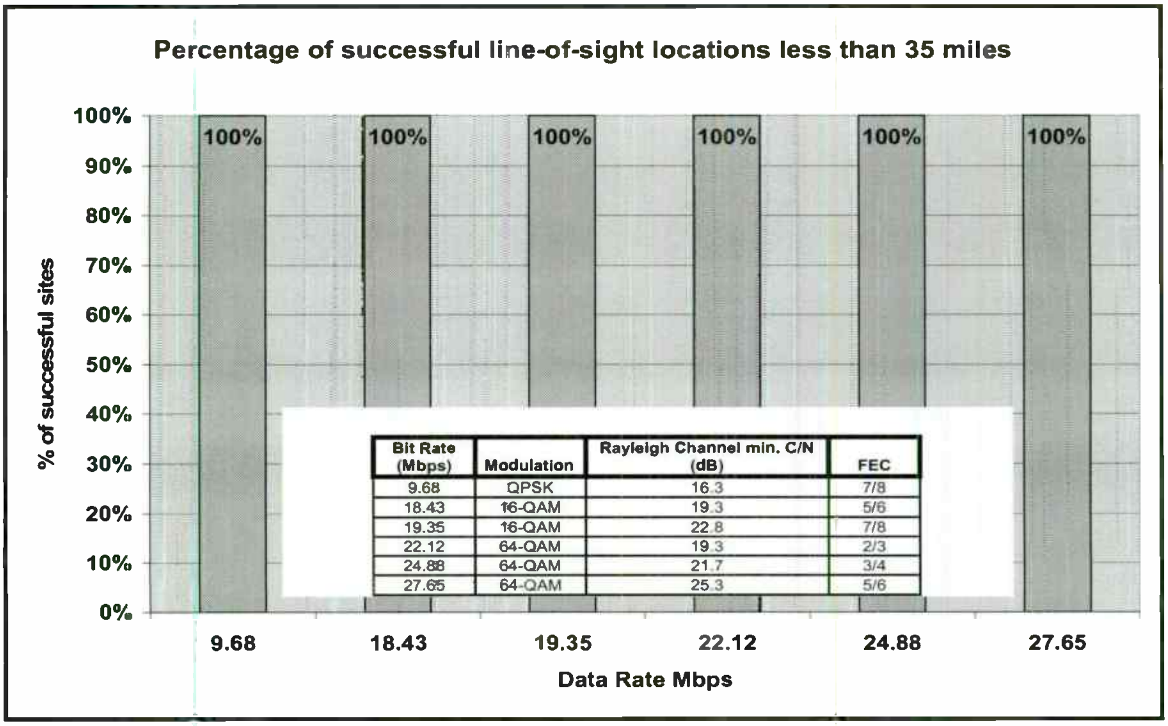
Figure 7 is a bar chart, which plots the percentage of successful sites for LOS locations within a 35-mile radius for the 6 data rates under test. A 100% success rate was achieved for the 6 data rates with only 1 Watt of transmitted power.

Bounce Locations

Figure 8 is a bar chart of the percentage of successful sites that required a bounce of the transmitted signal off adjacent buildings to reach the Empire State Building. The maximum transmission distance for any bounce site was 27.5 miles.

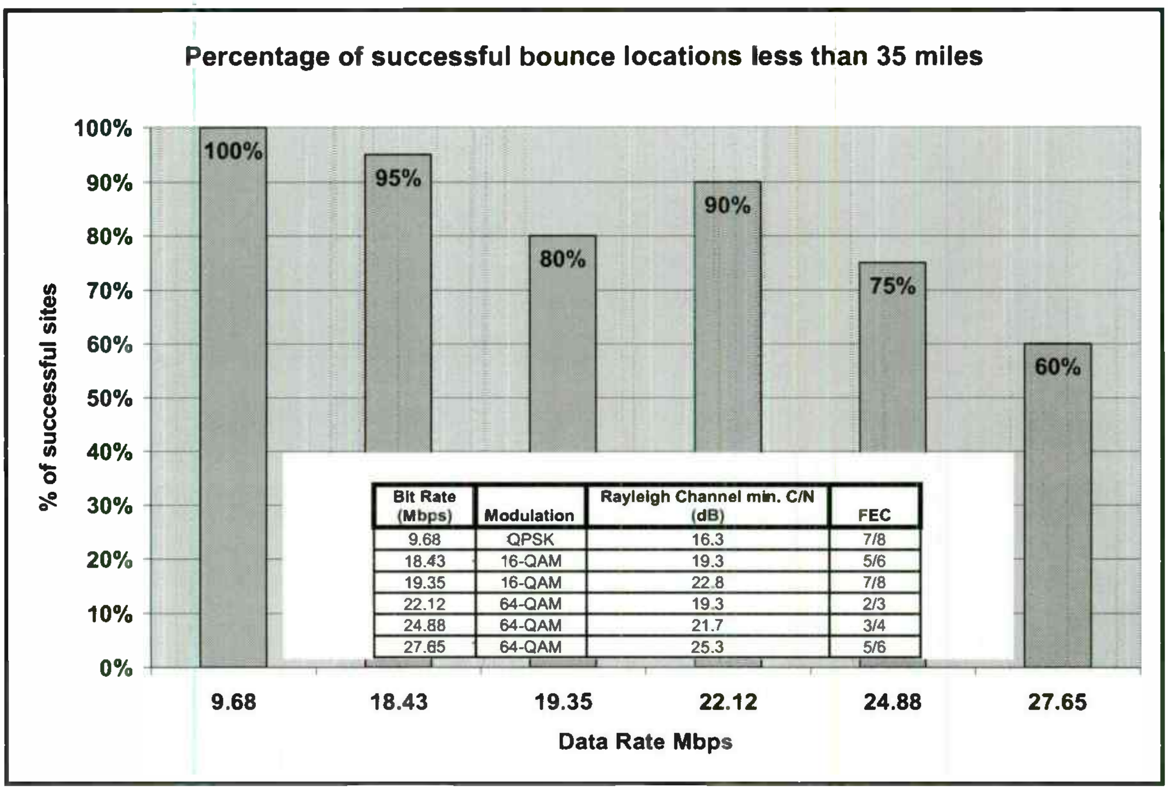
The baseline data rate of 9.68 Mbps, which was used for sighting the antenna, had a 100% success rate for all bounce locations. The 18.43 and 22.12 Mbps signals had comparable success rates of 95% and 90% respectively. This is expected since these two modulation schemes have the same minimum C/N requirement of 19.3 dB for a Rayleigh channel (Table 3).

The 19.35 Mbps signal had a receive success rate of 80%. This lower success rate is attributed to the higher C/N noise requirement of 22.8 db for a Rayleigh channel. The 24.88 Mbps signal had a success rate of 75% and the 27.65 Mbps had a success rate for bounce locations of 60%.



Bit Rate (Mbps)	Modulation	Rayleigh Channel min. C/N (dB)	FEC
9.68	QPSK	16.3	7/8
18.43	16-QAM	19.3	5/6
19.35	16-QAM	22.8	7/8
22.12	64-QAM	19.3	2/3
24.88	64-QAM	21.7	3/4
27.65	64-QAM	25.3	5/6

Figure 7 Percentage of LOS Locations Less Than 35 Miles



Bit Rate (Mbps)	Modulation	Rayleigh Channel min. C/N (dB)	FEC
9.68	QPSK	16.3	7/8
18.43	16-QAM	19.3	5/6
19.35	16-QAM	22.8	7/8
22.12	64-QAM	19.3	2/3
24.88	64-QAM	21.7	3/4
27.65	64-QAM	25.3	5/6

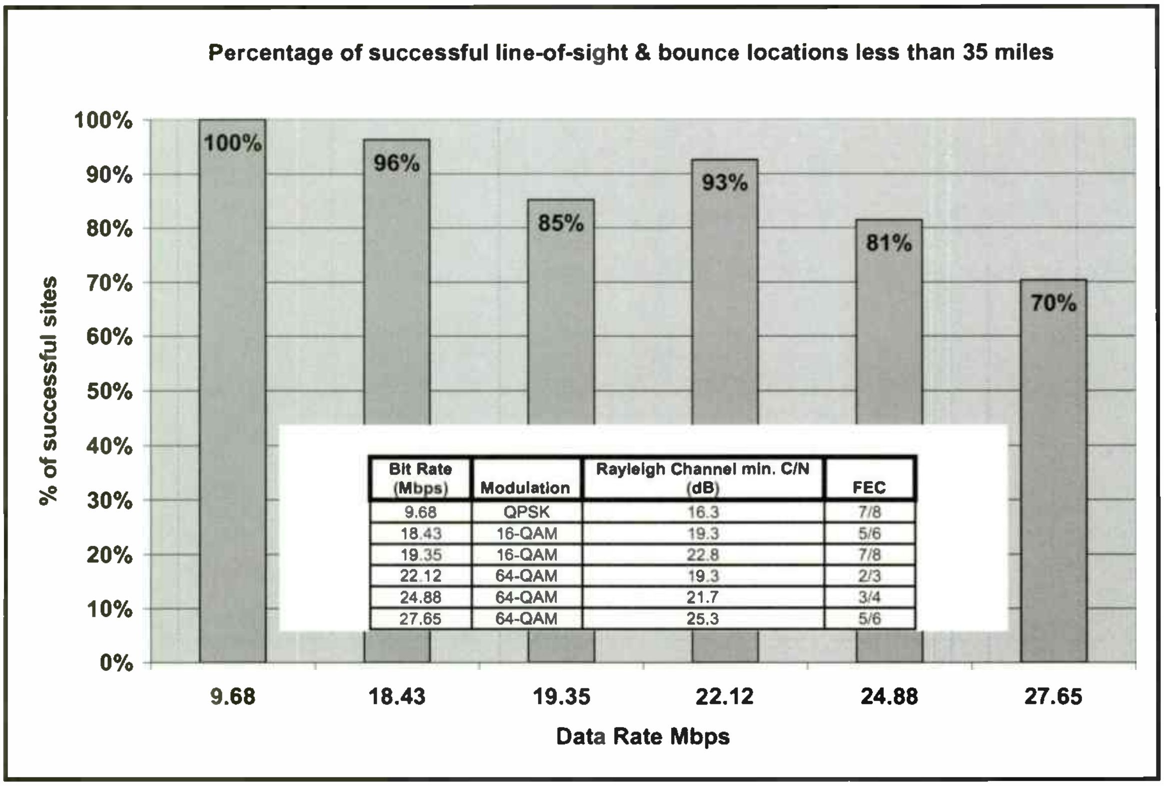
Figure 8 Percentage of Successful Bounce Locations Less Than 35 Miles

LOS and Bounce Location Analysis

Figure 9 is a bar chart of the total successful locations for both LOS and bounce sites. The baseline data rate of 9.68 Mbps had a 100% success rate. The 18.43 and 22.12 Mbps signals had a success rate of 96% and 93% respectively. The 19.35 Mbps signal had a success rate of approximately 85%. This lower success rate is again attributed to the higher C/N required for this modulation scheme. The 24.88 Mbps signal had a 81% success rate and the 27.65 Mbps signal had a 70% success rate.

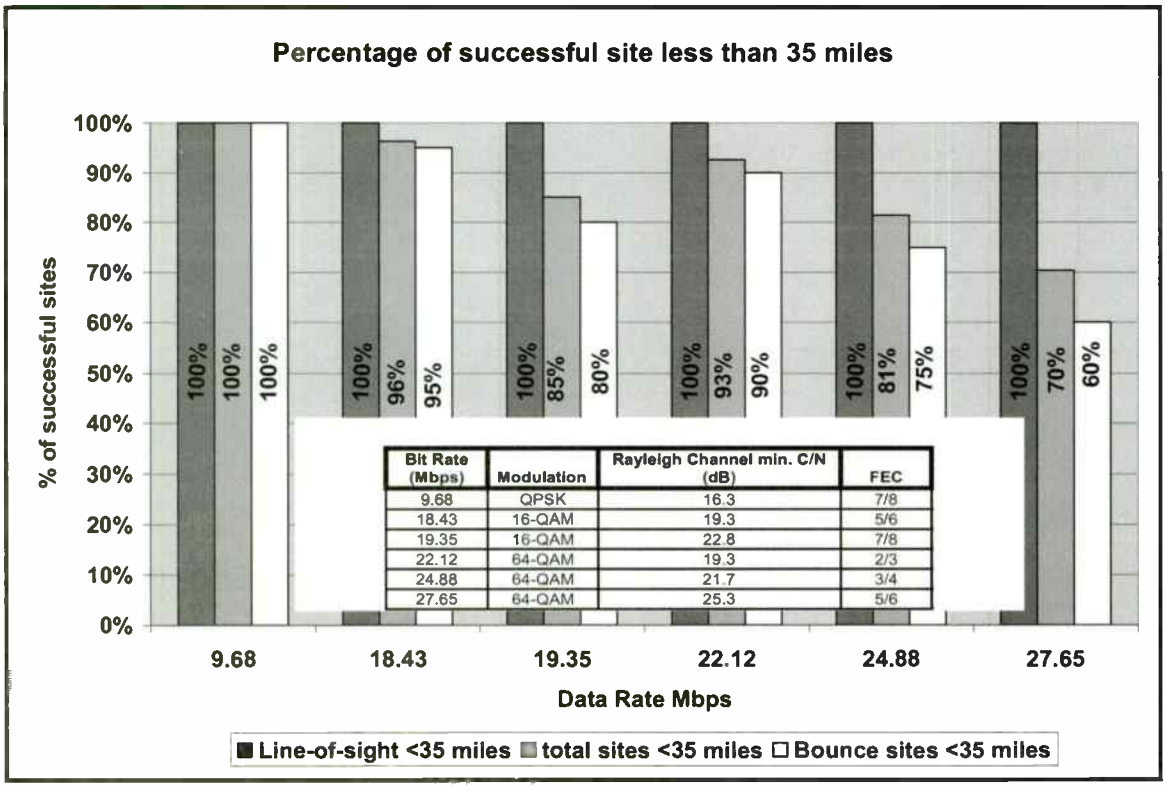
Summary of All LOS and Bounce Sites

Figure 10 is a summary bar chart of the total successful locations for both LOS and bounce sites less than 35 miles.



Bit Rate (Mbps)	Modulation	Rayleigh Channel min. C/N (dB)	FEC
9.68	QPSK	16.3	7/8
18.43	16-QAM	19.3	5/6
19.35	16-QAM	22.8	7/8
22.12	64-QAM	19.3	2/3
24.88	64-QAM	21.7	3/4
27.65	64-QAM	25.3	5/6

Figure 9 Percentage of Successful LOS and Bounce Locations Less Than 35 Miles



Bit Rate (Mbps)	Modulation	Rayleigh Channel min. C/N (dB)	FEC
9.68	QPSK	16.3	7/8
18.43	16-QAM	19.3	5/6
19.35	16-QAM	22.8	7/8
22.12	64-QAM	19.3	2/3
24.88	64-QAM	21.7	3/4
27.65	64-QAM	25.3	5/6

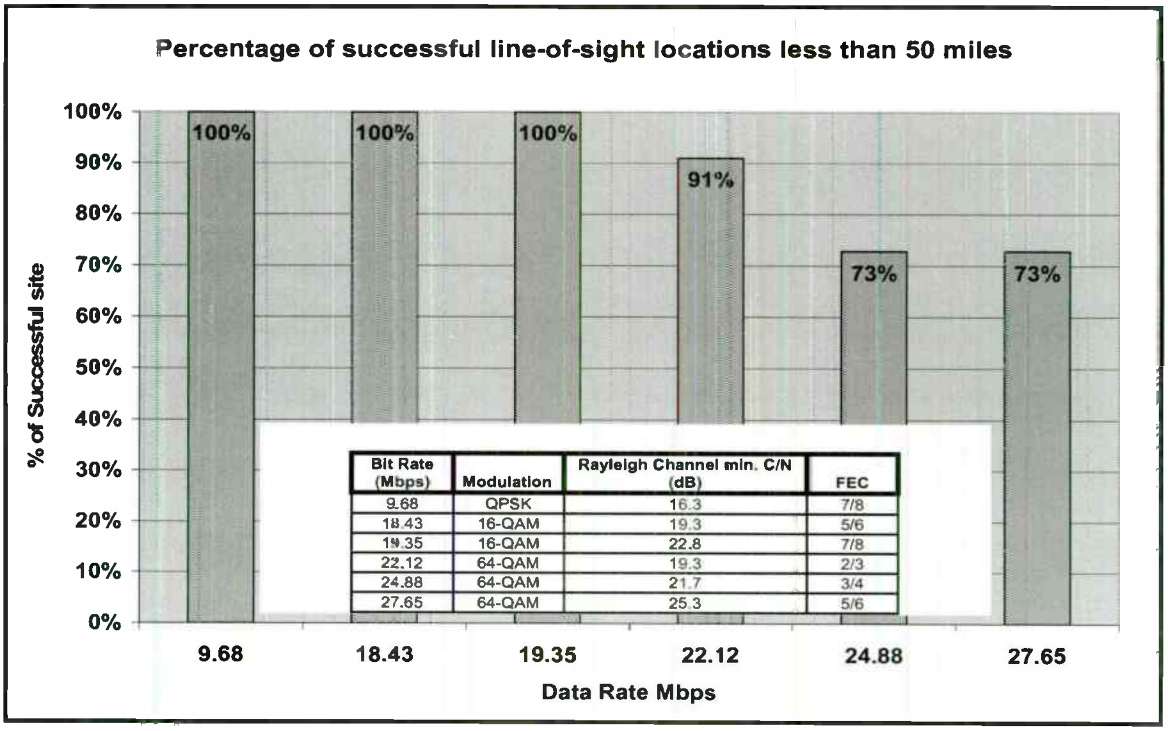
Figure 10 Summary of the Percentage of Successful Sites Less Than 35 Miles

Locations Within a 50 Mile Radius
To determine the distance limitations of each modulation format, transmissions were extended out to a 50-mile radius. Data was collected and analyzed for LOS and bounce sites within this radius and is presented below.

LOS Locations
Figure 11 is a bar chart that plots the percentage of successful LOS locations to a radius of 50 miles. A 100% success rate was achieved for the 9.68, 18.43, and 19.35 Mbps data rates at 1 Watt of transmitted power. As expected, bit rates above 22 Mbps show a distance limitation of less than 50 miles using 1 Watt of transmitted power.

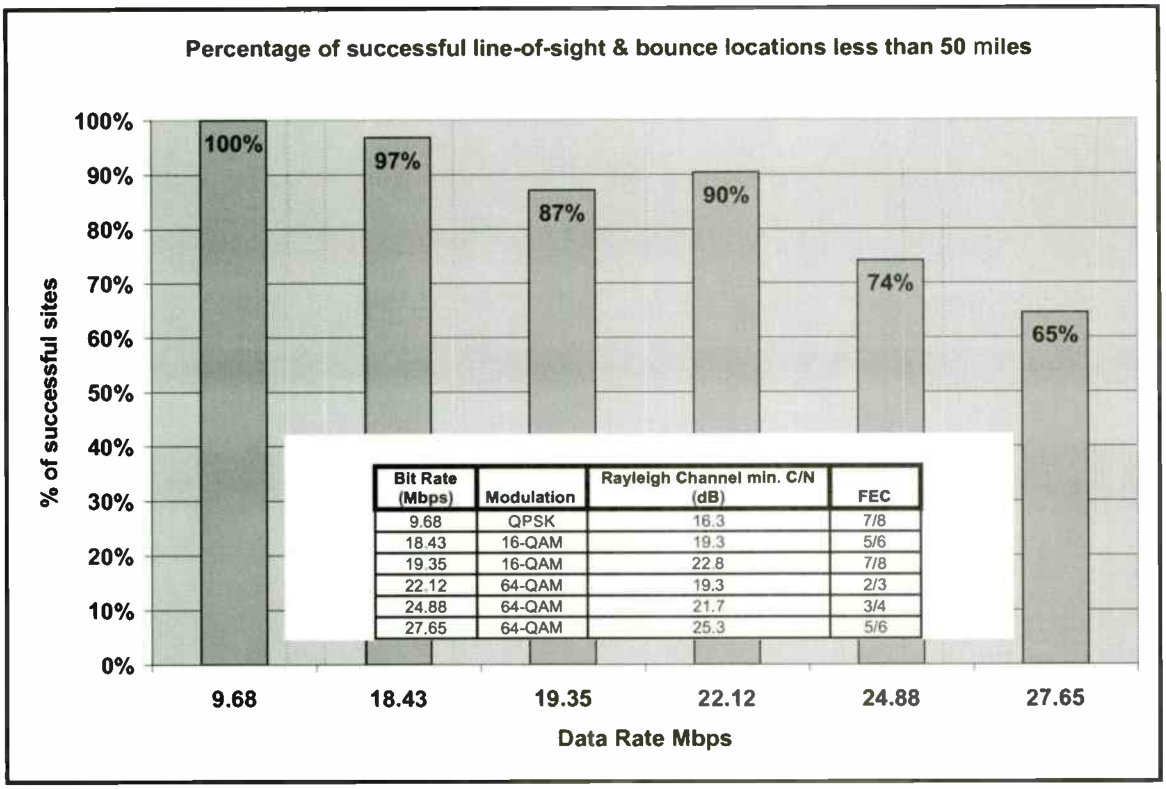
Bounce Locations
The furthest bounce site was 27.5 miles therefore the bounce location analysis is the same as that for the 35-mile radius case described in Figure 8

LOS and Bounce Location Analysis
Figure 12 is a bar chart of the total successful locations for both LOS and bounce sites. The baseline data rate of 9.68 Mbps had a 100% success rate. The 18.43 and 22.12 Mbps signals had a success rate of 97% and 90% respectively. The 19.35 Mbps signal had a success rate of approximately 87%. This lower success rate is again attributed to the higher C/N required for this modulation scheme. The 24.88 Mbps signal had a 74% success rate and the 27.65 Mbps signal had a 65% success rate.



Bit Rate (Mbps)	Modulation	Rayleigh Channel min. C/N (dB)	FEC
9.68	QPSK	16.3	7/8
18.43	16-QAM	19.3	5/6
19.35	16-QAM	22.8	7/8
22.12	64-QAM	19.3	2/3
24.88	64-QAM	21.7	3/4
27.65	64-QAM	25.3	5/6

Figure 11 Percentage of Successful LOS Sites Less Than 50 Miles



Bit Rate (Mbps)	Modulation	Rayleigh Channel min. C/N (dB)	FEC
9.68	QPSK	16.3	7/8
18.43	16-QAM	19.3	5/6
19.35	16-QAM	22.8	7/8
22.12	64-QAM	19.3	2/3
24.88	64-QAM	21.7	3/4
27.65	64-QAM	25.3	5/6

Figure 12 Percentage of Successful LOS and Bounce Locations Less Than 50 Miles

Figures 13 and 14 are plots that demonstrate the effects of distance on the different modulation techniques for LOS and bounce conditions. The plots show the success or failure of a given LOS or bounce site vs. distance. As expected the 18.43 Mbps signal and the 22.12 Mbps signal were able to achieve most successful distance performance due to the minimum C/N requirement. Additional plots for the other bit rates that were tested are listed in the Appendix Figures A1 through A4.

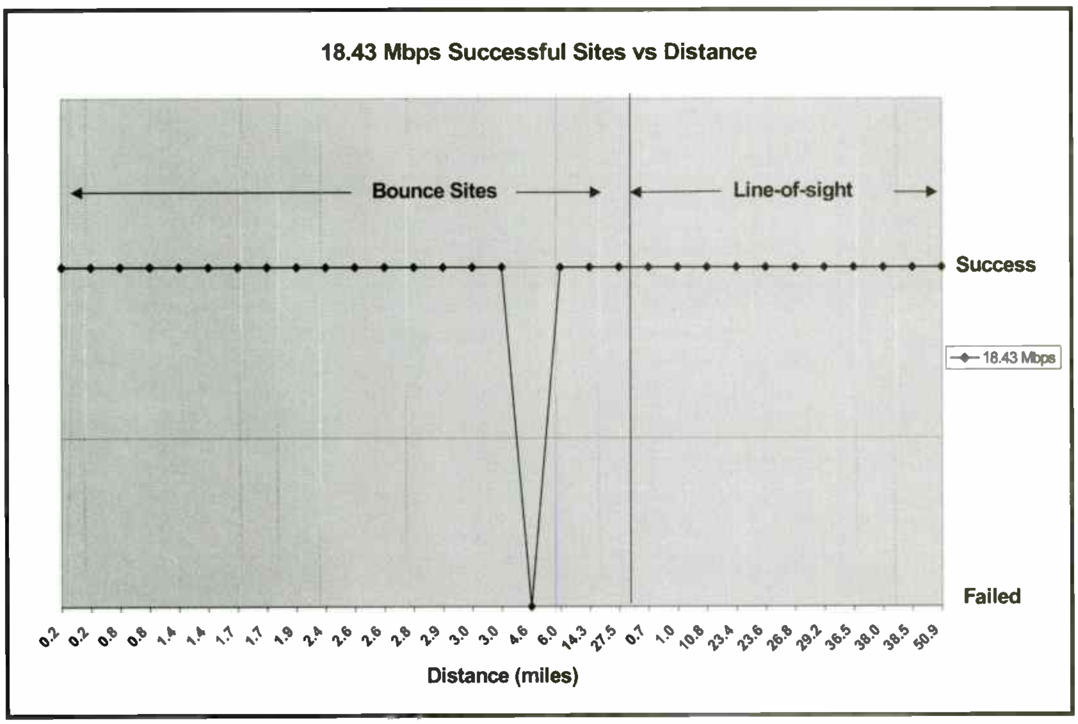


Figure 13 Successful Sites vs. Distance at 18.43 Mbps

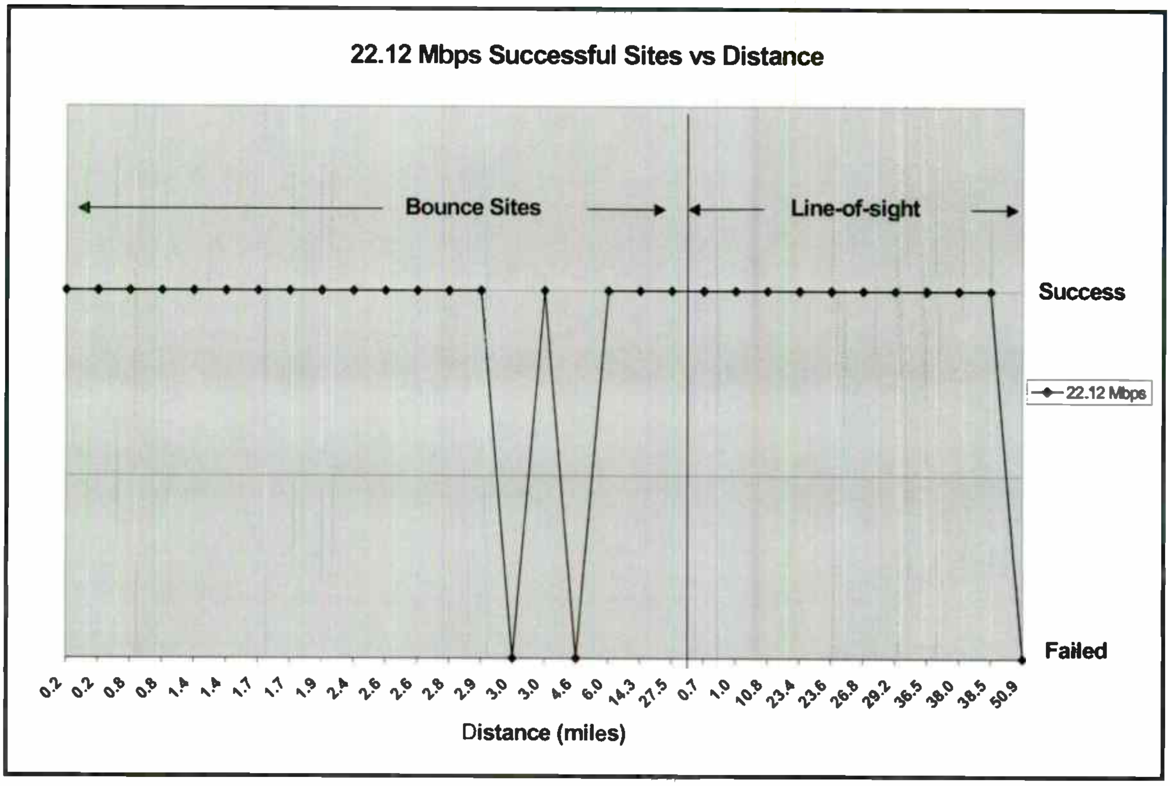


Figure 14 Successful Sites vs. Distance at 22.12 Mbps

C/N Margin Performance

C/N Margin performance for a given site is dependent upon the channel characterization that is assumed and whether a site is LOS or bounce. A Rayleigh channel characterization assumes the highest C/N for a given modulation format and therefore provides the most conservative estimate of margin performance. Figures 15 and 16 are margin plots for the 18.43 and 22.12 Mbps cases and show comparable margin performance. Plots of the other data rate cases that were tested are listed in the Appendix Figures A.5 through A.9. Both the bounce and LOS locations are plotted with increasing distance from left to right.

FEC Performance

The effect of FEC on the received signal is to improve the received signal bit error rate. This can be clearly seen in Appendix Figures A.9 through A.13. Note: the bit error rate of 2×10^{-4} required for QEF reception is indicated by the dark horizontal line. Figure 17 shows the distribution of sites with their associated bit error rate after FEC.

HDTV Data Rate Test Site Failure Analysis

A site was deemed to have failed when there were more than three audio/video disturbances or "hits" during a 5 minute period or when the receivers were unable to obtain a carrier lock, which resulted in no audio video signal.

A failed site could be the result of two possible conditions, insufficient RSL and failure to achieve the minimum C/N ratio for the modulation scheme. There was generally good correlation between a failed site and failure to achieve the minimum C/N or RSL of less than –75 dB. Figure 18 shows the number of failed sites for each data rate and the number of failed sites that did not achieve the minimum C/N or RSL of –75 dB.

High Data Rate 35 Mbps Tests

Figures 19 and 20 are plots of the prototype 35 Mbps non-COFDM modulation test results. LOS sites less than 35 miles performed with 100% success. Bounce sites less than 35 miles were successful 15% of the time. Total successful transmissions less than 35 miles were 37%.

LOS performance for sites less than 50 miles was 72%. Bounce performance for 50-mile sites is identical to that of the 35 miles case as no bounce sites at a distance greater than 35 miles were performed.

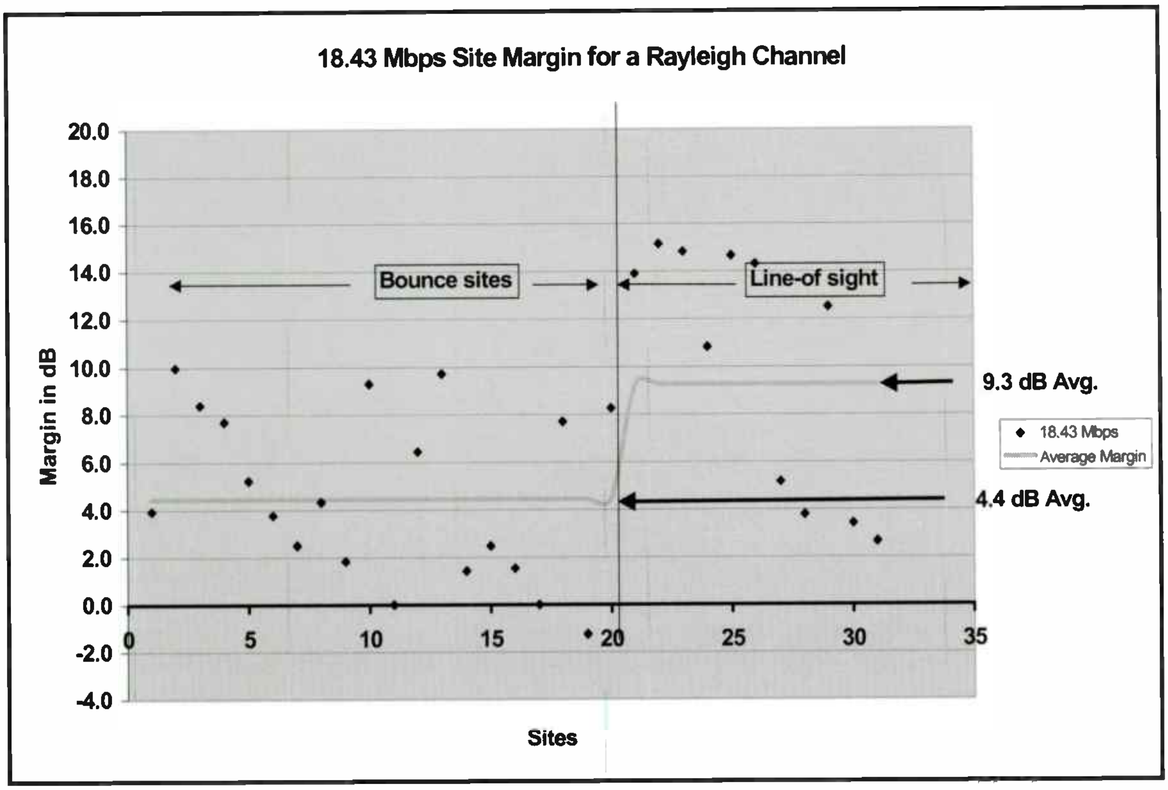


Figure 15 Margin Performances at 18.43 Mbps

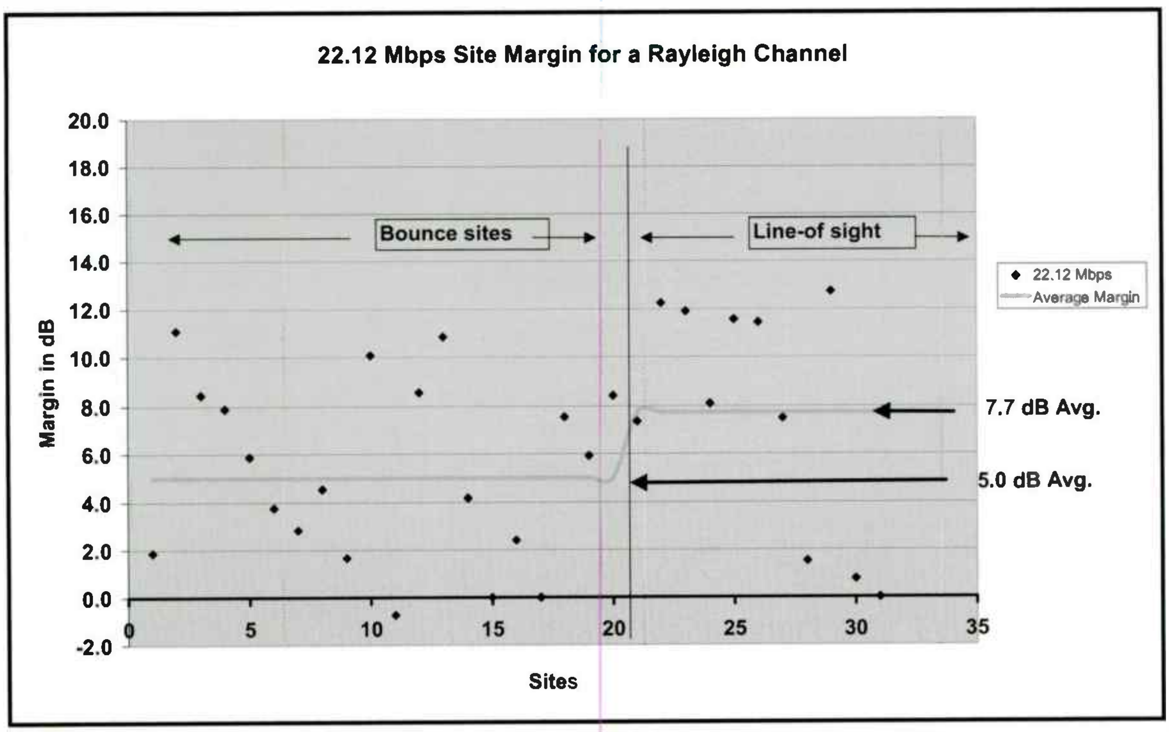


Figure 16 Margin Performance at 22.12 Mbps

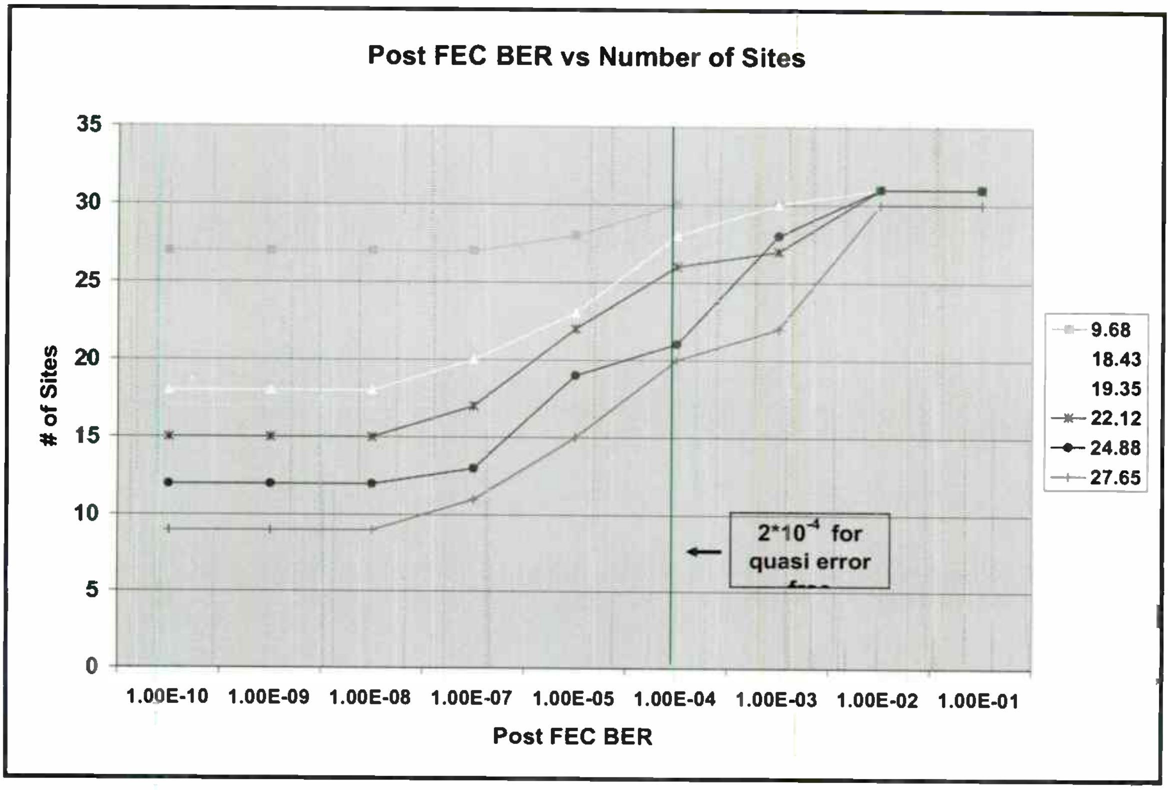


Figure 17 Post FEC BER vs. Number of Sites

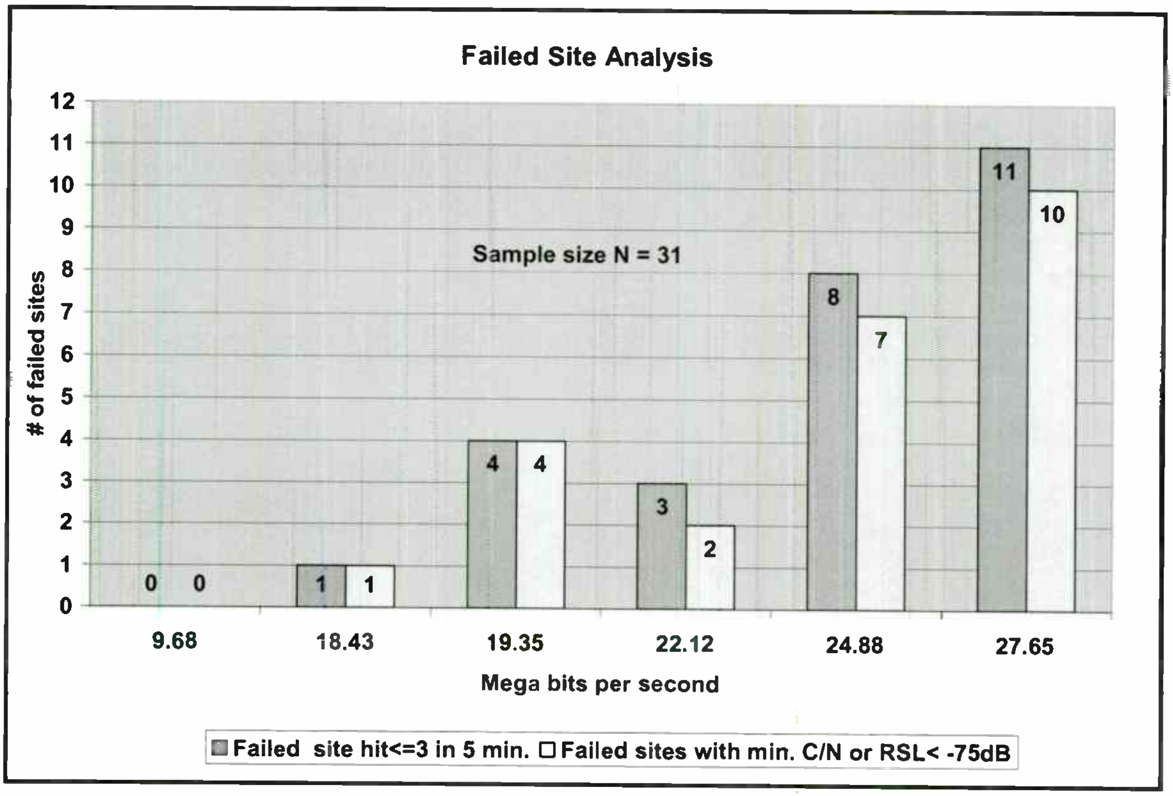


Figure 18 Failed Site Analysis

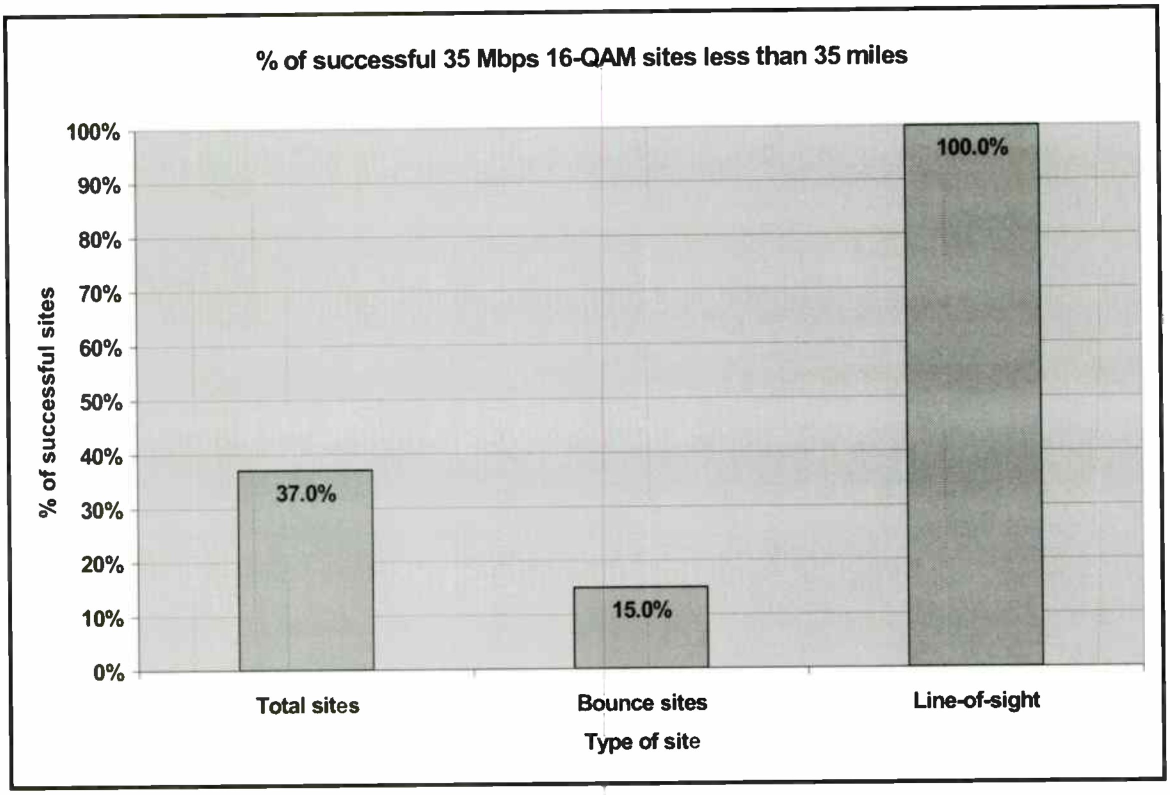


Figure 19 Successful 35 Mbps 16-QAM Sites Less Than 35 Miles

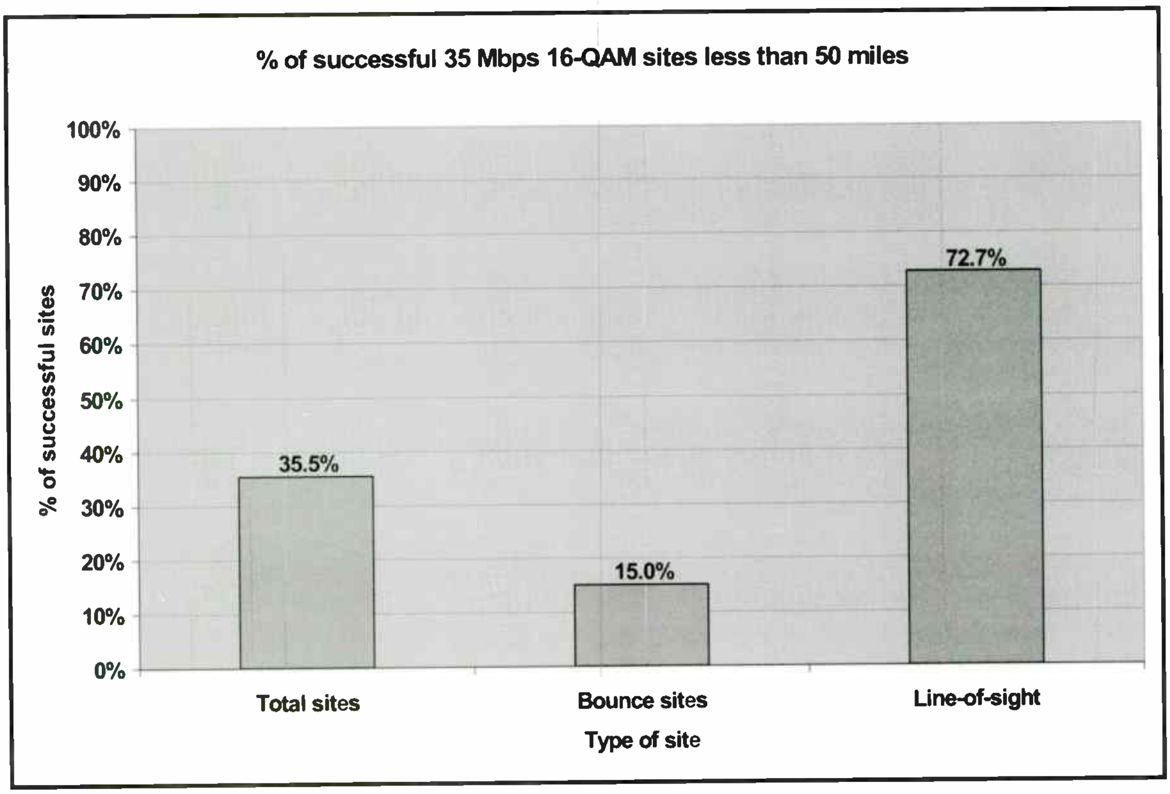


Figure 20 Total Successful 35 Mbps 16-QAM Sites

Adjacent and Co-channel SD Testing

The first tests were with one van (MW 9) at the Javits Center, while the second van (MW 4) transmitted from Weehauken, NJ, approximately three miles away, then at one mile away from ESB at the United Nations and Gracie Mansion on the upper east side of Manhattan. For this test, 8 MHz pedestal, QPSK was used at an FEC of 1/2, GI of 1/8. This corresponds to a data rate of 5.53 Mbps. Video was decoded without anomalies.

Three vehicles drove to Mahwah, NJ a town approximately 25 miles W-NW of ESB.

Tests were conducted using 8 and 6 MHz "pedestals", QPSK and 16-QAM modulation at various settings of FEC and GI with data rates from 4.23 to 14.51 Mbps. Adjacent channel tests were performed with two vans, upper and lower adjacent channel tests were performed using three vehicles. Performance was acceptable.

After establishing a proper link at the Meadowlands Sports Complex in NJ, a LOS path, 6 miles from ESB, one van drove to Mt. Arlington, NJ, a distance of 36 miles from ESB. Combinations of modulation, and data rates of 4.23 to 5.33 Mbps, were tested. Again performance was acceptable.

The adjacent channel tests assessed whether two digital ENG carriers could operate simultaneously.

The first test involved two trucks transmitting 8 MHz DVB-T pedestals in 12 MHz adjacent channels. Both signals were received with an error rate of $1x10^{-7}$ after FEC.

The second test involved two trucks operating 8 MHz DVB-T pedestals in 12 MHz adjacent channels with one truck transmitting with a directional antenna and the second truck transmitting an Omni antenna. Both signals were received with an error rate of $1x10^{-10}$ after FEC and no observed video disturbances for a period of five minutes.

The third test involved both trucks operating 8 MHz DVB-T pedestals on the same 12 MHz channel with geographically diverse receive antennas located on the east and wise side of the Empire State Building. Both signals were received with an error rate of $1x10^{-7}$ after FEC with no visible errors during a five-minute period.

The fourth test involved three trucks operating 6 MHz DVB-T pedestals located side-by-side. (i.e. one at the upper end of a 12 MHz channel and two in the next highest channel). All three were operated at equal carrier levels followed by unequal carrier levels that differed by 6 dB. All three signals were received with an error rate of $1x10^{-7}$ after FEC with no visible errors during a five-minute period.

RESULTS

Many of the most difficult ENG locations in the canyons of New York City; such as Gracie Mansion, United Nations, City Hall, Lincoln Center, Wall Street Area, World Trade Center provided reliable HD-ENG transmission paths using a variety of data rates. The success rate of these tests indicates the HD-ENG using an 8 MHz DVB-T pedestal can provide reliable transmission paths for both line of sight and bounce locations. Because high definition signal quality is of paramount importance, all six of these data rates should be loaded as pre-sets in the transmission encoders and modulators. This will afford the broadcaster the greatest flexibility and the highest quality for electronic newsgathering.

D-ENG transmission with adjacent and co-channel signals present will not present operating issues.

Since the 9.68 Mbps signal produced reliable reception for both bounce and line of sight locations, it can be used as a baseline for initial set-up and antenna positioning. For the LOS locations all six of the high definition data rates ranging for 18.43 Mbps to 27.65 Mbps produced a success rate of 100%. For bounce locations, the optimum data rate for the highest quality HDTV transmission is 22.12 Mbps, which achieved a 90% success rate for locations less than 35 miles. While the 18.43 Mbps rate produced a 5% higher success rate, the 22.12 Mbps and 18.43 Mbps rates only differed by one successful site when the transmission distances were less than 35 miles.

On average, the 22.12 Mbps data rate provided a 5.0 dB margin between the received signal level and the minimum C/N level for a Rayleigh channel for bounce sites and 7.7 dB margin for LOS locations. In most locations, the one-Watt transmit power level provided adequate RSL.

The 35 Mbps, 16-QAM (non-COFDM), transmission system should only be used for LOS locations and is not appropriate for links that involve bounce transmission paths to reach the reception antenna.

CONCLUSION

Microwave transmissions in the 2 GHz ENG band using the DVB-T (COFDM) transmission system with data rates from 18.43 Mbps to 27.65 Mbps can provide reliable High Definition Television links for electronic newsgathering. This range of data rates provides the broadcasters with a variety of HD quality levels that can be employed for "easy" LOS locations or "difficult bounce" locations with no direct LOS to the receiving antenna. The tests determined that the optimal operating point for the highest quality HDTV and best C/N is

22.12 Mbps and is the recommended operating point for HD-ENG.

While 9.68 Mbps using MPEG-2 encoding will not provide acceptable HD video quality for live action images, it can reliably be used for sighting transmit and receive antennas when first establishing the HD link. Recent developments in low latency MPEG4 encoders indicate that 9.68 Mbps could be used to deliver an equivalent quality of a 19 Mbps MPEG-2 signal, allowing the 9.68 Mbps to be used for HDTV. CBS is planning additional field testing to fully evaluate MPEG-4 encoding at this data rate.

ACKNOWLEGDGEMENTS

CBS Engineering would like to acknowledge and thank Jeffrey Birch, and the staff of WCBS-TV, New York, who assisted in these tests – Steve Pair, Peter McGowan and Rich Brody. A special thanks to John Cleary and Jim Fleischman, both of WCBS-TV ENG field operations, whose many years of expertise provided the necessary real world guidance the tests demanded.

Finally, the tests would not have been possible without the assistance and equipment of Microwave Radio Communications and the NTT Electronics Corporation.

REFERENCES

1. COFDM Standards and Bandwidth Testing for 12 MHz Broadcast Auxiliary Spectrum Channels, Prepared by Microwave Radio Communications, for Ad Hoc 2 GHz Reallocation Committee Meeting, Washington, DC, June 18, 2003
2. COFDM Bandwidth Testing for 12 MHz Broadcast Auxiliary Spectrum (BAS) Channels, Prepared by Microwave Radio Communications, September 2003
3. Section 2.0- Overview of Digital vs. Analog Operation, Report prepared by Microwave Radio Communications, January 10, 2005
4. Federal Communications Commission, Rules, Part 74.602, Frequency Assignment
5. ETSI EN 300 744 V1.5.1 (2004-11) Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television

APPENDIX

List of Tables

Table A.1 Test Sites for HDTV Data Rate Tests
Table A.2 Test Sites for Adjacent and Co-channel

List of Figures

Figure A.1 Successful Sites vs. Distance at 9.68 Mbps
Figure A.2 Successful Sites vs. Distance at 19.35 Mbps
Figure A.3 Successful Sites vs. Distance at 24.88 Mbps
Figure A.4 Successful Sites vs. Distance at 27.65 Mbps
Figure A.5 Margin Performance at 9.68 Mbps
Figure A.6 Margin Performance for 19.35 Mbps
Figure A.7 Margin Performance for 24.88 Mbps
Figure A.8 Margin Performance for 27.65 Mbps
Figure A.9 FEC Performance for 18.43 Mbps
Figure A.10 FEC Performance for 19.35 Mbps
Figure A.11 FEC Performance for 22.12 Mbps
Figure A.12 FEC Performance for 24.88 Mbps
Figure A.13 FEC Performance for 27.65 Mbps

Site #	LOCATION	DISTANCE TO EMPIRE (MI)	AZIMUTH FROM EMPIRE (DEG)	AZIMUTH TO EMPIRE (DEG)	B/LOS
21	377 E33 ST NYC	0.7	121	301	LOS
10	TIMES SQUARE	0.8	30	210	B
9	ROCKEFFELER CENTER	0.8	29	209	B
1	JAVITS CENTER	1.0	308	128	LOS
29	UNITED NATIONS 1	1.0	87	267	B
30	UNITED NATIONS 2	1.0	87	267	B
17	COLUMBUS CIRCLE	1.4	11	191	B
16	TAVERN ON THE GREEN	1.7	15	195	B
11	LINCOLN CENTER	1.7	7	187	B
18	TRUMP PLAZA	1.9	357	177	B
15	COLUMBUS AVE AT 79 ST	2.4	14	194	B
31	100 CENTRE ST	2.4	200	20	B
12	WEST SIDE MARINA	2.6	2	182	B
13	MADISON AVE & 84ST	2.6	33	213	B
32	WORLD TRADE CENTER	2.8	212	32	B
6	GRACIE MANSION	2.9	49	229	B
14	PARK AVE AT 92 ST	3.0	34	214	B
24	.SO STREET SEAPORT	3.0	199	19	B
19	32 125 STREET	4.6	29	209	B
3	YANKEE STADIUM	6.0	28	208	B
20	NEWARK AIRPORT- TERM B	10.8	248	68	LOS
2	PRESBYT HOSP	14.3	64	244	B
4	TIFFANY DINER	23.4	341	161	LOS
5	TAPPAN ZEE BRIDGE	23.6	15	195	LOS
25	SEA BRIGHT, NJ	26.8	179	359	LOS
23	PLAINVIEW	27.5	84	264	B
7	EXIT 9 NJ TNPKE	29.2	230	49	LOS
26	ASBURY PARK, NJ	36.5	182	2	LOS
28	ROBERT MOSES PARK, LI, NY	38.0	103	283	LOS
27	BRIGHTWATERS, LI, NY	38.5	94	275	LOS
22	LIE EXIT 63	50.9	82	263	LOS

TEST SITES FOR DATA RATE TESTS

Table A.1 Test Sites for HDTV Data Rate Tests

TRANSMIT LOCATION	DISTANCE TO ESB (MILES)	AZIMUTH FROM ESB (DEGREES)	AZIMUTH TO ESB (DEGREES)	LOS OR BOUNCE SITE
WEEHAUKEN, NJ	2.6	331	151	LOS
JAVITS CENTER, NEW YORK, NY	1.0	308	128	LOS
GRACIE MANSION, NY, NY	2.9	49	229	B
UNITED NATIONS, NY, NY	1.0	87	267	B
MAHWAH, NJ	25.0	340	160	LOS
MEADOWLANDS SPORTS CMPLX, NJ	6.0	313	133	LOS
MT. ARLINGTON, NJ	36.0	291	110	LOS

Table A.2 Test Sites for Adjacent and Co-channel

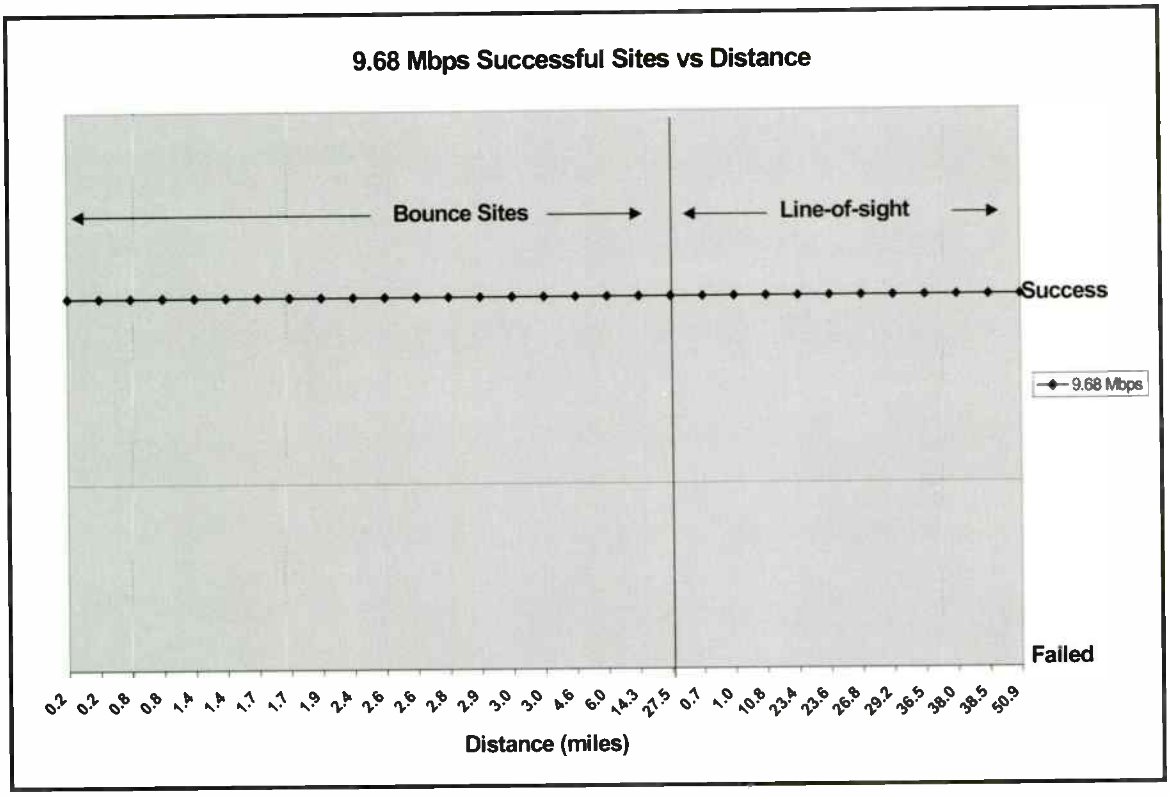


Figure A.1 Successful Sites vs. Distance at 9.68 Mbps

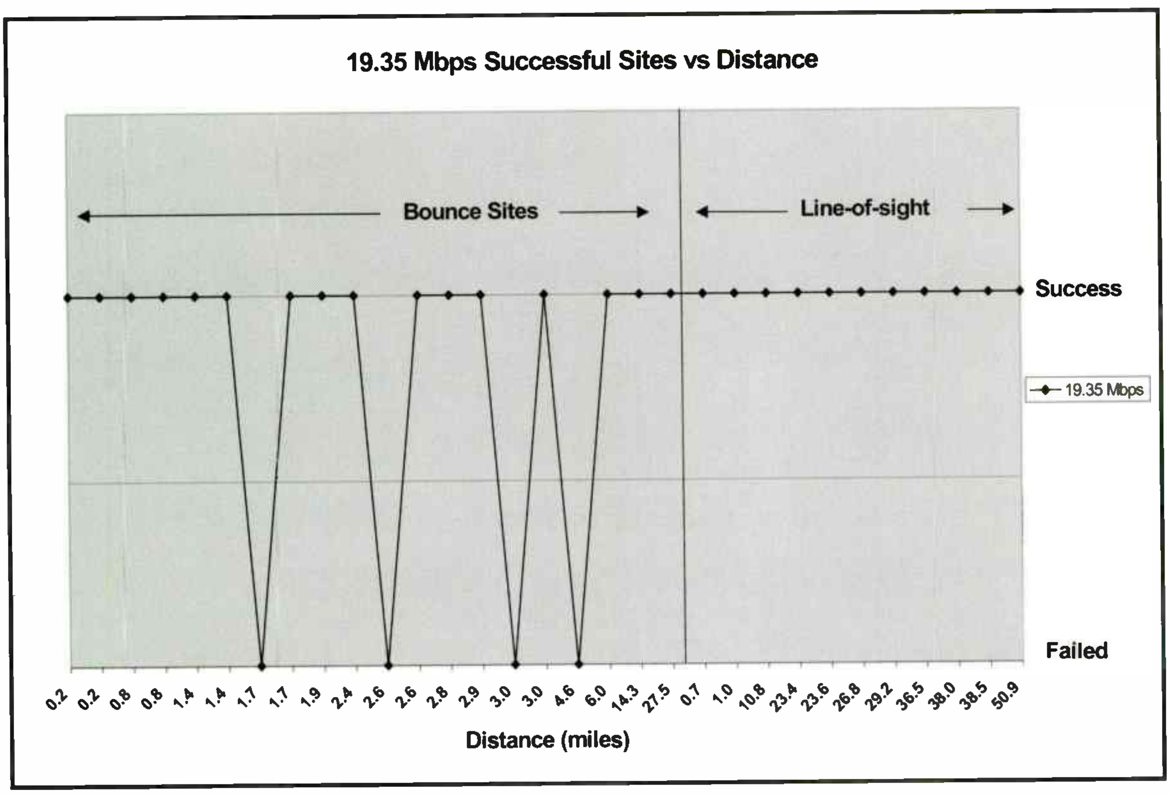


Figure A.2 Successful Sites vs. Distance at 19.35 Mbps

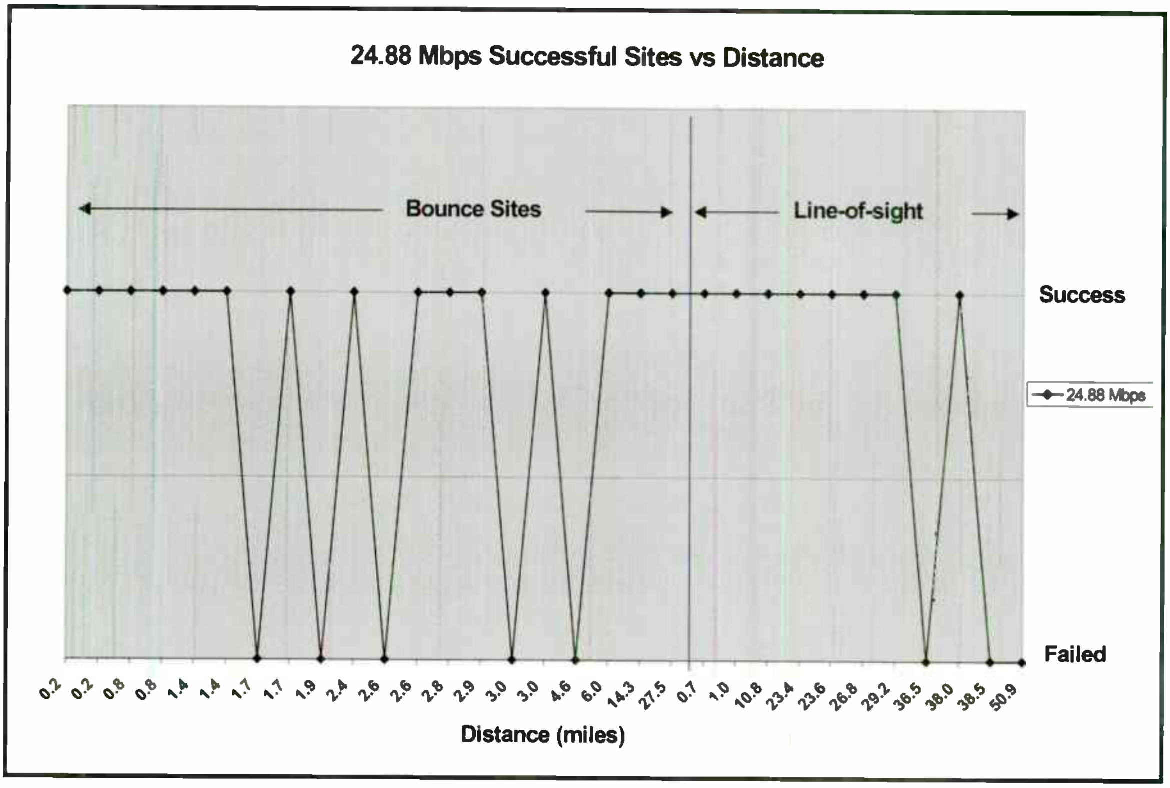


Figure A.3 Successful Sites vs. Distance at 24.88 Mbps

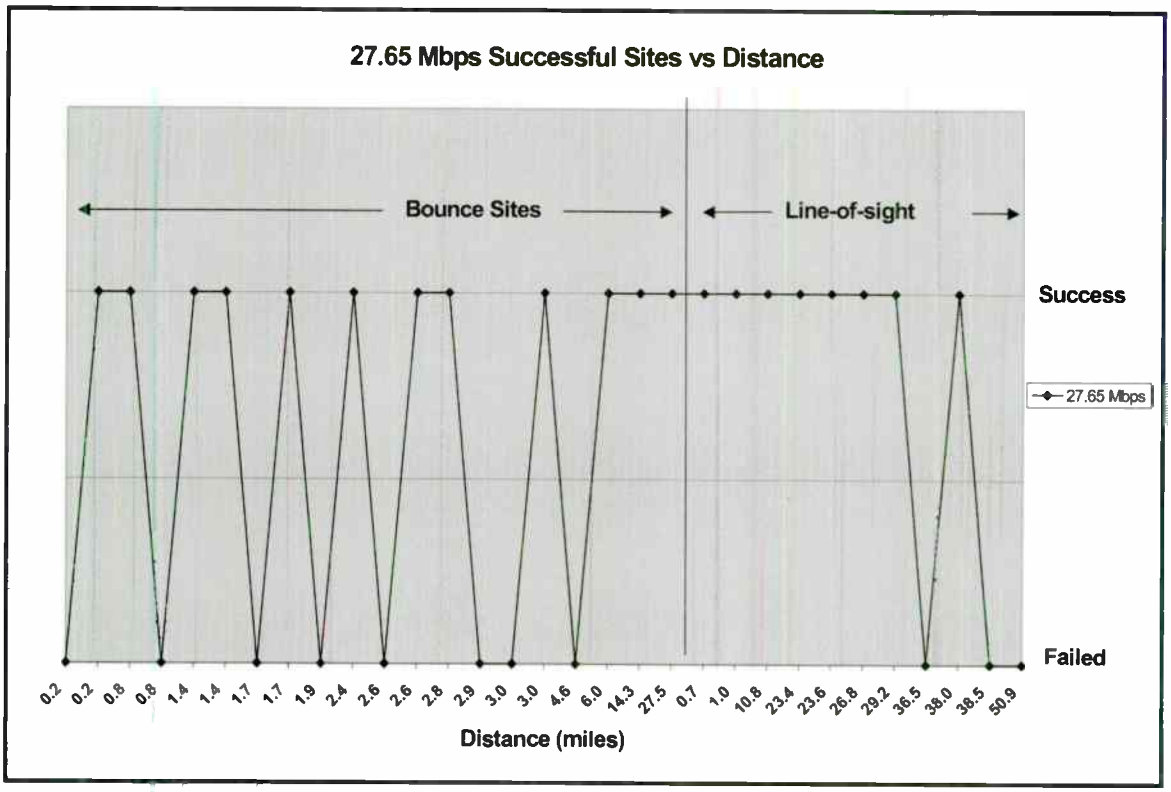


Figure A.4 Successful Sites vs. Distance at 27.65 Mbps

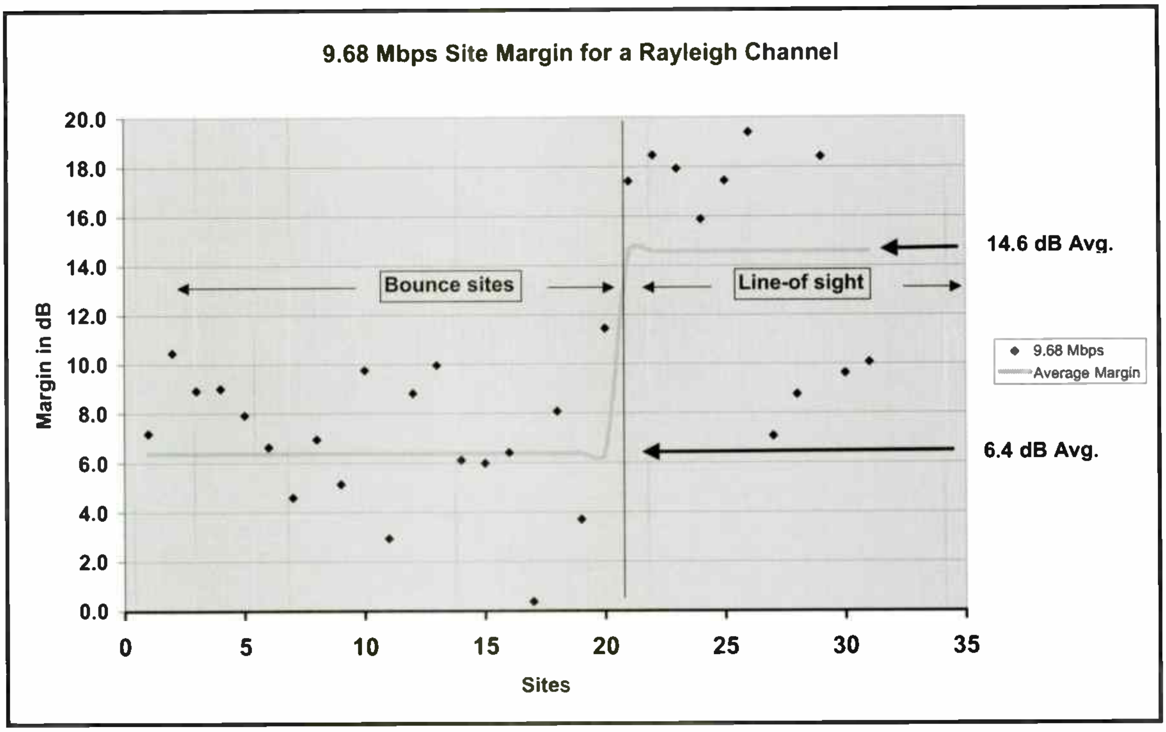


Figure A.5 Margin Performance at 9.68 Mbps

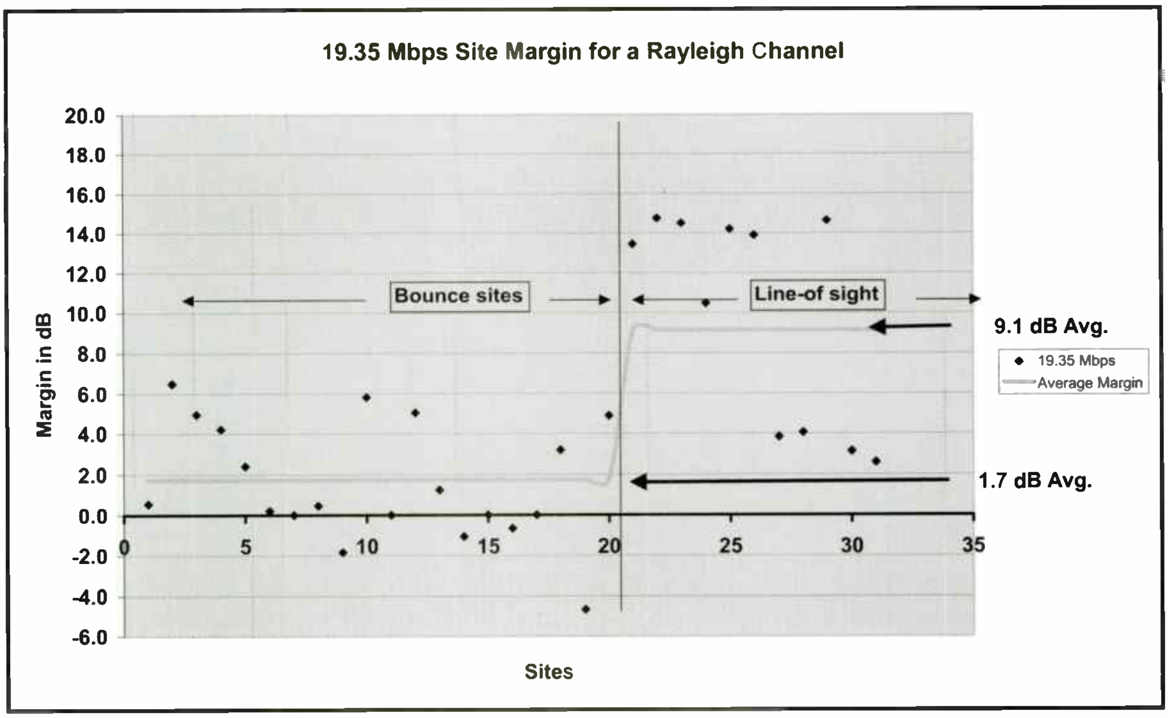


Figure A.6 Margin Performance for 19.35 Mbps

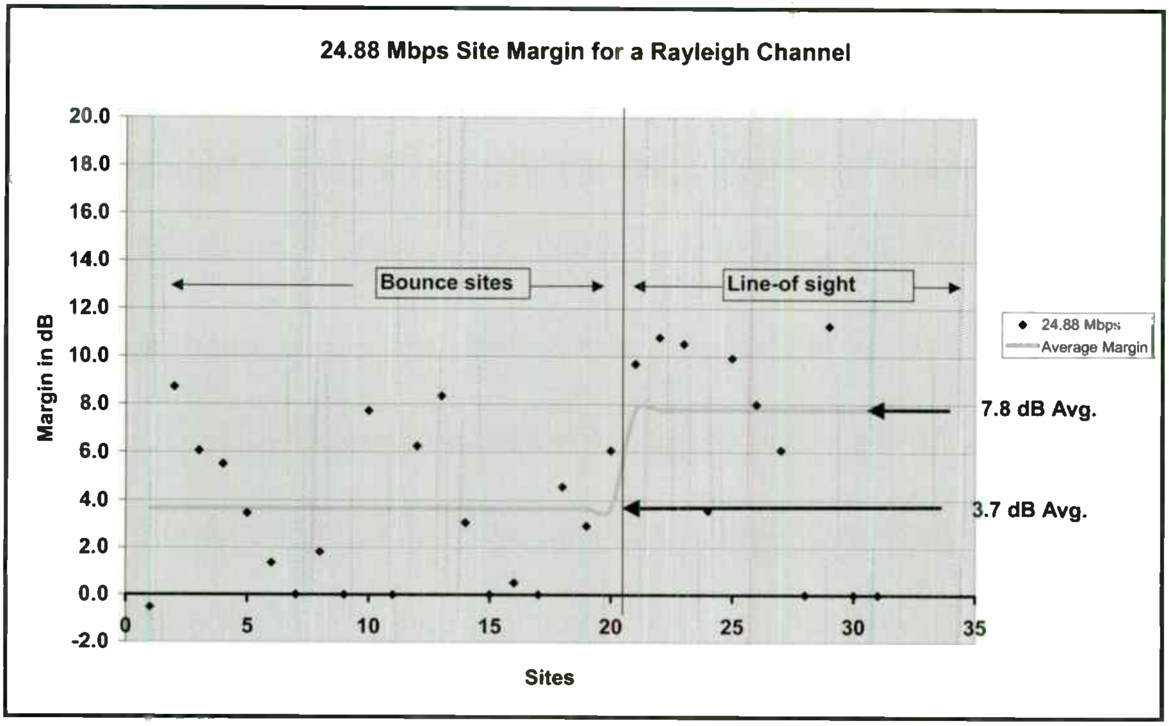


Figure A.7 Margin Performance for 24.88 Mbps

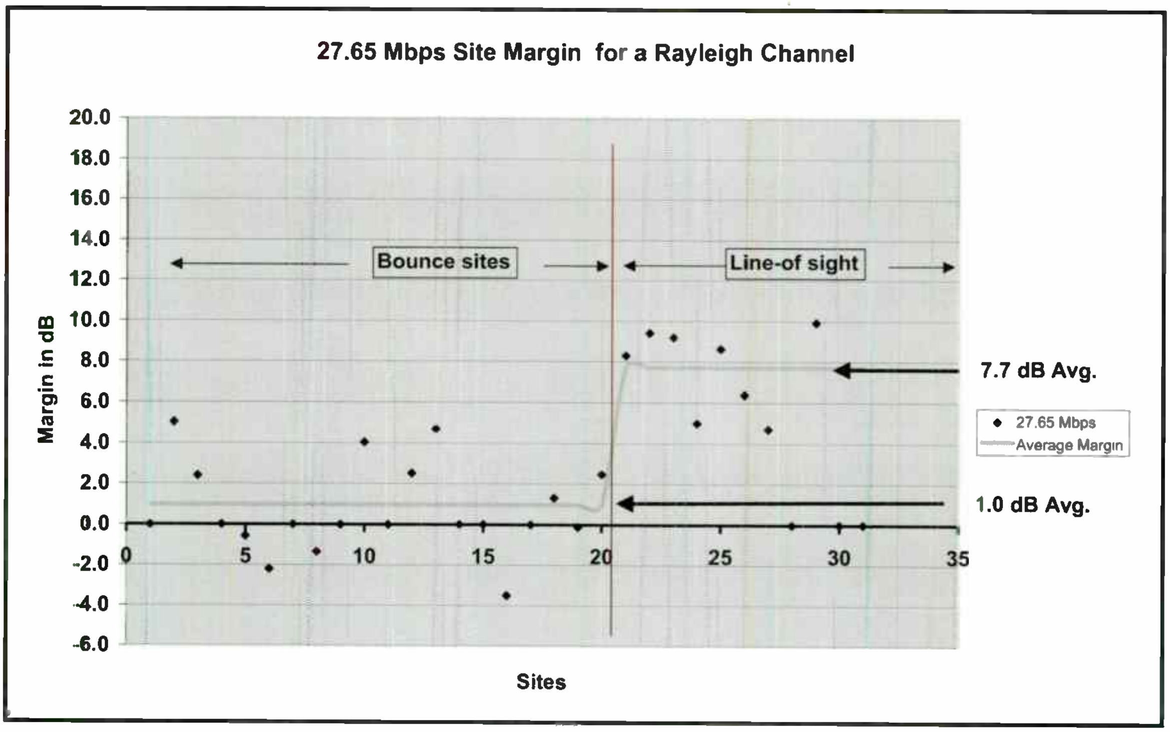


Figure A.8 Margin Performance for 27.65 Mbps

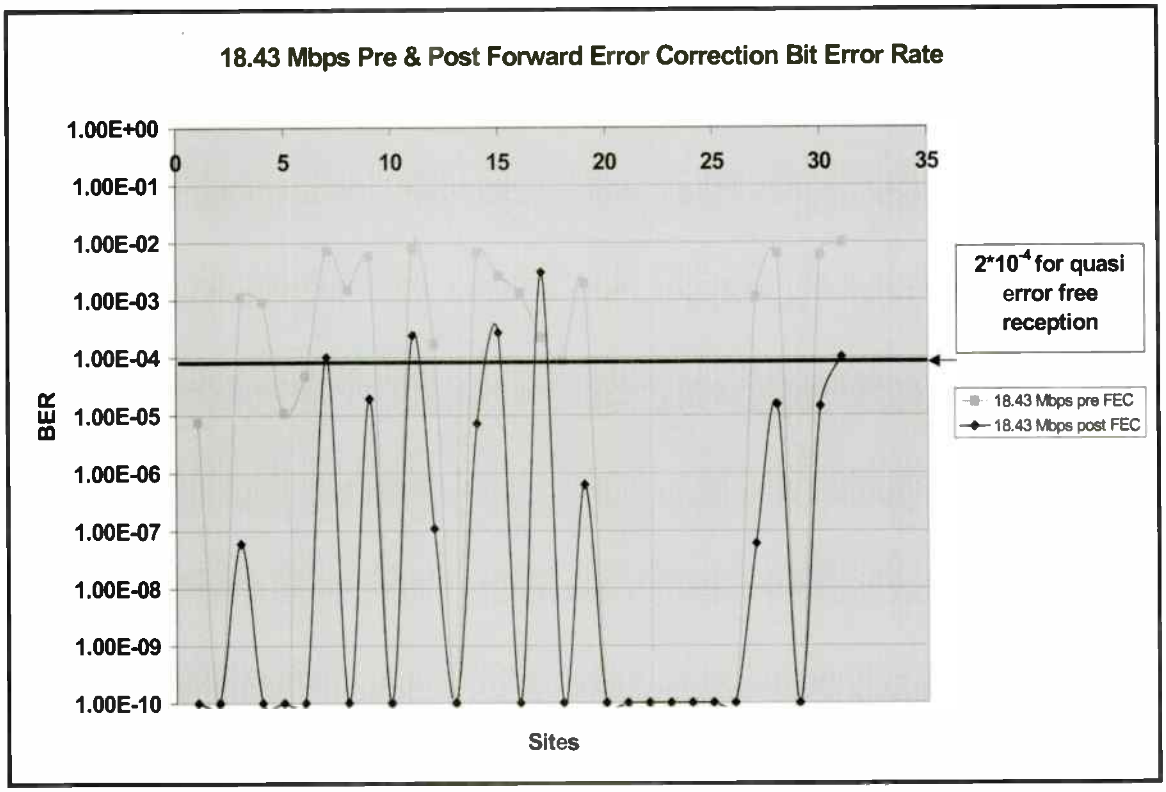


Figure A.9 FEC Performance for 18.43 Mbps

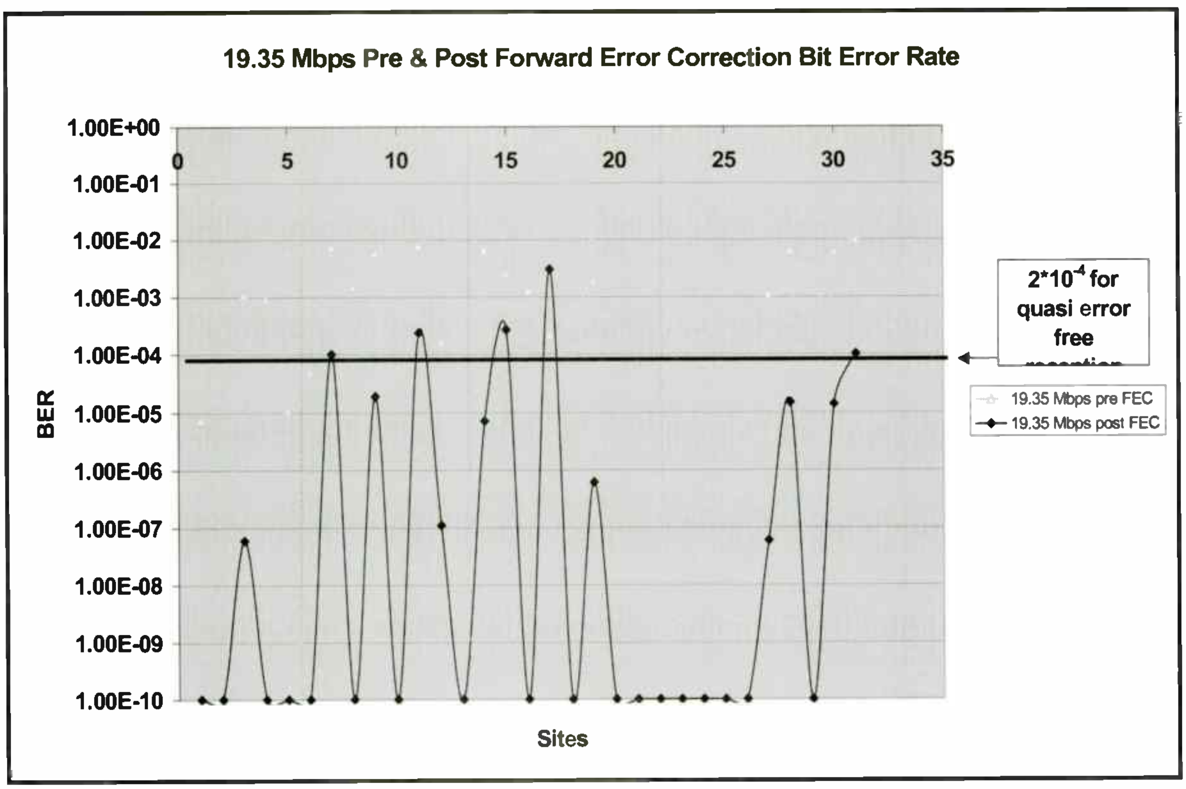


Figure A.10 FEC Performance for 19.35 Mbps

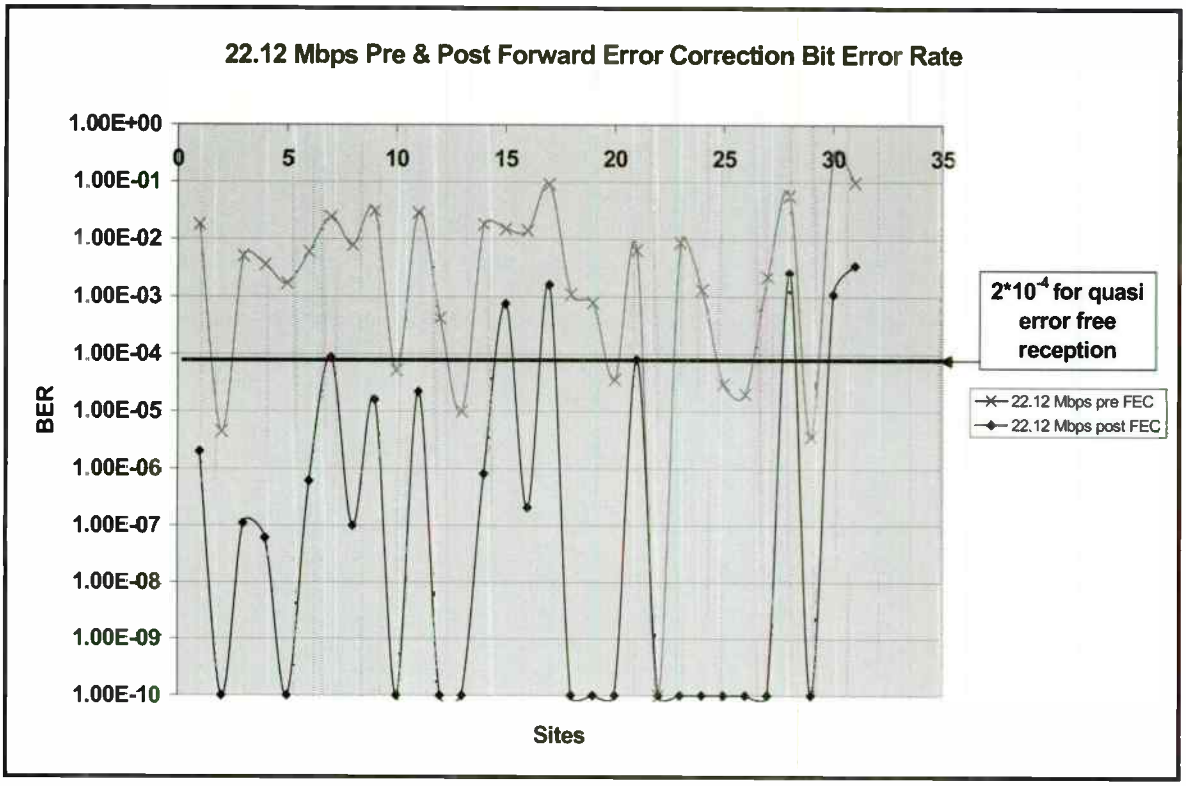


Figure A.11 FEC Performance for 22.12 Mbps

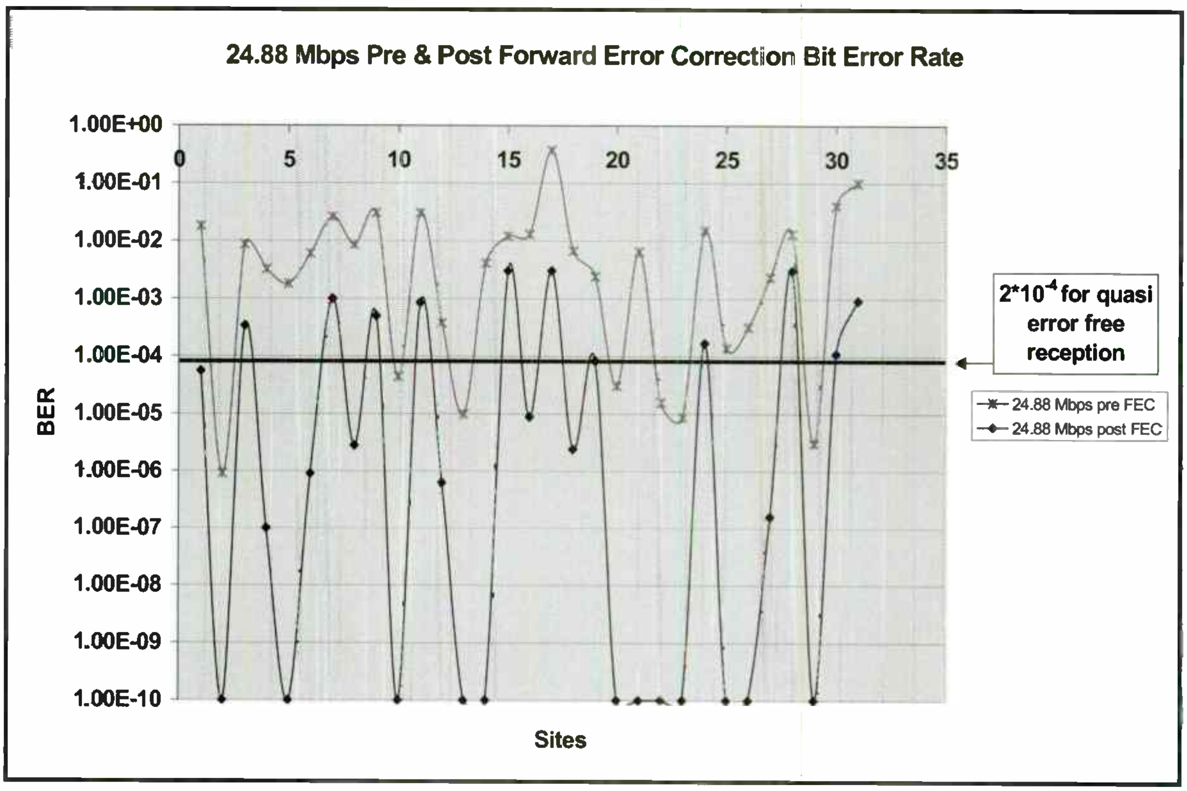


Figure A.12 FEC Performance for 24.88 Mbps

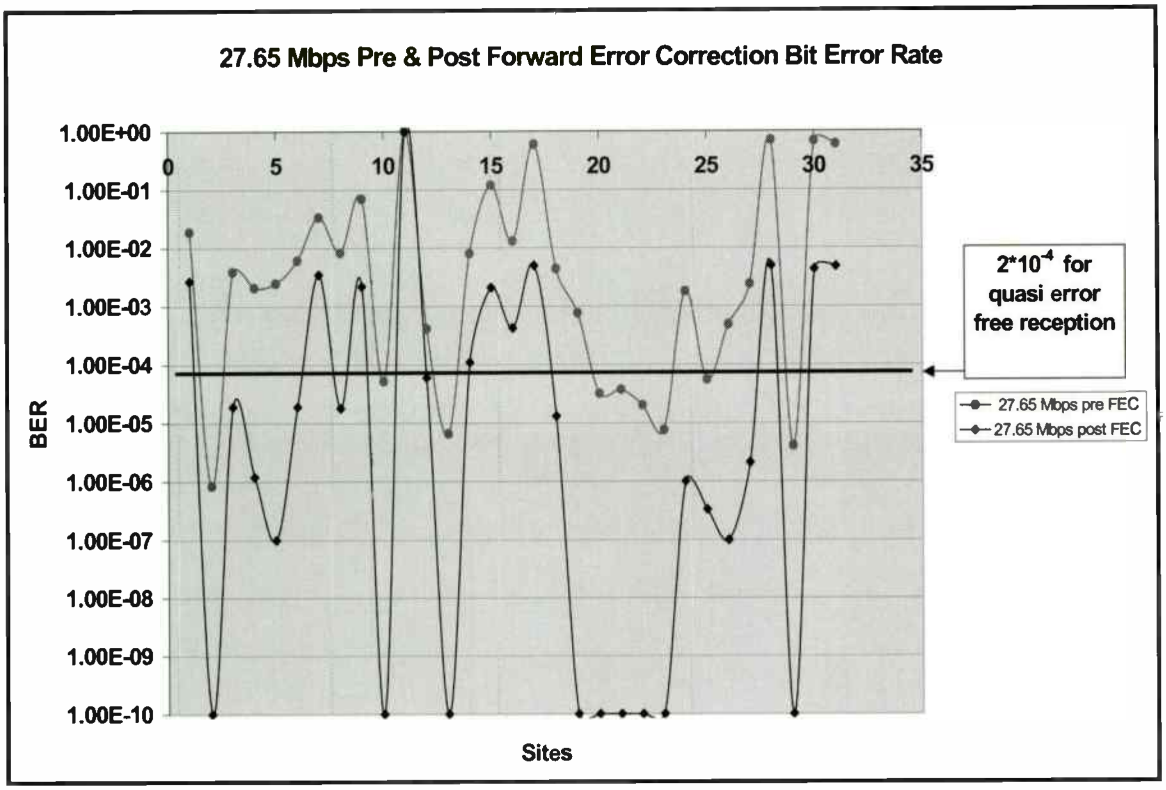


Figure A.13 FEC Performance for 27.65 Mbps

NBC UNIVERSAL'S NEW IPTV DISTRIBUTION SYSTEM

Robert Goldfarb, Director, Affiliate and Distribution Projects
Shadrach Kisten, Director, Technology Service and Support
NBC Universal, 30 Rock, New York

INTRODUCTION

Towards the end of 2006 NBC Universal began to plan for a large-scale effort to consolidate many of our duplicative operations, a plan called TV 2.0. Among the largest components of this plan was the consolidation of News Operations which included the relocation of the MSNBC cable operation from its former home of 10 years in Secaucus, NJ to the landmark headquarters of NBC at 30 Rock in New York City.

This relocation effort included the construction of new studios, production and support facilities, the relocation and re-design of a multi-antenna earth station facility (which could be the subject of an entirely separate paper) and just as importantly, the relocation of over 400 News, Operations and support personnel joining an existing staff of over 300. The resulting 700-plus individuals now will share much of the same space. An important goal was the elimination of most CRT TVs and managing the placement of flat-screen monitors throughout the new space.

At both the former MSNBC facility and the existing operation at 30 Rock, production staff, as well as management, relies heavily on some key support tools. In order to produce the news for multiple networks with dozens of news shows on a daily basis our operation utilizes remote feeds from nearly 200 satellite and fiber provided sources. These feeds, in addition to several dozen internal production sources, all need to be monitored throughout the production areas of the facility.

THE ANALOG PLANT

Prior to the consolidation effort, both facilities had relied upon a common technology to distribute these monitoring signals; analog cable technology. In fact, because of the large number or sources, both facilities actually employed two analog cable (MATV) systems per facility, as each MATV system could only carry a maximum of 125 channels over the 850 MHz plant. This limitation was mostly driven by the existing CATV tuners in most of the older CRT TV sets deployed throughout the facility. [i]

In the case of 30 Rock, the two systems were separated largely by their function. The "main" MATV system is distributed throughout the NBC occupied areas of the building to almost every office and operational area. This system carries most of the local broadcast channels, common basic cable services, as well as many internal production feeds. There are about 120 active analog channels. About two years ago we began to distribute about a dozen compressed HD internal production signals via QAM modulation overlaid on the analog plant. This was quite successful, except it required the use of a digital-ready QAM tuner either built in to a TV set or a dedicated clear-QAM compatible set top box. Compatible set top boxes with both clear-QAM and NTSC tuners remain an elusive commodity.[ii]

The second MATV system was designed solely to distribute the 120 or so Standard Definition remote sources used by News and Sports production. This system had a far limited distribution reaching only technical production areas such as studios, control rooms and editing facilities within the building. In both cases, the system capacities were topped out. With the integration of the MSNBC operation to 30 Rock, there would be some 50 additional remote sources to distribute with no way to squeeze them into the existing architecture.

WHAT TO DO?

We considered several options to our initial challenge. Building a third analog MATV system would have been the least expensive approach and would allow for reusing the existing inventory of CRT TV sets. However, a new coaxial plant would need to be built alongside the existing two coax runs requiring an A/B/C RF switch at each TV. The greatest limitation of this architecture is that there is no potential for growth. It would be limited to Standard Definition analog signals only.

The next approach would be to digitize some or all of the services on the second existing MATV plant allowing increased capacity for new services. The initial benefit would be reuse of the existing coax plant. This option was strongly considered early on because of our previous experience with QAM distribution[iii]. However, this plan would require the use of digital ready QAM tuners at every TV location as well as upgrading the coax plant to reach new locations that were previously not passed. Additionally, not unlike conventional cable operators, we would be burdened with the task of managing a hybrid analog/digital facility to support the existing analog users until a coordinated transition to all digital could be accomplished.

Additionally, several of the biggest limitations of both of these options are: A) lack of portability, that is, the signals would only be viewable within the physical topology of the coax plant – 30 Rock; and, B) many of the "hard-core" users depend on viewing multiple simultaneous sources in common locations. Typically, this had been accommodated by stringing multiple 9" or larger CRTs across desktops, shelves, walls, wherever they could find space for screens. That was no longer a viable option as we were now challenged with many users sharing common space. This meant that physical real estate (per person) was a limitation.

THE SOLUTION

Our solution will utilize an IPTV technology to distribute these monitoring sources. Using a software based decoder, users can access multiple services simultaneously on their workstation screens side by side with other conventional business applications such as email, web-browsing, news editing, etc.

By utilizing the corporate LAN we can potentially provide this service to any user with access to the network both within and outside of the building. This is a great win because the LAN already exists in every area of the facility and requires very little additional wiring. However, none of these benefits come without cost. I'll detail some of these challenges further on.

Our small team of broadcast and IT project managers interviewed a number of different IPTV technology vendors looking for the best solution which included end-to-end support, in other words, we were looking for a single vendor solution. Among our laundry list of requirements were good quality images at high compression rates (important because we are sharing the LAN); ability to support both software based and set top box decoders (STBs) with minimal middleware; ease of encoder management; flexibility with regard to multicast configuration; and, scalability to increase sources, both SD and HD.

We selected a vendor that met almost all of our objectives. We selected Optibase. While Optibase did not manufacture the STBs, they would manage the distribution of select models through our relationship, including those from Amino. While there are several IPTV vendors who manufacture both encoders and complimentary STBs, frequently these required rather sophisticated (and expensive) middleware solutions that were less desirable to this application and beyond the scope of the project. Additionally, most of those vendors had no integrated software decoding support.

NBC has significant experience with MPEG based compression encoding and significant experience with network engineering but neither were a substantial knowledgebase to undertake this project without a "Proof of Concept" test. Optibase was most cooperative with this test and we were able to successfully demonstrate a workable solution on a limited test network.

THE ART OF MULTICASTING

Among our concerns would be the ability to successfully multicast across the network. This is no trivial effort and, in fact, can be considered the "long tent-pole" of a successful deployment. We were fortunate that much of the critical network hardware at 30 Rock was of recent enough vintage to support the multicast requirements. Some of the equipment at many of our satellite locations is not up to par and will require hardware or software upgrades. As such, external deployment is limited, for the time being, to a few manageable locations.

We are employing a state-of-the-art multicasting scheme using IGMP v.2 through our Cisco network. IGMP (Internet Group Management Protocol) allows a given stream to be viewed or "joined" by a large number of users on a given network segment without increasing the required bandwidth. This is far more efficient than a "point-to-point" unicast transmission and certainly more desirable than a UDP "broadcast" transmission which would surely flood a network with unnecessary traffic.

IGMP alone is not enough to ensure robust distribution of multicast streams through the network. A key function of multicasting is the "rendezvous point", a routing methodology that establishes virtual connection points in the network. While not terribly well-documented, we discovered that operating rendezvous points in the PIM (Protocol Independent Multicast) – Sparse mode was essential to a robust network. The sparse mode allows the stream to follow a specific route (or tree) to a specific receiver. Dense mode, to the contrary, tends to flood the network with packets until a path is established, then pruning unwanted routes.

This brings us to – bandwidth. Bandwidth was an absolutely essential driver of this project since we would not have a dedicated "pipe" or network for the stream traffic. We decided early on to share the corporate LAN and in that regard it is critical that we occupy as lean a footprint as possible. Not only to be good neighbors with other applications and traffic fundamental to our business, but to ensure room for future growth as well.

We established that our network backbone within 30 Rock and several of our key operating locations was a minimum Gig-E. In some cases there are larger capacity trunks but we assumed the minimum case for design purposes. Now, this does not include the access layer, where most connections to desktops are 100Mbps. Although this is rapidly changing as new hardware is deployed. In most cases you simply cannot find hardware with less than Gig-E connections. Fortunately, the bandwidth bottleneck is at the multicast point of origin in our case and not the receive end of the link.

Since we started out with the requirement to deliver 200 standard definition services over the network, we planned to keep our footprint below 500Mbps. This meant fairly aggressive compression. While many users would view images on small windows on their workstation, the images still need to be viewed in full via STB on large-size flat-panel displays. Although the images are meant for monitoring purposes as opposed to production elements, reasonable image quality was important. We selected a MPEG-4 H.264 compression engine that yields very satisfactory results. So much so that we decided to set each service for about 1.7 Mbps for a total footprint of about 340Mbps. Well within our targeted "half-Gig" budget. There is a latency of approximately 4 seconds through the system but this was found tolerable for our applications.[iv]

THE HARDWARE

The hardware architecture is a robust platform as well. The encoders are purpose-built and blade-based residing in enterprise class chassis that are connected to dual Cisco Catalyst 6500 class switches. These switches were selected for their ability to handle HD streams across their backplane, making HD upgrades a bit less complicated. The switches uplink, via 10 Gbit/sec fibers, directly to the core routers on the network backbone ensuring high reliability of the multicast origination.

Although the encoders and chassis communications can be configured for N:K redundancy, due to the monitoring nature of the application and available funding, it was decided to operate with single thread encoding hardware. (This is not unlike the lack of redundancy of the analog MATV system it replaces). The design supports future redundancy if so desired. All of the network equipment is redundant and an alternating feed pattern of transport streams from each chassis was used to provide a level of diversity into the redundant network.

All of the encoding hardware and chassis are configurable via a Java based management tool. All of the components communicate via SNMP allowing for full system health status monitoring and alarming. This management tool can be accessed remotely from anywhere on the corporate network. Video encoding parameters can be set individually for each service and include resolutions from D1 on down to QCIF. H.264 bandwidth can be adjusted from 20 Mbps down to 500Kbps. GOP rates are configurable as well. In our case we are feeding all inputs as SD-SDI with analog processing occurring upstream although analog inputs are available to the encoders if desired. HD-SDI will be employed for HD encoders when we expand there.[v]

DECODING OPTIONS

As I mentioned earlier, NBC utilizes several methods of decoding the streams; software based decoders and set top boxes. In both cases, the requirement for good quality image reproduction, audio/visual lip sync and SD/HD compatibility in MPEG-4 were givens. The software solution was significant because of the volume of users and the need for flexibility of viewing streams. One solution that we are employing is a proprietary viewer from Optibase called EZTV. This is Java based and runs in most current browsers such as Internet Explorer and FireFox. It is an ActiveX applet that allows for the concurrent display of multiple simultaneous streams with the added benefit of a centrally managed channel guide.

A user simply selects the desired source from a channel list then chooses the viewing window and the stream begins. Audio is managed by highlighting the desired window. While this approach offers a convenient way to access the streams, the window sizes are limited in scalability, cannot be separated and it adds significant unused area to the desktop that may impede viewing other applications. While we are working with the vendor on a modified version of this app that fills some of these additional needs, we have also deployed a current version of VLC, a popular freeware video decoder that is compatible with H.264. While VLC offers the individual window scalability and simplicity that's desired by our users, it is an open-source code that is not maintained by a vendor and does not offer the centrally managed channel guide.

The VLC application is presently supported on the MAC OS as well. The EZTV web browser version is supposed to run on MAC OS but is in the process of being debugged.

During our proof of performance testing we discovered a critical gating factor to the performance of decoding multiple streams. The more streams in use, the greater the burden on the CPU. It is essential to have sufficient CPU RAM and video RAM, but most importantly a dual-core or greater CPU made the biggest impact to smooth multiple stream decoding. We typically spec a current model PC with a minimum of 3 GB of CPU RAM, 512Mb of video RAM and a 3 GHz dual-core CPU. This allows for four or more simultaneous streams to run uninterrupted. Of course, more streams would require more silicon.[vi]

The channel guide ties the multicast stream addresses to a textual description of the stream channel for easier user access. The guide can be updated and manipulated in real-time and remotely, allowing management to add or delete services as well as change or update channel description immediately, on-the-fly, without the need to restart user's sessions. The guide data and applet management is supported from a centrally located IIS web server that is dedicated to this system.

SET TOP BOXES

The other decoding tool that we use is the STB. We decided to use a low-cost consumer grade unit from Amino that did not require sophisticated middleware to manage. The box has a small footprint, traditional consumer connections (composite and component video and stereo audio on RCA connectors) as well as an HDMI connector. Of course it also has an RJ-45 network connection. Every STB requires a dedicated connection to the local access switch and providing a sufficient number of ports is imperative. The units are compatible with H.264 and are capable of decoding HD (1080i or 720p) when we add those services.

Because we elected to avoid complicated middleware, we do not support such features as conditional access or channel guides on the STBs. Since the channel line-up is shared between the software decoders and the STBs the channel list is only as far as the nearest PC.[vii] These features are able to be integrated should we choose to upgrade to them.

An interesting hurdle that we have encountered with the STBs comes from a fairly uncommon application. We have typically used dedicated analog NTSC cable tuners in most of our control rooms and edit/production areas to monitor the incoming feeds. The tuners used were rack-mountable with a keypad on the front panel to directly select the desired channel. Unfortunately, most consumer-style STBs do not offer this user interface leaving the only mechanism of channel selection to an infra-red remote control. This can be challenging to use when multiple STBs are closely collocated. Fortunately, this can be overcome by using a simple IP based remote control application that communicates to each STB (via IP or MAC address) over the LAN.

CONCLUSION

We believe the future of our IPTV facility is most promising. In the near term, we hope to expand the number of SD and HD inputs as well as building additional "encoding clusters" at some of our other production facilities, all with the common centralized control and channel guide management. We will certainly need to upgrade the WAN both with regard to multicast ability and raw bandwidth requirements so the content can be streamed to our other locations. We even foresee the deployment of "DVR" type applications, which are actually available today; we just need to overcome certain tech support and "rights and clearance" issues. In fact, it is certainly likely that this system will replace all that is left of the existing RF distribution plant and I welcome the opportunity to reduce maintenance and operations cost there.

In summary, with careful planning, detailed research and a dedicated team of individuals a successful IPTV system can be deployed to supplement and even replace traditional RF MATV facilities with greater flexibility, enhanced features and extended range.

I would like to thank several outstanding individuals whose contribution to this project was vital to its success:

Shadrach Kisten, Director, Technology Service and Support, NBC Universal.

Bharat Ananthakrishnan, Director, News and Production Services (TAM), NBC Universal

Colin Campbell, Enterprise Networking Architect, NBC Universal

Eli Garten, Manager, Technical Marketing and Professional Services, Optibase

---

[i] Most of the MATV plant is actually capable of passing 1 GHz, however, since we rely on built-in TV tuners instead of external STBs, the higher frequencies are of little value.

[ii] Initially, there were several inexpensive non-decrypting STBs that supported both NTSC and QAM signals, however, the market for this combination proved of little value as the traditional cable STB suppliers have dominated this segment.

[iii] An additional benefit would be an enormous amount of dedicated bandwidth if we were able to provide 100% digital service immediately, however, we would continue to need much of the bandwidth to support legacy analog TVs for some time.

[iv] The latency can be a challenge for our News production team, especially with live two-way interviews. In many cases, the crew will rely on the original baseband production feeds instead of the monitor feeds.

[v] Of course, HD signals will require increased bandwidth, even using H.264 compression. We estimate about 3 to 4 Mbps per service. Depending on the number of services, the network bandwidth must be carefully managed.

[vi] A quad-core CPU with 8 GB of RAM could display over 20 simultaneous streams.

[vii] The STBs also support a "home channel" that is effectively a browser homepage that can be coded with an HTML channel list.

Digital Radio Summit

Monday, April 14, 2008
3:00 PM – 5:00 PM

Chairperson: Barry Thomas
Lincoln Financial Media, Atlanta, GA

Bandwidth & Frequency Allocation Issues in International Digital Radio AM & FM Broadcasting
Chuck Kelly, Nautel, Ltd., Hacketts Cove, NS, Canada

New Standards and Codecs for European Digital Radio Broadcasting
Olaf Korte, Fraunhofer IIS, Erlangen, Germany

Mobile Coverage Optimization by Polarization Diversity in VHF and UHF Propagation
Myron D. Fanton, Electronics Research, Inc., Chandler, IN

Does Your Yotta Byte?
Andrew Janitschek, Radio Free Asia, Washington, DC

BANDWIDTH & FREQUENCY ALLOCATION ISSUES IN INTERNATIONAL DIGITAL RADIO AM & FM BROADCASTING

Charles W. Kelly Jr.
NAUTEL LIMITED
HACKETT'S COVE, NS CANADA

As Digital Radio is implemented around the world, broadcasters face a key decision: Which digital radio system is right for my country? The answer to that question is complex, and encompasses many technical, regulatory, financial and political issues. This paper does not aim to solve the question, but to examine one facet – the way existing band plans and allocation tables in the current AM and FM bands lend themselves to various technologies.

By limiting the scope of this paper to digital systems which operate within the existing AM and FM bands we do not discount DAB, DMB, DVB-T, DVB-H and ISDB-TSB which require additional spectrum to be allocated. In fact, these technologies are not mutually exclusive to systems which operate within the existing AM and FM bands, and should a country implement one of the systems they will still be faced with the question of whether and how to digitize the stations within the AM and FM bands.

AM IN-BAND SYSTEMS

There are two systems which are available for AM Digital Radio use, Digital Radio Mondiale (DRM) and HD Radio™. While there are profound differences between them in bandwidth requirements, and flexibility, they are both OFDM systems which use a number of sidebands to carry the digital information.

DRM as it exists today, is designed as a LW/MW/SW system, and has considerable flexibility in how it may be configured to meet various bandwidth, payload and robustness criteria. The DRM Broadcasters' User Manual downloadable from www.drm.org is a wealth of information about the various options that DRM makes possible.

In general implementation, however, DRM is normally installed with a 9 kHz bandwidth (carrier +/- 4.5 kHz), on a single channel, or 18 kHz bandwidth (carrier +/- 9 kHz) which utilizes one channel plus half of the two adjacents. There are options for using adjacent channels for an analog AM broadcast, but this isn't in common use, as some tests have indicated that the analog signal must be greater than 16dB above the DRM field strength at the receiver to avoid interference from the DRM signal into the analog receiver.

There are also proposals to implement a SSB AM signal and use the other half of the channel bandwidth (formerly occupied by the other sideband) – but this is predicated on synchronous detection of the AM signal, and in general, shortwave receivers are much more likely to feature synchronous detectors than are AM radios.

HD Radio was developed to meet the unique needs of the US broadcast market – and is based on the 10kHz AM channel spacing found in North and South America. In the US, the AM band is full of stations, and there aren't a lot of open frequencies to put a stand alone digital signal. On the other hand, adjacent frequencies are not normally allocated within the same area, so HD Radio was designed to keep the AM analog signal essentially unchanged (just band limited to 5kHz audio response) and to utilize the spectrum beneath the analog signal, plus on the two adjacent channels for OFDM sidebands to pass the digital component.

HD Radio fits the needs of broadcasters in the US, because it allows full simulcast of the legacy analog signal as well as the new digital signal. Eventually, when digital receivers make up enough of the receiver installed base, the analog signal may be shut down, and the HD Radio signal will be reduced in bandwidth to 9.6kHz (+/- 4.8kHz).

HD Radio reduces the effect of digital to analog on-channel interference because the sidebands above the carrier frequency are out of phase to those below the carrier frequency and are identical – thus they cancel in the analog receiver. While this is helpful for the on-channel signal, it does not affect the interference to adjacent channels.

There is no question that HD Radio can create significant interference to co-channel and adjacent channel signals, and depending on the IF bandwidth of the receiver, on the second adjacent signals as well. This has the effect of reducing the secondary coverage areas for a number of AM stations – in some cases, dramatically.

MARKET CONSIDERATIONS INFLUENCE DIGITAL STRATEGY

If the channel spacing in a country is 9 kHz, and there are free channels available for allocation, it is a simple matter to implement DRM. Many European countries fall into this category, as reduced AM listenership has resulted in some AM stations being taken off air – and these "dark" facilities may easily be re-purposed for DRM.

In some cases, however, HD Radio has been implemented successfully for market reasons. In Surabaya, Indonesia in 2006, a religious broadcaster bought three adjacent AM channels and installed an HD Radio AM station on the center frequency. He broadcast the same programming on the analog and digital transmissions, and used the analog channel to promote the purchase of HD Radio receivers, some of which are made in Indonesia. In this way, they provided higher quality audio programming to their loyal listenership, while having just one transmitter, one antenna and paying a much less expensive license fee from the government than an FM station would have cost.

Figure 1 shows how a standard HD Radio AM signal fits into a 9 kHz spacing band.

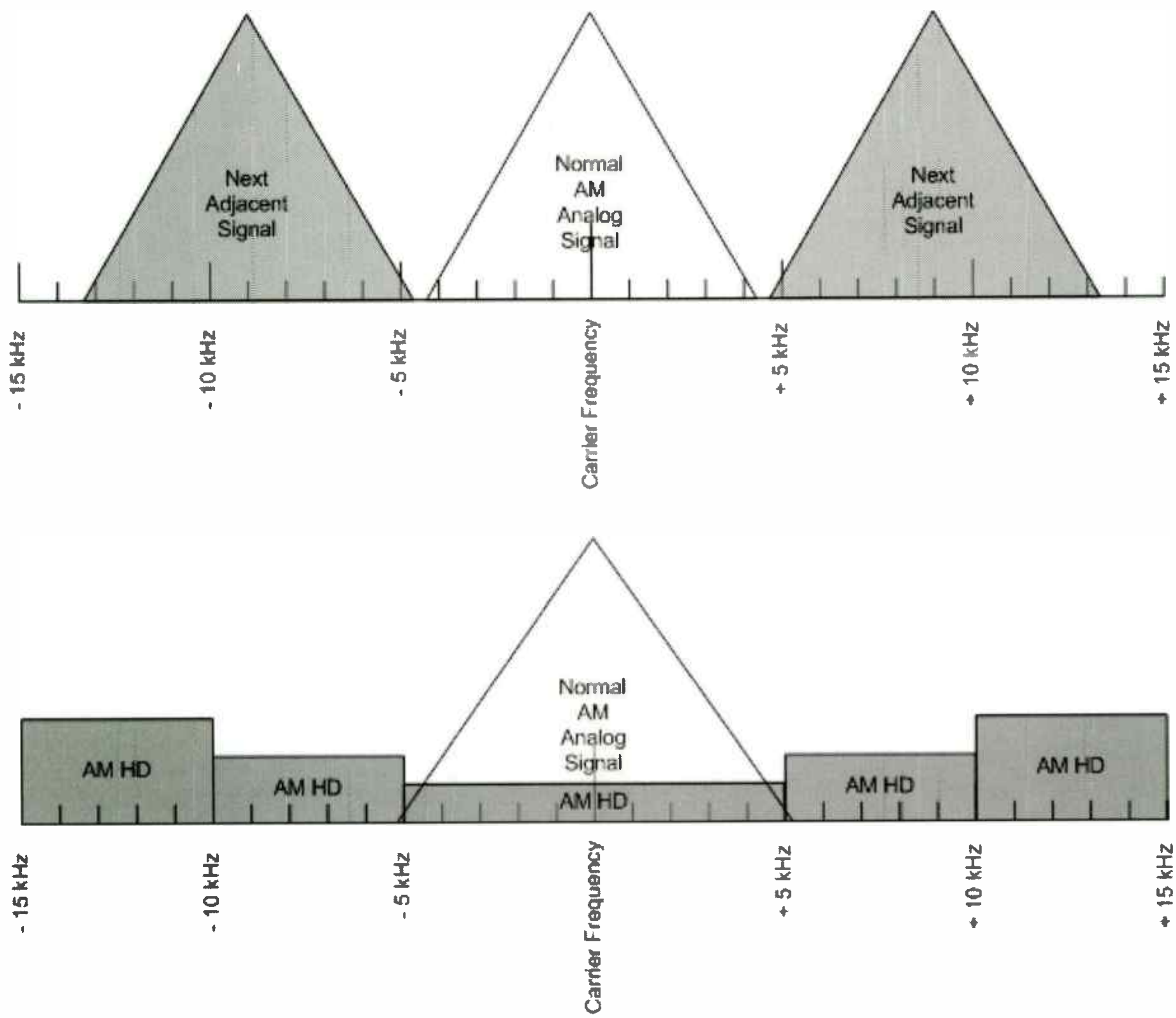


Figure - 1

Where the AM band is fully utilized, and spare channels are not available to construct new digital only DRM channels, HD Radio is a possibility. An example is the Philippines, which like all of Asia, has AM channel steps of 9 kHz. Figure 2 is a graphical analysis which presumes each station is running HD Radio in Manila, the most populous metropolitan area, and with the most crowded spectrum in the country.

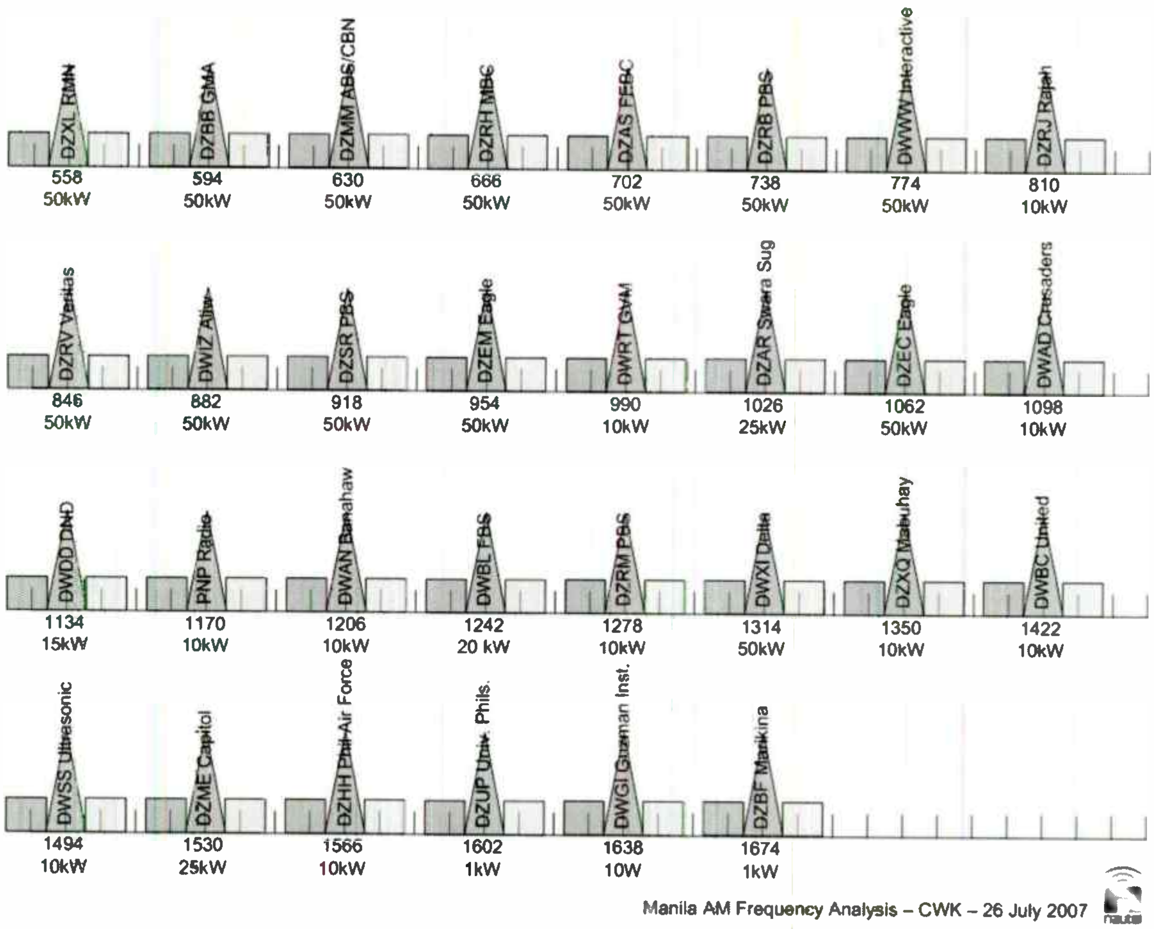


Figure - 2

As can be seen in Figure 2, the channel spacing within the Metro Manila area is 36 kHz between stations and a guard band of 6 kHz exists between the AM HD sidebands.

While Figure 2 is compelling, there is additional analysis needed before any station is approved for use with HD Radio AM. As these are existing stations co-channel interference to other stations on the same frequency has been presumably been known for some time. With the HD sidebands extending to +/- 15 kHz from carrier however, both of the first adjacent channels will be impacted, and to a lesser extent, the second adjacent channels will also be affected. Figure 3 is a graphical analysis of the physical locations of first and second adjacent stations in the Philippines relative to one 50kW AM station.

Figure - 3

As can be seen from Figure 3, the nearest station on the first or second adjacent channels to 558 kHz in Manila, is DZMQ, a 5kW station on 576 kHz in Dagupan City, nearly 200kM away from Manila. Calculating the power in each of the AM HD Radio sets of sidebands, approximately 375 watts from the 50kW Manila station falls within the 2nd adjacent channel of DZMQ – nearly 200kM distant.

A more serious problem exists with 1062 kHz – figure 4 below:

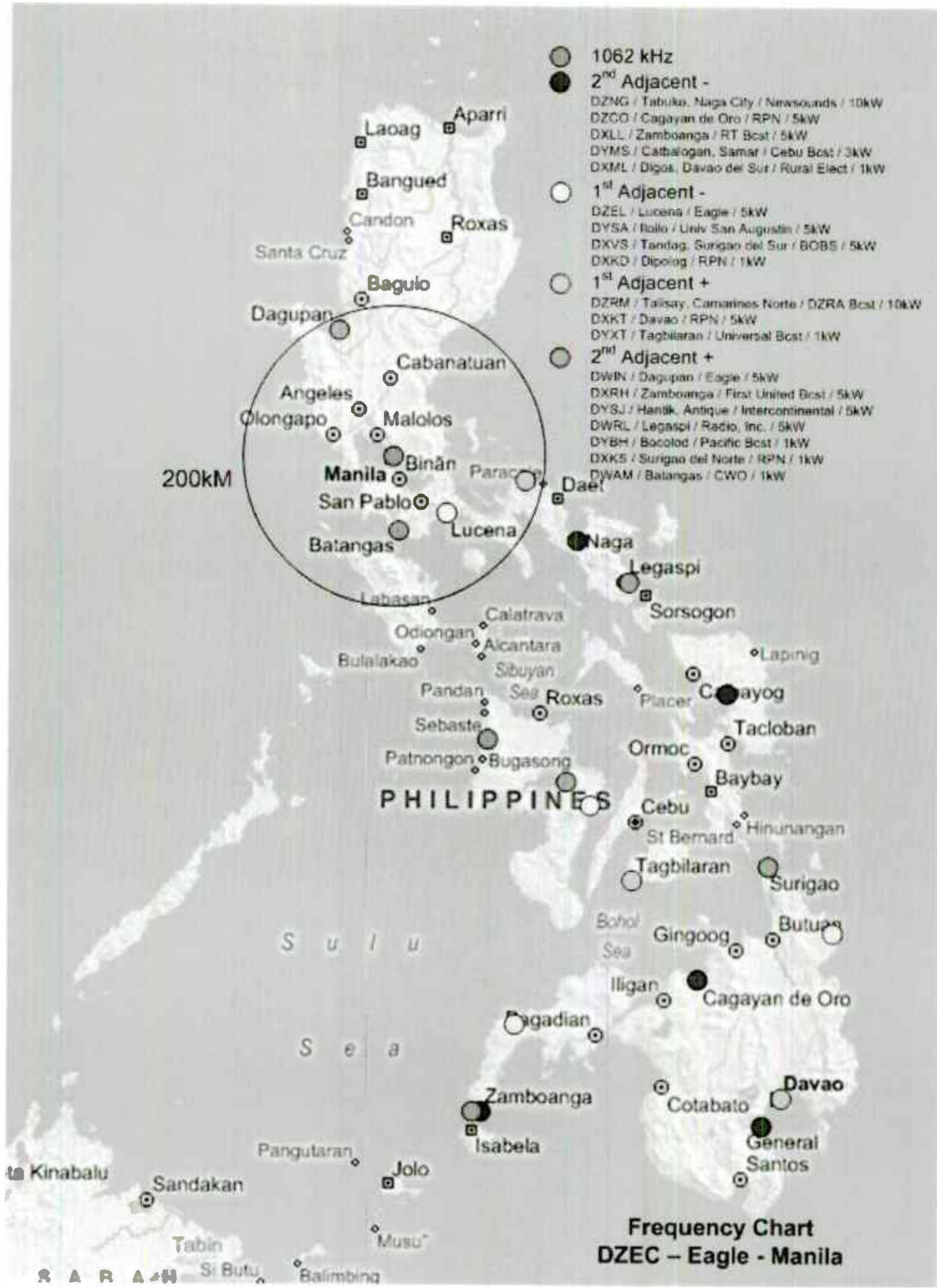


Figure - 4

In Figure 4, we see that there is a 1st adjacent station, DZEL 5kW, in Lucena on 1053 kHz. This station, only 100kM distant from Manila, would receive nearly 1,130 watts from the HD Radio sidebands of 1062 kHz. It is likely that some secondary coverage areas of DZEL would be affected, and conversely, HD Radio coverage of DZEC in the Lucena area could be affected. It is important to understand that the digital sidebands above the carrier and the mirror image ones below the carrier carry the exact same information – thus, if interference garbles the ones below the carrier, the receiver can fully demodulate the program audio – with reduced robustness.

In adaptation of digital radio in these circumstances, there are likely to be compromises to secondary coverage areas, but the tradeoff is that within the primary service areas, the broadcaster is able to deliver dramatically improved audio performance, as well as text data.

<u>FM In-Band Systems</u>

There are three Digital In-Band systems – and they are all dramatically different. HD Radio works a lot like the AM HD Radio system, with OFDM carriers on either side of the analog FM carrier, so both the analog and digital signals are on a single FM channel. DRM+/DRM120, currently in development, are digital only options, with the OFDM carriers occupying 96 or 100 kHz of FM spectrum. FMeXtra utilizes OFDM carriers added to the composite baseband of a conventional FM analog signal – and thus the bandwidth of the FM signal does not appreciably change. Interestingly, FMeXtra may be operated on an existing FM HD Radio station – thus increasing the possible total digital payload.

At this writing, HD Radio is currently on the air in over two thousand stations, FMeXtra is on the air at around 100 stations, and DRM+ is in testing stage at several stations.

Figure 5 is a graphical depiction of the FM band in Seoul, Korea – showing that FM HD Radio is compatible with all existing Seoul stations. In addition, it can be seen that many DRM+ channels could also be fit into the band.

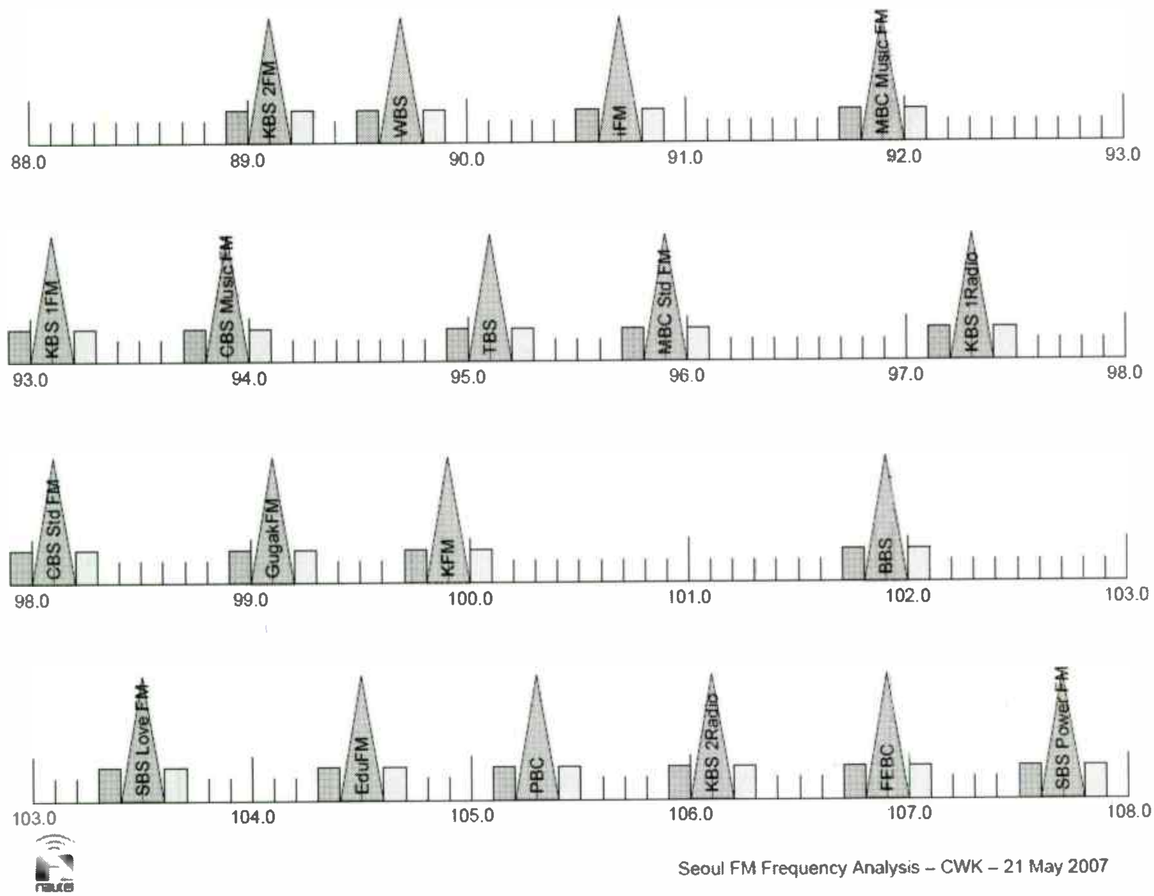


Figure - 5

The FM spectrum in Taipei, Taiwan, is depicted in Figure 6 below. Between 88 and 89 MHz, there are four educational FM stations which are spaced by only 200 kHz. As can be seen, the HD Radio sidebands are on top of each other, and it is likely that poor coverage would result, with HD Radio reception possible only where one station has a significantly higher field strength that the adjacent channels.

Between some of the stations, such as the ones at 89.3 MHz and 87.7 MHz, there is only 400 kHz separation. Theoretically, this will work, however the reception of one of the sidebands may be impacted if the immediately adjacent sideband is significantly higher in received strength. This condition is not unlike blanketing interference in analog FM.

Note that in HD Radio FM, the sidebands above the analog carrier and the ones below the carrier are carrying the same information, while there may be interference in one set of sideband – it may be possible for the receiver to properly decode the HD signal on just the clear set of sidebands, albeit with less robustness.

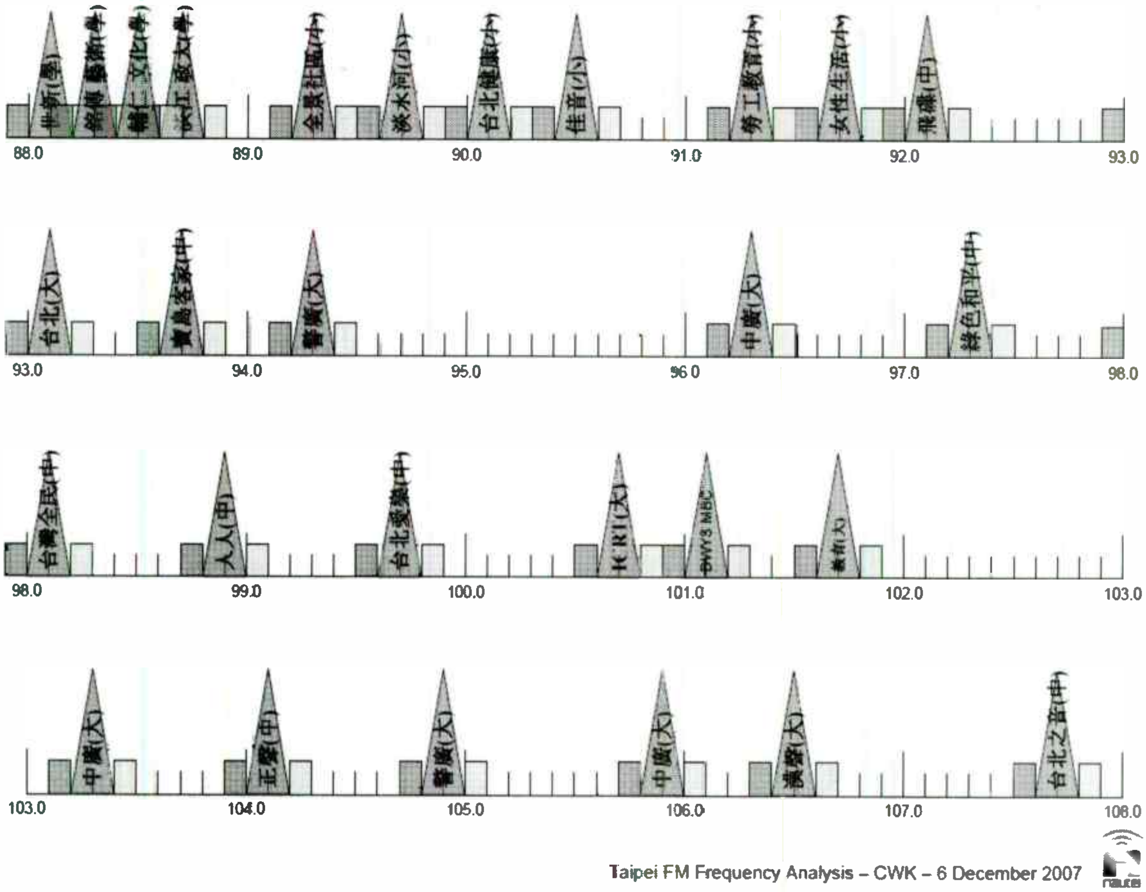


Figure - 6

In some cities, stations on the FM band are spaced too closely for either HD Radio or DRM+. For instance, the metropolitan Istanbul, Turkey market has FM stations virtually every 200 kHz from 87.5 to 108 MHz. In this situation, it may be that the only workable in-band digital system would be FMeXtra.

Summary

The selection of a digital radio standard encompasses much more than simply the technical capabilities of the systems. An examination of frequency usage, both within the market, as well as far field, can reveal how a successful transition can be made with minimal interference.

Acknowledgements

The author is indebted to:

Mr. Tim Hardy, Nautel Limited, Halifax, NS Canada

Mr. Glen English, Pacific Media Technologies, Cooma NSW Australia

HD Radio is a registered trademark of iBiquity Digital Corporation.

NEW STANDARDS AND CODECS FOR EUROPEAN DIGITAL RADIO BROADCASTING

OLAF KORTE
Fraunhofer IIS
Erlangen, Germany

ABSTRACT

While HD Radio is being deployed in the U.S., Europeans are deploying new versions of DAB and DRM and also introducing new systems such as DMB and DVB-H.

This paper will explain the background of these systems and their operational appeal to both public and private broadcasters, contrasting them to the efforts being made to introduce HD Radio to Europe.

The fragmented European market makes it difficult to introduce a common digital radio standard for all of Europe. While DAB is a great success in some countries, others have switched off their transmissions and are evaluating alternative solutions.

New satellite broadcasting systems for mobile reception will also be introduced within the next years and also Internet Streaming and IP based transmission for mobile reception will become more and more important.

The main focus of this paper will be on terrestrial broadcast systems.

CURRENT SYSTEMS IN USE

The Classic: DAB

The original goal of "EUREKA 147 DAB" was the replacement of analog FM, though it is used in different frequency bands (Band III 174-230MHz and L Band 1452-1492MHz). The system was officially launched at IFA 1995 in Berlin.

DAB supports single frequency networks for covering large areas and bundles several independent programs into one "Ensemble". A DAB transmission has a bandwidth of 1.5 MHz and contains typically 6 to 10 audio programs that are encoded with MPEG-1 Layer II audio coding.

The combination of Layer II with MPEG Surround has recently been added to DAB to enable the backwards-compatible transmission of high quality multichannel audio. MPEG Surround offers the broadcaster a way to broadcast both multichannel audio and stereo simultaneously with a high quality and nearly the same bit rate as stereo only. This is important as it saves capacity – and money – on the transmission channel.

Besides audio, a wide variety of additional data services, used as program associated data (PAD) or program independent stand-alone data (N-PAD) can be used in parallel. Typical examples for these data service types are textual information services like Dynamic Label and Journaline®, Electronic Program Guide (EPG), Slideshow or Broadcast Website. For telematics services the well known RDS-TMC (Traffic Message Channel) can be used as well as the new TPEG services which give much more flexibility with regards to telematic information services. For high precision location based services the transmission of Differential GPS is possible. In Germany there are several cities that use DAB for the transmission of text information, pictures and video clips for information displays installed in public transport. Other proprietary data services are also in use via DAB.

Conditional Access can be applied for pay services or closed user groups but is rarely used up to now. This is expected to change with the commercial introduction of telematic services in the "TISA" project, which is based on the results of the "Mobil.Info" project.

For various reasons, DAB is a great success in some European countries like UK and Denmark, while it is not present in others like Sweden that have switched off DAB. In the first years of DAB, the lack of cheap receivers was the main problem. Later on, in some countries like Germany the restriction of transmission power caused major problems for in-home reception as well as for mobile use. Another problem was that there were not enough attractive services on air to attract potential listeners, because of the lack of frequencies. Most of these issues have been solved in the meantime in those countries that still have DAB on air. Germany now plans a "Big Bang" in 2009 with a massive rollout of new additional services over whole Germany.

In recent years the technical standard has been extended by the two major amendments DMB and DAB+. Both did not change anything in the basic channel coding scheme. This makes it possible to have a mixture of DAB, DAB+ and DMB within the same DAB Ensemble.

Visual Radio and Mobile TV: DMB

The first DAB extension named DMB was presented by BOSCH Blaupunkt in 1995. The idea was to use the DAB system for transmission of TV programs into high speed trains and cars. The system was based on MPEG-2 video coding and did not reach a commercial status.

One decade later the Korea developed and introduced Digital Multimedia Broadcasting (DMB) as an extension to DAB for Mobile TV applications. For video coding MPEG-4 AVC (H.264) is used. For audio coding MPEG-4 HE-AACv2 or in the Korean market BSAC for is used. For the implementation of data services BIFS (Binary Format for Scenes) is part of the DMB standard, but no BIFS profile has been defined up to now. Within one DAB Ensemble a mixture of DMB, DAB and DAB+ services can be carried.

DMB is a great success in Korea. The European market is still under development but the availability of equipment and cheap receivers is promising.

In Germany DMB has been introduced and is used for Mobile TV and also for Visual Radio which in fact uses a very low frame rate for the video part to transmit still pictures. The DAB and DAB+ equivalent for this application is the use of audio in combination with a slideshow as PAD service.

France wants to introduce DMB as a unique system also for radio-only services. Their main argument is, that in contrast to DAB+ the receivers are already available. But the use of audio without a video component is not covered by the current DMB standard and therefore an extension is required. A first approach for such an amendment has been rejected by WorldDMB, as DAB+ offers a more efficient solution for radio-only services and DAB+ receivers will soon be available for the market.

Other countries in Europe have started trials or plan to introduce DMB. There is also an intensive discussion about potential competition of DMB with DVB-H. DMB offers less services in parallel but network costs are lower. So also a combination of both - DVB-H for urban regions and DMB for selected programs with country-wide coverage - could make sense.

State-of-the-art audio coding for DAB: DAB+

This extension of the DAB standard was named DAB+ and was approved end of 2006.

To open new markets which require state-of-the-art bandwidth efficiency WorldDMB adopted HE-AACv2 to be used as an alternative audio codec instead of MPEG-1 Layer II. This enables broadcasters to transmit 2-3 times more audio services with HE-AACv2 at a bit rate equivalent to one MPEG-1 Layer II audio service. From the beginning DAB+ will also support the transmission of multichannel audio by using MPEG Surround.

The first commercial receivers will appear soon and Australia will be the first country using the DAB standard only with DAB+. Other countries already using DAB are planning to setup new services in DAB+ in order to offer a larger variety of services.

For radio DAB+ has some advantages in comparison to a potential DMB radio only mode as discussed in France. First DAB+ includes all kind of data services that are defined for standard DAB are also valid for DAB+. Secondly DAB+ requires less overhead than a DMB radio only would cause.

As already mentioned for DMB, DAB+ can be used within the same Ensemble together with DAB and DMB.

For markets where DAB still has difficulties, such as Germany, the introduction of DAB+ has two sides. On the one hand it makes it possible to increase the service capacity by the factor 2 to 3. On the other hand, the use of DAB+ is restricted to new receivers. This is mainly an issue for the car industry which needs several years to update their models to the DAB+ standard, whereas the highly integrated DAB receivers in already delivered cars cannot be updated to DAB+.

Digital AM: DRM

The development of Digital Radio Mondiale (DRM) started in 1995 and was finalized in 2005. It is designed for the digitization of the AM bands (Frequencies below 30MHz: LW, MW, SW) to improve the audio quality and add convenient data applications and service information features. The bandwidth of one DRM channel is backwards-compatible to the analogue channel spacing (4.5, 5, 9, 10, 18, 20kHz) and therefore DRM offers only very low bit rates for its applications. So HE-AACv2 was selected to provide a good audio quality. For some applications even lower rates have to be used and for speech only additional speech coding schemes CELP and HVXC are available. Also data services can be used in DRM. All kind of data services used in DAB can be used also in DRM under the restriction of a much lower bit rate.

Currently Digital Radio Mondiale is mainly used by shortwave broadcasters such as BBC, Voice of America, Radio France, Voice of Russia, or Deutsche Welle. Their main target is large coverage areas. Medium wave transmissions are also performed for local broadcast stations.

Digital Radio Mondiale still lacks cheap mass market receivers. Only some models are available up to now. The use of DRM in mass markets like Russia and China will help to solve the receiver problem.

SYSTEMS IN INTRODUCTION OR UNDER TEST

Mobile TV and Radio: DVB-H

DVB-H is part of the DVB standard and can be used in combination with DVB-T as well as on a separate transmitter network. Transmissions have already started in several European countries. Whereas Italy is already on-air with regular services, test transmissions were started for example in Germany and Austria.

DVB-H uses MPEG-4 AVC (H.264) for video coding and HE-AACv2 for audio coding. Currently also the use of MPEG Surround is standardized for DVB-H within the DVB-AVC group. In contrast to DMB, radio services with audio only are also possible.

This makes DVB-H also a competitor to the DAB family. On the other hand, for a full replacement of DAB it would require additional standardization of the data services that are part of DAB and which are required, for example, for automotive environments such as TMC, TPEG or Journaline®.

Digital FM: DRM+

DRM+ is an extension to expand the use of DRM for the frequencies above 30MHz up to 120MHz. The standardization is expected to be finalized until 2009 and introduction to the market is planned for 2011.

DRM+ offers a much higher data rate than DRM, up to 186kbps. This enables high quality audio with HE-AACv2 including MPEG Surround for multichannel as well as the whole range of additional data services known from DAB.

In the future, DRM+ could be the solution for small local broadcasters with individual coverage areas that therefore do not fit into a DAB Ensemble structure. In order to be independent of large network operators and other broadcasters, these broadcasters have an interest in operating their own small transmitter station independent from their competitors.

America in Europe: HD Radio

HD Radio has originally been designed for the U.S. market for the digitization of the FM and AM (only MW) bands by the U.S. company iBiquity. For the important transfer period from analogue to full digital it offers a hybrid mode that makes it possible to transmit analogue and digital simultaneously via the same transmitter. The digital part offers at least one digital audio service – called "HD-1" - which is similar to the analogue audio content. Seamless switching between analogue and digital is possible. Additional audio programs – called "HD-2", "HD-3" and so on are possible. Besides that data services can also be realized.

HD Radio also offers a full digital mode for situations where no analogue simulcast is required. This mode offers more capacity as the hybrid mode.

HD Radio is not an open standard like DAB/DAB+/DMB, DRM/DRM+ and the DVB family. Most of the implementation is done by iBiquity, which has drawbacks but also advantages. One the one hand only iBiquity can modify the core software for bug fixing or functional extensions on the other hand it is in the commercial interest of iBiquity to push HD Radio and they are doing a lot for its success.

In Europe the interest in HD Radio is based on the fact, that it is the only available solution for the digitization of FM without disabling the existing analogue services. There is no discussion about using HD Radio also for AM as for this band DRM is already the accepted standard.

Due to the narrower channel spacing in Europe some experts expect interference problems for the analogue FM in cases where adjacent channels are occupied by other transmitters.

Switzerland was the first country in Europe that has started HD Radio transmissions.

Currently, HD Radio FM test transmissions are also taking place in other European countries in order to test the compatibility with analogue FM in Europe. Germany has just started test transmissions with HD Radio. Especially in eastern European countries, there seems to be high interest in using HD Radio.

In September 2007 the European HD Radio section has been founded in order to harmonize and coordinate the introduction of HD Radio in Europe.

New satellite broadcast systems

With the development of the new ETSI-SDR standard the basics for the next generation of hybrid broadcast systems have been fixed. These systems use satellite broadcasting in combination with terrestrial repeaters and will enable European wide mobile reception similar to the already well-established XM Radio and Sirius systems in the U.S..

The improvement in channel coding schemes makes the new systems much more efficient which is very important for the European market since the different languages used mean more programs need to be transmitted to cover Europe with attractive content.

The operators of the new European broadcast will be able to adapt their system with regards to bandwidth and quality of service. For audio coding state-of-the-art audio coding schemes like HE-AAC v2 will be used.

Italy will be the first country where Worldspace offers its new services in cooperation with car manufacturer FIAT at the end of 2008/beginning of 2009, then followed by France, Germany, Spain, UK, Turkey and Poland. In each of these countries more than 50 channels will be available.

The new services will not only include audio broadcasting but also cover all kind of data services and multimedia services like time shift, recording, EPG and information services like Journaline®.

The competitor of Worldspace in Europe is ONDAS which recently has signed a deal with Nissan to provide car receivers for their fleet. ONDAS will provide more than 150 multi language programs and additional push and store services for audio and video as well as data services for stationary and mobile usage.

The new satellite broadcast systems are driven by commercial operators which makes it obvious that pay models similar to the U.S. will apply for the use of their services. This can be difficult for Europe as the public broadcasters already offer attractive content for free. Also, a large number of private broadcasters are available throughout Europe. Therefore, the situation cannot directly be compared to the U.S. with XM and Sirius but recent market research shows that even in Europe there is significant interest in the advertisement-free and European-wide services that Worldspace and ONDAS want to offer.

OUTLOOK

In Europe, the decision on which digital radio standards will be established is still open. DAB is well established in Great Britain and also Denmark seems to be a success for it. Sweden switched off DAB some years ago and is looking for other solutions.

The important German market is less clear. Some German states such as Bavaria are pushing DAB, DAB+ and also DMB while other states such as Lower Saxony are more focused on DVB-T and DVB-H.

France has selected DMB for Digital Radio. For small broadcast stations also DRM will be available. So if listeners from other European countries want to listen o digital radio they need to have a DMB receiver and/or a DRM receiver, whereas French listeners coming to Germany will need a combined DAB/DAB+ receiver.

Commonalities with regards to audio coding and data services between DAB/DAB+/DMB and DRM/DRM+ make the use of these systems attractive. Operators can expect similar requirements for their services over the different bearers and receiver manufactures can build a common service layer in their receivers for a lot of application types. Depending on its requirements, a broadcaster can select the appropriate system without restricting to system specific application types. Combined receivers supporting at least DAB/DAB+/DMB will be available soon and DRM/DRM+ are already in discussion by chipset makers and receiver manufacturers to be added to such platforms in the future.

The DAB family is suitable for large broadcasters such as the German public broadcasters as they can fill a whole DAB Ensemble with their program bouquet. For large private broadcasters this is the same.

For local public and private broadcasters in urban regions also DAB can be the right solution, if there are enough other local broadcasters that fit into the same coverage area to fill a complete DAB Ensemble.

For small local broadcasters and in situations where a DAB ensemble cannot be filled with enough services, DRM and DRM+ are a good alternative. One single transmitter or a small number of transmitters in single frequency networks can be used for a single or a small number of programs. Using DRM or DRM+ instead of DAB or DAB+ still offers the choice of most or all of the application types that are known from the DAB family.

HD Radio is a separate issue but seems to have the charm of easily upgrading the existing FM networks to digital with the HD Radio hybrid mode.

Positive reports of HD Radio tests are well known from Switzerland. The results from other test sites like Germany will come soon and show if the HD Radio system is compatible with the narrow FM channel spacing in Europe. HD Radio is currently the only available solution that enables digital simulcast within the FM band. In the future DRM+ may be an alternative for this, but it is not yet available, whereas HD Radio is already on the market including receivers for stationary and mobile use. Assuming the channel spacing does not cause problems, and also assuming that iBiquity opens the system for the required European data services, HD Radio could become a success in Europe.

Up to now no common Europe-wide terrestrial broadcast system is visible. For local services this is not a big issue, but for service providers that want to offer their services over the whole of Europe it is. This is valid, for example, for car manufacturers that want to offer their customers a reliable service over all of Europe. The only system that could offer European-wide terrestrial coverage independent from countries' individual broadcast infrastructure is DRM in shortwave mode, but it offers only very limited data rates.

This makes the use of upcoming satellite broadcast systems for European very promising. The economic viability of such a system depends on how many Pan-European services will come in the future and how many customers are willing to pay for such services.

REFERENCES

DAB/DAB+/DMB:
www.worlddab.org

DRM/DRM+:
www.drm.org

DVB-T/DVB-H:
www.dvb.org

HD Radio:
www.ibiquity.com

WorldSpace :
www.WORLDSPACE.com

ONDAS:
www.ondasmedia.com

MPEG Surround:
www.mpeg-surround.com

Journaline:
www.journaline.com

RDS-TMC:
www.rds-tmc.com

TPEG:
www.tpeg.org

MobileInfo:
www.mobile-info.org

BIFS:
www.chiariglione.org/mpeg/faq/mp4-sys/sys-faq-bifs.htm

MOBILE COVERAGE OPTIMIZATION BY POLARIZATION DIVERSITY IN VHF AND UHF PROPAGATION

Myron D. Fanton, PE
Electronics Research, Inc.

ABSTRACT

Orthogonal polarizations in RF transmissions from a broadcast antenna provide increased channel capacity and improved channel performance. In the same channel frequency band, two independent signals may be transmitted and received with isolation provided by dual-linear or dual-circular polarizations. This technique has wide application in communications: satellite communications, terrestrial microwave, cellular phone networks, and broadcasting.

Transmitting antennas that provide polarization isolation for channel combining are analyzed and found to be successful and efficient in applications such as the IBOC FM scheme and DTV simulcast applications. However, the UHF and VHF broadcast propagation channel may not preserve the polarization sense and purity of the transmitted signal, in some cases degrading reception. The possibility of overcoming these problems and application of the multiple-input multiple-output (MIMO) channel is discussed.

The use of orthogonal polarizations to increase channel capacity for VHF and UHF broadcasting is analyzed. Dual-polarized receive antennas have been found to dramatically increase mobile signal reception and virtually eliminate disconnects from the base station, and the application to broadcast DTV and IBOC transmission is made. The application of dual-polarized channels in Distributed Transmission networks and low power broadcast repeater networks is also described.

INTRODUCTION

The transmission channel presented to electromagnetic waves propagating in broadcast environments and broadcast frequency bands possesses field components and polarizations that are statistically uncorrelated [1]. A simple plane wave travelling in free space has fields that are completely described by Maxwell's Equations. In the real propagation environment, even the electric and magnetic field components are uncorrelated at a single receive location [1]. These uncorrelated field components may be received by a practical antenna and employed to eliminate deep fades in one particular component [2]. This characteristic is of high importance in an application where a deep fade can drop a mobile device from the network or causes a receiver to re-synchronize. Beyond eliminating deep fades, the multi-mode channel may be used to increase capacity employing MIMO transmission [3], [4].

THE POLARIZED PROPAGATION CHANNEL

From rainbows to radio propagation, the interaction of electromagnetic waves with the environment is polarized. The waves emanating from a UHF broadcast antenna mounted beside a cylindrical mast interact differently. In fact, as shown in Figure 1, where a null occurs in the vertical polarized radiation pattern, a peak generally occurs in the horizontal. Much of the propagation data, simultaneously recording vertical and horizontal polarization indicate that deep nulls do not occur simultaneously [5].

The major propagation effects, line-of-sight (LOS) reception, diffraction over obstacles, reflections, and to some extent, refractions are all polarization sensitive. Clearly LOS transmissions make use of independent channels on isolated polarizations to double capacity in, for example, terrestrial or satellite microwave links. Terrestrial broadcast is certainly not a LOS environment.

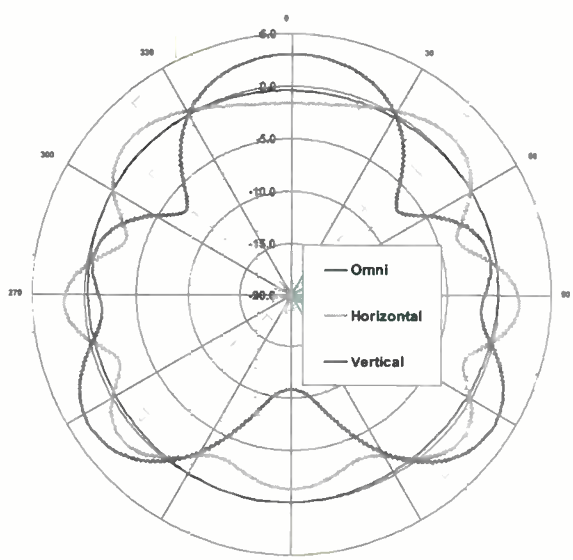


Figure 1: Polarization Diverse Scattering

Reflection phenomena are polarization sensitive, with much of the reflectors in the propagation environment parallel or perpendicular to the surface of the earth. Waves polarized parallel to the reflecting surface (say horizontal) are reflected at all angles of incidence while perpendicular polarized waves (vertical) exhibit larger variations. Vertically polarized waves have a minimum reflection coefficient at 3° and 15° for sea water and land respectively [5].

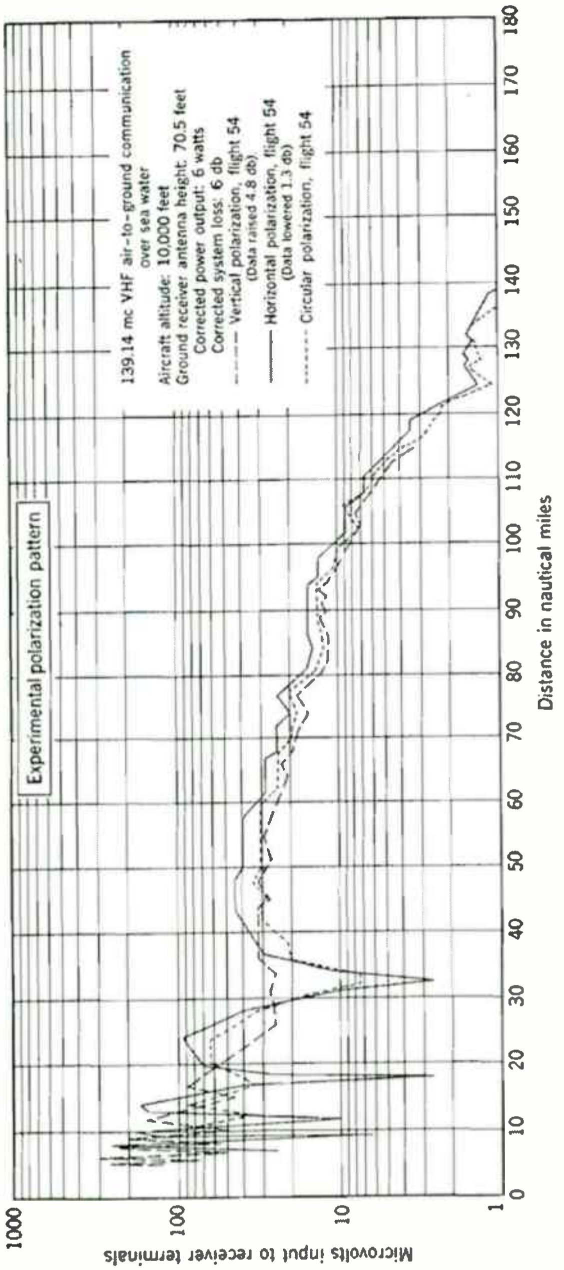


Figure 2: Dual Polarized VHF Propagation From Ref. [5]

Upon reflection from any flat surface, elliptically polarized waves reverse sense of rotation (say from right-hand to left-hand) [106]. So upon the first reflection from a mostly conductive surface, the vertical and horizontal polarized wave begins to be independent. Data from Reed & Russel [5] shows, Figure 2, that when a deep fade occurs in one polarization it does not occur in another. This is true of VHF and UHF frequencies, and is a phenomena observed in the UHF band by Lee in [2].

Diffraction phenomena are typically modeled and discussed without reference to polarization. In fact, the root of a great deal of skepticism about the utility of polarization diversity in broadcast applications is the feeling that the signal is de-polarized at the receiver, or all polarization information must be lost. Even if the polarization states, or put in the abstract jargon of MIMO: multiple input modes, are unknown, they can be reconstructed with great gain by simple receiver circuits [1]. Many diffraction situations are polarized, affecting independently the amplitude and phase of each polarization [7].

POLARIZATION DIVERSITY AND MIMO

Having asserted that the VHF and UHF broadcast propagation channel has uncorrelated polarization components, it has at least diversity implications and at most, dual mode capacity. To obtain diversity gain, a dual-mode receiver and diversity circuit schemes and algorithms may be employed at the receiver [1]. More elaborate direction-of-arrival algorithms have been proffered [8].

The far-field coupling is another parameter used to describe the isolation of two transmit modes, or orthogonal polarization states [9]. The available diversity gain varies with antenna characteristics, namely the output power correlation, related to the far-field coupling. In fact, Lindmark and Nilsson summarize cross-polar discrimination characteristics for urban and sub-urban environments, 0-12dB, stating that it provides the required, uncorrelated Raleigh channels for diversity gains of 9-11dB. These diversity gains vary with the azimuth pattern and polarization properties of the antenna. Vaughan and Bach Anderson report similar gains in UHF mobile systems [10]

MIMO systems spread the information across multiple modes in transmission and receive. The application is often a dual-polarized antenna: horizontal and vertical, left and right-hand circular, or +/- 45° slant. The mode of the receive antenna need not be identical to the transmit antenna. Because the modulation scheme is tied to a multi-mode, say dual-polarized antenna, the encoder and transmitter deliver information to multiple antenna ports. A receiver equipped to interpret orthogonal signals must also be employed. The UK trials of a UHF band MIMO system showed double channel capacity over the same coverage area [3].

CONCLUSION

The UHF and VHF broadcast propagation channel has uncorrelated polarization components, it has at least diversity implications and at most, multi-mode capacity. Diversity schemes involve more complex receive antennas, receiver circuitry, and receiver logic. For the additional complexity, order of magnitude system gains are possible as well as the elimination from the mobile channel of the long-duration, deep fades. MIMO transmission further exploits the channel, doubling its capacity at the expense of complexity at the transmitter and receiver. This additional capacity can be employed in mobile systems with forward error correction or tiered modulation schemes.

REFERENCES

[1] J.D. Parsons, *The Mobile Radio Propagation Channel,* John Wiley & Sons, Inc., New York, 2000.

[2] Lee, W.C.-Y, "Theoretical and Experimental Study of the Properties of the Signal from an Energy Density Mobile Radio Antenna," IEEE Trans. Vehicular Technology, vol. 16, pp. 25-32, Oct. 1967.

[3] Boyer, J., et. al., "MIMO for Broadcast – Results form a High-Power UK Trial," BBC Research White Paper WHP157, online http://www.bbc.co.uk/rd/pubs/whp/index.shtml, Oct. 2007.

[4] Mitchell, J.D., et. al., "A Dual Polarisation MIMO Broadcast TV System," BBC Research White Paper WHP144, online http://www.bbc.co.uk/rd/pubs/whp/index.shtml, Dec. 2006.

[5] H.R. Reed, C.M. Russell, *Ultra High Frequency Propagation,* John Wiley & Sons, Inc., New York, 1953.

[106] C.A. Balanis, *Advanced Engineering Electromagnetics,* John Wiley & Sons, New York, 1989.

[7] S. A. Attwood, ed., *Radio Wave Propagation,* Academic Press, New York, 1949.

[8] Zhang, L., et. al., "Obtaining Diversity Gain for DTV by Using MIMO Structure in SFN," IEEE Trans. Broadcasting, vol. 50, pp. 83-90, Mar. 2004.

[9] Lindmark, B., Nilsson, M., "On the Available Diversity Gain from Defferent]Dual-Polarized Antennas," IEEE Journal on Selected Areas in Communications, vol. 19, pp. 287-294, Feb. 2001.

[10] Vaughan, R., Bach Anderson, J., "Antenna Diversity in Mobile Communications," IEEE Trans. Vehicular Technology, vol. 36, pp. 149-172, Nov. 1987.

DOES YOUR YOTTA BYTE?

Andrew P. Janitschek
Radio Free Asia
Washington DC

ABSTRACT

Terms like megabyte and gigabyte are part of our everyday language. Others like 1080p, HD Radio, HDMI and Voice over IP can also be classified as part of our daily vernacular in the radio business, but only a few years ago they were virtually unknown to most. This is a brief study of 'the now' in broadcasting language but also a glimpse at the not-too-distant future where the unknown terms of today will soon be part of our everyday speech.

The focus of engineers continues to be on larger storage capacity, faster data transmission, smarter broadcasting and greener engineering. Just in terms of data storage, most of us are only now poised to think in terms of terabytes and petabytes. Data transmission rates of gigabits-per-second will give way to terabits-per-second. What other technological advances will change the way we live, work and speak? Let us take a moment to examine some of the new broadcast technologies and, of course, the associated terms we will hear, and use in the future.

WHERE DO YOU COME FROM?

The English language is in a constant state of evolution. Many words common in the 1800s and 1900s are no longer considered useful in 21st century. According to the Global Language Monitor, there are approximately 1-million words in the English language. A personal vocabulary of 7,500 words is considered necessary in order to write and speak fluently today. As you can guess, many words in use today will cease usefulness as others find their way into our every day usage. While a spittoon is certainly not 'in vogue' any more, when was the last time you heard someone ask for an ash tray? Sure, these are not technical terms, but they have seen their day. When was the last time you heard someone talk about dropping a needle on a record? If you are like most, it has been a while.

Let us take a moment and examine the word 'radio.' We have all been touched by this technology throughout our personal lives and our careers. Radio was originally called 'wireless telegraphy' and was shortened to 'wireless.' The word 'radio' was first used in 1897 by French physicist Edouard Branly when he wrote about radioconductors. The term 'radio' is based on the Latin word 'radius' which means 'the spoke of a wheel; a ray of light.' The use of the word 'radio' as we know it today appeared in a 1907 article by Lee de Forest, but creation of the word is credited to a marketer by the name of Waldo Warren. In 1912, the US Navy adopted the use of the word 'radio' and by the time commercial broadcasts began in the 1920s 'radio' became the common word used to refer to the actual transceiver. It is interesting to note that the term 'wireless' has been revitalized and has taken on a new meaning. One hundred years ago, 'wireless' was used to describe radio, but today 'wireless' is used as a reference to wireless networking and internet connections.

With most professions, there is a fair amount of jargon. Jargon is simply a group of special words that are specific to a profession, group, or activity. If you sail, you would likely be familiar with jargon like jib, halyard, and leeward. If you are a philatelist, or stamp collector, terms like perforations, hinges, and possessions take on new meanings. Computer related jargon continues to grow as technology improves and the needs of users change quickly. Can you remember the first time you heard of Microsoft Vista or XP? At first, they were strange, but today most of us use them as commonplace terms.

According to the Global Language Monitor, the most commonly used jargon in the world is the term 'O.K.' If you remember your history lessons, O.K. is the abbreviation for Old Kinderhook; this was the nickname of our 8th President, Martin Van Buren. Kinderhook is a direct reference to the town of his birth in New York. In Van Buren's 1840 campaign for re-election, his slogan was 'Martin Van Buren is O.K'. Today, the term 'O.K.' is globally used to mean 'all is well.'

In their 2007 list of the '10 Most Confusing Yet Frequently Cited High Tech Buzzwords' the Global Language Monitor published the following:

1. *iPod* - Apple's portable media player.
2. *Flash* - As in memory; also software from Adobe.
3. *Cookie* – Package of text sent from a server to a web browser then sent back each time the browser accesses that server.
4. *Nano* - Smaller version of Apple's iPod.
5. *Kernel* - Central component of most computer operating systems.
6. *Cell* - As in cell phone and being within a cell in order to use a cell phone.
7. *Plasma* - Video display technology; more on this in a moment.
8. *De-duplication* - Similar to data compression, but looks for redundancy of very large sequences of bytes across very large comparison windows. The first stored version of a sequence is referenced instead of being stored again.
9. *Blu-Ray* – One of two primary DVD formats for high definition video playback; HD DVD is the other format. More on these in a moment.
10. *Megahertz* – In computing, it refers to the clock speed of the central processing unit; it is the electrical voltage that changes from low to high and back again at regular intervals.

As a matter of reference, here are some of the words that made their list in 2005:

- HTTP
- Voice Over IP (VoIP)
- Megapixel
- Plasma
- Robust
- WORM
- Emoticon

STORAGE

In 1991 I purchased my first desktop computer for use at home; it was a Packard Bell with a 286 megahertz (MHz) processor with a 40 megabyte hard drive for storage. By today's standards, the processing speed and storage capacity is laughable, but it does give us some perspective of how far we have come in relation to the norms for data storage. Fifteen years ago we were talking of tens and hundreds of megabytes but in 2008, we seem ready to retire the term gigabytes in favor of terabytes and petabytes as our engineering paradigms continue to evolve along with technology. As costs for storage dropped and the demand for greater storage capacity rose, we witnessed the inclusion of terms like bits and bytes into most every language on the globe. We were also participants and spectators to the inclusion of terms like kilobyte, megabyte, and gigabyte as these larger storage levels were incorporated into our everyday lives. While the use of the terms terabytes and petabytes are achieving greater importance, it is apropos that we look at storage from the bit up to the yottabyte. For most, this is a simple review of data storage and the terms associated with specific data storage capacities, but it does offer a perspective of how much data is stored at levels we have used, and have yet to use.

A bit is the basic form of data upon which all others are built. A bit can be one of two digits; a 0 or a 1. As a binary digit, a bit is not only the basis of all digital communications but digital storage. Data transfer rates are usually described in terms of bits per second (bps), but for the sake of this discussion, we will focus on the use of bits as data storage units. Computers have always used the binary system and will likely continue to do so in the foreseeable future. When stringing a series of bits together, we achieved a byte. When first used, a byte was made of various strings of bits, but it is generally accepted that eight bits comprise one byte. For those of you who prefer, a byte is also called a word.

As the size of documents and computer programs grew, the amount of needed storage space also grew. When approximately a thousand bytes of space was needed on a hard drive, this was called one kilobyte. The true definition of one kilobyte in the binary system is 1,024 bytes, but it is usually used to mean approximately one thousand bytes. In decimal systems, a kilo means 1,000, but a kilobyte is also known by its binary system designation of 2^{10} since this is the power of 2 which is closest to one-thousand. Kilobyte is often shortened to K or KB, though the Institute of Electrical and Electronics Engineers (IEEE) recommends using the small letter *k* for a kilo in the decimal system while using a capital *K* when referring to a kilo when used in the binary system. Exploring the differences between the decimal and binary systems is interesting and can be pursued further through publications available in the NAB Bookstore, your library, or on the Internet.

Demands for more storage continued to rise over time. To meet that demand, devices capable of storing megabytes, or thousands of kilobytes, were produced. The need for one thousand kilobytes (actually 1,024 kilobytes) brought us up to a storage level of one megabyte; 1,024 megabytes took us further to one gigabyte while the need to store 1,024 gigabytes started the use of one terabyte as the next generation of storage terminology. One terabyte hard drives are readily available today through major electronic retailers for no more than a few hundred dollars each. Companies like Cisco already work with petabytes of storage capacity, or roughly 1-quadrillion bytes per petabyte. In 2005, Cisco's storage for their Marketing, Sales and Human Resources departments was over two petabytes; last year they reached 11 petabytes. How long before Cisco is dealing with storage area networks with a combined total of exabytes or more?

The following list sums up storage capacities and their associated terms; some of which we have yet to adopt for everyday use:

- 1 Kilobyte (K or KB)
  or approximately one thousand bytes
  or 1,024 bytes
  or 2^{10}

- 1 Megabyte (MB)
  or approximately one million bytes
  or 1,048,576 bytes
  or 2^{20}

- 1 Gigabyte (GB)
  or approximately one billion bytes
  or 1,073,741,824 bytes
  or 2^{30}

- 1 Terabyte (TB)
  or approximately one trillion bytes
  or 1,099,511,627,776 bytes
  or 2^{40}

- 1 Petabyte (PB)
  or approximately one quadrillion bytes
  or 1,125,899,906,842,624 bytes
  or 2^{50}

- 1 Exabyte (EB)
  or approximately one quintillion bytes
  or 1,152,921,504,606,846,976 bytes
  or 2^{60}

- 1 Zettabyte (ZB)
  or approximately one sextillion bytes
  or 1,180,591,620,717,411,303,424 bytes
  or 2^{70}

- 1 Yottabyte (YB)
  or approximately one septillion bytes
  or 1,208,925,819,614,629,174,706,176 bytes
  or 2^{80}

WiMAX

WiMAX is not necessarily something many of us are anticipating but it does have the capability of changing much of what we do in our daily routines. WiMAX enables the delivery of last mile wireless broadband access as an alternative to wired broadband like cable and DSL. WiMAX provides fixed, portable, and soon, mobile wireless broadband connectivity without the need for direct line-of-sight with a base station.

In a typical cell radius of two to six miles, WiMAX systems are expected to carry up to 40 Mbps per channel, for fixed and portable access applications. This is literally enough bandwidth to support hundreds of businesses with T-1 connectivity simultaneously and thousands of residences with DSL-equivalent speeds. WiMAX on a mobile network should provide up to 15 Mbps of capacity within a typical cell radius. WiMAX is also expected to become part of a standard load on laptops so that urban areas and cities can become 'metro zones' for portable outdoor broadband wireless access.

To summarize, the bandwidth and reach of WiMAX make it suitable for these potential applications:

- Connecting Wi-Fi hotspots with each other and to other parts of the Internet.

- Providing a wireless alternative to cable and DSL.

- Providing high-speed data and telecommunications services.

- Providing Internet connectivity from various sources. As part of a business continuity plan, a business can have a fixed and a wireless Internet connection, especially from unrelated providers. If connectivity is lost by one provider, the business will not likely be affected by a same-service outage.

- Provide connectivity wherever you go.

Will WiMAX be the 'killer ap' that closes the books on satellite radio? Doubtful, but it could provide more competition to the likes of Sirius and XM Radio. Only time will tell.

RADIO GOO GOO

HD radio is here! The FCC's Second Report and Order on DAB issued in 2007, made it possible for broadcasters to begin FM In Band On Channel (IBOC) multicasting and Extended Hybrid operation without permission or notice. Additionally, it allowed for AM IBOC broadcasting at night. IBOC gives stations the capability of being able to transmit digital and analog radio signals simultaneously on the same frequency. Needless to say, HD Radio has swept the US and is now available in all major radio markets. As of early 2008, Ibiquity.Com showed North Dakota as the only state without HD radio broadcasts. iBiquity Digital developed HD Radio technology and has helped changed how we listen to radio with more channels, improved sound quality, and new services. Look for more signs this year on the use of IBOC for datacasting for businesses. Other competing formats included FMeXtra and Digital Radio Mondiale (DRM).

Created by Digital Radio Express, FMeXtra uses subcarriers within the existing signal and therefore allows stations to use exiting equipment too. With FMeXtra, there are also no royalties to pay. A recipient of Radio's World's 2007 Cool Stuff Award, Digital Radio Express' Aruba FMeXtra Receiver handles both analog FM broadcasts as well as FMeXtra digital programs. The other competitor for digital broadcasting is Digital Radio Mondiale (DRM), developed by a non-profit global consortium. DRM is designed to work over the AM broadcast bands and provides great gains in audio quality for shortwave broadcasters. With the use of MPEG-4 codecs, DRM can fit more channels at a higher quality into a given bandwidth.

A TELEVISION OR A MONITOR?

When discussing the purchase of a new television or monitor for computer work, we are invariably talking of flat screens these days. Sales of LCDs and plasma displays have skyrocketed leaving the older cathode ray tubes, or CRTs, in a waning position in the 'new purchase market.' In 1976 I worked in a personnel department in New Mexico where we processed data by typing information into computer terminals connected to a mainframe computer located in another building. The terminals were similar to old teletype machines where the data was typed onto a wide roll of paper and each character was placed there with a loud CLACK. We also used IBM cards to update low-priority data overnight along with reels of paper tape; all with holes punched into them representing specific information.

In time these gave way to the use of television-like screens called the cathode ray tube, or CRTs. In the late 1970s, CRTs were the greatest thing to hit our profession as we dropped the use of paper for displaying data and simply let the new system display the information on a monochrome CRT.

Originally designed in 1897 by German scientist, Karl Braun, the CRT consists of a vacuum tube and an electron gun. The electron gun works at right angles with the fluorescent coating on the inside of the television tube. The gun generates a beam of electrons that moves back and forth and when the electrons from this beam strike the phosphor dots that are on the inside of the tube, they light up and create the active sections of the screen. All these active sections together create a complete image as the beam draws a complete set of lines from the top to the bottom of the television screen. The technology is still employed by many televisions and computer screens still in use today, but that is changing. Many people are opting to give up their CRTs for flat liquid crystal displays, or LCDs, and in some cases, PDPs, or plasma display panels.

The concept of using liquid crystals dates back to 1888, but it was in 1972 when the first LCD was produced in the United States. If retail sales numbers for the 2007 holiday season are any indicator, LCD displays have alredy replaced CRTs as the preferred standard for new televisions. While there are definitely advantages to using LCD monitors, they do have their drawbacks:

- When hooked to a computer, LCDs produce sharp images only when using their native resolution (when the input matches the native resolution, the LCD produces its optimal display); CRTs are capable of displaying multiple video resolutions without producing any noticable artifacts. Some resolutions work well if they are exact multiples of smaller image sizes. For example, a 1600×1200 LCD will display an 800×600 image well, as each of the pixels in the image will be represented by a block of four on the larger display, without interpolation. Since 800×600 is an integer factor of 1600×1200, scaling will not affect the image. As of early 2008, only the very best LCDs can approach the contrast ratios of plasma screens while most LCDs still lag behind.

- LCDs typically have longer response times than CRTs and plasma displays. LCDs experience ghosting when images quickly change. For example, when moving the mouse quickly on an LCD, sometimes you see multiple cursors.

- LCDs have a limited viewing angle when compared to CRTs and PDPs. The overall effect of this is that fewer people are able to view the same image; most laptop computers suffer from this same problem. While fewer people can view an LCD, it does require less power thus extending the life of the batteries and reducing the overall consumption of electricity.

- LCD monitors usually break easier than CRTs, especially since the screen of the LCD display lacks the thick glass front so common on CRTs.

Plasma display panels were invented in 1964 and began as monochome units. They are flat panels that we normally see used for large video displays. The cells of a PDP are held between two plates of glass contain a mixture of neon and xenon gases. The mixture changes into plasma when the cells are exposed to electricity; this activates the phospors causing them to emit light. While a PDP provides crisp pictures, there are also issues with the use of a PDP:

- Over time a PDP can suffer from burn-in. The problem is that the phosphor compounds used for producing light lose their luminosity with use. If a PDP is used for a computer program and is constantly powered to display information, over time areas become visible and do not go away, even after powering off the PDP.

- The quality of the images displayed gradually decline over time reducing the sharpness and clarity of the picture.

- When groups of pixels are forced to display high light levels, especially white, for long periods, this can create an image on-screen that will remain viewable until the power is cycled on the PDP or after running some random display on the PDP.

Since their early days though, makers of PDPs have develolped ways to reduce the problems of image retention through the use of image washing systems and pixel orbiters.

One technology that has yet to take off until recently is Laser TV. In 2007 there was some talk about laser television but nothing ever came of it. Laser TV is a projector technology using lasers instead of incandescent lamps to create light. Those that support Laser TV say it uses less power and allows for lighter sets than even LCDs offer. with "bulbs" that never burn out. It was in early 2006 when reports on the development of a commercial Laser TV were first released. At the Consumer Electronics Show (CES) 2008, Mitsubishi unveiled their first commercial Laser TV; it is a 65-inch model that is fully HD capable. The reviews from those that were able to view the Laser TV system were good. Those that saw the clips from popular movies felt the color was phenomenal, but also to the point of looking artificial. Laser TVs are expected be on sale in time for the 2008 holiday season. On a final note, China's SYCO demonstrated a 120-inch Laser TV in 2007 which is the world's largest Laser TV to date.

Whether using a desktop computer or a laptop today, the majority of the world still relies on a visual display of data through the use of some type of electronically powered screen or display. One final note, the world's largest plasma display is currently a 150-inch unit made by Panasonic and was one of the many new innovations shown earlier this year at the 2008 Consumer Electronics Show (CES); standing 6 foot by 11 foot, it is expected to initially retail for $150,000. This behemoth easily overshadows the 108-inch LCD that was unveiled at the 2007 CES by Sharp.

The choices are there, one only need compare prices and determine what is best for you. For my money, I still prefer the Sony Trinitron E540 21-inch CRT for my computer work at the office and at home.

VIDEO GA GA

In the global competition for high definition video playback, there have been two main competitors: Blu-Ray and HD DVD. Consumers that stay on the leading edge of technologies are helping drive the market. These are some of the same people that purchase cameras with ever higher megapixels though they fail to notice any difference between a picture taken with the new camera and one taken with an older camera capable of fewer megapixels. Why does anyone care about high definition DVD formats? One reason is that the two standards are incompatible. Eventually, one of the two would have to fall while the other inherited the DVD legacy. Earlier this year, Warner Home Video announced that it was dropping support for the HD DVD format and would release future video products only in Blu-Ray. This is considered a major win for companies like Sony who backed Blu-Ray; Sony included Blu-Ray playback capability in its PlayStation 3 entertainment systems. While there are winners in this competition, there are losers too. For example, Microsoft decided to provide HD DVD capability in its Xbox 360 gaming systems. Another was Toshiba who reacted to Warner Home Video's announcement by slashing the prices of their HD DVD players in an effort to attract new users to the HD DVD format.

Not Your Parent's Google

Earlier this year, Truveo, the leading video search engine, announced that its index had exceeded 100-million online videos. Based on current growth rates, Truveo expects their index to reach one billion searchable videos by 2009. Since 2004, Truveo has operated one of the most comprehensive video search engines on the Internet. Truveo's search engine continuously searches the web tracking videos that come online. Is Truveo the same as YouTube? Yes, in many way, but Truveo is not limited to a few specific sources. YouTube was created in February 2005, as a 'consumer media company' for people to search, view and share original videos worldwide using the Internet. YouTube was purchased by Google in October 2006. Though YouTube and Google have a large inventory of videos, they only search themselves while Truveo not only searches YouTube and Google, but multiple other sources.

Analog To Digital

Let us take a look at the pending demise of analog television scheduled for February 2009. The allotted spectrum used to broadcast television is granted by the government. In 1996 it was doubled so a station could have analog and digital transmissions. The spectrum freed up by the loss of analog signals in 2009 will be auctioned off while the rest will be used for national security and emergency communications.

If you still have an analog television at this time, you will need a converter box to receive a digital signal and view it in the analog world. Converter boxes will be less than $100 each. To encourage the transition of our fellow citizens to digital, the US government is providing an incentive in the way of $40 coupons to be used towards the purchase of the digital-to-analog converter boxes though the coupons are only available in limited supply. One only needs visit www.dtv2009.gov to request the coupons. Again, the converter boxes are only necessary if you are viewing over-the-air broadcasts on an analog television; if you pay a monthly subscription to a provider, a converter box will not be needed.

Most of us will buy a digital TV soon if we do not have one already. The best selling TVs are those capable of high definition TV (HDTV). In the United States, HDTV displays are 720p, 1,080p and 1,080i. These numbers tell you how many lines of information the signal holds. The "p" stands for a non-interlaced progressive scan. With progressive scanning, the lines displayed on your TV appear smoother to the eye. The letter 'i' in 1080i stands for interlaced or non-progressive scan. The numbers 720 and 1080 represents the number of lines of vertical resolution. With HDTV, assume we are viewing a widescreen aspect ratio of 16:9, implying a horizontal resolution of 1920 pixels and a frame resolution of 1920 x 1080 or just over 2-million pixels; the same applies to 1080p.

One lesson we learned in our family about 2-years ago is that if you have an HDTV, it does not mean you are watching HDTV programming. You must ensure you are also receiving an HDTV signal too. Besides the over-the-air option for HD programming, most cable and satellite providers, along with many phone companies, like Verizon, will provide you this programming through a monthly subscription.

If you have seen the DirecTV commercials with Jessica Simpson, then you have heard of 1080i. How many really know what 1080i is? When Ms. Simpson encourages you to buy a DirecTV high definition (HD) system, it is all about the quality of high-definition television. In Jessica's words, "Hey, 253 straight days at the gym to get this body and you're not going to watch me in DirecTV HD? You're just not going to get the best picture out of some fancy big screen TV without DirecTV. It's broadcast in 1080i. I totally don't know what that means, but I want it." This is true for most consumers; we're not sure what it is, but we do know it is better than normal TV and we want it!

High Definition Multimedia Interface

HDMI stands for high definition multimedia interface. According to the official website, www.HDMI.com, it boasts of "one cable – one standard." They also say HDMI is "the future ready way to connect HD." HDMI provides a simple-to-use interface between any suitably equipped audio/video source, like a DVD player, an A/V receiver or monitor over a single cable. You can still find systems that simply use three RCA cables to bring in two channels of audio and the video feed, but that is changing quickly.

HDMI states their standard "supports standard, enhanced, or high-definition video, plus multi-channel digital audio on a single cable. It transmits all ATSC HDTV standards and supports 8-channel, 192kHz, uncompressed digital audio and all currently-available compressed formats (such as Dolby Digital and DTS), HDMI 1.3 adds additional support for new lossless digital audio formats Dolby® TrueHD and DTS-HD Master Audio™ with bandwidth to spare to accommodate future enhancements and requirements." HDMI has become the standard digital interface for high definition television.

HDMI is also the interface being used to join computer and consumer electronics devices. The advantages of HDMI over existing analog video connections like S-Video, composite, and component video is that it provides quality signals, is easy to use and is HD ready.

FINALE

There are many terms like gigabyte and terabyte we use regularly and without hesitation. While this was an overview of some terms we use and some that may find their way into our daily lives, we have only scratched the surface. As we continue to improve operations through larger storage capacity, faster data transmission, smarter broadcasting and greener engineering, we will certainly adapt new words to our vocabulary just as we have in the past.

The challenge for engineers is to stay abreast of technologies as they change and affect us. We also know that it is impossible to stay 'on top' of everything. You can still keep ahead of the game by attending the NAB as often as possible, continually reading professional publications and email lists, and communicating with your co-workers and peers. Good luck!

New Technologies for Radio Listening

Tuesday, April 15, 2008
9:00 AM – 11:30 AM

Chairperson: Steve Fluker
Cox Radio / Orlando, Orlando, FL

RF Simulcasting over IP Networks

Junius Kim, Harris Corporation, Mason, OH

Practical Considerations of Radio Broadcast Operations in an Arbitron PPM™ Market

Larry Paulausky, Greater Media, Inc., Bala Cynwyd, PA

Consumer Ratings of Impaired Audio at Various Signal/Noise Ratios

John Kean, NPR Labs – National Public Radio, Washington, DC
Ellyn G. Sheffield, Towson University, Towson, MD

*Data Services for Digital Broadcasting

Alexander Zink, Fraunhofer IIS, Erlangen, Germany

*Affordable IP Based Remote Monitoring and Control of Transmitter Sites

Johannes Rietschel, Barix AG, Zurich, Switzerland

*Paper not available at the time of publication

RF Simulcasting over IP Networks

Junius Kim
Harris Corporation
Mason, Ohio

INTRODUCTION

RF simulcasting uses multiple, geographically disperse RF transmitters operating on the same carrier frequency, modulating the same program material. By using multiple transmitters, geographic RF coverage area is expanded. The region where a RF receiver can pick up multiple signals feeds from multiple transmitters is the overlap region. In an audio broadcasting application, a RF receiver in this region will simultaneously demodulate audio programming carried on multiple RF carriers. In this region, the audio modulation should be closely phase aligned from the multiple transmitters to provide the best receive quality.

The audio program material can be transported from the studio site to multiple transmit sites using STL (Studio to Transmitter Links) over a telecommunication network. To keep program material alignment, a simulcasting system STL should keep a constant, precise, and well known delay between all the studio and transmitter sites. IP (Internet Protocol) has emerged as a cost effective and flexible networking technology for transport of STL program material. The IP network delay is dynamic and can change over time due to route changes, changes in router characteristics or changes in link characteristics.

This paper outlines a system for audio simulcasting over an IP network. The system discussed uses a precision absolute time reference provided by GPS (Global Positional System). Using this reference, the system can measure the STL delay between the studio and transmitter sites. The system uses this information to set a programmable digital buffer delay to reach a target delay. The buffer delay changes are smooth and hitless resulting in no noticeable disturbance of the audio program material. The system is constantly measuring STL delay and automatically correcting for any changes in network delay.

BACKGROUND

Simulcasting

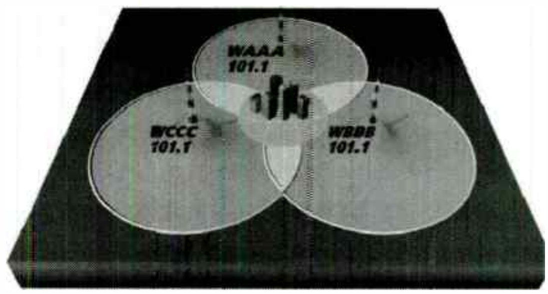


Figure 1 - Simulcasting System

Broadcasting from two or more nearby transmitters on the same frequency can lead to reception problems in the overlap areas – the areas in which the RF signal level from multiple transmitters is similar in strength. Figure 2 depicts the contours of relative signal strength from a two site system. In the overlap area, the relative power levels differ by less than 6 dB. The Harris SynchroCast technology makes this type of broadcasting possible by integrating information from GPS satellite receivers with precision and automatic digital delay management.

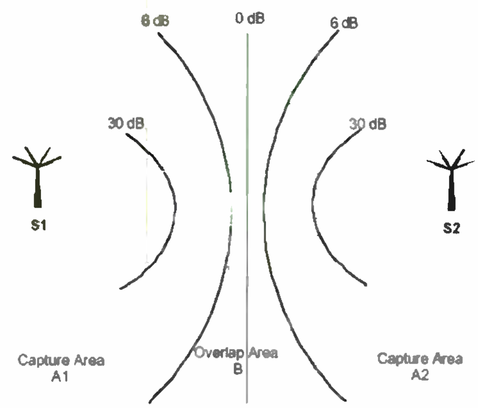


Figure 2 - Overlap and Capture Areas

Effects of Same-Frequency Simulcasting

Simply broadcasting the same signal from two nearby transmitters on the same frequency produces a cascade of effects. Listeners located closer to any one of the transmitters where the signal is significantly stronger (usually 6 dB or greater) will hear only the closer transmitter due to the "capture effect" of an FM receiver. Listeners in moderate signal areas will hear one or the other transmitter and may transition between them. Listeners located in "equal signal" areas or fringe areas can experience serious reception problems where transmitter footprints overlap.

For simulcasting to work effectively, the broadcast signal from each transmitter must arrive at the receiver at a precisely controlled time. A signal leaving the studio will be subject to two delay factors: uncontrolled studio transmitter link path delay and the "flight time" in the air from the transmitter to the receiver, over two different paths. The arrival time of each signal at the receiver can differ significantly, causing distortion, echoes and other artifacts. In addition, each transmitter's local oscillator frequency will be slightly different, causing phase errors between carrier frequencies in the overlap area.

GPS Satellite Technology

The advent of GPS satellites has created an effective method for synchronizing transmitters. A GPS receiver can deliver a precise timing reference to the studio and to each transmitter site in the simulcast system. Carrier RF frequency and program audio timing at all transmitters can be locked to the GPS timing standard, reducing or eliminating unwanted artifacts at the listener's receiver.

In FM broadcasting, the effects of multiple transmitters, each broadcasting a locally generated 19 kHz stereo pilot in a simulcast system, must be considered. Proper decoding of the stereo L-R signal depends on the accurate reception of the stereo pilot. Just as the transmitter carrier frequencies in a simulcast system must be locked to a GPS delivered reference, the FM stereo pilot should be locked to a master reference as well.

Migration to IP Networks

Previously, the FM simulcasting systems were typically limited to usage over PDH (Plesiochronous Digital Hierarchy) network link types like T1 or E1 transmission links. Recently, the Harris SynchroCast system in conjunction with the Harris NetXpress system has been successfully adapted for simulcasting usage over packet switched networks like IP.

T1 or E1 circuits are dedicated point to point "nailed-up" circuits. Although these circuits have low-delay (typically < 10 mS) and the reliability is very good, these circuits are inflexible in their configuration and tend to be expensive as measured as cost per bit in relation to packet switched network technologies.

In contrast to T1/E1, IP networks are designed to be a shared resource. Many different types of data along with many hosts can be converged onto a single network connection from a service provider. IP networks offer the possibility of highly flexible, low-cost audio transport, and when properly managed can offer the degree of reliability required for use in professional audio contribution and distribution networks.

SYNCHROCAST

Broadcasting from two or more nearby transmitters on the same frequency can lead to reception problems in the overlap areas. For effective simulcasting, the end-to-end audio delay for each STL link in the simulcast system should be the same. This results in a receiver located in the overlap region receiving exactly the same signal at exactly the same time from multiple transmitters.

To keep audio alignment, SynchroCast will keep a constant, precise, and well known STL delay between all the studio and transmitter sites. The STL delay being defined as the end-to-end audio delay.

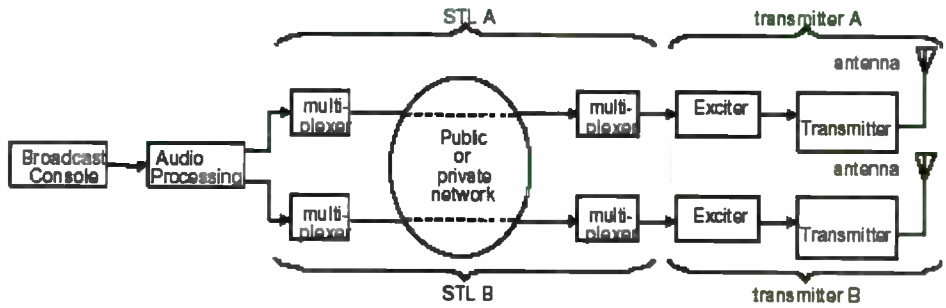


Figure 3 – Simulcast System

In SynchroCast, at the studio, a GPS derived timing marker is sent along with the program audio over the STL links to each transmitter site. At the transmitter site, this timing marker is received and decoded and compared to a local GPS derived timing reference to determine the actual STL delay. Once this delay measurement is established, SynchroCast calculates and introduces a precisely controllable delay, causing exact alignment of the audio signals at all transmitter sites. In addition, the GPS receivers at all transmitter sites provide a clock reference signal to each transmitter exciter, locking all of their carrier frequencies to the same satellite-delivered timing reference.

The timing marker consists of a 24-bit counter in 100 nS units. The timing count has a value between 0 and 10,000,000 (1 Second). The 100 nS resolution and 1 second maximum count come about because GPS receivers commonly have 10 MHz and 1 Hz clock outputs.

The system operates automatically once the initial installation and alignment is complete. It continually monitors the delay of each STL link, keeping the total STL delay constant, even if the actual path delay changes. This can occur, for example, if the IP network path gets rerouted to an alternate path due to network considerations.

Simulcasting over IP Networks

A simulcasting system using an IP network must perform the following functions:

- Encoding and decoding of real-time program audio
- Transport of program audio over IP via a process of packetization
- Sending a GPS referenced timing marker from the studio site to the transmitter site
- Establishment and maintenance of timing across the IP network

Circuit Emulation Service (CES) technology has emerged as a method to transport Time Division Multiplexing (TDM) trunks containing real-time applications such as audio, across IP networks. This technology is sometimes referred to as pseudo-wire, as it emulates the TDM circuit across a packet network using virtual IP tunnel or path. These emulated services can be implemented using a gateway device that provides for an inter-working function (IWF) between TDM and IP networks. The primary benefit of this technology is the cost and simplicity of deployment to support all types of existing TDM applications without the need for complex protocol inter-working functions.

Figure 4 – Reference Model for CES over IP

Figure 4 shows a reference diagram for supporting TDM applications over IP networks using CES.

Within an IP packet network, there is no inherent transport of timing information between hosts or end-points since packets are discontinuous in time. Achieving and maintaining synchronization is a critical aspect of achieving high quality circuit emulation over IP service. The algorithms and methods for achieving this under various network and bandwidth constraints are not standardized and can thus become a differentiating factor between products of different vendors.

In a CES application, the gateway device must transport timing, establishing the same TDM frequency at both ends of the circuit. The consequence of a long term mismatch in frequency is that the packet queue of the Gateway device at the egress of the network will either fill-up and overflow or empty and underflow. The direction depends on whether the regenerated clock is slower or faster than the original. This will cause loss of data and degradation of service.

The variation in the inter-packet arrival time at the receiving gateway is caused by network jitter. The paths in an IP network are connectionless and statistically multiplexed with other sessions. The amount of network jitter depends on how the network has been engineered and how many "hops" or routers must be traversed.

A well engineered IP network can be designed to control the network variation, while an unmanaged network, such as the public Internet, can produce large amounts of jitter. The gateway's architecture absorbs this variation in delay by providing a jitter buffer. The jitter buffer adds additional delay to the end-to-end service. Thus, the Gateway must provide for a flexible configuration of this parameter so that it can be optimally engineered to work in a variety of IP networks, from extremely well managed to the public Internet.

Figure 5 shows a diagram of a jitter buffer located at the transmitter site. Each transmitted IP packet in the CES "stream" will have a sequence number in its packet header. Upon reception, the sequence number in the received packet header will be examined to identify early, late, lost or out-of-order packets. If not too early or too late, the packet will be placed into the jitter buffer according to its sequence number. The packet shown at the "bottom" of the jitter buffer is "played out" to a TDM bus. It is processed by a packet to TDM conversion engine which plays out the data in real-time. After playout, the packet can be discarded. Once converted to TDM data, an audio decoder processes the data for usage by an RF exciter.

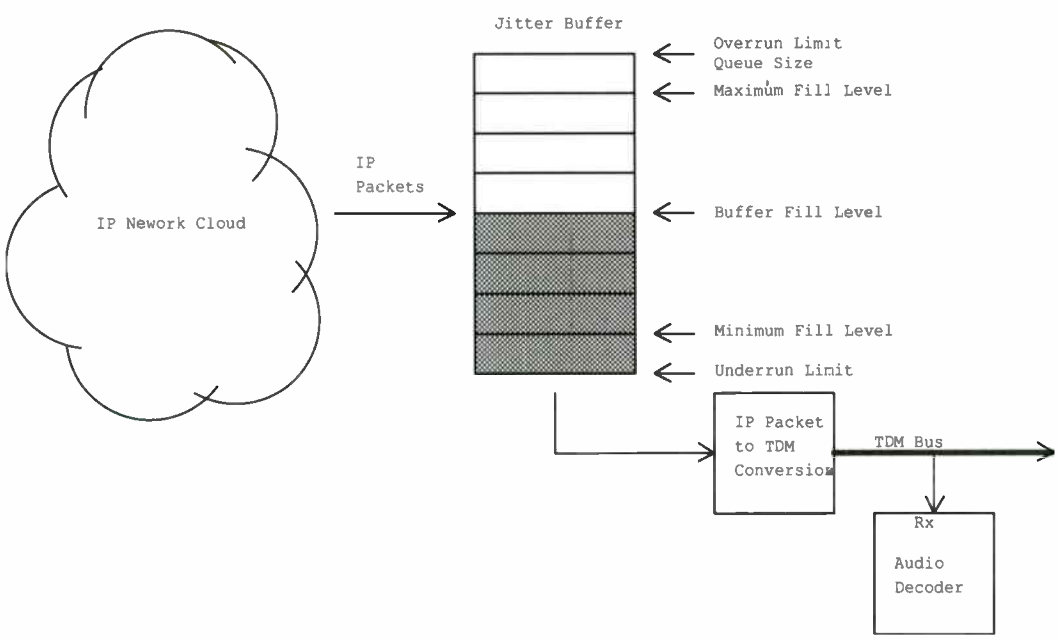


Figure 5 – Jitter Buffer

The jitter buffer has a characteristic very useful for simulcasting in that it provides a controllable delay. The delay can be changed in a hitfull or hitless manner. To change the delay hitfully, data can be artificially added (addition of delay) or removed (reduction in delay) to/from the jitter buffer. This type of delay change has the advantage to being instantaneous with the obvious disadvantage of resulting in a discontinuity of the dataflow. To change the jitter buffer delay hitlessly, the playout rate or bit rate of the TDM bus can be changed. Under normal condition – static delay, the bit rate of the TDM bus at the CES gateway receiver is the same as the bit rate of the TDM bus of CES gateway transmitter located at the studio site. To change the delay hitlessly, the TDM bit rate difference between the CES receiver and transmitter is made nonzero. For example, if the playout rate is made faster than nominal, the jitter buffer will begin to deplete data. The buffer fill level shown in Figure 5 will decrease and thus the delay will decrease in direct proportion to the fill level. Conversely if the playout rate is slower than nominal, the jitter buffer will begin to increase the amount of data it holds and the delay will increase.

When making hitless delay changes, the jitter buffer characteristics must be closely monitored to prevent an overflow or underflow condition from occurring. In reference to Figure 5 for example, if the playout rate is indefinitely altered from nominal, at some point in time the jitter buffer fill level would be greater than the overrun limit or less than the underrun limit resulting in an overflow or underflow condition. During an overflow/underflow condition the jitter buffer is reset to the nominal fill level – typically half the jitter buffer size. An overflow/underflow results in discontinuity of the dataflow.

By changing the playout rate from nominal, the delay is changed without any loss of data since the continuity of the dataflow is maintained. Also note the delay change granularity can be extremely small – fractions of a bit. That is one could chose to make a small change in the playout rate for a small period of time resulting in delay change that is less than a TDM bit cell.

As a practical matter, the hitless delay change method is more useful for a simulcasting system than the hitfull method since quality of audio playout is more important than delay convergence speed.

The maximum jitter buffer delay is a function of the jitter buffer size. The jitter buffer size is a function of the number of entries in the jitter buffer and the size of packets. Figure 5 shows a simple jitter buffer with eight entries, but the jitter buffer could be made as large as the memory on the CES gateway device allows.

The minimum jitter buffer delay is close to but not zero. The IP packet to TDM conversion engine shown in Figure 5 needs some minimal amount of data to work against.

Measuring the STL delay (end-to-end audio delay) and managing the jitter buffer delay is the essence of an effective simulcasting over IP system.

Implementation

Harris SynchroCast system in conjunction with the Harris NetXpress system has been adapted for simulcasting usage over packet switched networks like IP (see Figure 6). The NetXpress system provides a CES gateway function along with an audio encoder/decoder function. The SynchroCast system provides a means of precision measurement of the end-to-end audio delay.

SynchroCast measures the IP Network Delay + Jitter Buffer Delay + CES Processing Delay and communicates this information to a microprocessor. This microprocessor implements an algorithm for changing the jitter buffer delay to reach a customer specified target delay.

Target Delay = Network Delay + Jitter Buffer Delay + CES Processing Delay + Delay Error

Delay Error = Target Delay - (Network Delay + Jitter Buffer Delay + CES Processing Delay)

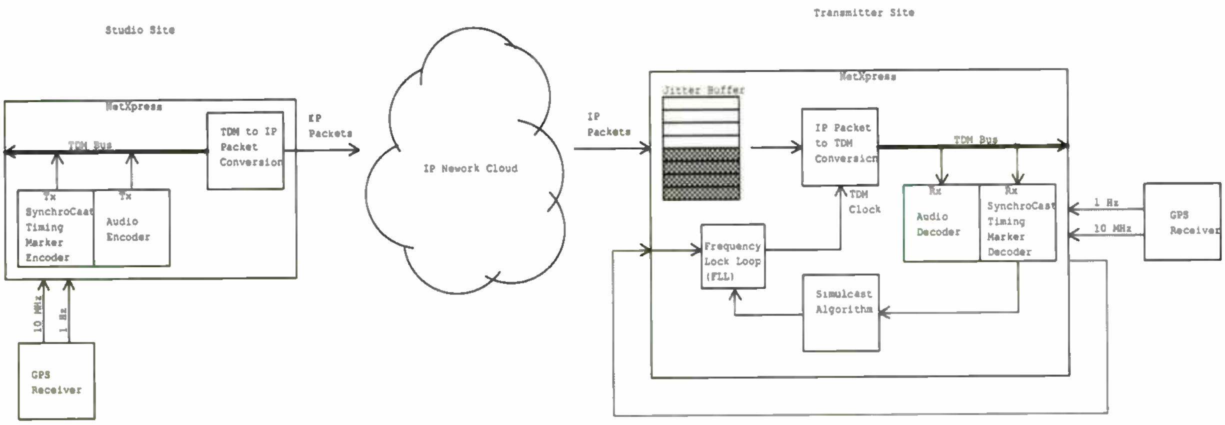


Figure 6 – SynchroCast for NetXpress

Note(s):

1. The CES processing delay is defined as the delay in converting IP packets into constant bit rate TDM data.

2. In NetXpress, the CES processing delay is minimal, 1 or 2 TDM frames (125 to 250 uS).

The target delay is configured by the customer. It is a 24-bit value in 100 nS units.

Given the target delay and delay error, the desired jitter buffer delay can be computed by the simulcast algorithm. The target jitter buffer delay is reached by altering the playout rate (TDM bit rate) for some period of time. For example, if the desired jitter buffer delay is 1 mS greater than the existing jitter buffer delay, the algorithm can decrease the nominal TDM clock by 50 PPM for 20 seconds to achieve the target jitter buffer delay.

Jitter Buffer Delay Change (Sec) = Time (Sec) × Frequency Offset (PPM) ÷ 1,000,000

By altering the TDM bit rate to effect delay changes, delay changes are hitless.

The algorithm implements the function of a digital feedback loop control element. The algorithm keys off delay measurement information from the SynchroCast system and uses the jitter buffer to try and null any delay error. This process of measurement and correction is continuous and automatic.

The simulcast algorithm limits the jitter buffer queue excursion to less than the absolute limits (maximum and minimum fill levels shown in Figure 5) to prevent overflow and underflow. Also, some margin is added to allow for network packet jitter.

At the STL system at the studio site, no unique simulcast algorithm is required. All adjustments and measurements are performed at the transmitter site.

At both sites, CES gateway devices are configured to use external timing provided by SynchroCast. So under static delay conditions, the CES gateway TDM timing is referenced the GPS receivers at each site.

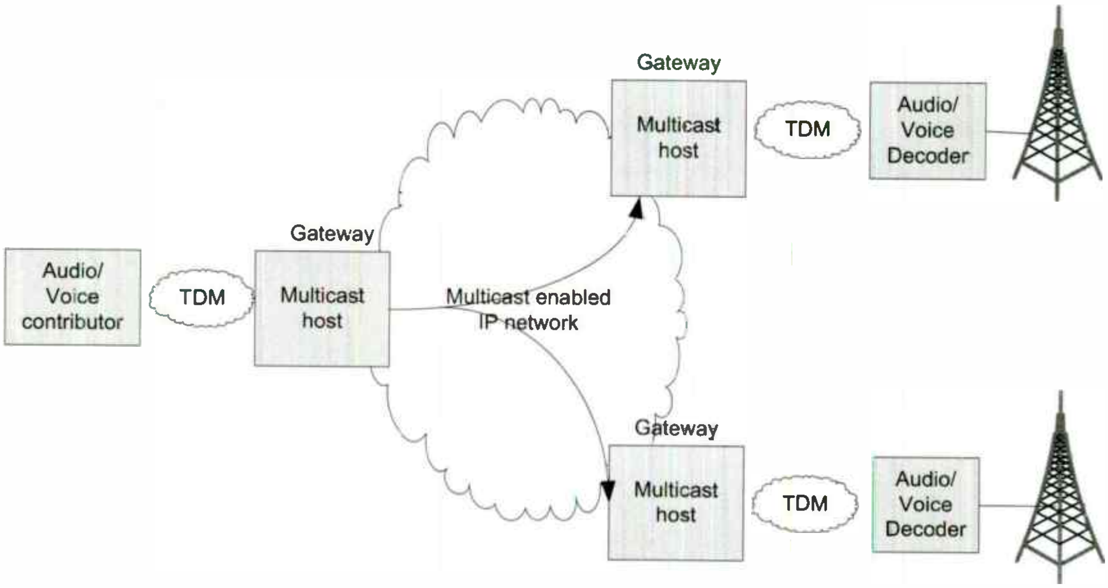


Figure 7 – IP Multicasting

Frequency Lock Loop (FLL)

NetXpress has a Frequency Lock Loop (FLL) which allows a precision offset to be added or subtracted from the TDM bit clock. The FLL uses the GPS clock for its reference. The magnitude of the offset is dependent on the delay change required. The maximum offset is limited to +/-100 PPM to minimize distortion of the program audio decoding which is timed from the TDM bit rate. This maximum offset is used when making large (> 1 mS) delay changes. Very large delay changes can take a long period of time to complete. For example a 100 mS delay change can take 2000 seconds or over 33 minutes to complete.

IP Multicast

One key advantage to using IP as underlying network transport vs. T1 or E1 is more flexible network topologies can be implemented. For example, using IP multicast, one system at the studio site can multicast to many NetXpress systems located at many transmitter sites (see Figure 7). This is in contrast to the T1/E1 systems which are point to point.

IP Stream Considerations

In a CES application, the generated IP packets are a "stream" of periodic same sized packets. The simulcast program audio and timing marker should be multiplexed onto the same IP stream. This way the audio and delay time information are on the same IP packets and will take identical routing paths over the IP network and have an identical network delay.

The maximum measurable delay is a function of the slowest clock provided by the GPS receiver, typically 1 second. The target delay must be less than this maximum and the Network Delay + Jitter Buffer Delay + CES Processing Delay must be less than this as well.

To enforce the maximum delay limitation, the simulcast algorithm on the NetXpress will not allow the jitter buffer delay be to greater than this maximum.

Maximum Jitter Buffer Delay = Jitter Buffer Size × TDM Payload Per Packet × TDM Frame Period

Network Considerations

The SynchroCast system tries to maintain a constant STL or target delay. Over IP networks the network delay may vary due to route changes, changes in router characteristics, or changes in link characteristics. A network delay change will cause the SynchroCast system to respond by altering the jitter buffer delay to maintain the target delay. As discussed earlier, a change in jitter buffer delay takes some period of time to complete. That time is proportional to the delay change. While making a delay change, the simulcasting quality in the overlap region of multiple RF transmitters is degraded. So the quality of the simulcasting is directly related to the delay consistency of the network. In general, a high quality MPLS network will maintain a more consistent network delay than an unmanaged network.

The IP network delay must be less than the maximum measurable delay. If the network delay exceeds this, the delay measurement will "wrap" and the actual delay error will be greater than the measured delay. Note that typical IP network delays are less than 50 mS.

Delay Resolution

The delay resolution is a function of the fastest clock provided by the GPS receiver, typically 100 nS. Delay changes and measurements can be made to within this granularity. This resolution defines the accuracy of locating the simulcasting overlap region. The geographical point in which the audio from two transmit towers are in phase will move 1 km for every 5.364 uS of delay variance, so 100 nS provides an accuracy of 18 meters.

CONCLUSION

Simulcasting on multiple overlapping transmitters on the same frequency can provide broadcasters with significant advantages in increased coverage and lower operating costs. The effectiveness of same-frequency overlapping transmitters depends on accurate synchronization of the carrier frequencies and broadcast audio. Only when this timing is precisely controlled can the listener enjoy seamless reception with a minimum of artifacts.

Audio transport is migrating towards IP based packet switched networks. IP offer the possibility of a highly flexible, converged network for many service types including low-cost audio transport. Emulated services, such as CES, allows audio circuits to be bridged between locations by providing a pseudo wire tunnel across a provider's IP network. Using a precision time reference provided by GPS, it is possible to measure STL delay and use this information to set a programmable jitter buffer delay in the CES gateway to reach a target delay. Ideally such a system should constantly and automatically measure STL delay and correct for any changes in network delay. Buffer delay changes should be smooth and hitless so as to result in no noticeable disturbance of the audio program. Using these methods, effective simulcasting over IP networks is possible.

REFERENCES

1. *"Delivering Professional Quality Audio Services over IP Networks*", Harris Corporation White Paper, 2005

2. *"Using Synchronized Transmitters for Extended Coverage in FM Broadcasting*", Harris Corporation White Paper, 1999

3. Keyur Parikh, Junius Kim, *"TDM Services over IP Networks*", Milcom Conference, October 2007

PRACTICAL CONSIDERATIONS OF RADIO BROADCAST OPERATIONS IN AN ARBITRON PPM MARKET

LARRY PAULAUSKY
Greater Media, Inc.
Bala Cynwyd, Pennsylvania

ABSTRACT

Arbitron, Inc.'s new Portable People Meter (PPM) electronic ratings system completely replaces the traditional "diary" system that had been in place for decades. PPM ratings are based on the recovery by portable pager-sized listening devices carried by survey respondents of unique subaudible identifying codes which are added to each audio channel by broadcasters.

It is essential that broadcasters in PPM markets develop systems that can reliably detect encoding problems as soon as they occur and that will immediately notify technical responders of problems. This paper will describe the system developed to monitor the twenty-two unique program streams originated by Greater Media's five-station cluster in Philadelphia, and will provide tips and insights to the PPM encoding process that will help readers plan their own installations as PPM technology is rolled out to their markets over the next few years.

BACKGROUND

Beginning in the spring of 2007, Philadelphia was the first US market to "go live" with official published audience estimates based entirely on PPM data. Prior to then, Arbitron had been running field tests of its encoding equipment and PPM devices in test markets including Philadelphia, Wilmington, and Houston for several years.

Not counting the data collection and analysis systems internal to Arbitron's home office, the PPM system consists of several field components: encoders, respondent PPM devices, and confidence decoders.

Encoders

The PPM system requires broadcasters to place Arbitron-supplied encoding boxes in the audio air chain of each program stream. The encoders analyze the program audio and opportunistically add a subaudible coded message to the content. The message uniquely identifies the program stream and includes a timestamp.

Respondent PPM Devices

Panelists in the survey carry small PPM devices which monitor the audio environment and which can identify PPM encoding if it is present. Arbitron says its devices can reliably detect PPM encoding after only a few seconds of exposure to the encoded stream, even in noisy acoustic environments. Headphone listening is accommodated through the use of "loop-through" jacks; external source equipment (for example, a walkable radio) is connected to the PPM device with a short jumper, then the listener's headphones are plugged in to the PPM device itself. The PPM devices have their own internal clocks, and maintain a record of decoded audio streams throughout the day.

The PPM devices also contain sensitive motion sensors, and can add a notation to their internal logs when they are not moving. This notation can be used to exclude "listening" results that might have been recorded while a device was not actually being carried by its survey respondent.

When a respondent returns their device to its battery charger at the end of the day, the device downloads its log to a centralized data collection controller in the home. Multiple respondents in the same home may carry PPM decoders. At specified regular intervals, the centralized data collection controller automatically "phones home" to Arbitron to transfer all of the listening information it has collected.

Confidence Decoders

Arbitron makes available dedicated rack-mounted decoder boxes which are programmed to respond to the presence or absence of a particular encoding stream. By connecting these decoders to an off-air feed, the end user can be warned when encoding has failed.

PLANNING

Arbitron Equipment

The encoders and the confidence decoders are the only components of the PPM system whose operations are under the broadcaster's control. While their deployment is relatively straightforward, the number of these boxes required to support the program streams generated by a cluster of stations can be surprisingly large.

Start by considering how many separate programs in your facility will require Arbitron encoding. Our facility, with one AM station and four FM stations, various HD and streaming offerings, and backup air chains for every program, required no less than 44 separate Arbitron encoders, as detailed in Table 1.

Station Program	Main Air Chain	Aux Air Chain
WPEN (AM)	1	1
WPEN Stream	1	1
WMGK (FM)	1	1
WMGK Stream	1	1
WMGK HD1	1	1
WMGK HD2	1	1
WMGK HD2 Stream	1	1
WMMR (FM)	1	1
WMMR Stream	1	1
WMMR HD1	1	1
WMMR HD2	1	1
WMMR HD2 Stream	1	1
WBEN-FM	1	1
WBEN-FM Stream	1	1
WBEN-FM HD1	1	1
WBEN-FM HD2	1	1
WBEN-FM HD2 Stream	1	1
WJJZ (FM)	1	1
WJJZ Stream	1	1
WJJZ HD1	1	1
WJJZ HD2	1	1
WJJZ HD2 Stream	1	1
TOTAL	22	22

Table 1. 44 Arbitron Encoders Required!

Arbitron has 1RU studio encoders available in both analog and AES audio In/Out models, and each model can be ordered with or without a local time code connection port.

Note however that even if a local time code is provided, the local time code will *not* reset the time of the Arbitron encoder. The Arbitron unit always uses its own internal clock-calendar, which cannot be set or modified by the end user. The only purpose of the local time code is to provide a basis of comparison for the unit's internal clock-calendar. If the two times differ by an inordinate amount, a front panel alarm message will be generated.

Even if your facility does not have dedicated backup air chains for every single program stream, consider that it is nevertheless advisable to install backup Arbitron encoders for every stream. These units can be wired in series, with one unit actively encoding and the other unit in bypass. With such a backup unit already installed in this "hot standby" configuration, a failure of the main encoding unit can be quickly remedied. Without a backup encoder, it would be necessary to wait for a replacement unit from Arbitron – during which time your facility would not be receiving any PPM listening credit.

Next, consider how many confidence decoders will be required. For most organizations, a reasonable guide-line would be "if it is important enough to encode in the first place, it is important enough to make *sure* it is *being* encoded". In our case, we elected to obtain confidence decoders for all 22 of our program channels.

The confidence decoders have XLR audio inputs for connection to encoded audio (preferably, from an off-air or post-transmission source), and a D-connector which provides a dry contact closure indicating that encoding is successfully being detected.

The confidence decoders also come equipped with a rear panel RS-232 output, which outputs a text stream at periodic intervals, indicating either that encoding is being detected, or that encoding is not being detected.

With a list of desired equipment in hand, as your market moves up the queue to be converted from diary measurement to PPM measurement, it will soon be time to talk to the folks at Arbitron. In the case of the Philadelphia market, Arbitron's engineers were in contact with market engineers commendably early in the PPM rollout process. Encoder and decoder orders were placed many months before the system was scheduled to go online, and equipment showed up in plenty of time for both main and backup units to be installed and thoroughly tested.

Verification of proper encoding in many cases could be done over the telephone. A local engineer could simply hold his telephone receiver up to a speaker playing various examples of encoded audio, one program at a time, and the Arbitron engineer at the other end of the line could verify the encoding was correct. It was even possible to leave a voicemail recording of local audio for the Arbitron engineer to listen to and verify the encoding the next day.

Other Equipment

As if the prospect of the sudden delivery of sixty-six boxes of equipment isn't daunting enough to you and to your receiving department, it is worthwhile to examine what other equipment will be needed to successfully monitor your system.

Recall that since Arbitron's PPM encoding process analyzes the program and adds encoding energy where it won't be audible, then it follows that if for some reason a program stream is silent, PPM encoding cannot occur. The predictable result is that a station will not get ratings credit for any period that it has silence on the air.

Given this, it is worth evaluating all air chains to be certain they can successfully and automatically recover from silence problems. Many devices exist that can monitor a program air chain and switch to a backup source (an emergency CD, for example) when silence is detected, and their deployment is highly recommended in a PPM environment.

Secondly, an off-air silence sense is a useful tool for technical responders to evaluate whether Arbitron decoder alarms are in fact due to an Arbitron encoding problem, or are simply due to silence on the air. This is especially important for channels which may not be easily monitored by an on-call technical responder, such as HD or web streaming channels. In our case, we elected to replace a haphazard existing system of silence monitoring with new, centrally located silence sense devices, one for each program stream.

Next, monitoring requirements should be addressed. (See Figure 1.) Each Arbitron confidence decoder (and the associated silence sense for that channel, if one is used) needs to have a dependable source of off-air or post-transmission audio. Analog channels each require a dedicated analog tuner. HD and HD2 channels each require dedicated HD tuners. Web streaming channels each require a dedicated computer "tuned" around the clock to the stream in question. Your facility should plan to acquire these monitoring sources, or if you already have them, plan to extend audio from them to the Arbitron confidence decoders and silence senses. Do not overlook that a stack of individual computers used to monitor your web streams will probably require consumer-to-pro level audio interface converters, a local rackmountable pullout monitor and keyboard, and a large, remotely accessible KVM switch.

Figure 1. Analog and IBOC tuners, with dedicated webstream monitoring CPUs below and silence monitors at right.

Lastly, there is the need for remote control monitoring equipment to which all the dry alarm contact closures from Arbitron and from silence sense equipment can be connected. Such a remote control should be capable of permanently logging error conditions and then generating several types of alarms, from e-mails and pages to POTS line voice callouts. It is also useful to consider a system which can evaluate alarms and make appropriate callout decisions – for example, problems with Arbitron encoding on a main analog channel are top priority and get sent to all on-call engineers, while problems with stream encoding get sent to the IT department on-call staff.

Few existing facilities have the spare rack space to swallow the significant amounts of new encoding, receiver/monitoring, confidence decoding, silence monitoring, and remote control equipment that is required by implementation of PPM measurement. Space for new racks may need to be found, and power and cooling needs should also be evaluated.

OTHER CONSIDERATIONS

Air Chain Location

It is well to give some thought to where in the program chain the Arbitron encoders will be located. Arbitron recommends that they be installed after a pre-limiter, but before final audio processing and before EAS encoding equipment. Note that this means a station would not be generating Arbitron encoding while its EAS device was running tests, forwarding alerts, etc. (assuming the EAS encoder is in the air chain and not an input to the console).

Note that a station using a broadcast delay system would wish to install the Arbitron encoder after the delay, so that the encoding would not be disturbed by an obscenity "dump".

Program Archiving

Another consideration is the storage and forwarding of Arbitron encoded audio. One case is where a daily talk show wants to produce a "podcast" of the day's program for its listeners to hear online or via iPod. Arbitron has held that encoded audio that is monitored by a PPM device within a week of original broadcast will be considered as heard at the time of the original broadcast. This is like time-shifting.

Recall that the encoded material is timestamped, and that the PPM devices carried by respondents also have internal clocks. A PPM device hearing such a podcast would log the disparity in time, i.e. that it is hearing audio encoded with a Tuesday at 9:30am timestamp, although it is aware the actual time of day that listening is occurring is Wednesday at 4:00pm. When the listening is analyzed ultimately at Arbitron's home office, the listening would count as if the listener had heard the program live. Given this scenario, it would make sense to have a means to record post-Arbitron encoded audio for use in producing podcasts.

There is another case, however. Suppose a station has a high-profile and highly rated morning show, which likes to run "Best of" episodes on holidays and vacations. Since these episodes would probably play much later (months or even years after they were originally recorded), a means should exist to record this audio before Arbitron encoding is applied.

Remote Broadcasts

Because of the reality of lengthy broadcast obscenity and IBOC blend delays, stations which conduct elaborate remote broadcasts often do not provide an off-air PA feed at a remote venue but rather send a special locally mixed live feed to the remote site. Consideration should be given that public attendees at such an event would not be exposed to Arbitron encoded audio in this case, and if any happen to be PPM survey participants, their listening would not be counted.

Live Monitoring

Stations which use neither obscenity delay nor IBOC delay may question whether the processing being done by the Arbitron encoder in evaluating where to add encoded energy results in enough throughput delay to cause problems with live off-air headphone monitoring.

Very early versions of Arbitron encoders were a concern in this regard; some air personalities were very sensitive to the post-Arbitron encoded sound and found it somewhat unpleasant. However, our experience was that Arbitron worked attentively to address this issue, improved their coding algorithms very significantly, and effectively solved this problem in early field testing many years ago.

INSTALLATION

The Arbitron encoders are easily installed. XLR connectors are provided on the rear panel for audio in and out connections. The device is equipped with a hard relay bypass, so that if power is removed from the unit, a relay releases to directly connect the input to the output connectors.

A front panel key-operated switch provides a similar function; turning the key to the "Bypass" position also connects the input and output connectors. Software inside the unit will also force a bypass if certain error conditions are detected.

The unit has a front panel LED that is green when all is normal, and which changes to flashing red when various faults occur. The unit also has an LCD screen which provides messages about the device's current state (Figure 2).

Figure 2. A portion of the Arbitron encoder front panel

Aside from the input, output, and time code connectors, the back panel has a DB-9 connector with dry contact closures to indicate the status of the Bypass relay.

A vital step in the setup process is to verify with Arbitron that the encoders all work as intended. Consider that in a market-wide rollout, hundreds of boxes will need to be processed, shipped and installed; you would like to know for certain that the box in your station's air chain and that will be responsible for your station's ratings is the correct one, and that it is working perfectly.

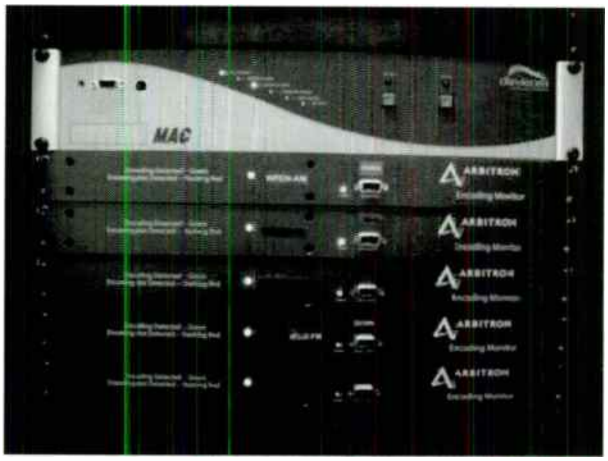


Figure 3. Remote control system and five Arbitron confidence decoders

The confidence decoder installation is just as straightforward. Plug line level audio from the correct air monitor on XLR connectors in to the box, and the front panel LED will almost instantly change from flashing red to steady green.

The leftmost rack contains the monitoring devices: analog tuners, HD tuners, and rack-mounted CPUs for webstream monitoring. These are arranged in frequency order from top to bottom, with the analog tuners first, the HD tuners next, and the CPUs at the bottom. Since monitoring devices for our entire facility are now centrally located, we also located a remote node from our routing switcher to this rack, bringing access to all of these feeds (many of which were never commonly available before) to all routing switcher destinations in our facility.

The middle rack contains one-half of the remote control system, and all of the dedicated silence senses, one per monitored program stream, with each arranged to the immediate right of its corresponding audio source.

Finally, the rightmost rack contains the other half of the remote control system and the Arbitron confidence decoders, again, each one arranged to the right of its corresponding monitoring source. Walking in to the room, it is easy to see at a glance the status of the various devices.

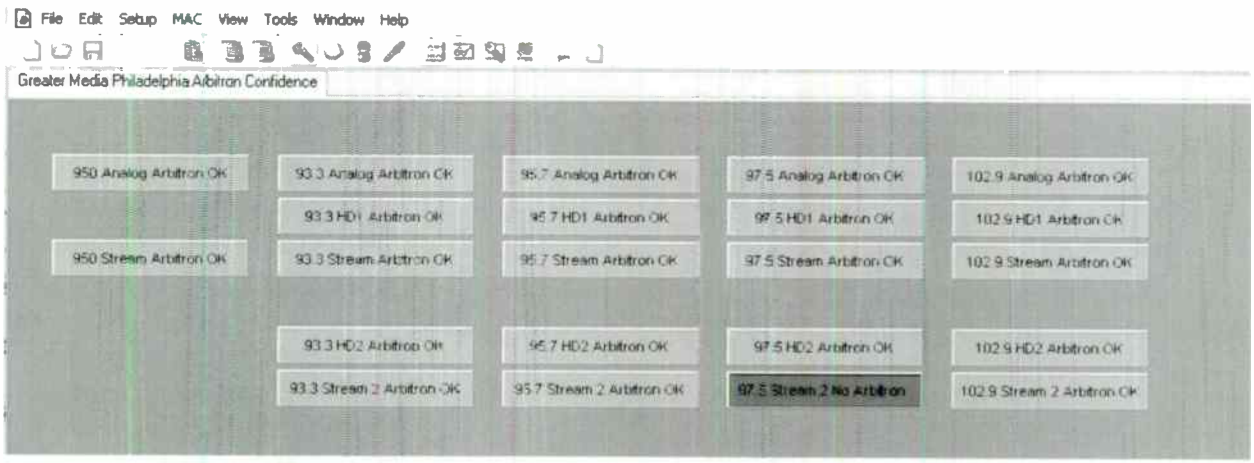


Figure 4. Remote control screen showing Arbitron confidence status

A confidence decoder box will only successfully decode its dedicated companion encoder. For example, WJJZ's decoder will not show a green light if it is hooked to WMMR's off-air monitor, even if WMMR's encoding is working properly. Arbitron says their decoders have an internal three-minute delay before a loss of detected encoding results in an alarm. (See Figure 3).

OPERATION

Greater Media elected to acquire additional space in our building's penthouse mechanical area to hold the three new racks full of Arbitron confidence monitoring equipment we installed.

Blank spaces were left in the rack at appropriate locations for the eventual installation of HD3 monitoring equipment.

The wiring from the devices in each rack was brought out to punchdown blocks which are mounted on painted plywood and attached to the back wall. Audio outputs from the monitoring devices, audio ins and outs from the routing switcher remote node, dry status contacts from the silence senses and Arbitron confidence monitors, and status, telemetry and control contacts from the remote control system are all easily managed and immediately accessible.

The remote control system continuously monitors the various dry status contact closures of the system, and responds on an alarm condition. (See figure 4.) E-mails and pages are sent, and the system can place POTS calls and speak in English to an on-call respondent.

IMPROVEABLE ASPECTS OF THE SYSTEM

A short time after the date we needed to have our system up and running, the manufacturer of the HD tuners we had purchased announced a new product. Rather than a single radio in a 1RU housing, the new product would have three radios in a 2RU housing. For a given station, its analog, HD1, and HD2 tuners could be grouped into this smaller package. Also available with the new package was an integral silence sense for each tuner, brought out to dry contacts! This new radio would have saved four rack units per station, and would have cost considerably less than buying the individual components.

Using less rack space would have also meant that we could have provided some blank space in between units. The DSP-based HD tuners run quite warm to the touch, especially when they are packed in to the rack one atop another.

We have added a remote power cycle feature to each of the tuners in the rack, such that if the remote control notices a silence sense indication of failure on a particular tuner, it can issue a momentary pulse that will cycle the power for that tuner. In such a case, if the silence problem were due to a receiver lock-up condition, the reboot would cure the condition before the Arbitron confidence decoder's three minute time out, preventing nuisance alarms.

Keeping computers running and happily connected to a live webstream is an ongoing issue. Nuisance alarms occur when a machine spontaneously disconnects from its host and comes up silent. We respond to these alarms by logging in via VNC from a remote location and manually restarting the affected computer.

Other audio monitoring points have been considered, such as the post-encoded audio right before it leaves our office on the way to the stream host provider. These alternatives were rejected, because if used, they would be unable to alert us to problems that might develop with the streaming host itself.

A worthy goal would be to develop a small program that could run on each CPU that would respond to either an external contact closure, or perhaps to a serial port string, and which would cause the computer to close all programs and restart. Upon restarting, the computer can automatically connect to the appropriate webstream. If restarting cures the problem, the remote control system could make a log entry and simply send an e-mail concerning the issue. If restarting fails to cure the problem, then the remote control system could take additional action to page the on-call engineer.

SUMMARY

Arbitron's PPM system is a radical change in audience measurement methodology. Careful planning and thoughtful consideration of its many ramifications, as well as the dedication of resources of time and equipment, are suggested to successfully integrate its operation in a broadcast facility. Particularly important is the development and installation of a reliable monitor system that can immediately alert the operator when there has been a failure involving PPM encoding equipment.

SOURCES

The author gratefully acknowledges the use of information contained in these Arbitron publications:

Studio Grade Encoder: 1U Analog Interface. Operations and Field Service Manual.

CBET Encoder Monitor. Operations and Field Service Manual.

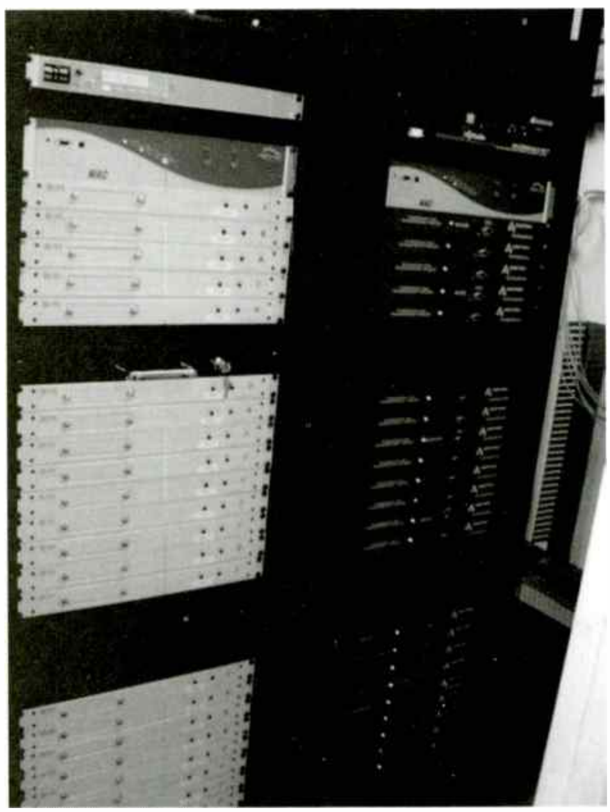


Figure 5. Silence senses and Arbitron confidence decoders

CONSUMER RATINGS OF IMPAIRED AUDIO AT VARIOUS SIGNAL/NOISE RATIOS

Ellyn Sheffield
Towson University

John Kean
NPR Labs - National Public Radio

Dan Schwab
NPR Labs- National Public Radio

ABSTRACT

The purpose of this study was to explore consumers' reactions to audio samples containing varying levels of impairment and to determine the level at which they would no longer continue to listen. Three types of interference were tested at five signal-to-noise ratios using a network-weighted quasi-peak measurement. Results indicate that consumers were sensitive to noise, and claimed they would no longer continue to listen in signal-to-noise conditions of less than 25 to 30 dB.

BACKGROUND

A cornerstone of radio allocations is understanding that a particular RF signal-to-interference (protection) ratio yields a desired audio signal-to-noise ratio. In the U.S., however, minimum audio SNR standards for FM stereo reception are not available. The purpose of this test was to determine how consumers would rate audio samples with various types and levels of noise impairment and at what audio SNR consumers would turn off the radio because of the impairment. Three different kinds of interference were tested at five different levels. Seven different audio clips were used to simulate all possible types and styles of broadcasts.

The audio noise meter chosen for this study complies with the ITU-R 468 standard, which combines a quasi-peak reading audio voltmeter with a frequency-weighting curve to objectively measure audio noise similar to the human ear. This instrument, sometimes called a "psophometer" is widely used when measuring noise in audio systems, especially in the UK and European countries.

Most audio engineers are familiar with the A-weighting curve, which is said to reflect the 'equal-loudness contours' derived initially by Fletcher and Munson (1933). However, these curves relate only to the subjective loudness of pure tones, not noise. Developments in the 1960's, spread by audio tape recording and FM broadcasting, indicated the need for a better weighting curve. The ITU-R 468 curve was developed to better match the ear's response to low-level noise. As shown in Figure 1, the curve rises at a 6 dB/octave rate to 6.3 kHz, where it has 12 dB of gain (relative to 1 kHz). From here, it quickly attenuates high frequencies at approximately 30 dB/octave.

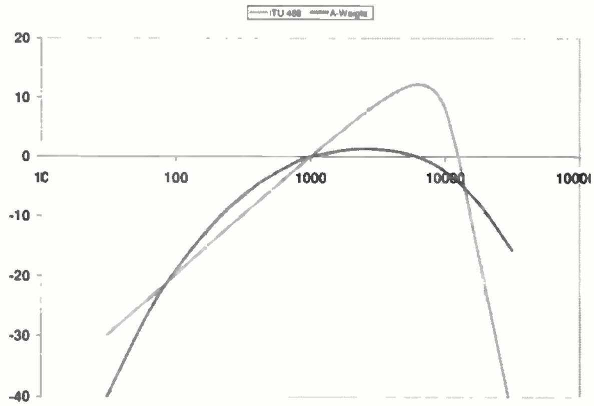


Figure 1- Frequency response curves for ITU 468 (red) and A-weighting (blue); vertical scale in dB, horizontal in Hz

It is important to note that the ITU 468 specification uses a very special quasi-peak rectifier with carefully devised dynamics (A-weighting uses RMS detection). Rather than having a simple 'integration time' this detector requires implementation with two cascaded 'peak followers', each with different attack time-constants carefully chosen to control the response to both single and repeating tone-bursts of various durations. This ensures that measurements on impulsive noise take proper account of the ear's reduced hearing sensitivity to short bursts. The ITU 468 measurements are referred to herein as a weighted quasi-peak signal-to-noise ratio (WQPSNR).

STIMULI

Seven audio samples were processed with three types of interference (pink, white Gaussian, and USAI pulsed) at five WQPSNR levels (45, 40, 35, 30, and 25 dB). Additionally, an unimpaired reference sample of each clip was included for a total of 112 clips presented to consumers. Audio samples lasted between 15-20 seconds. They included a low density music selection (e.g., Edward Gerhard's "If I Feel/In My Life"), a medium density music selection (e.g., Jimmy Buffet's

"I don't know and I don't care"), a high density music selection (Fleetwood Mac's "Go your Own Way"), a female-voice commercial for NPR's "Fresh Air", a male-voice commercial for NPR's "Morning Edition", a passage of female speech from NPR's "All Things Considered", and a passage of male speech from "All Things Considered".

In order to generate audio samples for listening sessions, the reference samples were passed through a Telos Omnia processor for standard broadcast level processing and compression. The composite analog stereo FM signal of the Omnia was modulated by a Hewlett Packard 8647A signal generator that served as the "desired" signal source. The resulting signal was then mixed with lower first-adjacent channel RF signals to generate varying degrees of audible impairments to the desired signal. These interfering signals were produced by the combining the FM generator with a hybrid (HD Radio) signal, produced by a Harris Dexstar exciter at the standard 1% power ratio. The analog FM signal was modulated as follows:

1. Modulated with audio pink noise and set to a level that caused an FM modulation monitor to indicate 100% modulation peaks at about ten second intervals.
2. Modulated with white Gaussian noise and set to a level that caused an FM modulation monitor to indicate 100% modulation peaks at about ten second intervals.
3. Modulated by pulsed-USASI noise at a level that caused an FM modulation monitor to indicate 100% modulation peaks.

A Pioneer model VSX-D814 home theater receiver was used to receive the resulting signal. The audio samples were then recorded directly to an audio CD-R disk.

PARTICIPANTS

Thirty listeners (12 males and 18 females) between the ages of 21 and 65 were recruited for this consumer test. Fifteen of the participants (4 males and 11 females) were employees of National Public Radio. They were contacted through a mass email to all National Public Radio staff. The other 15 participants (8 males and 7 females) were recruited via a posting on websites.

TEST ENVIRONMENT

Testing was held in Broadcast Studio 5A at National Public Radio Headquarters in Washington D.C. A Dell Dimension GX870 with a Samsung Sync Master 765MB monitor were used to run custom testing software to administer the test. A Creative Sound Blaster Audigy-LS sound card converted the wave files to analog audio, which were sent to an APHEX Model 124A audio interface unit. This output of the interface unit was carried by balanced XLR audio cables to the stereo pair of Mackie HR 824 self-amplified professional monitor speakers that provided the audio stimulus for the test.

TEST METHODOLOGY

Prior to testing, participants were read instructions which explained the testing procedure. Participants were tested individually. They were allowed to set the volume on the first trial of the test, after which the volume remained constant. Participants were presented with one audio clip at a time and after listening to the clip, they were asked 3 questions:

1. On a scale of 1-6 (excellent, good, fair, poor, fair, failure), please rate the overall audio quality of the sample
2. On a scale of 1-4 (extremely annoying, slightly annoying, not annoying and no noise), rate how annoyed you were with background noise
3. Given the audio quality and background noise, would you keep the radio on or turn it off?

RESULTS – OVERALL QUALITY

Table 1 shows the MOS scores by signal-to-noise ratios, divided by Interference Type and Sample Type. Analysis of variance (ANOVA) indicates that there was a main effect of signal-to-noise ratios (Level), a main effect of Sample Type (commercial, music and speech) and a significant interaction. There was no meaningful difference among types of background interference noise. Therefore, interference noise types were collapsed for further analysis and presentation.

Speech was the most susceptible to interference. At the worst interference level (25 dB) speech was rated at .94, slightly below "bad", compared to 1.16 for Commercials and 1.97 ("poor") for Music. Ratings gradually got better as signal/noise ratios improved. Rating differences between 45 and 40 dB appeared, but were fairly small. Sharper differences occurred at 35 dB, and in the speech genre mean scores begin to drop below 3.0, numerically equivalent to under "fair". By 30 dB all genres are clearly negatively affected, with participants rating all audio below 3.0.

Table 1: Overall Quality Scores

Interference	Level	Commercial	Music	Speech	Total
NRSC	-45	3.8	4.1	3.7	3.9
	-40	3.8	3.9	3.1	3.6
	-35	3.2	3.6	2.3	3.1
	-30	2.5	3.0	1.6	2.5
	-25	1.5	2.2	1.0	1.7
Pink	-45	3.8	4.2	4.0	4.0
	-40	3.5	3.8	3.5	3.6
	-35	2.7	3.4	2.4	2.9
	-30	2.0	2.7	1.5	2.1
	-25	1.0	1.8	0.9	1.3
WGN	-45	3.8	4.0	3.9	3.9
	-40	3.4	3.9	3.0	3.5
	-35	2.6	3.6	2.3	2.9
	-30	1.7	2.8	1.7	2.2
	-25	1.0	2.0	0.9	1.4
	Unimpaired	4.0	4.2	4.6	4.3

RESULTS – BACKROUND NOISE

Tables 2, 3 and 4 show the number of participants expressing annoyance at each signal-to-noise level. Notice that for all audio samples, at 25 dB an overwhelming majority of people express that they heard extremely annoying or annoying background noise. At 30 dB the majority of participants now express hearing slightly annoying or annoying background noise. In the speech genre, however, a majority still report hearing annoying or extremely annoying noise. At 35 dB, the picture changes substantially. Now, for commercials only 22% report that the noise was extremely annoying or annoying. For music, only 15% report the noise as extremely annoying or annoying. For speech, most affected by interference, 37% still report the noise as extremely annoying or annoying. In all cases, the MOS and Annoyance scores are negatively correlated – the more annoying the background noise, the lower the MOS.

Table 2: Percentage of participants claiming that they heard background noise – Speech

	Speech					
	-45	-40	-35	-30	-25	unimp
No noise	0.39	0.12	0.02	0.00	0.00	0.92
Not annoying	0.41	0.36	0.13	0.02	0.00	0.05
Slightly annoying	0.18	0.36	0.48	0.19	0.05	0.03
Annoying	0.02	0.16	0.29	0.56	0.33	0.00
Extremely annoying	0.01	0.01	0.08	0.22	0.62	0.00

Table 3: Percentage of participants claiming that they heard background noise – Commercials

	Commercial					
	-45	-40	-35	-30	-25	unimp
No noise	0.38	0.26	0.11	0.03	0.01	0.40
Not annoying	0.40	0.42	0.30	0.11	0.00	0.50
Slightly annoying	0.18	0.24	0.37	0.31	0.14	0.05
Annoying	0.03	0.08	0.16	0.37	0.37	0.03
Extremely annoying	0.01	0.00	0.06	0.18	0.48	0.02

Table 4: Percentage of participants claiming that they heard background noise – Music

	Music					
	-45	-40	-35	-30	-25	unimp
No noise	0.60	0.51	0.37	0.21	0.07	0.76
Not annoying	0.29	0.29	0.29	0.20	0.13	0.17
Slightly annoying	0.10	0.16	0.20	0.27	0.22	0.07
Annoying	0.01	0.04	0.11	0.23	0.31	0.01
Extremely annoying	0.00	0.00	0.04	0.09	0.27	0.00

RESULTS – LEAVE-ON RATES

With regard to leave-on rates, both MOS and annoyance scores were correlated. MOS to Leave-on was positively correlated, and Annoying to Leave-on was negatively correlated. Table5 shows the signal-to-noise ratio of interference and the percentage of people claiming they would keep the radio on. Notice that for the 45 and 40 dB levels almost all listeners would keep the radio on. At the 35 dB level, the percentage of listeners reporting that they would leave their radios on is still very high. However, once again speech is the most affected by the interference as evidenced by the percentages of people keeping the radio on. At the 30 dB level of interference, all three categories of genres are negatively affected by the interference with the majority of participants claiming they would turn off their radio.

Table 5: Percentage of Participants claiming that they would leave the radio on

Interference	Level	Commercia	Music	Speech	Total
NRSC	-45	100%	99%	98%	99%
	-40	98%	94%	90%	94%
	-35	97%	88%	65%	84%
	-30	65%	74%	25%	58%
	-25	37%	49%	10%	34%
Pink	-45	97%	100%	100%	99%
	-40	95%	96%	100%	97%
	-35	80%	91%	78%	84%
	-30	58%	72%	40%	59%
	-25	20%	43%	22%	30%
WGN	-45	98%	98%	100%	99%
	-40	98%	96%	93%	96%
	-35	78%	90%	70%	81%
	-30	55%	78%	48%	63%
	-25	23%	49%	22%	34%
	Unimpaired	98%	98%	100%	99%

CONCLUSIONS

In this study we examined listeners' attitudes towards audio that was recorded at different signal-to-noise ratios. We first asked them to rate the audio using a modified ITU-R recommended MOS scale. We then asked them to identify background noise and rate how annoying it was. Finally we asked them to tell us the point at which they would turn off their radio, given the background noise in relationship to the desired signal.

We found that participants were sensitive to background noise, as shown by the increasing displeasure as more noise was inserted on the desired signal. This was particularly apparent for speech, which allows more background noise through than dense audio, such as processed music. Mean opinion scores showed that at approximately 35 dB WQPSNR people were becoming no longer satisfied with audio, rating it "fair". At 30 dB listeners became so discontent that they rated what they heard as "poor".

Participants rated noise slightly more favorably, with the greatest number of complaints coming at 30 dB. Although at 35 dB a majority of participants heard noise, the largest percentage claimed it was "slightly annoying" rather than "annoying" or "extremely annoying". At 30 dB, however, an overwhelming majority began to strenuously object to the background noise, especially when listening to speech.

Interestingly, listeners were more likely to complain about the audio in their quality ratings than they were willing to turn the radio off. Although listeners rated audio at 35 dB WQPSNR as fair (3.0), over 80% claimed they would continue to listen to the program. Percentages were particularly high when they were listening to music and commercials. At 30 dB, when listeners were now claiming that the audio was poor (approximately 2.2), approximately 60% were claiming they would leave the radio on. Thus, participants seemed most prone to changing their behavior (turning off the radio) when noise reaches a level typically heard at 30 dB or 25 dB WQPSNR.

REFERENCES

Fletcher, H. and Munson, W. A., (1933) Loudness, its definition, measurement and calculation" in J.Acoust. Soc Am.5, 82-108

ACKNOWLEDGEMENTS

We would like to thank Jan Andrews and Matt Burrough of NPR Labs for their technical support. Further we thank the folks at NPR for participating in this test.

Video Content Creation & Manipulation

Tuesday, April 15, 2008
9:00 AM – 5:00 PM

Chairperson: John Turner
Turner Engineering, Mountain Lakes, NJ

File Formats in Television Archiving and Content Exchange
Peter Thomas, Blue Order Solutions AG, Kaiserslautern, Germany

***Watermarking and Fingerprinting: The Wave of the Future**
Andy Nobbs, Teletrax, New York, NY

Non-Real-Time Services
Richard Chernock, Triveni Digital, Princeton Junction, NJ

***All-Digital Media: Best Methods for Integrating and Distributing**
Robert Lerner, Gefen, Chatsworth, CA

BXF - The Promise of Reduced Costs and Increased Revenues
Chris Lennon, Harris Corporation, Mason, OH

Seam Carving for Video
Mike Knee, Snell & Wilcox Ltd., Havant, UK

An Integrated, File-based Production Workflow for HD Television: Expected Impact and Challenges
Luk Overmeire, VRT Medialab, Ghent - Ledeberg, Belgium

From MXF to SOA
Ernesto Santos, MOG Solutions, Maia, Portugal

Migration to All-IP Infrastructures for Distribution of Broadcast Services
Tom Lattie, Harmonic Inc, Sunnyvale, CA

Forensic Marking for HD VoD and Broadcast Services
Pascal Marie, Thomson, Cesson-Sévigné, France

From Camera to the Home -- Managing Aspect Ratio through the Production and Distribution Process
Larry Thaler, NBC-Universal, New York, NY

***Leveraging IT Technologies and Concepts to Enhance HD Sports Programming**
Luis Estrada, IBM, Atlanta, GA

***Viewer Contribution – Dealing with Massive Media**
Fred Fourcher, Bitcentral, Irvine, CA

The State of Broadcast Automation
Sid Guel, Broadcast Automation Consulting, San Antonio, TX

---

*Paper not available at the time of publication

FILE FORMATS IN TELEVISION ARCHIVING AND CONTENT EXCHANGE

Peter Thomas
Blue Order Solutions AG
Kaiserslautern, Germany

ABSTRACT

As broadcasters migrate to digital file-based acquisition, production, distribution and archiving, a plethora of questions continue to arise. One is the selection and the operational use of file formats in television. What is the best file format for my overall operation? Should I use embedded descriptive metadata, and if yes, where and which? Which benefits and penalties are associated with a certain file format selection? Are there interoperability issues when exchanging content between systems or enterprises?

This paper tries to provide answers to some of such questions and some guidelines for the decision making process, starting by reviewing file format related requirements for archiving and file exchange. Then, embedded metadata is discussed for archiving and file exchange, and an opinion is given on the usefulness of descriptive embedded metadata in these scenarios. In addition, the advantages and disadvantages of using MXF vs. Quicktime as an archive and file exchange wrapper are discussed, and a proposal is made when to use what. The paper concludes with a summary set of recommendations for selecting file formats and using embedded descriptive metadata.

INTRODUCTION

As broadcasters migrate to file-based production, distribution and archiving in television, the question of the selection of the right file format is becoming of vital importance for the successful and efficient implementation of the desired new file-based workflows and business processes. Key business requirements for such a file format are:

- Supports all core processes in television without need for time-consuming transformations;
- Minimum number of transcodings and re-wrappings outside of core processes;
- Has a credible future;
- Is backwards compatible to older versions.

Core processes include file import and signal ingest, post-production, playout and archiving (Figure-1). File import typically requires a transcoding of an arbitrary format to the house standard archive format. Non-playout delivery, such as file exchange or distribution via Web, IPTV, VOD or mobile, typically requires at least a transcoding step for delivery format creation.

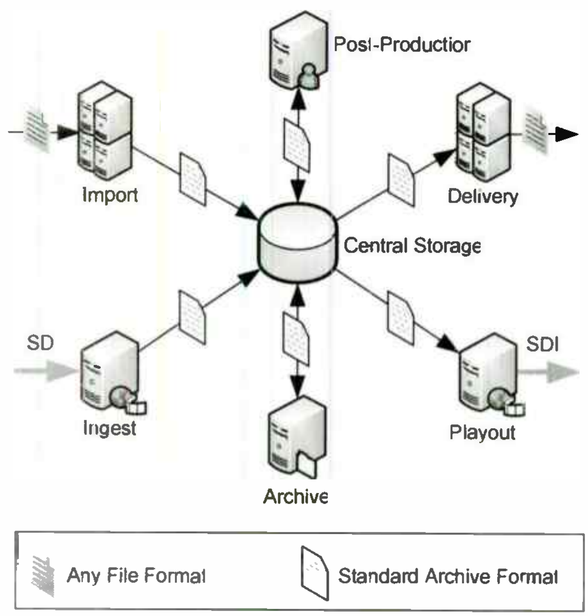


Figure-1

As depicted in Figure-1 a single standard archive format used in all core processes would optimize all processes. This ideal situation is not always achievable. Still, the selection of the optimum file format is a somewhat holistic approach, not to be dominated by one department or group, but to be taken as an enterprise-level business decision.

FILE FORMATS

A file format comprises two key elements which are mostly independent from each other

- The encoding format
- The wrapper format

For standard definition (SD) video material, popular formats are D10 [1] (especially D10 at 50 MBit) and DV/DIF [2].

For high definition (HD) video material, as of today no common denominator has emerged. Businesses may use different encoding formats for different business processes, but should, wherever possible, avoid transcoding, as it introduces a generation and thus reduces quality. The digital archive must be able to transparently archive and restore any format, without any change.

The most useful wrapper formats in television archiving today are MXF OP 1a [3, 4] and Quicktime [5]. In post-production, also MXF OP Atom [6] is used. For MXF, the way how D10 and DV essence is embedded in the Generic Container [7] is also standardised by SMPTE [8, 9]

FILE FORMAT REQUIREMENTS

Requirements for Television Archiving

For use in television archiving, the file format has to meet a number of requirements. With respect to the encoding format, the major requirements are:

- Well-known and standard encoding scheme;
- All encoding parameters, including bit rate, defined;
- Widely supported in the industry;
- As a minimum, ensured compatibility between the products used in-house for acquisition, production, delivery and archiving;
- Allow for partial restore;
- A more than one transcoding system exists that supports the format as a source.

For the wrapper format, the following requirements primarily apply:

- Open and well documented;
- Ideally standardised;
- Widely supported in the industry;
- Long term perspective for being supported by numerous vendors ("future proof");
- Avoid lengthy re-wrap procedures in time-critical business processes (especially in production);
- Allow for partial restore;
- Allow for play while record;
- Lightweight – very little overhead compared to the payload;
- Mappings for the selected encoding formats exist and are well documented;
- Support for embedded technical metadata.

Please note that, on purpose, support for embedded descriptive metadata is not listed as a requirement for the file format for archiving. The reason for this will be provided in the section on embedded metadata below.

Requirements for File Exchange

Requirements for the encoding and wrapper format for file exchange between businesses (B2B) or between businesses and consumers (B2C) are somewhat simpler.

For the encoding format, it is important that the format meets the specifications as agreed upon in the service contract.

For the wrapper format, it is important that

- The wrapper format meets the specifications as agreed in the service contract;
- The wrapper format can embed both technical and descriptive metadata.

In file exchange you want to embed a subset of the metadata known to your business in order to meet the requirements of your service contracts, but you may not want to give away all information you manage internally. Hence, the way from the archive to delivery typically includes two processing steps:

- Creation of the desired wrapper and encoding format via transcoding;
- Embedding a selected set of metadata into the wrapper.

The requirement to exchange in MXF OP 1a does not imply that a business has to archive in MXF OP 1a. The processing effort to create a new wrapper is small compared to the frequently required encoding format changes, and metadata updates will also be standard procedure – hence, the file is transformed anyway.

For file exchanges within your business you want to avoid any form of transformation to the utmost extend, as they are time consuming and, in case of encoding format changes, introduce generation losses.

A transformation that often can not be avoided is when businesses use MXF OP 1a as the wrapper format for ingest, delivery and archiving, but the post production system requires OP Atom. However, this transformation is fast and lightweight, does not introduce any generation losses as it does not touch the encoded essence, and hence can be readily accepted. A disadvantage is that the tools that perform this re-wrap may induce additional cost.

EMBEDDED METADATA

General Considerations

The fact that modern file formats allow the embedding of technical and descriptive metadata is an enticing feature, and will be an enabler for a number of future workflows in file-based distribution and exchange of content – and potentially also in various production workflows.

However, there still are some caveats. Whilst MXF provides an elegant means to embed metadata and specifies DMS-1 [10] as a metadata framework, there is little to no recommendation available that would permit the tying down of the entire semantics of a metadata element. Using SMPTE metadata dictionary unique keys [11, 12, 13], the semantics of an attribute itself can be defined. However, a fully specified *standard reference data model* would be required to complete the semantics – tying the attribute to a specific entity in this model.

If such a data model would be available, an organisation could map its own data model to this reference model and would thus ensure that the semantics of the metadata embedded in the file is clearly articulated. If an organisation just maps their own proprietary data model to DMS-1, the result is just as proprietary as the data model – no other organisation has knowledge of the semantics and hence can not reliably interpret the data. This has been made quite obvious in, e.g., the 2004 workshop on the use of DMS-1 at IRT (Institut fuer Rundfunktechnik), where businesses presented the way they implemented their respective metadata in DMS-1 – the results of all of these efforts were incompatible.

Furthermore, changes in the corporate data model that may be required for valid business reasons may change the naming or, worse, the semantics of certain attributes. This may render the file level metadata in files created prior to this change at least partially incompatible with the company's own catalogue.

Hence, at this stage, using file level descriptive metadata can only be recommended in very well defined scenarios, and only for transient operations, such as file exchange between two systems or two organisations for which the following is unambiguously specified at the time of the exchange:

- Set of metadata to be exchanged;
- Full semantics of this metadata.

For exchange between businesses, the set may differ depending upon the type of transaction and the relationship between those businesses. Between systems and products, the set may differ depending upon the releases or versions of the respective products.

Today, experience gained in several projects shows that the exchange of descriptive metadata between existing systems and products is at least challenging. Even for technical metadata, difficulties still exist. For MXF, examples are the interpretation of timecode which can live in various locations and may differ between these locations, or the handling of the frame rate for interlaced video (frame or fields). Upcoming releases of the MXF standard will provide clarification on many issues for technical metadata, which will continue to improve interoperability between products. For descriptive metadata, recommendations and standards have yet to be defined to enable interoperability.

File-Level Metadata in Archives

The usefulness of descriptive metadata in files that are to be archived is, unfortunately, questionable. At the first glance, there seems to be an advantage in maintaining at least parts of the descriptive metadata in the file when archiving:

- The contents of the file can be identified even when no database is available that would reference the file;
- In case of a loss of the database, basic information can be restored from the file.

In order to qualify the first point, we have to understand that, in a digital archive, hundreds of thousands of files will reside on IT storage systems, primarily data tape storage vaults. There simply is no way that a user could find a file via exploration of file level metadata, as this metadata is not searchable. Only when maintaining metadata in a database, or as an index in a search engine, can users search and find content.

As a potentially richer set of metadata is present in the database of the Media Asset Management System (MAM, [13]) or TV catalogue, using the search functions of these systems is the only sensible way to search for a file. In the case where some files that just happen to be somewhere on some disk or other, again a search in the MAM or the catalogue using the unique filename provides much faster and richer information about the content of the file.

For the second point, the IT industry agrees that the right way to protect a database is using standard IT database backup. In the event that a database fails, restore is typically a matter of, at maximum, hours, if using a single data tape drive. In case of an archive of 100,000 hours of content in DV50 with 8 audio tracks (48kHz, 24bit), it would take close to 260 days of restore time using a single LTO-4 drive, and even if the process would use 10 drives in parallel, it would still take almost four weeks to retrieve and analyse the files – assuming that all files have to restored entirely, which would be required if there is metadata associated to the timeline. There are ways to optimize this procedure by introducing certain restrictions (e.g., only header metadata is stored in the file), but still the process would be, to say the least, inconvenient. So using the files as "belts and suspenders" for database loss is not really feasible – it is more economic to invest in standard IT protection mechanisms and apply related best practices.

The second problem with descriptive metadata in archived files also stems from the fact that, due to technical limits of other storage technologies, the primary long term storage technology applied to television is digital data tape. Here, it is very difficult to apply updates to metadata hosted in the file. You cannot simply overwrite the header; you have to retrieve the entire file, apply the change to the header, re-write the file to tape as a new file, and mark the former version as invalid. The space that the invalid version occupies on the data tape has to be recovered by tape defragmentation, which typically involves copying all valid files from the original tape to a new tape, and release the original tape into the pool for re-use. In short, this procedure would involve too much overhead to be applied in real-world applications.

Moreover, some businesses hold a second copy of each data tape on-the-shelf as a backup for content. For such copies, an update would require manual operator intervention to retrieve the data tape from the shelf and perform the procedure described above.

Another problem is that a file may reside in multiple, and potentially remote, locations at the same time, e.g., in local archive storage systems at remote sites, provided for faster access to frequently used content. Provided the deployed media asset management system is aware of all such locations, file header metadata updates have to be performed, as a "distributed transactions", on all those managed copies of the file. Otherwise not only the database and the primary archived file go out of synch, but also the files in the various storage locations would have different metadata. "Unmanaged" files go out of synch anyway.

Unfortunately, metadata updates are quite common in a database that is managing descriptive information. Even basic metadata, such as titles, is sometimes changed throughout the lifetime of a content element. Hence, there is a high probability that file header metadata in the archive is effectively outdated (metadata in the file and in the database go out of synch) and has to be quality controlled before is used. Quality control effectively means that, upon delivery of a file from the archive to a target that requires such file header descriptive metadata (e.g., for programme exchange or content delivery to external customers), the metadata would have to be updated. Hence, the use of such metadata is very limited.

If an update of the metadata is required anyway, it can be also be put in at delivery time as a standard process, independent of the prior existence of any metadata.

The situation may improve in the future, provided that the following criteria are being met:

- A standard reference data model exists;
- A mapping of the corporate data model to the reference data model exists;
- A set of metadata is selected for which corporate business rules prohibit any change after archiving of the files;
- Only this is embedded in the file.

Still, even if these criteria are met and strictly enforced, when delivering the file it will frequently be necessary to enrich the "core" metadata set with additional and current data from the database. So operationally, there still is no major advantage.

Metadata in File Exchange

Being able to embed metadata in files to be exchanged between businesses (B2B) or businesses and consumers (B2C) is very useful, as this allows the tight coupling of metadata and essence in delivery scenarios that are beyond the control of the delivering entity.

Within a business, file transfers and file exchange between systems and products are under control of the business. Hence, metadata exchange can also be accomplished in other ways that may be more efficient or less complex. Examples are partial database synchronisation, exchange of metadata via an enterprise application integration bus, or file level exchange via XML.

However, as soon as the delivery target is external to the business, file level metadata makes a lot of sense – it is not by accident that the "X" in MXF stands for "exchange".

Today, due to the lack of recommendations and standards with respect to reference data models that would uniquely define the semantics and the context of the metadata on attribute and entity level, partners that desire to exchange content with embedded metadata have to agree upon

- The set of metadata to be exchanged;
- Full semantics of this metadata.

The set may differ depending upon the type of transaction and the relationship between partners. From a system perspective, this means it must be possible to:

- Embed selected metadata into the file prior to

delivery, where the selection may depend upon who will receive the content and for what purpose. The embedding has to follow jointly agreed upon semantics;
- Read metadata from files received from partners, applying the jointly agreed upon semantics, and transfer the extracted metadata, in full or partially, to the internal database.

Whether the metadata now remains in the file when it is archived, or is deleted is of little importance. It is important to remember, though, that the metadata will have to be updated when the file is retrieved from the archive and delivered to another partner, as it may have been changed, augmented or updated to reflect house standards or recent events, or because the exchange involves a different set of metadata and/or different semantics.

DISCUSSION OF OPTIONS

Until recently, the obvious choice for the archive and exchange file format was MXF OP 1a. Even though some non-linear editing (NLE) systems prefer to internally use MXF OP Atom, this re-wrap does not introduce any major latencies and can be performed while transferring the file to the NLE or its shared storage system, if available, and back.

Most popular ingest and playout servers now support MXF OP 1a, non-linear acquisition formats deliver content in MXF (OP 1a for Sony XDCAM, OP Atom for Panasonic P2), transcoders and other important software tools support MXF – so even though incompatibilities still exist between the various individual implementations of MXF in the different products, basic interoperability between products along the production chain is given for most combinations of such products.

However, one product has very successfully entered the market that changes this picture – Apple Final Cut Pro (FCP). As of today, FCP does not natively support MXF – and there is at least no indication that this situation will change in the near future. FCP uses Quicktime – a format widely accepted in the IT industry, and a technology that has been created, and is being strongly supported, by Apple.

Businesses that decide to use FCP face the situation that all MXF-wrapped material has either to be re-wrapped before being delivered to FCP, or a separate Quicktime file pointing to the MXF file has to be created in order to make FCP believe it actually is working in Quicktime. Content created on FCP is rendered to Quicktime and requires re-wrap to MXF OP 1a before it can be used in an MXF environment. This increases system complexity, increases processing overhead, introduces latencies, adds cost, and increases risk. Hence, businesses that are using, or intend to use, FCP as the predominant editing platform may want to reconsider the choice of archiving format.

Requirement	MXF	QT
Open and well documented	✓	✓
Ideally standardised	✓	✕
Widely supported in the industry	✓	✓
Long term perspective for support by numerous vendors (“future proof”)	✓	✓
Avoid lengthy re-wrap procedures in time-critical business processes	✓	✓
Allow for partial restore	✓	✓
Allow for play while record	✓	✓
Lightweight – very little overhead compared to the payload	✓	✓
Mappings for the selected encoding formats exist and are well documented	✓	✓
Support for technical metadata	✓	✓
Support for descriptive metadata	✓	✓

Table-1

As Quicktime fulfils all of the requirements for a file exchange and archive wrapper format as given above (except being standardised, see Table-1), Quicktime would be a very valid archive format for such businesses. A prerequisite is that technology is implemented that uses Quicktime across the entire production chain (ingest, production, delivery and archiving), thus avoiding re-wrapping entirely. If this is not possible, because a technology choice requires using a non-Quicktime wrapper in a vital system used in the core processes, MXF continues to be the best choice. Also, it is absolutely indispensable that businesses opting for Quicktime as production and archive format still add re-wrap functionality to their environment in order to be able to both deliver and receive MXF wrapped content via file exchange.

CONCLUSION AND RECOMMENDATIONS

Based on the thoughts and opinions presented above, as a conclusion this paper gives the following recommendations:

- Use a single codec throughout all your core processes to avoid transcoding latencies and generation losses. For SD, use either D10 or DV, ideally at a video bandwidth of 50 Mbit/s. For HD, analyse your system landscape, revisit your investment strategy, and see whether you can identify a single codec that can be used in all core processes;

- Decide for a unique file wrapper format to be used throughout all of your core processes. Decide for Quicktime, if **all** of your systems involved in your core processes support Quicktime, otherwise decide for MXF OP 1a. Using MXF OP Atom in post-production is possible with very limited performance penalties for re-wrapping if MXF OP 1A is used as the standard wrapper elsewhere;
- Declare the selected combination of Codec and wrapper to be your house archive format and transcode all incoming material immediately to this format;
- Do not use descriptive metadata in files to be archived;
- Use descriptive metadata in file exchange between systems only after clarifying the capabilities of the individual systems used and mapping the respective data models and semantics of each system to your corporate data model. Then, for file exchange use the MAM system to embed the required metadata on-the-fly as part of the file transfer to or between systems. Consider easier alternatives for metadata synch between systems;
- Identify required descriptive metadata for file exchanges for B2B or B2C according to the individual service contracts you have with your respective business partner. Use the MAM system to embed the required metadata on-the-fly as part of the file delivery process.

REFERENCES

1. SMPTE 365M-2001 (Archived 2006) Digital Television Tape Recording – 12.65-mm Type D-10 Format for MPEG-2 Compressed Video - 525/60 and 625/50, SMPTE, U.S.A., 2004.
2. SMPTE 314M-2005 Television – Data Structure for DV-Based Audio, Data and Compressed Video – 25 and 50 Mb/s, SMPTE, U.S.A., 2004.
3. SMPTE 377M-2004 Television – Material Exchange Format (MXF) – File Format Specification, SMPTE, U.S.A., 2004.
4. SMPTE 378M-2004 Television – Material Exchange Format (MXF) – Operational Pattern 1a (Single Item, Single Package), SMPTE, U.S.A., 2004
5. QuickTime Documentation, http://developer.apple.com/documentation/QuickTime/, Apple, Inc., 2007.
6. SMPTE 390M-2004 Television – Material Exchange Format (MXF) – Specialized Operational Pattern "Atom" (Simplified Representation of a Single Item), SMPTE, U.S.A., 2004.
7. SMPTE 379M-2004 Television – Material Exchange Format (MXF) – MXF Generic Container, SMPTE, U.S.A., 2004
8. SMPTE 386M-2004 Television – Material Exchange Format (MXF) – Mapping Type D-10 Essence Data to the MXF Generic Container, SMPTE, U.S.A., 2004.
9. SMPTE 383M-2004 Television – Material Exchange Format (MXF) – Mapping DV-DIF Data to the MXF Generic Container, SMPTE, U.S.A., 2004.
10. SMPTE 380M-2004 Television – Material Exchange Format (MXF) – Descriptive Metadata Scheme-1 (Standard, Dynamic), SMPTE, U.S.A., 2004
11. SMPTE 335M-2001 Television – Metadata Dictionary Structure), SMPTE, U.S.A., 2004.
12. RP 210.8-2004 Metadata Dictionary Registry of Metadata Element Descriptions, SMPTE, U.S.A., 2004.
13. SMPTE 336M-2007 Data Encoding Protocol Using Key-Length-Value), SMPTE, U.S.A., 2004.
14. Mauthe, A. and Thomas, P.: Professional Content Management Systems – Handling Digital Media Assets, J. Wiley&Sons Ltd, Chichester, UK, 2004.

NON-REAL-TIME SERVICES

RICH CHERNOCK
Triveni Digital
Princeton Junction, New Jersey

ABSTRACT

Television viewers are increasingly becoming used to an "on-demand" world. DTV technology is rapidly changing to enable new consumption and distribution models – receiving devices contain persistent storage, personal media players are commonplace and inter-device connectivity is practical. These factors combine to allow a shift from linear TV viewing to on-demand consumption of content. One of the main enablers of this shift is the capability for Non-Real-Time (NRT) delivery of content – content that is delivered in advance of use and stored for access on demand.

NRT content includes "traditional" TV fare (presented in a customized and non-traditional way) as well as information not currently aimed at the TV (including content targeted to PCs, handheld media players or even commercial platforms).

Adding non-real-time services to traditional linear programming allows service providers to provide a more valuable package to their customers. This paper will introduce the basic concepts of non-real-time services, give illustrative scenarios and discuss technical requirements for building NRT services.

WHAT ARE NON-REAL-TIME SERVICES?

As the name implies, NRT services consist of non-real-time delivery of content – content delivered in advance of use and stored for later consumption by (or on behalf of) the viewer. Conventional (or linear) television services consist of streams of audio and video data that are consumed in real time. Historically, linear streaming is what is typically associated with television. However, when one examines the field of entertainment, it is apparent that many alternatives to linear television have appeared that compete for the viewer's attention. In recent years, people have turned to the Web, DVDs, personal media players, DVRs and VOD services as sources for entertainment. This has led to a growing desire for "everything on demand" or "I want it now". The viewers are clearly interested in the content, not in how it is delivered.

The adoption of NRT services will allow the broadcaster to satisfy the viewers desires for immediacy. Existing broadcasters have significant advantages for the delivery of NRT: high bandwidth wireless delivery of content, broadcast economics, local presence and (soon) the ability to deliver the same (or similar) content to mobile and handheld devices.

NRT HISTORY

The idea of NRT services is not new – under the name of data broadcasting, technologies for pushing content to receiving devices, standardization of some of the components and even commercial deployments have been available for many years. The majority of the current NRT type services have taken the form of B2B (Business to Business) and G2G (Government to Government) and have proven to be sustainable businesses. The early attempts at B2C (Business to Consumer) did not prove to be sustainable.

However, as mentioned above, the consumer environment has changed. Television viewers have become more accustomed to the "digital age". Entertainment is no longer a purely passive event (the stereotypical "Joe 6-pack") – viewers are used to interacting with web pages, DVDs, DVRs and the other alternatives. Similarly, the On-Demand concept has become familiar and expected. Persistent storage has become commonplace on receiving devices (driven by the spread of DVR capability). The broadcast DTV pipe, which is well suited for delivering these services, is now widespread and available in all DMAs. Standardization activity is well underway within ATSC, which will allow mixing and matching broadcast/receiving devices. Many of the factors that contributed to the lack of success of B2C NRT services in the past have been overcome and there is good reason to expect that these services are now becoming viable.

EXAMPLE NRT SERVICE SCENARIOS

There are a number of potential scenarios for NRT services that have been agreed upon by the broadcast industry. Below are examples, with brief explanations:

News, Weather, Traffic, Sports Clip Service

This scenario involves the delivery of short video clips to persistent storage in a receiving device for later viewing. These clips typically would involve national and/or local news, weather, traffic or sports information. To aid in user navigation, receivers may provide a menu of available data types (Stock prices, weather etc) or some sort of guide to select available clips. Criteria based filtering may be used to determine which clips to download and store. Services of this type may utilize different business models, including subscription based, advertising based or free. This type of service can also be viewed as the foundation for personalized television channels available on demand.

Two delivery scenarios are envisioned: 1) A modified "Push" scenario where the content is automatically loaded on the receiver, subject to user-defined filtering mechanisms or filtering determined by the receiver based on observed user preferences. 2) A "Pull" scenario where the user specifically requests content through some means.

Telescoping Ads

This scenario involves the delivery of content to a receiver to augment broadcast advertisements. Content in the form of video segments and/or "web pages" are pushed to the receiver as files, allowing the viewer to drill down to more detail for an ad he/she finds of interest. If the application is so designed and a return channel is present additional information pertinent to the individual user could be pulled. The primary delivery scenarios envisioned is a "Push" scenario where the content is automatically loaded on the receiver, subject to user-defined or other filtering mechanisms, which could include observed user preferences.

Long Form Entertainment Programming Downloading

This scenario involves the downloading of entertainment content to receivers for later viewing (for example, TV Episodes and movies – typically 20 minutes to 2 hours in length, commercials, advertisements, music, home shopping, 3D Video Services, etc.). Possible business models for these services include advertising supported, subscription based, and pay per view based. Advertising supported and subscription based services can include: movies, ethnic content channels, Network TV episodes, local sports, music and home shopping. Free local DTV broadcast channels can be blended together with the private long form entertainment to provide more value.

This service could be combined with the recording of DTV station channels to provide a digital video recorder capability for free to air broadcasts. Broadcast PSIP information can be used to schedule free to air channel recordings. Availability of storage on the target devices allows creation of schedules or playlists for playback of downloaded content enabling creation of "virtual live" channels. This playback of scheduled content from the target device storage moves the playout server to the target device and can reduce overall bandwidth requirements.

Targeted Advertising

This scenario involves services that will display pre-distributed targeted advertising, possibly self-selected (for example, a consumer selects the type of commercial, which then enables viewing of cached commercial content) or demographics driven (similar concept to DCC, but would not use switching, just access to the applicable cached content). Commercials that are more relevant to the viewer (targeted) are of more value to the advertiser - feedback on advertising usage/viewing could be an important added value component.

Download games from broadcast

This scenario describes a service where executable objects may be downloaded from a broadcast. The objects may be games, applications or other software. The business side may involve subscription models, pay-per models or even free downloads (for example, an advertising supported service). The objects may be used on the receiving device or may be intended to be moved to another device within the home network. Objects to be downloaded may be selected using preference filtering on the receiving device, selected from a pre-announced schedule or by some other means. Marketing and scheduling the broadcast of games may be different from other NRT programming.

Downloading Music

Music has been shown to be a high value media for both fixed and mobile scenarios. Although music bandwidth is significantly lower than video, there are still advantages to offering music download services over a broadcast channel, rather than point-to-point. Music can be broadcast to the receiver, using a subscription or advertising based model. Content that matches the user's interest may be saved in the local storage of the receiver. In this way, the user has a personalized library of content.

Downloading Web Content

In this scenario, high demand web content is pushed to the receiving device to allow local browsing. The content could be ad supported or could require a subscription for viewing. Examples of content that might be pushed to the device include content assets such as local news and local weather. For some of these sites, in addition to the main page, several of the sub-pages might be cached. For example, all the category pages might be cached for a news site.

It is useful in this scenario for the non-real time service to exist not in isolation, but as a cache for very high traffic. This would requires the terminal device to have the ability to initiate a point-to-point connection back to the service. If a point-to-point connection is not available, the user gets an extremely narrow (but still useful) view of the Internet. If there was a point-to-point return connection, the broadcast non-real time content serves the same function as an auto-updating cache of certain high traffic sites.

Telescoping Content

In this scenario, Web-based data or multimedia content is pushed to the receiving device to allow the user to view more detailed information about another piece of content. Content can be contextually relevant and/or based on user-defined profiles. The additional content will be pushed one-way as pre-downloaded files. Telescoping content can include wikis, show/broadcaster sites, movie/studio sites, additional clips or audio files complementing the main information, etc.

NRT BASICS

A set of requirements has been created that provides the functionality necessary to support the types of expected scenarios, as listed above. From this, a listing of base technologies has been created. It should be observed that a significant portion of these technologies have already been standardized by various recognized standards bodies.

NRT Technologies

- Content Delivery: The delivery of push content and metadata over the DTV broadcast pipe. The functionality needed includes how to encapsulate the content, transport it through an MPEG-2 Transport Stream, how to add FEC (Forward Error Correction) and how to establish and communicate a file structure for the content.
- Signaling: Metadata that tells the receiver information about NRT content that is currently being broadcast – how to find it in the transport and which pieces go together to form a coherent service.
- Announcement: Metadata that tells the viewer about NRT content that will be available in the future. Announcement information provides the basis for an Electronic Service Guide for NRT and includes such things as Name and Description of content, Ratings, Genre, Types of content and Content Duration.
- Content Referencing: Metadata that allows content to be accessed. Content referencing includes such things as establishing the file directory structure, file grouping for services, unique identifiers for files and folders and how to access files that may be remotely located (such as over the Internet, if a return channel is present).
- Content Types: What types of content can be used in NRT services (Video codecs, Audio codecs and graphics formats). In the absence of a selection of content types, a means of communicating the types used in an NRT service will be necessary.
- Content Access Restrictions: Conditional Access (CA) and Digital Rights Management (DRM). CA limits access to the broadcast content to those with the rights to receive it. CA is typically done via encryption of the content on broadcast and decryption via keys on receipt. DRM refers to protection and access control of the content once received.
- Receiver Targeting: Information used to associate broadcast NRT content with receiver types (filtering). This may include the receiver type (Fixed, Mobile, PC), user preferences, geographical location, age-group and so on.
- Ad Selection and Insertion: NRT can enable ad insertion localized on the receiving device – the same capabilities can be used to splice pushed content into live broadcast or create new "channels" of pushed content. Considerations of content/stream conditioning, specification of splice points, user metrics and synchronized graphics overlays are needed.
- Return Channel: While the presence of a return channel is not required for NRT services, there are a number of interesting possibilities if one is present (possibilities include full time or sporadic return channels).

ADDING NRT CAPABILITIES TO BROADCAST STATION ARCHITECTURE

Figure 1, below, shows a simplified view of the basic elements of a digital broadcast station. Uncompressed A/V television content enters from the left of the diagram and is compressed. The multiplexer mixes the streams of encoded A/V and PSIP into a single emission transport stream. This transport stream is then modulated and broadcast.

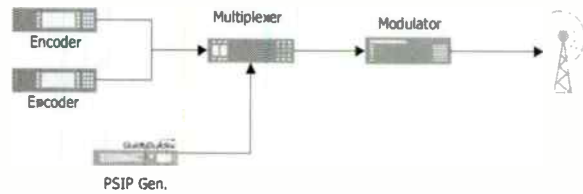


Figure 1 Typical DTV Broadcast Station

The modifications necessary to support NRT services are illustrated in Figure 2, below.

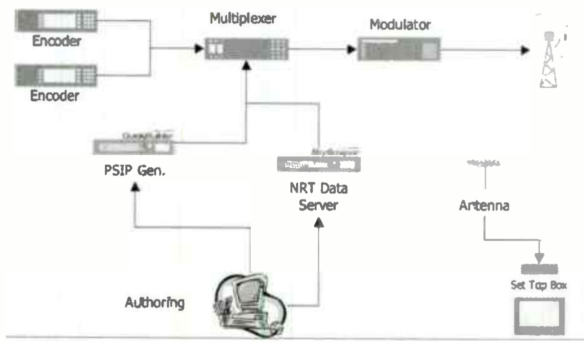


Figure 2 DTV station with NRT capability

The new functions that need to be added to a DTV station to support NRT are as follows:

- Authoring: There must be a means to create NRT content for broadcast. This function will be separate from the linear content flow through the station. Some form of authoring system will be needed to create the content that adheres to the NRT standards (along with the necessary metadata). Often, the authoring system will be located at the source of content creation, rather than at the broadcast station.

- NRT Data Server: Once created, NRT content would be moved to an NRT Data Server. This server would hold the content until the appropriate broadcast time and then encapsulate it for injection into the transport (if the content is not pre-encapsulated). Insertion into the broadcast transport would be done via the existing multiplexer.
- PSIP generator: Signaling and announcement metadata needs to be inserted into the broadcast stream for the NRT content. While this could be done through a separate server, it makes sense to add this function to the existing PSIP generator (the types of information needed are analogous to the PSIP information already being managed and inserted).
- Multiplexer: The multiplexer inserts the encapsulated NRT content into the broadcast transport stream. As long as a free port exists on the mux, no modifications in it's operation would be anticipated.

In addition, some new functions are needed on the receiver as illustrated below.

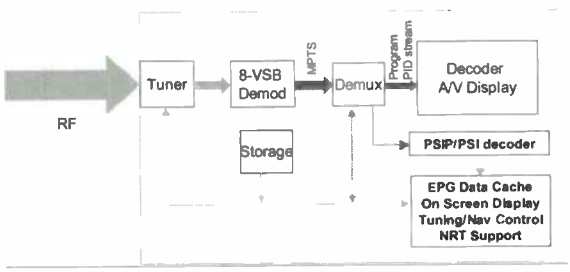


Figure 3 Receiver with NRT support

The additions to the receiver include:

- Persistent Storage provides a place to hold NRT content until it is needed. With the spread of DVRs (Digital Video Recorder) in the marketplace, large hard disks are becoming commonplace in receivers.
- Middleware support: The middleware on the receiving device must be able to deal with the NRT content being distributed according to the standard that is being developed.
- Enhanced Navigation: Today's receivers typically use a grid-style EPG for linear content (normal television fare), which may not be suited for NRT services. Some modifications will likely be necessary to support NRT.

As can be seen, while there are modifications necessary to both the broadcast plant and receivers to support NRT services, they are not drastic.

STANDARDIZATION STATUS

Work is progressing in ATSC (S13-1) to create standards supporting NRT services end-end. Participation includes broadcasters, equipment manufacturers and others. With a desire to not "re-invent the wheel", we are drawing upon technologies disclosed from existing implementations and existing standards as much as possible. The target goal (which currently appears to be achievable) for completing an NRT standard is the analog turn-off date in Feb, 2009 – with a desire to allow broadcasters to announce new and interesting services at that time.

SUMMARY

The time is ripe for deployment of NRT services. Viewers now understand the concept and desire "On-Demand" capabilities for entertainment content. Broadcasters want new revenue producing services available at analog turn-off. Business models are evolving that are likely to close. The addition of NRT capabilities to broadcast plants and receivers is practical. A standard for providing NRT services is well underway and expected to be available in time.

It is now time to start thinking about new services that will capitalize on these new capabilities.

BXF - THE PROMISE OF REDUCED COSTS AND INCREASED REVENUES

Chris Lennon
Harris Corporation
Denver, CO

ABSTRACT

Have I got a deal for you! This new technology you're going to put into your facility is not only going to save you money, it's going to pay for itself with the new revenue it brings in!

How many times have you heard that? How often has it actually turned out to be true? I'm guessing that your answers are predictable.

The adage "if it sounds too good to be true, it probably is" tends to be reinforced on an almost daily basis.

So, why are you still reading this? Was PT Barnum correct?

You're probably still reading because you're hopeful – optimistic, even. There are rare cases in which the reality meets or exceeds the promise, and it is those which make it worthwhile to sort through the noise of unfulfilled promises.

The first thing that will come to most peoples' minds when they hear BXF is that it sounds like technology. It is, but in this case, the technology is not the exciting part…it's what that technology can do to improve your workflow. Of course, you aren't in the habit of changing workflow for its own sake – you're looking for benefits. Are cost savings and promise of increased revenues realistic? Let's dig a little deeper, and then let you be the judge.

WHAT IS BXF?

The acronym BXF is short for Broadcast eXchange Format. It is an XML schema, which is intended to allow broadcast systems to more easily exchange three primary types of data.

1. Schedule-related data
2. Content-related data
3. Content movement instructions

It was originally created to solve the mess that existed between Traffic and Automation systems. Literally hundreds of proprietary, batch-oriented interfaces littered the landscape, making interoperability between systems a serious challenge.

After the formation of the group within SMPTE that was tasked with developing BXF (the standard is officially known as SMPTE-2021), it became apparent that it made sense to get Content Distribution representatives involved. At a high level, the messaging needed to support the movement of content metadata as well as content movement instructions seemed pretty similar, whether dealing with content within a facility, or content needing to be obtained from outside sources.

So, the basic membership of the group developed into representatives of the following interests (primarily):

- Traffic
- Automation
- Program Management / Listing Services
- Content Distribution
- Users

WHAT SPECIFIC PROBLEMS DOES BXF FIX?

The short answer is "many". However, there are some very specific areas that were targeted during the development of BXF as needing to be solved.

A Single Standard

Much of the problem that has plagued the industry for decades has been due to the fact that each manufacturer of each broadcast system found themselves having to reinvent interfaces each time they worked with another vendor. There was no standard, so anarchy reigned. While "custom" is nice for those who can afford it (who doesn't like custom cars, or the endless custom hot drinks at Starbucks?), it is often prohibitively expensive. In an increasingly lean market like Broadcasting, the cost of "custom" often cannot be justified.

So, simply creating a single standard that vendors can write to is a quantum leap.

Dynamic versus Batch

One of the real limitations of these traditional proprietary interfaces was that they were, without exception, batch in nature. What this means is that they were intended for a workflow in which System A does everything it needs, then hands the data "over the transom" to the System B. At this point, System A washes its hands of any responsibility and moves on.

While that's nice and simple if you're the manufacturer either of those systems, it doesn't reflect the reality of the workflow in a broadcast facility. Workflows are rarely purely linear in nature…often steps are performed iteratively. Take the construction of a broadcast schedule, for instance. The Traffic System at some point deems the schedule ready to publish to the Automation System. Once this is done, that's rarely the end of the story. Changes to the schedule are inevitable, and sometimes quite voluminous. With batch-oriented interfaces, there is no simple mechanism to enable messaging of those changes between systems. As a result, the changes must be accommodated outside of the interface, using manual steps.

Building dynamic integration between systems, in effect removing the "transom" referenced earlier, allows those systems to mimic the actual workflow that exists. This means that broadcasters have systems that fit their natural workflow, rather than having to fit their workflow around those of the systems (a recipe for disaster, or at least manual workarounds).

Support Files as well as Messaging

It would have been nice if we could have said that we were abandoning file-based exchanges altogether and going exclusively to a message-based architecture. However, as with most things, technological changes that are most effective are often accomplished by way of many small steps, rather than a single leap.

It was unrealistic to expect the wide variety and vintage of systems being used today to completely change their approach simply to accommodate BXF. Therefore, it was decided to structure BXF in such a way that it could be used effectively in traditional file exchanges, and in message exchanges.

In practice, the workflow that most broadcasters seem to prefer involves a combination of the two. Exchanges of data files make sense when a large volume of data must be sent at a specific time, for instance, when a log is deemed complete and ready for publication by a Traffic system. However, messages make far more sense than files when smaller volumes of data are to be exchanged on a more ad hoc basis.

It is possible that, as systems evolve, file-based BXF exchanges will become a thing of the past, and messages will be used for all data flow between systems. However, we are not currently at that point, nor will we be for a while yet.

Human Friendliness

The traditional interfaces which BXF replaces were always computer-friendly. However, human friendliness was not one of their attributes. The focus when many of these protocols were invented was on efficient machine-to-machine communications. Bandwidth and computing power was limited, meaning efficiency won out over readability. This is how we ended up with values encoded in binary coded decimal, hex, or even pure binary. Even pure text values were expressed in arcane shorthand (doesn't everyone know that a spot type of "R" is a promo? Sorry – "P" was already taken for PSAs).

Some questioned the importance of human readability when we began the BXF effort. However, the voice of the user community was loud and clear on this. They were tired of trying to open files in an effort to resolve problems only to be faced with a collection of data harder to decipher than a Doctor's handwriting!

The results were twofold. First, BXF was created using XML (eXtensible Markup Language), which imposed structure on the data, and allowed for helpful annotation within the file itself. Second, great pains were taken to ensure that data values included in BXF were verbose. If the data value is meant to indicate that the spot is a "promo", we use the value "promo". This way, even a relatively non-technical person can open a BXF file and have a chance of reading it, and perhaps even find the source of the problem that caused them to open the file in the first place.

Extensibility

As outlined in the previous section, XML was readily agreed to as the format that made sense for BXF. By its nature, XML is extensible.

Although great time and effort was spent on including everything that was felt necessary for effective communication amongst systems, the group acknowledged that there would be things that would inevitably be missed, and unanticipated industry requirements would arise after publication of the standard. For this reason, it was critical that extensibility be built into BXF. Failure to do so would result in a largely ineffective standard.

Extensibility is accommodated in two ways within BXF. The first and most straightforward way is via the nature of XML itself. Attributes and elements can be added to future versions of BXF without causing incompatibility. Applications that don't understand new elements and attributes simply ignore them.

The second way in which extensibility is achieved within BXF is targeted at more specialized, short-term requirements for extensibility. "PrivateInformation" elements and "##any" attributes can be found at virtually every level of the schema. These allow vendors to add in site or system specific elements and attributes to the base schema. This way, if there are things that BXF does not handle, but which are needed, vendors can still utilize the base BXF schema, and simply add in the elements and/or attributes they need. If these private extensions develop into general requirements, they can then be added formally into a future version of BXF.

HOW DOES THIS SAVE ME MONEY?

Everything covered thus far is well and good, but by this point, you are likely asking yourself when we'll get to the "good stuff"? Broadcasting is a business, and technology is not implemented for its own sake. There must be a compelling business case to justify the introduction of new technology into a facility.

Everyone's looking for ways to reduce expenses, while streamlining operations. BXF has much to offer in this area.

No more Post-It™ Notes

Those not involved in Traffic and Master Control operations are often shocked to find out how day of air changes to the broadcast schedule are accomplished today. In most cases, a Traffic or Operations person will hand write those changes onto a paper log, or simply slap a Post-It™ note onto the (hopefully) appropriate location on the log. The Master Control Operator will then have to take this information and manually enter it into the Automation System. This means that data was entered three times – once into the Traffic System, once onto the Post-It™ note or log, and once into the Automation System. Not only does this mean that three times the effort was expended, but there is three times the chance for the introduction of errors.

When you're dealing with spots valued at thousands or in some cases millions of dollars per thirty seconds, anything that can eliminate duplicate (or in this case, triplicate) data entry can have a serious impact on your bottom line.

Dealing with missing content

Just because content (be it commercial or program content) is supposed to be in-house well in advance of broadcast doesn't mean it is. Often, when the playlist is loaded into the Automation System, warnings pop up that some content which must be played in the coming hours cannot be found.

Prior to BXF, missing content often meant manual procedures. Phone calls, e-mails, and frantic last-minute scrambling was often needed to get that missing content onto a playout server. Sometimes, this worked, and the content arrived and was played. Sometimes, all this manual effort was for naught, and commercials were missed (resulting in lost revenue).

BXF's ability to allow systems to send content metadata, as well as content movement instructions (i.e. get that Toyota spot from DG onto our playout server) between systems enables the automation of a station's missing content procedures, reducing manual effort, and increasing the likelihood that the content will be found and made available for playout on time.

Simple Reconciliation

A largely unintended consequence of the BXF effort was something that was largely beneficial to the Accounting/Finance departments at broadcast operations.

The process of reconciling the as run log with the broadcast schedule has long been a painful and time-consuming one for broadcasters everywhere. This was largely due to the period of time during which the Traffic/Billing system and the Automation system didn't communicate. When changes in one are made, and are not made in the other, it is the job of Accounting/Finance to sort out the resulting mess, and "reconcile" the two.

If the Traffic and Automation views actually matched, reconciliation would become a breeze. This is where BXF helps. That period of disconnect between the two systems, when their views of the schedule diverge, is eliminated. Changes to the schedule can be made in Traffic, then sent down to Automation. Automation airs the events, which can then be automatically matched up with the original events scheduled by Traffic.

Suddenly, reconciliation becomes much easier. The only events that really need to be reconciled are those which were added or deleted by Automation without involvement or notification of Traffic. This may be limited to events that went wrong in the middle of the night, or on the weekend.

In truth, even those cases can be handled by BXF, which allows for event updates to be sent back to Traffic by Automation when edits must be made in the Automation System. Another approach is to require that all playlist edits be made in the Traffic System. While both of these are technically possible using BXF, it may be some time before we see either used commonly in the field.

More Hands-Off Master Control

Many Broadcasters ask "does BXF mean the end of the Master Control Operator"? The answer is not clear cut.

BXF certainly means reduced manual effort required on the part of the Master Control Operator. If Traffic was made a 24x7 operation, the job of maintaining the Automation playlist could, theoretically, be moved out of Master Control and into Traffic. However, the typical Operator does more than simply monitor and maintain the playlist.

There is also the reality that in most cases, Traffic will continue to be an 8-12 hour a day operation for the foreseeable future. In those cases, Traffic can assume responsibility for maintenance of the Automation playlist for up to ½ the day, but arrangements need to be made for the other ½.

So, could BXF eliminate the need for a Master Control operator maintaining the Automation playlist? Yes. Will this be an immediate impact of BXF? In many cases, it will likely be an evolutionary thing.

HOW DO I MAKE MORE MONEY USING BXF?

Cost savings are always popular, and often seem to be the focus when trying to justify new technologies. However, of equal importance can be ways in which a new technology can allow you to increase revenue.

Fewer missed spots

Missed spots have a direct and significant impact on the revenue of a station. A certain percentage of these missed spots are missed due to communications issues, either between Traffic and Automation, or between the station and the entity distributing the spots.

With its tighter integration of the systems involved, BXF can help to reduce the number of spots missed due to these communications issues.

Make Good Spots Quickly

When spots are missed, one way to recapture that revenue is to reschedule them as quickly as possible. Getting a spot rescheduled on the same day, or even in the same program/time slot in which it was originally scheduled to air, can mean no preemption and no costly makegood. There is also the case which today, often results in direct revenue loss, when a spot is missed on the last day of its run, resulting in a credit to the advertiser. This happens so often because Traffic/Billing is not notified of missed spots until the next day. With the dynamic messaging capabilities of BXF, Traffic/Billing can be notified immediately of a missed spot, and can reschedule it (again, using BXF) later that same day.

Support for Sponsored Secondary Events

Broadcasters have been seeking revenue outside of the 30-second spots that have traditionally constituted their inventory for some time now. Such non-traditional revenue in the traditional broadcast stream can be realized by allowing advertisers to purchase time during programs, or by moving time (and revenue) consuming promos out of coveted 30-second avails and over top of programming, using secondary events.

Traditional Traffic/Automation interfaces have not allowed for "sponsored secondary events", and workarounds have had to be created to enable this new revenue source. BXF was designed with this in mind, with the ability to associate as much advertiser-related metadata to a non-primary event as to a traditional 30-second spot.

This allows Broadcasters to fully capitalize on the potential for sponsored secondary events.

Painless Last Minute Selling

Associating the word "painless" and "sales" in the same sentence is something many Broadcasters will likely cringe at. However, BXF can make this process much simpler, which today is quite painful, and as a result tightly controlled or strongly discouraged.

BXF's dynamic nature means that formerly intensive manual effort is largely eliminated, making last-minute changes to the broadcast schedule fairly simple. Because the cost involved with making such changes is minimized, it is more likely that such changes will be permitted.

Allowing more last-minute selling allows you to capitalize on real-time developments and charge premium prices to allow advertisers to also capitalize.

CONCLUSIONS

It is important when you think of technologies such as BXF, that you focus not so much on their technical aspects, but on what they can do for you in three terms:

1. How does this simplify (and complement) my existing workflows?

2. Are there quantifiable cost savings associated with this?

3. Does this technology enable new revenues?

As this paper demonstrates, BXF was built to provide "yes" answers to all three of these questions. How you realize these goals is up to you and your chosen vendors. Being a new technology, not all of its potential will be realized immediately. However, it should provide comfort and ease decisions regarding whether or not to implement BXF in your facility, knowing that it was built from the ground up with these three goals in mind.

SEAM CARVING FOR VIDEO

MIKE KNEE
Snell & Wilcox
Havant, UK

ABSTRACT

Seam carving has caused quite a stir in the image processing world. It is a technique for content-aware image resizing in which the size and shape of regions of visual importance are preserved without resorting to cropping. With user interaction, unwanted objects can also be removed from an image with minimal effect on the remainder of the picture. The results on still pictures have been spectacular, but extending the idea to moving video presents significant challenges.

This paper describes how seam carving can successfully be modified and extended to work with moving video. The resulting algorithm is robust, computationally efficient and gives pleasing results. It can be combined seamlessly with dynamic reframing and conventional resizing to provide a rich toolkit for content-aware repurposing of video. Applications include post-production, conversion between HD and SD TV standards, aspect ratio conversion, repurposing for mobile devices and internet video.

INTRODUCTION

This paper presents a seam carving algorithm for content-aware resizing of moving video sequences. The main application considered here is aspect ratio conversion, for example from 16:9 to 4:3, but the techniques presented can readily be applied to other video sequence resizing and re-purposing tasks.

Conventional approaches to aspect ratio conversion include cropping (removal of information at the sides or top and bottom of the picture), letterboxing (adding black bars) or stretching or squeezing the picture to the required shape. These approaches all have drawbacks. Cropping removes information that might be important, letterboxing wastes precious screen space and loses resolution, and stretching or squeezing changes the shape of objects. The drawback of cropping can be overcome to some extent by a pan-scan process in which the preserved region of the picture is moved around smoothly to follow the most important content. This can be done manually or by an automated process which tracks regions of interest in the scene [1]. However, pan-scan can fail when there are objects of interest near both ends of the picture, for example in a motion picture dialog scene.

Seam carving is an exciting and novel technique for content-aware resizing of pictures. The technique is described by Avidan and Shamir [2]. The idea is to resize or reshape a picture by removing less important pixels. The integrity of the information that is retained is preserved by insuring that the pixels that are removed form connected "seams" which "carve" through the picture from top to bottom or from left to right. Spectacular results have been presented in which pictures can be gracefully and progressively resized by the repeated removal of seams. The technique can also be used for the expansion of pictures and for the removal of specific objects. When used for aspect ratio conversion, the disadvantages of conventional methods can be overcome. The full extent of the original picture can be preserved without changing the shape of the objects in the scene and without wasting screen area and losing resolution.

Seam carving would therefore be a highly desirable addition to the toolkit of techniques for resizing or aspect ratio conversion of moving video. However, seam carving as originally presented was developed for still pictures. It produces unacceptable artifacts when applied to moving sequences, because the seam carving decisions are made independently for each frame of the sequence.

This paper describes how seam carving can be adapted for moving sequences, using two new techniques. We refer to these techniques as **recursive energy weighting** and **map processing.** Recursive energy weighting addresses the fundamental problem of coping with moving sequences. Map processing improves the quality and flexibility of the seam carving process and brings two additional benefits. It can dramatically increase processing speed, which is particularly important for seam carving of HDTV sequences, and it can provide elegant solutions to the problem of combining horizontal and vertical seam carving for still pictures as well as for sequences.

We begin with a brief description of the existing seam carving algorithm and an explanation of its failure on moving video. We then present the new techniques, giving examples of their performance. Finally, we discuss remaining problems and opportunities for further work.

THE ORIGINAL SEAM CARVING ALGORITHM

This section gives a brief, informal description of the seam carving algorithm for still pictures. A much more detailed description is given in [2].

Suppose we wish to shrink a picture horizontally. Seam carving is applied repeatedly, shrinking the picture by one pixel width at a time. Each pass of the algorithm operates as follows. We calculate an **energy** or activity function for each pixel in the picture. Typically, this is the sum of absolute differences between the current pixel's luminance value and each of its four neighbors. We then find a **seam** of minimum energy extending from the top to the bottom of the picture. A seam is a set of connected pixels, one pixel per line, the connection criterion typically being vertical or diagonal adjacency. Figure 1 gives an example of a seam on a very small picture.

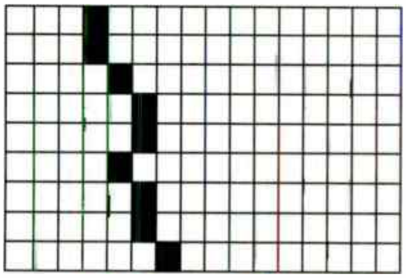


Figure 1 A seam

The energy of a seam is the sum of the energy values of the pixels in the seam. The minimum-energy seam can be found using a recursive technique in which we calculate best partial seams leading to each pixel on successive rows of the picture until we have a minimum-energy seam leading to each pixel on the bottom row. We then take the minimum of all the bottom-row results and back-track along the seam to the top of the picture.

Having found the minimum-energy seam, we simply remove all its pixels from the picture, shifting the rest of the picture into the gap to make a new picture one pixel narrower than before. This is the process of "carving" a seam from the picture.

A spectacular example of the power of seam carving is shown in Figure 2, in which a picture is shrunk horizontally by a factor of two, while largely preserving nearly all the important objects in the scene.

Seam carving may be performed horizontally, removing top-to-bottom seams to reduce the width of the picture, or vertically, removing left-to-right seams to reduce the height of the picture. If the picture is being shrunk in both directions, individual horizontal and vertical seam carving operations may be carried out according to a pattern or rule, or in a picture-dependent order based on minimizing the energy cost of the seam removal operations, as described in [2].

Figure 2 EBU test slide "Boy with Toys" before and after seam carving

Seam carving may also be used to expand pictures. In this case, seam carving is first used to reduce the size of a picture by the same number of pixels as the desired increase. Then, for every pixel that is removed in the seam carving process, a new pixel is instead added to the original picture by interpolation. Thus, low-energy areas that would have been removed are in fact doubled in size.

SEAM CARVING ON MOVING SEQUENCES

The seam carving algorithm described above does not work well on moving sequences. Why not? Consider a sequence with a detailed, moving object that is being tracked over a relatively blurred background by a camera. In any particular picture in the sequence, seam carving (to reduce the picture width, say) will work well. Seams will pass to the right and left of the object (because the object has more energy than the background), the object will be preserved and the lowest-detail parts of the background will be removed. Suppose that in one picture, 40 seams are removed to the left of the object and 50 to the right. In the next picture in the sequence, the background will have changed and the seam carving process will make different decisions. Even with quite a small change to the background, the process might well find 42 seams to the left and 48 to the right. In the next-but-one picture, we might have 39 seams to the left and 51 to the right. This will have the effect of moving the object several pixels from one picture to the next, producing severe motion judder. This effect is what happens in practice. The resulting sequence is completely unacceptable, even if only a few seams are being removed from a sequence with slow, uniform motion.

RECURSIVE ENERGY WEIGHTING

The solution we propose for seam carving of moving sequences is to introduce **recursive energy weighting.** We record the position of each seam in a picture. In the following picture, we apply a weighting or bias to the energy function to reduce the energy in the vicinity of the corresponding seam in the previous picture. In our experiments, we have used a simple additive bias function as shown in Figure 3. In this example, a value of 20 (based on an 8-bit gray scale) is subtracted from the energy of all pixels in the seam and progressively lower values are subtracted from pixels that are less than 5 pixels away from the seam. In the frame immediately following a scene change, no weighting or bias is applied.

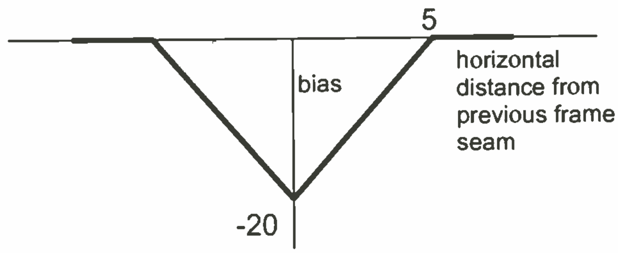


Figure 3 Energy bias function

Figure 4 shows the benefit of recursive energy weighting on a slowly moving sequence (shown inset with seams marked in red). The graph plots the relative horizontal position of the central mast of the ship for the source sequence and for the sequence reduced from 16:9 to 4:3 aspect ratio by seam carving without (marked "Spatial") and with recursive energy weighting.

Motion compensation

Recursive energy weighting solves the problem of seam carving on moving sequences very effectively if the motion is slow, as in the above example. However, problems remain with faster motion because the energy weighting fails to track moving objects. We therefore introduce **motion compensation** into the recursive energy weighting calculations. Motion vectors from one frame to the next are measured for each pixel in the scene. Any reliable method of motion estimation may be used. In our experiments, we used a dense motion estimation algorithm based on phase correlation [3]. Each seam is then projected into the next picture according to the motion vector. The resulting set of pixels (which might not now be a valid seam in itself because of variations in the motion vectors) forms the basis for the calculation of the recursive energy bias function.

Figure 5 illustrates the benefits of motion compensation on a critical sequence. The first two images show the seams 7 and 27 frames from the start of the sequence when recursive energy weighting is used without motion compensation, and the next two images show the same frames with motion compensation. Without motion compensation, the seams to the right of the faces fail to follow the faces as they move to the right, so the eyes and mouth begin to be crossed by seams. With motion compensation, the preserved areas follow the faces better.

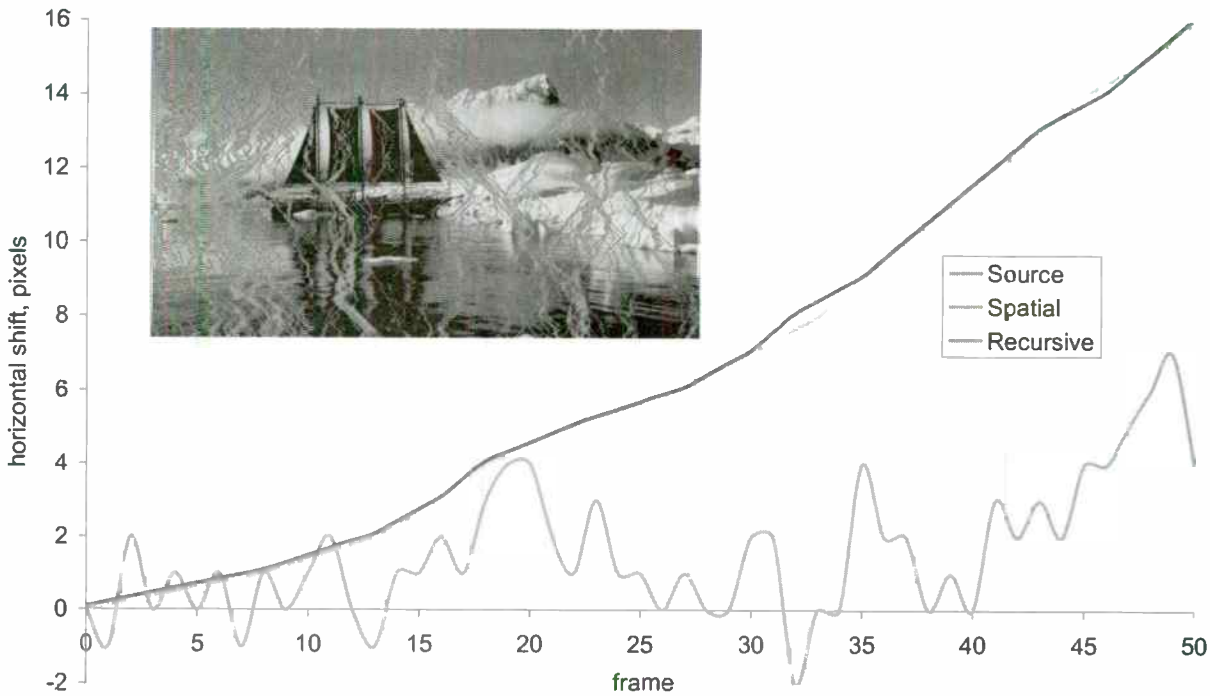


Figure 4   Benefit of recursive energy weighting

Figure 5   Benefit of motion compensation

Motion compensated recursive energy weighting allows seam carving to be performed effectively on a wide range of moving sequences. However, as for still pictures, seam carving cannot be applied to all moving picture material with equally good results. It is very difficult to prevent objects from changing shape as the algorithm progresses through a sequence. Another problem comes from sudden small shifts in regions of the picture from one frame to the next, which inevitably occur from time to time even if motion compensated recursive weighting is being used. These problems can be effectively tackled using the second technique presented in this paper, **map processing**.

MAP PROCESSING

Map processing is a powerful way to perform and to enhance the seam carving process, both for still pictures and for moving sequences. In order to understand this technique, we need to take a philosophical step back from the seam carving algorithm as presented so far. Seam carving performs two functions. It **analyzes** the picture (by finding seams) and it **processes** the picture (by removing seams). These two functions are intimately related in the original seam carving algorithm, because a seam is removed before the next one is found. In map processing, we seek to separate the two functions into an analysis phase and a processing phase. We shall see that this approach has many benefits.

The use of seam carving to expand a picture has already given us a hint of this approach. The analysis phase consists of applying seam carving to reduce the size of a picture. Then the processing phase consists of observing where seams were removed, and interpolating new pixels into the original picture along the seams. The analysis phase can be made to produce a function or lookup table linking output pixel addresses to input pixel addresses. There are two ways of describing this function. One way is to consider it as a **projection** from the input picture to the output picture: for every input pixel, where does it go to in the output picture? The other way is to consider it as a **map** for the output picture: for every output pixel, where does it come from in the input picture? Both ways have their advantages and disadvantages, and in any case it eventually becomes necessary to transform between them. In this paper we are using the map-based approach, so our function's domain is the output space and its range is the input space.

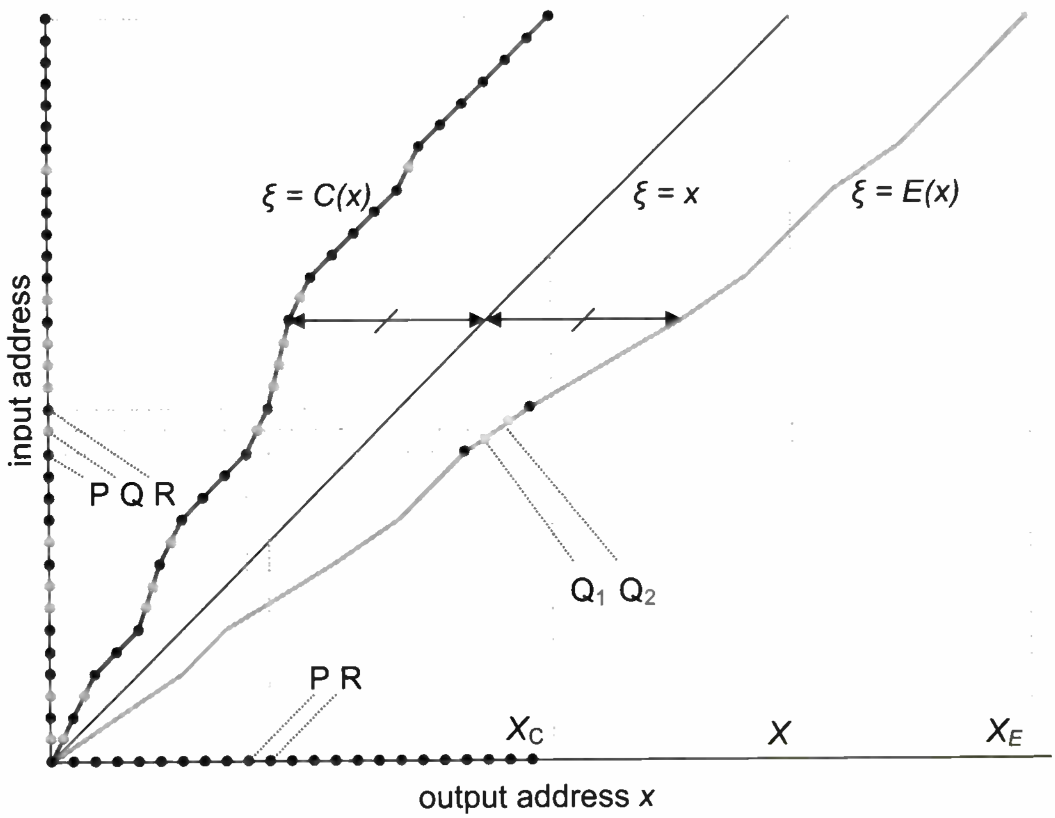


Figure 6 Seam carving contraction and expansion maps

For example, the curve $\xi = C(x)$ shown in Figure 6 represents a horizontal seam carving map for one line of a picture. The x axis represents output pixel addresses and the ξ axis represents input pixel addresses. The points on the axes represent individual pixels. The curve is constructed by incrementing x for each step in ξ, except where ξ represents a pixel in a seam. For example, pixels P, Q and R in input space are mapped to pixels P and R in output space, because pixel Q belongs to a seam. The curve $\xi = C(x)$ gives us, for each pixel position in the output picture, the address in the input picture where the pixel value is to be found.

The curve $\xi = C(x)$ can be transformed to a second curve $\xi = E(x)$ to generate an expanded picture. For example, pixel Q in the input picture, instead of being removed, generates two new output pixels Q_1 and Q_2. Figure 6 shows that the resulting curve $\xi = E(x)$ can be derived geometrically from $\xi = C(x)$ by horizontal reflection about the line $\xi = x$. The function $y = E(x)$ can then be used to generate the expanded line.

In generating and applying the curve $\xi = E(x)$, we have effectively detached the analysis phase (generating $C(x)$ by seam carving) from the processing phase (deriving $E(x)$ and using it to generate the final output). We shall now see how this approach can be taken much further to increase the efficiency, effectiveness and flexibility of seam carving for still pictures and for moving sequences.

Map transformation

In the example above, the original seam carving map was transformed to create an expansion map. There are many other transformations that we could apply. In the next part of this paper we shall consider three possible transformations: *mixing, scaling* and *smoothing*. We shall then consider how map processing can be used to combine horizontal and vertical seam carving.

Map mixing

For expansion. we applied a geometric transformation to $C(x)$ to produce $E(x)$. Mathematically, we performed a simple operation between the inverse functions:

$$E^{1}(\xi) = 2\,\xi - C^{1}(\xi)$$

This can be generalized to a variable mix between y and $C^{1}(\xi)$:

$$E^{1}(\xi) = (1-\alpha)\xi + \alpha\, C^{1}(\xi)$$

By varying the parameter α, we can choose any degree of expansion or contraction based on the original seam carving map. For example, if the original map was generated for a contraction from 720 pixels to 540 pixels, then $\alpha = -1$ would give an expansion from 720 to 900 pixels, while $\alpha = 0.5$ would give a contraction from 720 to 630 pixels. How would this "indirect" contraction to 630 pixels differ from direct use of seam carving to contract to 630 pixels? The answer is that the indirect approach is a kind of "dilution" of the seam carving process; more seams are identified but each is only half-removed. The indirect approach turns out to be somewhere between direct seam carving and linear rescaling of the picture. We have therefore derived a neat way to control a "mix" between seam carving and linear rescaling. In moving sequences where direct seam carving produces undesirably harsh results, the use of map mixing may give us a suitable compromise.

Map scaling

The use of a map allows us to decouple the size of the picture used for analysis from the size of the picture being processed. For example, an HDTV sequence could be downconverted to a much smaller size for seam carving analysis and the resulting map function upconverted for application to the original sequence. This leads to a significant reduction in processing time, which is proportional to the number of pixels in the picture. In our work, we have downconverted 1920 x 1080 HTDV pictures to 480 x 270 for seam carving analysis, bring a 16-fold savings in processing time. The potential problem of seams being "wider" with respect to the original picture size, and the resulting increased coarseness of the seam carving process, can be solved by map smoothing, which is described next.

Map smoothing

The "coarseness" of seam carving, particularly when the analysis phase is carried out on small versions of the picture, is a particular problem for moving sequences. Sudden shifts of picture information by as little as one seam's width are highly visible and disturbing, and these occur even when recursive energy weighting significantly increases consistency from one frame to the next. The answer is to apply smoothing (low-pass filtering) to the map function before the final result is generated. We have found that the best results are obtained by filtering the map function in all three dimensions – horizontal, vertical and temporal – using simple, separable filters, each with three or five taps.

Using maps to combine horizontal and vertical seam carving

We now present two ways in which maps can be used to combine horizontal and vertical seam carving operations. In the first approach, horizontal and vertical seam carving are performed in series, while in the second approach they are performed in parallel.

For seam carving in series, we may choose in which order to take the two directions of seam carving, or we could interleave the two directions as described in [2]. In this example, we will perform horizontal seam carving first. This will lead to a map function for the whole picture:

$$(\xi_H, \upsilon_H) = \boldsymbol{H}(x, y) = (h(x, y), y)$$

We then perform vertical seam carving on the resulting picture, generating a second map function:

$$(\xi_V, \upsilon_V) = \boldsymbol{V}(x, y) = (x, v(x, y))$$

A combined map can be generated very simply by cascading the two map functions, so

$$\begin{aligned}(\xi_C, \upsilon_C) &= \boldsymbol{C}(x, y) \\ &= (h(x, v(x, y)), v(x, y))\end{aligned}$$

For seam carving in parallel, independent horizontal and vertical seam carving operations give us map functions for the whole picture:

$$(\xi_H, \upsilon_H) = \boldsymbol{H}(x, y) = (h(x, y), y)$$

$$(\xi_V, \upsilon_V) = \boldsymbol{V}(x, y) = (x, v(x, y))$$

This time we combine inverse map functions:

$$\begin{aligned}(x_C, y_C) &= \boldsymbol{C}^{1}(\xi, \upsilon) \\ &= (h^{-1}(\xi, \upsilon), v^{-1}(\xi, \upsilon))\end{aligned}$$

from which we calculate a combined map function:

$$\begin{aligned}(\xi_C, \upsilon_C) &= \boldsymbol{C}(x, y) \\ &= (h^{-1}(\xi, \upsilon), v^{-1}(\xi, \upsilon))^{-1}\end{aligned}$$

The inversion of map functions is in fact the projection operation common in computer graphics. In this case it is pixel addresses that are projected from one space to another, though in the final stage the inversion of the map could be avoided by projecting pixel values from the input picture to the output picture.

In summary, map processing allows us to detach the seam carving operation, which analyzes the picture, from the actual processing of the picture. It allows us to reduce the processing time for seam carving and to optimize the quality of the output picture. It also provides a mechanism for combining horizontal and vertical seam carving and for gracefully combining seam carving with other operations such as linear rescaling and dynamic reframing.

RESULTS

The techniques described above were combined together into a single algorithm and tested on many sequences, including some 6,000 frames of the most interesting and critical material from a 24Hz progressive HDTV version of the motion picture "Mission Antarctique". Our main goal was aspect ratio conversion of this material from 16:9 to 4:3. Several insights were gained from these tests. The original seam carving algorithm, applied picture by picture, produces unwatchable results, whereas recursive energy weighting produces smoothly moving results whose remaining temporal inconsistencies are largely removed by map smoothing. The addition of motion compensation is beneficial, particularly in sensitive regions such as faces, where changes of shape due to erroneous propagation of seams are particularly disturbing.

Map processing was also used to combine horizontal and vertical seam carving, in order to "spread the pain" of seam carving between the two dimensions. Our aspect ratio conversion goal was achieved by combining a mild degree of horizontal contraction with an equally mild degree of vertical expansion. Specifically, working on a 480 x 270 downconverted version of the source, horizontal seam carving was used to generate a map for horizontal contraction to 416 x 270. In parallel, vertical seam carving generated a map for contraction to 480 x 228, which was transformed to a map for expansion to 480 x 312. The two maps were then combined to produce a map for resizing to 416 x 312. This map was linearly upconverted to produce a map for aspect ratio conversion from 1920 x 1080 to 1440 x 1080. For the smaller-picture results shown below, the map was also linearly downconverted to 360 x 270.

The combination of horizontal and vertical seam carving was found to have some benefits and some drawbacks. Limiting the contraction horizontally helped to mitigate some of the fundamental problems that occur when objects traverse the screen. Vertical expansion exploits the fact that areas of water or sky near the top and bottom edges of the picture are ideal for seam carving. However, vertical seam carving does lead to the possibility of spurious vertical motion being introduced, which can be disturbing in landscape scenes with mainly horizontal motion.

Figure 7 shows a source picture, the horizontal and vertical seams and the resulting rescaled picture.

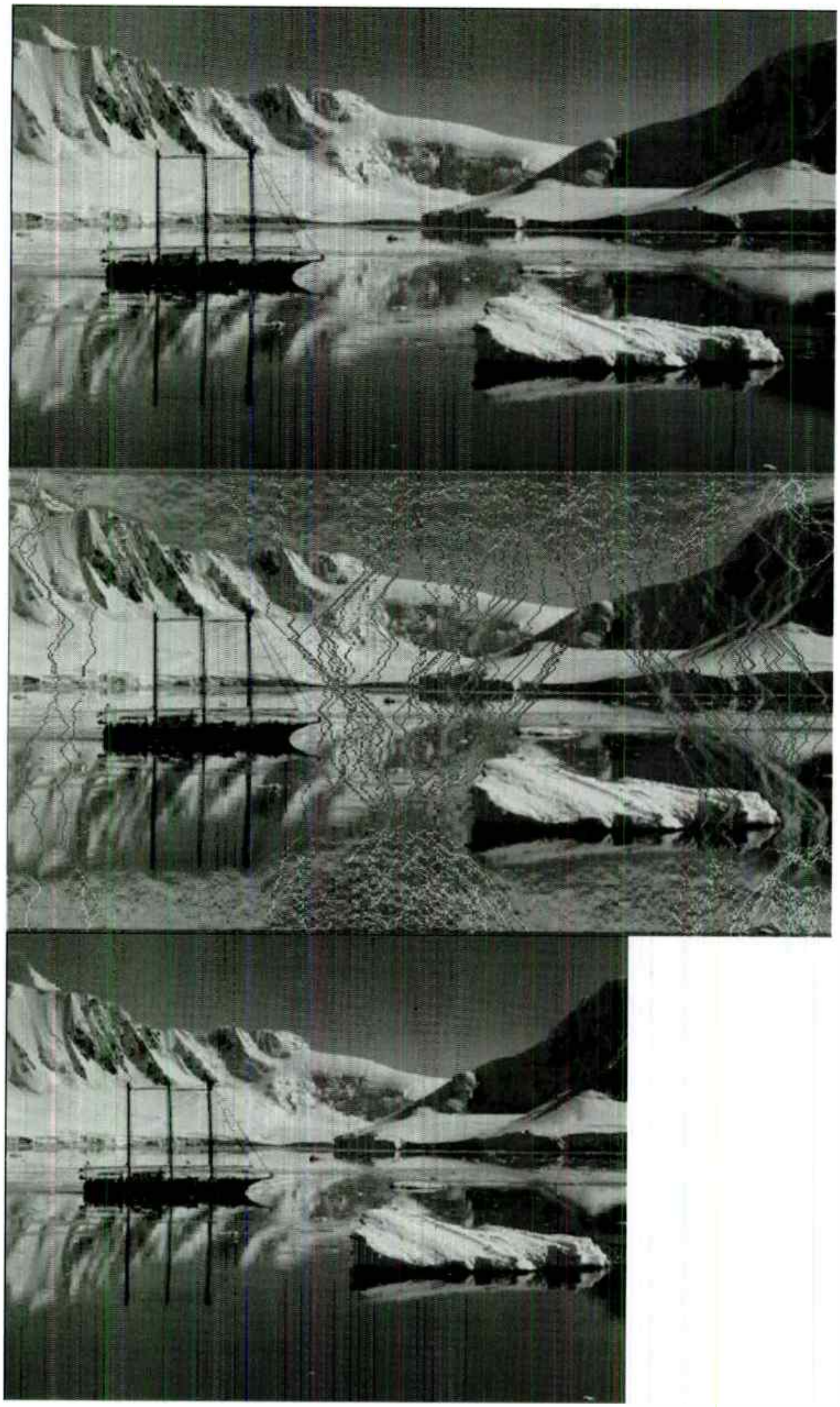


Figure 7   Combination of horizontal and vertical seam carving

DISCUSSION

In general, the algorithm works remarkably well on natural scenes with detailed foreground objects or people. The foreground is preserved, and the inevitable distortions in the shapes of landscape elements are much less apparent than they would be on foreground objects.

The remaining problems with the algorithm fall into two categories. First, there are problems to do with the nature of the source material, which could be solved using known techniques. For example, in many scenes from "Mission Antarctique" the foreground is relatively dark and can sometimes lose detail, leading to low energy values and the possibility of attracting seams and therefore being distorted. Smooth foreground objects, for example penguins and diving suits, can also end up narrower than they should be. In both cases, the problems could be overcome by algorithms for detecting regions of interest based on color and/or motion, for example as described in [1].

The second category of problem is more fundamental and is concerned with what we actually desire when we remove areas from moving scenes. To give one example, what should happen to a steadily panning background while the camera is tracking a foreground object? If the algorithm is working correctly, seams in the background will be tracked until they leave the edge of the picture, but what should happen then? Those seams can either "bunch up" near the edge of the picture, "reappear" at the other end of the picture, or be "set free" for reallocation to another part of the background. The first two outcomes seem to be the most desirable, as they tend to result in cropping, and there is an argument to say that cropping is perfectly acceptable in a background pan because it only serves to bring forward the disappearance of information that is about to disappear anyway, or to delay the appearance of information at the other end of the picture.

The current version of our algorithm leads to "bunching up" of seams and this appears to be perfectly acceptable, as illustrated in Figure 8, which shows pictures from a panning sequence at 60-frame intervals. This example also illustrates some of the overall benefits of seam carving. If cropping were used instead, the two cyclists would never be shown together, and if the picture were squeezed, the wheels in the side-on view would no longer be round. Seam carving shortens the space between the cyclists and removes background that is about to leave the picture.

As a further example, what should happen to background that is revealed or obscured as an object passes across it? Inevitably, the amount of detail in the background will change, and while our recursive techniques help to smooth out the effect of those changes, there will inevitably be the occasional "rupture" when the recursion fails, and the motion of the foreground object will become disturbing. Seam carving allows objects to be preserved by giving their pixels high energy values, but this does not preserve the <u>position</u> of objects. It may be necessary to impose a higher-level constraint on the whole process, whereby the motion of key objects in the resized picture is forced to be related in a smoothly varying manner to their original motion.

FURTHER WORK

Following on from the discussion above, further tests are needed to incorporate algorithms for detecting regions of interest, modifying the energy function accordingly. Beyond that, it would be desirable to concentrate on the more fundamental problems of removing material from moving sequences. Ultimately, the goal may be to move away from the concept of seams altogether and toward an approach where the map function is considered as a two- or even a three-dimensional "membrane" which is distorted smoothly by a measure of importance of the picture information.

Another area of potential further work is to extend the object removal techniques presented in [2] to moving sequences. This would necessitate the reliable tracking of objects, in addition to the techniques described here.

CONCLUSIONS

In this paper, we have presented techniques by which seam carving, previously demonstrated to good effect on still pictures, may be extended to moving sequences. Recursive energy weighting, which may be motion compensated, helps to insure temporal consistency of the process and removes the worst motion-related artifacts. Map processing, by mixing, scaling and smoothing, can be used to reduce processing time, to provide a controlled compromise between content-dependent and global rescaling and between horizontal and vertical processing, and to further reduce unwanted temporal effects. The resulting algorithm fulfills its objectives and has produced some pleasing results, but there is plenty of scope for further work before content-aware resizing could be successfully applied universally and automatically to moving sequences.

Source Seams Result

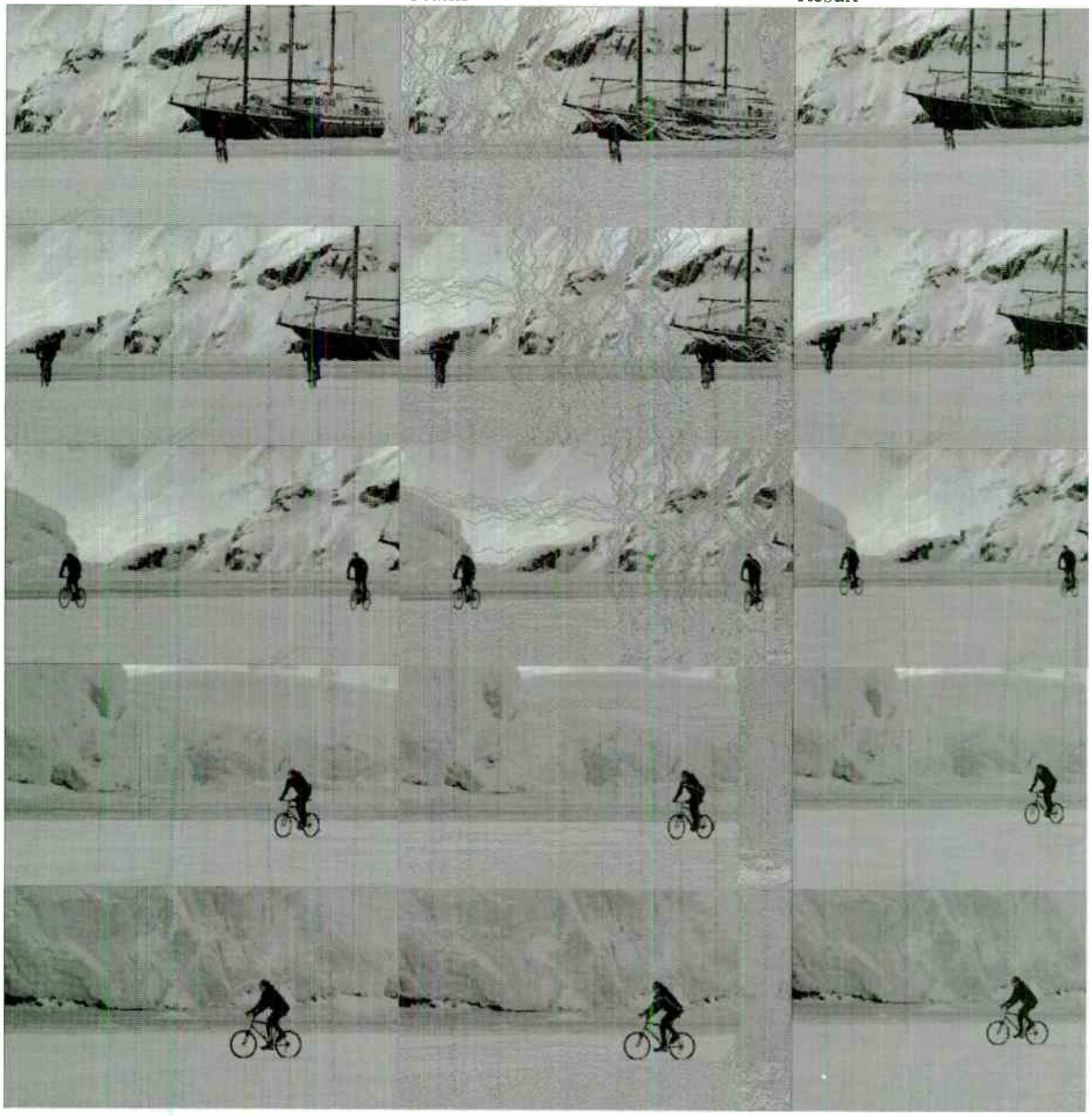


Figure 8 Seam carving on a panning sequence

REFERENCES

1. Knee, M.J. 2006. Image segmentation algorithms for video re-purposing. CVMP, London, 29 November 2006.
2. Avidan, S and Shamir, A. 2006. ACM Transactions on Graphics, vol. 26, no. 3, SIGGRAPH 2007
3. Kuglin, C D and Hines, D C. The phase correlation image alignment method. Proc. Int. Conf. on Cybernetics and Society, San Francisco, September 1975, pp 163-165.

ACKNOWLEDGEMENTS

The author thanks his colleagues for helpful discussions, and the directors of Snell & Wilcox for permission to publish this paper

An integrated, file-based production workflow for HD television: expected impact and challenges

LUK OVERMEIRE, MATTHIAS DE GEYTER, TOM BECKERS, STEVEN VAN ASSCHE, LUC ANDRIES
VRT Medialab
Ghent, Belgium

ABSTRACT

The Flemish public broadcast company in Belgium, VRT, has introduced the *Digital Media Factory*, which brings about a fully integrated file-based workflow for news and mainstream media production in Standard Definition. The architecture consists of multiple layers: storage and network infrastructure, production tools, information flow (data model) and business process layer. It is organized around a central media asset management system, which is carefully integrated with added-value craft tools based on open standards (MXF, AAF, P/Meta, SOAP, …). While a substantial part of storage, interoperability, compression and file format issues have already been resolved for SD, new challenges turn up with the forthcoming transition to High Definition Television. The R&D department VRT Medialab investigates the expected impact of HD on the final production architecture from a multi-layer perspective: HD-ready storage infrastructure, new HD and low resolution compression formats, further refinement of file format strategy and SOA- and BPM-based production workflows. This paper presents a case study on the technical challenges of a gradual transition to a file-based HD production architecture, building on the experience for SD. The reflections presented are valuable for all technical practitioners facing the introduction of a file-based workflow for SD and HD production.

OVERTURE

The Digital Media Factory (DMF) has introduced an integrated media production architecture for radio, television and on-line at VRT, while successfully accomplishing the transition from tape-based to file-based workflows. This has brought about a profound paradigm shift for the core media production infrastructure and processes. It has delivered distinct benefits such as increased use of general-purpose IT equipment for media production, concurrent engineering, editing and media processing while ingest or transfer, added-value information exchange, enhanced workflows, … Clearly, the scope of such an architecture was very challenging and the process of project execution and delivery to a large user community has yielded a lot of very valuable hands-on experience and built-up technical knowledge. A key aspect of the Digital Media Factory revolves around the efficacious integration of its different components. This implies effective exchange and synchronization of media and related metadata between the subsystems. Interoperability is achieved by applying open standards such as MXF and AAF, P/Meta and by adopting the principle of a Service Oriented Architecture (SOA). It turned out that the relative immaturity of media production tools regarding integration API's, data models, standards adherence and documentation was a major obstacle to overcome, together with existing interpretation ambiguities and the large degree of freedom in MXF which lead to interoperability issues.

While the current implementation for Standard Definition television is continuously evolving into an optimized, full-fledged system, HD television in all its different aspects is on the brink of breakthrough. Despite the considerable progress accomplished regarding SD-related storage and interoperability issues, it is expected that the advent of HD and its coexistence with SD will bring along new challenges and will inevitably have a number of technical and architectural ramifications as well.

This paper first explains the architectural principles of the Digital Media Factory taking the postulated guiding principles for SD as a starting point. The expected architectural impact and challenges related to the gradual transition into a file-based HD media production environment are explored, taking into account the unremitting evolution of standards, technology and user requirements. Different aspects of an HD-enabled DMF are discussed in depth from a multi-layer perspective: storage and infrastructure, media production systems, information exchange and software integration. A special focus is given to the choice and application of compression and file formats, requirements for central media asset management, efficient integrations with specialized work centres and SOA-based production workflows. From this dilation, a realistic step-by-step approach for transition to HD is suggested and the related technical challenges are elaborated, taking into account the lessons learned for SD and heeding best practises and the longer term architecture. More information on the work done and ongoing research projects related with the investigation of these topics can be found at [1].

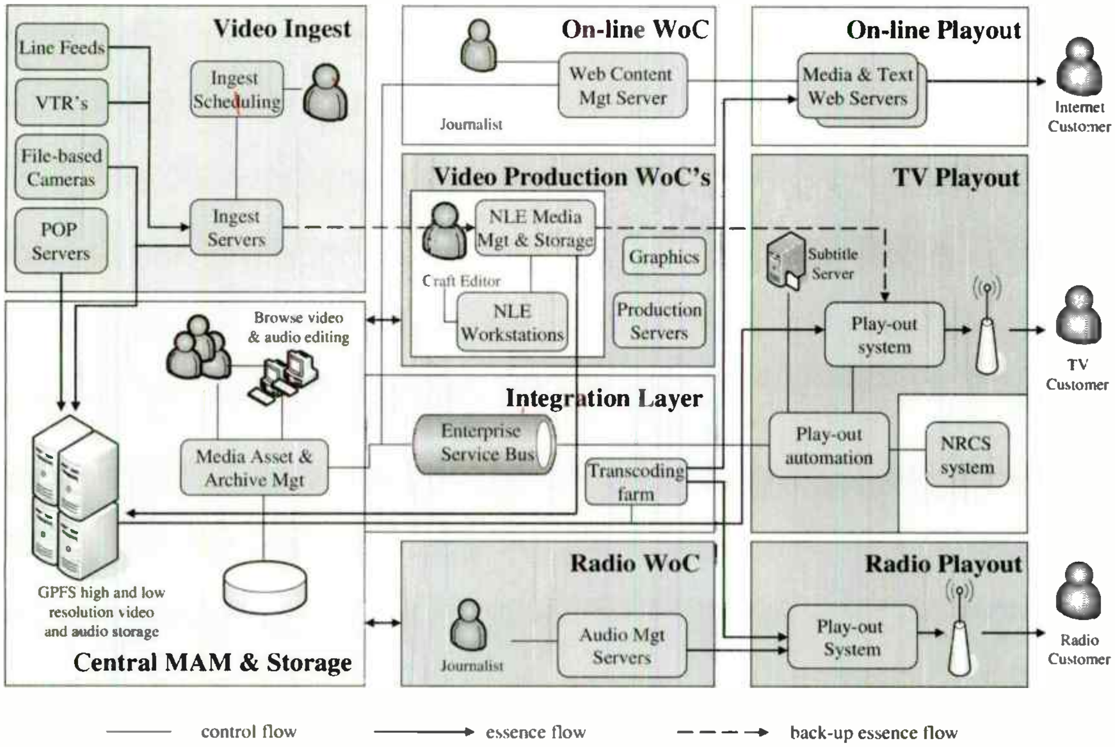


Figure 1 - Overview of the Digital Media architecture

PRINCIPLES AND ARCHITECTURE

An architectural overview of the Digital Media Factory is depicted in Figure 1.

File-based workflow paradigm

The underlying philosophy of the Digital Media Factory is to realize an end-to-end, fully integrated file-based workflow, from acquisition (i.e. file-based cameras and production servers) to play-out and archive [2]. Traditional tape-based workflows are replaced by content-centric non-linear workflows, where different operators can perform different production steps at the same time [3]. The file-based paradigm enables cross-media reuse and repurposing. Media files are preferably stored in their contribution or production format, while keeping transcoding and file transfers to an absolute minimum. Each new investment is oriented towards the file-based philosophy.

Integrated and layered architecture

The Digital Media Factory is based on an open and layered architecture. The different layers (storage and network infrastructure, production, information and business service layer) are loosely coupled. In order to facilitate the integration and coupling of the different system components and domains, the *single sourcing* principle - i.e. one vendor solution for each functional area - is applied as much as possible. Open standards such as MXF and AAF [4], P/Meta, SOAP and XML have been adopted to ensure interoperability. More specifically, MXF OP1a has been selected as the standard file format, wrapping essence in either DV25 or D10 (IMX 50) for SD.

Work centre and central media asset management

The boundaries of the Digital Media Factory are demarcated at the transition between the real-time infrastructure and the file-based production environment. It consists of different, specialized work centres (WoC) such as audio and video editing, subtitling, and graphics which are connected to a central media asset management (MAM) system (Ardome by Ardendo). The MAM system serves as a media-aware hub and repository for production and archived material. The Ardome system is the main access point to media essence for a large media production user community (e.g. journalists, program assistants), by offering search, retrieval, browse and simple editing functionality on low resolution material. By contrast, the work centres are typically specialized, optimized for a particular task and targeted to a smaller group of craft users (e.g. video editors).

Separation of essence and metadata

Metadata is gradually enriched throughout the different steps of the production lifecycle of the related media content. Notwithstanding the ubiquitous application of MXF as the storage and interchange format in the DMF and its inherent capability to wrap media and metadata, the latter is strictly confined to a few technical parameters, such as time code, video and audio format. As a general rule, media and metadata are handled separately in order to ensure the integrity in case of parallel usage of metadata in different parts of the process. Synchronization of media and metadata between MAM and work centres is handled by the (SOAP-based) integration layer.

Use of standard IT equipment

Generic IT technology and tools are used as much as possible. Flexibility and programmability are some of the obvious advantages. An optimal selection and fine-tuning of mostly IP-based components to the specific needs of media production and processing, ideally result in both higher performance and lower cost. For the central storage architecture, IBM's GPFS clustered file system has been selected to build up a high-capacity file server and storage environment. Specialized equipment is mainly utilized in work centres such as editing and play-out automation.

PRODUCTION MODELS AND APPLICATION AREAS

In principal, two types of media production can be distinguished: *item-based* cross-media production on the one hand and *project-based* cross-media production on the other hand. In item-based production, the emphasis lies on simultaneous, short-term, high-volume media production and reuse by a large number of users in a heterogeneous environment (television, radio, on-line). Typical examples are news, sports and cultural programs. In this type of production, automation is an indispensable requisite and integrations between the central MAM system and work centres are limited to the level of a media item or a container of media items and simple timeline information (such as EDL). In this context, the DMF model as depicted in Figure 1, with a prominent role for the central MAM system, prevails. As the broadband Internet medium is ideally suited for item-based distribution, it becomes more and more the primary outlet instead of a derivative of television or radio production.

Project-based production typically relies on smaller teams of people and is generally more spread in time and thereupon less time-critical. Typical examples include drama and documentary production. In this case, complex timeline-based integrations between the different components in the production workflow are imposed. As a consequence, the role of the work centres becomes foremost, as shown in Figure 2 by combining multiple workflow steps (e.g. ingest, editing, sonorisation, graphics) before handing over to the central system. Although cross-media reuse remains important, the focus rather lies on a specific medium (e.g. television) in this case.

From a qualitative perspective, one can identify the following application areas of television production:

- high-end production: landmarks, drama
- mainstream production: soap series, documentaries
- news and sports
- video journalism (consumer cameras)
- user generated content

Each of these areas corresponds to different requirements regarding quality, compression, storage, etcetera.

From this discussion, it is clear that continuously increasing diversification in media production is inevitable and that for each particular area efficiency can only be guaranteed by clear rules and technical guidelines, optimized infrastructure and integrations and distinct application specifications.

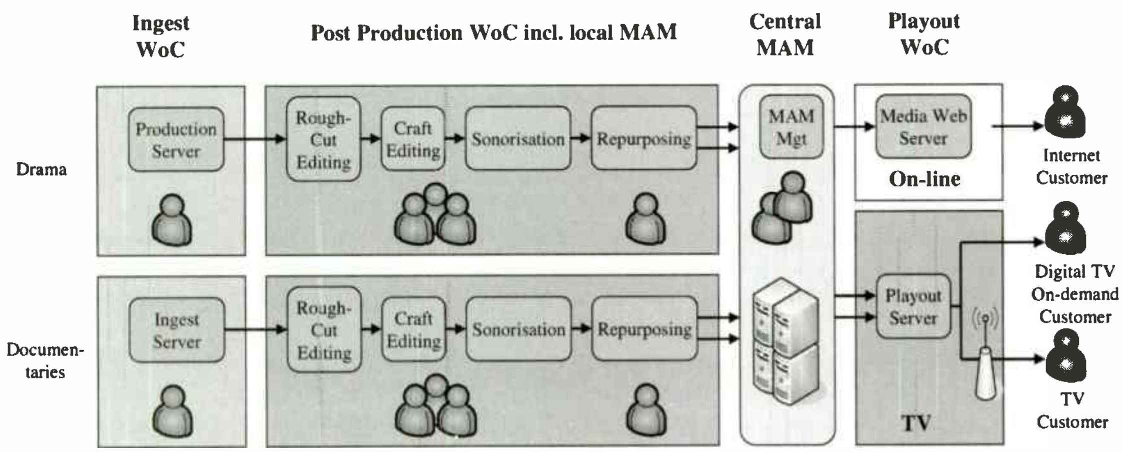


Figure 2 - Island-based architectural model

EXPECTED IMPACT OF TRANSITION TO HD

Transition plan

This year and the subsequent years are predestined to reveal the definitive breakthrough of High Definition Television production at VRT. The transition from SD to HD will occur gradually and is planned to be completed early next decade. This summer, sports events such as the Tour de France and the Olympic Games are planned to be broadcast in HD. In addition, the first full-HD broadcast of documentaries and a multi-camera drama series are anticipated this autumn. In the latter case, server-based capturing and editing in a HD work centre will be deployed first; integration with the central MAM system will follow subsequently. Live events in smaller, middle-sized studios and HD compatibility of the central media platform are planned in the course of 2009. Later on, large studio productions and HD News are likely to follow. Clearly, the aim is to carry out a feasible, deliberate and gradual extension of HD production workflows in the coming years. Roughly speaking, the challenges concerning introduction of HD are three-fold. First, the necessary investments in real-time equipment (e.g. cameras, studios, production servers, final control room) should be performed. Next, the Digital Media Factory has to embrace HD-based workflows in all its different aspects such as storage and network infrastructure, media file formats and integrations between systems. The third concern comprehends the distribution of the produced HD material to the end-user.

HD-related evolutions and prospects

Unarguably, with the advent of HD television, the following tendencies regarding media production file inventories can be identified:

- multiplication of storage and bandwidth requirements with approximately a factor of four compared to SD. In order to turn the file-based paradigm into advantage, a processing rate of up to four times real-time is considered an indicative, "rule of thumb" target figure for complex multi-step workflows with a limited number of head-tail transfer operations. An example of such a workflow is the intake of file-based camera material in the central MAM system (including low resolution and key frame generation) and subsequent forwarding to an editing work centre.
- proliferation of related media instances: inventory management of different versions (e.g. HD, SD, browse, Web, mobile) will be required.
- proliferation of HDTV formats: 720p50, 1080i25, 1080p50. Different aspects play a role in the final choice: format support in equipment throughout the end-to-end media production workflow, glass-to-glass quality preservation, expected future path, program content, content contributor, ... One should be careful not to be trapped by the *law of the hampering headstart*.
- proliferation of high resolution formats: different types of optimisations regarding HD compression formats emerge. On the one hand, optimisations for minimum storage and maximum quality are typically applied in camcorders, archive, play-out and distribution solutions. Example formats are MPEG-4 AVC I-frame only and long GOP, MPEG-2 long GOP and JPEG2000. Optimisations for media processing and quality preservation on the other hand, are employed in post production editing suites, e.g. Avid's DNxHD and Apple's ProRes format. These new advanced formats are not fully optimized yet and only start to outperform legacy formats such as DVCProHD and MPEG-2 and each necessary transcoding step between them causes quality loss.
- proliferation of low resolution formats: application-specific optimizations will appear; for instance, edit-friendly low resolution media for desktop editing.
- proliferation of wrapper formats: due to the anticipated diversification, both on the level of variety in work centres, application areas and repurposing opportunities, the entrance of Quicktime and more complex MXF flavours is expected, e.g. different operational patterns (OP-Atom, OP1b, OP2b), inclusion of subtitles, media analysis information and application-specific metadata in MXF.

By poignant contrast, a quest for HD-capable media storage and network infrastructure, minimal file transfer, transcoding and rewrapping and an increased importance of quality control is anticipated at the same time.

From a broader perspective, the following trends can be added likewise:

- proliferation of storage requirements for different applications: the processor-intensive transcoding and rendering tasks bring along different storage requirements than central storage of HD, SD, browse and Web material, secondary storage, or editing on generic storage.
- concept of distributed media processing: a typical performance bottle-neck in media production workflows concerns transcoding operations. The inevitable explosion of transcoding needs and the corresponding processing power requirements (four times real-time likewise) urge for performance-effective solutions. More concretely, both hardware and software solutions should be optimized per task and tuned to each other. As an example, the use of CELL processors in a grid-based architecture can be mentioned as a promising technique to cope with multimedia-specific demands in this context [5].

- service oriented architectures (SOA) for media production: due to increasingly complex workflows, reusable services and horizontal integrations are gaining rapidly in importance. SOA provides for modelling, automation and monitoring of media services, possibly including human interaction.
- intelligent media analysis and indexing: in order to enable efficient search and retrieval mechanisms for ever increasing quantities of media content, efficient metadata extraction and indexing algorithms are essential. These techniques will further contribute to metadata enrichment and hereby further encourage process automation.

Expected impact on architecture

Demanding a full-blown central MAM-based integrated architecture for HD to be ready for use on a short term is not a realistic expectation. At present, no system solution on the market can meet all requirements regarding bandwidth and storage capacity in a cost-effective way. HD-enabled work centres exist, but are not yet offering quadrupled transfer and processing speed, and storage capacity for instance.

In case of large transfers, network overload already looms for SD and the situation obviously deteriorates in an HD architecture. An intelligent traffic management system to control transfers between different system nodes becomes more and more essential. Moreover, Ethernet network transfer rates of 1 Gbps still prevail for standard server equipment. This will inevitably introduce a *threshold effect* for HD processing, especially in case of multi-stream handling such as conform operations, hardly surpassing real-time processing due to this.

Another infrastructure-related issue of particular interest relates to efficient restore mechanisms for HD in case of disaster recovery. Present mechanisms typically require an excessive amount of time.

For reasons of bandwidth and current product support amongst others, the future-oriented 1080p50 HDTV format is too premature to be a viable option on a short term. From a technical standpoint, provided sufficient support throughout the complete workflow is available, 720p50 seems the most suited candidate at present, both for production and distribution [6]. By avoiding the application of the 1080i25 HDTV format, inter-format conversions (e.g. deinterlacing in Flat Panel Displays) become obsolete and future upgrading to 1080p50 is more straightforward. However, commercial arguments can influence the final choice.

The choice of HD compression format in production is a different story. Unfortunately, the *war of formats* has not yet been concluded. At present, it is unlikely that one of the new, advanced formats will prevail shortly throughout the overall workflow. As explained before, it is rather expected that at least two types of formats will settle: camera and post production format. In addition, diversification in editing work centres may even increment the number of formats. Provisionally, one-off conversions to the play-out format might be required as well. A very extensive analysis of the HDTV compression formats available on the market has been conducted at EBU, with a special focus on HD production and post-production scenarios [7]. In a next phase, the impact of transcoding will be investigated. On a short term, the usage of legacy formats such as DVCProHD appears to be a very viable makeshift, for it provides a reasonable degree of quality and sufficient support is already available.

With respect to file formats, one can expect further refinement of the present MXF and AAF strategy in VRT DMF: adequate usage of different operational patterns (e.g. OP Atom in case of stand-alone editing or possible editing on generic storage), optimized conversion scenarios between MXF and Quicktime, advanced quality control, the entrance of process-dependent MXF application specifications ... An application specification constrains the number of options associated with the use of MXF in a facility and hereby restricts the different interpretations and related interoperability issues to an absolute minimum. Moreover, the focus regarding MXF and AAF has shifted from the "bits" to the "workflows". In order to enable efficient and automated repurposing of media material to many different platforms in a broadcast environment, the Advanced Media Workflow Association (AMWA) has introduced the MXF Mastering Format during NAB 2007 [8]. It facilitates the creation of different versions from one or more media *master* files by handling the individual media components (video, audio) separately. The concept can be applied both physically (i.e. structured as a directory of files) or virtually (e.g. in a MAM database). The big challenge here lies in the conditional implementation and appropriate deployment in each application area (MAM, play-out, ...). Finally, a certain degree of interoperability between different post production editing work centres should be achievable based on the AAF Edit Protocol. Such integrations will preferably rely on coarse-grained, configurable services built upon the freely available AAF SDK, e.g. execution of media transfers between work centres based on unique identifiers inside an AAF file [9].

Support for HD in the central MAM system typically occurs at variable levels, from mere storage capability for the selected format(s) to a full-blown, integrated service platform embracing complex HD workflows and HD-capable integrations with work centres. Besides, the emphasis for work centres is likely to shift even more towards added-value, compound workflow scenarios in order to avoid superfluous file transfers. In this context, editing of an SD proxy on central storage (i.e. outside the work centre) might be an appropriate alternative in specific production cases.

Undoubtedly, the area promising the largest margin of further maturement is the information and metadata layer, with manifold, possible improvement opportunities in the different steps throughout the production chain, i.e. from camera acquisition to archive annotations. One of the most significant added benefits of integrated, file-based media production lies in the effectual exploitation of metadata enrichment. Related herewith, the core integration component of the *Service Oriented Architecture* in the current implementation of VRT DMF is the Enterprise Service Bus (part of IBM WebSphere ESB). At present, the focus is mainly on application integration and data exchange (i.e. media and related metadata) rather than on real functional integration. Further refinement of information and enriched metadata will leverage automated business processes and business scenarios requiring human interactions in the future. Typical examples of potential SOA-based services in a broadcast environment include: ingest, transcoding, quality control, non-linear editing (NLE) integration.

Finally, the comprehensive, but one-off transition to an inherently, all-digital HD production and distribution environment will establish the technological baseline for further growth to even higher resolution formats such as 4K and UHDTV. Comparable to IT technology and in contrast with the preceding *steady-state* PAL system, this evolution is likely to happen as a continuous movement.

REALISTIC APPROACH AND TECHNICAL CHALLENGES

The far-reaching technical challenges related to the introduction of HD in an integrated, file-based media production system stand in sharp contrast with the considerable extent of short-term expectations. Therefore, the architecture will be gradually refined starting from the essential components and minimal integration complexity.

From an architectural viewpoint, a trade-off will be required between accepting a certain degree of independent HD work centre islands on the one hand and forging ahead with stressing the centralized architecture for HD workflows. The optimal trade-off will depend on the respective production models. The island-based approach brings along some simplifications, making it easier to extend HD production step by step, especially at an early stage. However, it impedes seamless inter-system transfer and possible reuse by a large user community. This might be acceptable though in a limited number of use cases.

Related to storage infrastructure, one can conclude from the discussion earlier an increasing need for "*fit-for-purpose*" storage clusters, as different requirements apply for central storage of respectively HD, SD and browse media, secondary storage, CPU-intensive tasks such as transcoding and rendering, and possible generic storage for central editing. In addition, optimized software solutions should be fine-tuned in correspondence with these different types of clusters. Clearly, the challenge is to further refine the present file server SD storage cluster environment in VRT DMF in the context of HD.

While the choice of the HDTV format is expected to follow the EBU recommendation, i.e. 720p50, a decision regarding the selection of one or more HD compression formats in the respective application areas is less straightforward. In some isolated, advanced production cases, proven legacy codecs such as DVCProHD turn out to be the most convenient option, as an intermediate solution. If an intermediary scenario based on legacy codecs will be necessary and for which time span, depends on a lot of factors: product-specific support, standard adherence, subjective quality, bandwidth requirements, format agility (more specifically in the most demanding production steps) and licensing models. There exists a risk that commercial considerations regarding licensing could drive the overall market into the direction of HD format fragmentation, hereby augmenting the need for transcoding and corresponding extra licensing costs.

Taking into account the lessons learned for SD, the same "*keep it as simple as possible*" approach will be followed initially for MXF-wrapped HD. Extensive testing based on the core application specification for HD in VRT DMF, will be performed systematically in preparation of the go-live. It is expected though that the interoperability efforts accomplished for SD will pay off for HD. Particular care will be paid to contingent, non-constant bit rate compression formats, such as MPEG long GOP or the upcoming AVCHD format, which typically require more complex interleaving and index tables.

From the core MXF application specification, application-specific extensions or adjustments will be put in place subsequently, in order to cope with the different functional requirements in an increasingly heterogeneous environment. From there, quality control can be further extended to encompass media compression formats and even content-specific checking.

Regarding MXF complexity, inclusion of removable user-specific descriptive metadata in MXF is expected to arise mainly in MXF import and export scenarios, i.e. when media enters or leaves the Digital Media Factory. In the context of automated repurposing of raw and finished media material, the application of the concept of the above-mentioned MXF Mastering Format inside VRT DMF should happen gradually and well thought-out. The central MAM and play-out system are amongst the most likely candidates for adopting this advanced paradigm, albeit inherently different in nature, i.e. on the database level inside the MAM system, respectively on the file structure level in the play-out system. Obvious advantage scenarios in a play-out environment are: last-minute addition of extra language track, swift replacements of credits or content, ...

Due to the increasing diversification, new work centres (such as the Apple editing suite) appear, which are not by definition adherent to the ubiquitous MXF format. In this case, the challenge is to attain in due time a comparable degree of MXF support and functional integration capabilities at the interface level as the work centres already in use.

With respect to the development of a genuine Service Oriented Architecture in VRT DMF, upcoming efforts will initially concentrate on additional integrations, such as a rewrapping and transcoding services in many-to-many integration scenarios. On the longer term, automated business processes are envisaged. The ingest process is a suitable kick-off use case, as it is believed to have uniform, reusable characteristics for many broadcasters. A concise, high-level BPMN (Business Process Modeling Notation) model of the ingest process has been drawn up in the EBU P/CP workgroup, by considering the related production tasks of the members of the group. The MSAG (Media Services Architecture Group) workgroup within AMWA intends to further refine these process descriptions and implement them partly as a Proof of Concept, desirably by conferring with EBU. Once proven its viability, other services are likely to follow along, such as transcoding, quality control, NLE integration. In the latter case, both mixed *central MAM-NLE* and *back end production-distribution* combined workflows should be contemplated.

CONCLUSION

In the coming years, VRT's Digital Media Factory for SD will gradually evolve into a fully integrated, HD-enabled production environment. Although a firm base for further enhancements and extensions has already been established, this evolution still remains a huge challenge in the different layers of the architecture. The impact on the architecture and a realistic step-by-step approach in each of these areas has been presented. Particular care has to be taken that short-term objectives are implemented in accordance with the envisioned long-term architecture. As such, the true benefits of the file-based paradigm can be guaranteed in future HD production environments. The extensive transition efforts with regard to HDTV will pave the way for next-generation systems such as UHDTV.

REFERENCES

[1] http://medialab.vrt.be

[2] J. Hoffman, P. Soetens and M. De Geyter, "An integrated file-based production workflow: the Digital Media Factory", International Broadcast Conference Proceedings, September 2007.

[3] A. Mauthe, P. Thomas, "Professional Content Management Systems", West Sussex, England, Jon Wiley & Sons, 2004.

[4] N. Wells, B. Devlin and J. Wilkinson, "The MXF Book – Introduction to the Material eXchange Format", Elsevier, 2006.

[5] L. Lurng-kuo et al, "Video Analysis and Compression on the STI Cell Broadband Engine Processor", Multimedia and Expo, 2006 IEEE International Conference on Volume, Issue , 9-12 July 2006 Page(s):29 - 32

[6] H. Hoffmann, "Quest for finding the right HD format: a new psychophysical method for subjective HDTV assessment", International Broadcast Conference Proceedings, September 2007.

[7] EBU P/HDTP, "New HDTV Studio & Acquisition Compression Analysis (Sony, Panasonic, Thomson, Avid & Test plan)", EBU technical publications, EBU BPN 076-079 & Supplement 1, December 2007.

[8] http://www.amwa.tv

[9] M. De Geyter, N. Oorts, L. Overmeire, "File-based broadcast workflows: on MAM systems and their integration demands", SMPTE Technical Conference, October 2007.

From MXF to SOA

ERNESTO SANTOS

MOG Solutions, Portugal

ABSTRACT

The professional media industry realized years ago the urgent need to move towards Information Technology based solutions in order to reduce costs, improve product quality and reduce time-to-market. The first step was to define a standard way to represent audiovisual material in computer systems. As a result the MXF file format was born.

The ability to hold both the audiovisual material and its related Metadata, makes MXF an ideal vehicle for the Content, as it traverses the different stages of the Content lifecycle. Systems can feed on that Metadata to automate decisions, and add to that Metadata to convey information to other automated systems. Therefore, the workflow itself is becoming automated, enabling humans to focus more and more on creativity. However, in order to orchestrate all the systems in such a collaborative environment, usually spanning across computer networks, two kinds of interfaces are required: data and behavior.

As the data interface consolidates, with MXF becoming an established technology, initial work on behavior interfaces confirms this as the logical next step. These interfaces expose the Service(s) that a system Provides as well as the ones it needs to Consume in order to contribute to the collaborative environment.

Service Oriented Architecture (SOA) is all about that. More than a technology, it is a concept and it is bound to have a tremendous impact in the media industry. This document explores the key aspects on how to build on the consolidation of MXF file-based production, to design the next generation SOA-based workflows.

FILE-BASED SYSTEMS

The move towards tapeless environments, building on the file based paradigm, made MXF a widely adopted technology across the globe.

Although initially pushed forward by the TV industry, other industries are rapidly adopting MXF since it applies to any system that needs to produce, manage and distribute content. Examples are the Motion Picture industry, Medical applications and the Military industry. Each of these areas has their own requirements and, even inside the same area, different organizations with different ways of working will use MXF differently. The fact that the MXF toolset can be put together in a variety of configurations that easily adapt to different scenarios, is therefore a selling point for this technology.

In the TV industry, MXF adoption is starting right from the camera. In fact, today all major ENG[1] camcorders already record in MXF format. The fact that each vendor targets a specific workflow, leads to different MXF flavors being adopted in different camera systems [1]. The storage mechanisms for these camera systems also varies and, within the storage media, XML files are used to represent additional metadata, that helps manage the content wrapped inside the MXF files.

Once this content gets to editing suites, another technology, AAF, joins the party to help manage authoring information such as effects and transitions. Please refer to [2] for a discussion on these technologies. And if we go into the actual compression formats and Metadata schemes which are used, quite a few more acronyms will rapidly pop up.

Now, if the growing number of acronyms makes it difficult even for Engineers to keep up with these new technologies, it makes life even harder for those that have to work with these formats every day, regardless of which role they play in the content lifecycle. Why are Users that much exposed to all this technology? Because although the content is now in files, the process of handling them is still very much manual. Instead of carrying tapes around, feeding VTRs and pushing buttons, you now need to transfer files around and click buttons on a GUI[2].

With more and more functionality being added to software packages that process these files, more and more you need to know details about these different technologies in order to operate increasingly

[1] ENG – Electronic News Gathering

[2] Graphic User Interface

sophisticated GUIs, to have your audiovisual material go from "box" to "box"[3] as you build your Content.

In order to address these issues, there is an increasing move towards automating the integration between systems making technical details transparent to Users. However this actually means automating parts of the workflow, and that can become a problem very quickly since each organization tends to work (at least) slightly differently from the previous one.

Therefore manufacturers are finding themselves in a situation where they spend an increasing amount of effort designing custom integrated solutions, which end up not being cost effective to either vendors or users. More than that, often the result falls short of user's expectation. Vendors are rapidly learning that you should not disregard the inertia of a User who, all of a sudden, sees a lot of technology invading their life and trying to change the way he or she has worked for years.

HANDLING INERTIA

For a lot of media professionals these days, instead of carrying tapes around, you carry disks and memory cards around. So quite often, MXF becomes yet another format in a different kind of tape. That actually made it quite easy for MXF to enter the market.

However, with all that information inside that "new kind of tape", there is a lot more that you can do with it. And since media professionals are used to all kinds of automated tools from the Internet world, that they even use at home every day for their holiday pictures and videos, they start to expect more at their job sites as well. They expect something that helps them focus on the creative part of the job.

The problem is that in order to be able to build a product, a manufacturer needs to make something that is usable by many in order to make it economically viable. So it becomes a challenge to find the workflow's least common denominator.

Sooner or later the design is directed towards systems that are **flexible**, in a way that these can be configured to adapt to the User's own workflow, and **extensible** in a way that enables the addition of functionality that is very specific to one User or one profile. And once some of the fundamental tasks in the everyday operation in a facility become automated, with much of what is going on happening in the background, **reliability** also becomes even more of a key issue since users are less aware of what is going on under the hood.

Of course you can achieve a lot of the above with custom development but that often raises **cost-effectiveness** issues beyond the possibilities of the most Users' budget.

Budget limitations also often dictate that the switch to these technologies needs to be gradual. So such systems often start in specific departments inside organizations and then expand, as experience grows. This turns **scalability** into a key issue for any solution that gets deployed. And with the requirement of scalability, usually comes the need to meet target **performance** indicators. Especially with systems such as the ones in broadcast industry, that handle high volumes of high quality (and often low-compressed) Content.

The way a User perceives performance will very much depend on its role in the system. Therefore, more and more the usual performance requirement will be superseded by broader and more more sophisticated **QoS**[4] requirements.

Finally, as these systems expand, they will rapidly become the automated foundation of any business they support. Therefore, Users that are already less and less keen to vendor lock-in, will continue to increase the push towards **open standards.**

The previous paragraphs list quite general requirements. However, vendors often find themselves with such broad statements as the ultimate requirements list from potential customers. The reason is that we are getting into such a novel and highly specialized area of technology, and so different from conventional broadcast technology, that Users find it difficult to communicate their needs.

It follows from the above that manufacturers need to focus more on mapping the technology to the User and not the other way around, since technology will no longer make the sale. The end result will.

Users will focus more and more on what they do best, which is to create innovative Content, and expect the tools that they use to... just work!

---

[3] Here the word "box" is used to mean product (software and/or hardware based) to emphasize that you move content from a product that is designed by one manufacturer to a product that is potentially designed by another independent manufacturer. Therefore you are then open to problems that might arise in case these manufacturers have different interpretations of interface protocols.

[4] Quality of Service

BOTTOM-UP DESIGN APPROACH

We also need to take into account that automated workflow systems will only continue to appear if they make sense from a business point of view. Manufacturers need to find a way to clearly delimit the bounds of the "box" they are providing as well as the added value it will bring to the Users. However, since the User will more and more be expecting for the solution to be easily deployable, the manufacturers will also continue to increase their interest on open standard interfaces. After all, these will enable the connection of their "box" to the other "boxes" that make up the solution for the User.

This is happening already today, as "boxes" are gradually being wrapped with layers of software that somehow automate the way these connect to other "boxes" in the system. Therefore design is starting from the functionality provided by the products, and in a way that makes them usable in a variety of workflows. However, that may not be adequate to a number of those use cases.

This bottom-up approach to design leads to systems that are not optimal but do allow you to control costs and build the system as you go. The problem is that as the system gets bigger, you may find out that the architecture is not scalable or just doesn't adapt to the actual intended workflow. And by then you start having to deal with legacy... Nevertheless, this approach as the advantage of getting systems deployed quickly. It also teaches us lessons on ways to follow and options to avoid.

This approach is becoming popular these days as more and more "boxes" are being wrapped with layers of software that provide a standard interface, often by making use of the Web Services concept.

WEB SERVICES

This key technology was build on top of XML's popularity and currently enjoys widespread adoption. It has gradually become the reference for distributed systems on the Internet where low bandwidth and high latency demanded document based messaging interfacing, rather than the more conventional object-oriented network interfacing, usually associated with technologies such as CORBA [3].

It must be said that CORBA itself does not limit you in any way to that model, and is perfectly suitable to document based messaging as well. However, once you move towards a document based interface, your system will tend to a design that can live without the high performance characteristics of CORBA and, in that scenario, the ease of use and therefore low cost development and deployment of XML based Web Services [4] using technologies such as SOAP [5], quickly become attractive.

Web Services ease of use comes mostly from the large number of tools available in the market that, integrated with all major development environments, enable even low skilled developers to quickly develop and deploy Web Services based applications.

Since in the end the service requests across the network get mapped down to XML messages being exchanged by end points, Web Services are inherently platform independent. Therefore services running on different platforms, and implemented even in different programming languages, are able to communicate seamlessly[5].

For all of the facts above, but also in some measure for the usual hype that surrounds all fashionable technologies at one point or another, more and more we will see professional media market products provide Web Services interfaces that expose the behavior interfaces.

The word behavior is used in this document to encompass both the situation where the Web Service is controlled by and external entity such as another service or a GUI as well as the situation where a Web Service interacts with other parts of a system via an Event Channel[6]. In this case the Web Service might react to events received, or emit events in order to promote an effect in the some other part of the system.

The mechanisms that enable the communication in a Web Service environment, including the format of the XML messages travelling the network, is defined by organizations such as the W3C, ensuring this technology stands on open standards and recommendations. Nevertheless, additional levels of standardization need to be developed, to further refine the interfaces taking into account the professional media market.

STANDARD INTERFACES

Some steps have already been taken towards the definition of these specialized Web Service interfaces. Shortly after the publication of the MXF standard, SMPTE identified the need to make it easier to interface MXF to some of the IT technologies that are most commonly used in Metadata handling in a variety of industries.

---

[5] A feature that is also key in CORBA, although in that case a binary representation for the messages is used.

[6] For the reader familiar with the Distributed Systems field, this is meant in the sense of Message Oriented Middleware

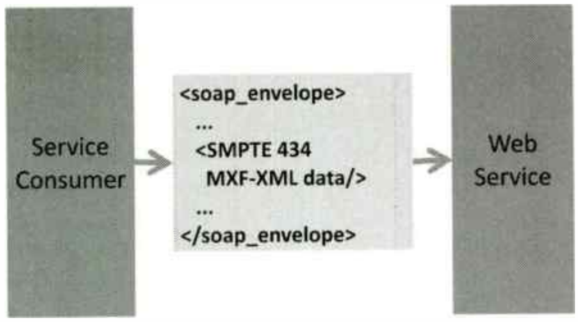


Figure 1 – Using SMPTE 434 MXF-XML

The first outcome of this effort was SMPTE 434, also known as MXF-XML [6]. With this format one can create an XML representation of the Metadata, which is usully binary encoded in an MXF file. This is perfectly suited to travel inside the XML messages such as SOAP messages, which are exchanged between Web Services.

The logical next steps in standardization will for sure be on the behavior interfaces, their syntax and semantics. In the remainder of this document specific techniques are discussed which will help identify candidates for standardization.

SERVICE ORIENTED ARCHITECTURE

Will this move towards Web Services lead us, by itself, towards the currently acclaimed Service Oriented Architectures? Not by itself. SOA is a much deeper concept than just exposing product's functionality as Web Services.

SOA builds on Object Oriented Technologies by establishing the concept of contracts between entities which reflects the expected dependencies between those entities. However it takes a loose coupling approach, since SOA services have much coarser granularity than the classical Objects in the Object-Oriented world.

The loose coupling led to document based message interfacing with more information being exchanged on each service request. This eventually evolved into the concept of the actual message being an "independent unit of communication" [7], which can actually include information on how it should be routed and/or processed.

Such an amount of information included in a message makes it a powerful mechanism, but also demands that you really understand the workflow in order to build the services that route/process the appropriate information accordingly.

So taking into account these boxes that will provide services to a broadcast infrastructure, and the ongoing standardization of their data interfaces (as more and more functionality gets added to the MXF standards suite), how do we design the behavior interfaces?

And how do we connect them together in a way that automates a variety of workflows in which these may be involved in?

Actually following this line of thought guides us into bending the workflow until it fits the boxes. Therefore, we risk ending up trying to adapt Users to the technology. Ideally, we would do it the other way around, following a top-down approach.

SOA LAYERED DESIGN

A top-down design approach is the best way to lead us to a true Service Oriented Architecture. In such a design process, you analyze the Users's organization in terms of their internal processes, in order to come up with the actual Services that you need, and the messages that flow between them.

These Services usually capture the business logic of the organization and therefore are considered to sit at a Business Services Layer.

Then you look into the design of these Services, often through delegation to lower level Services. The later are usually much more delimited in scope and often re-usable in a number of scenarios.

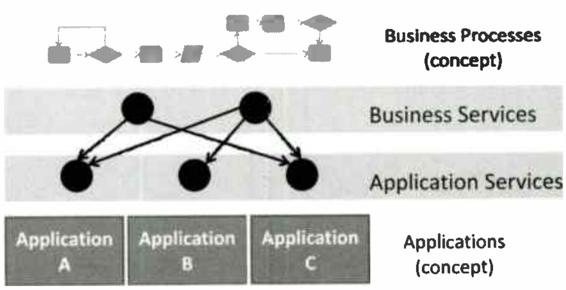


Figure 2 - SOA layers

These lower level Services therefore constitute de interface of very specific applications and are, therefore, usually considered the building block of the Application Services Layer.

Usually this top-down approach demands for a considerable analysis effort, which may not fit most budgets. In [7], an Agile design approach is recommended that basically proposes an iterative process whereby, by applying a sequence of top-down iterations, one gradually improves/enhances the system with working and functional intermediate delivery stages.

In reality, in the professional media market, the early stages of this Agile approach are bound to follow a bottom-up approach. This because there is a high amount of legacy "boxes" that needs to be taken into account in any deployment. Also because the novelty of the new file based technologies present so many options in terms of workflow, that Users and Manufacturers are still trying to converge towards the most adequate workflows.

Finally, the relative inexperience of this industry in terms of Distribute Systems, is bound to present us with several cases of evolution by trial and error, which can still be managed if one starts from a known reference, that is, existing "boxes".

Of course these initial stages of bottom-up approach have their down-side as well. Since while designing you are not seeing the "big picture", there is a high risk that you will reach a completely inadequate architecture at some point, and actually have to revert the process at very high costs.

ADAPTER SERVICES

Following the above, in these early stages of SOA in the professional media market, we are bound to see more and more usage of a specific kind of Application Service called Adapter. This is a kind of service that wraps an existing system, usually a legacy system, and exposes its functionality on the Web Services network.

Let's consider an example. It follows from the introduction in the market of new file-based camcorder technology, that one of the key processes today is the ingesting of content into editing systems, possibly with shared storage.

The actual workflow process will vary from organization to organization and even in the same organization it may vary depending on the kind of production. For example, is the prime concern how to quickly get news to air? Is it the ability to select relevant content from a high shooting ratio production for a reality show?

Such variety of workflows in this specific area of ingest alone, would lead to a incredibly costly top-down design approach. Therefore, manufacturers starting with a bottom-up approach by deploying Adapters that wrap existing boxes as Web Services and deploying additional generic services such as transcoders and intranet search engines.

Some of the conclusions that follow result from the author's own work in this area, contributing to the design and roll-out of MOG Solutions TOBOGGAN product line.

ADAPTER CANDIDATES

Camera systems and Editing suites constitute excellent candidates for Adapter services. The operation of these systems revolves around a number of concepts that can be reused across multiple workflows. The concept of Clip, EDL and proxy, among others, are all common aspects for which you find equivalents in a number of products.

Therefore, deploying Adapter Services in machines to which these devices attach using protocols such as USB and IEEE1394, enable these devices and their Content to be efficiently shared as services on a network. In addition, the above common aspects, can be exposed as common interfaces abstracting (to some degree) the other Services in the system from the specific camera system being used.

Should these same Adapter services include functionality to initiate transfers into Editing suites? Not at all. This would make these services aware of the workflow and immediately limit their re-usability. In order to enable these services to remain workflow agnostic, they should only expose the Content and functionality available in the camera system.

The same holds true for Editing suites. These also need to be exposed as Web Services, and at the time of this writing, NLE vendors actually already started following that path. Should these services be able to trigger transfers from camera systems? Again, such functionality would tie these services to a specific workflow, and also limit their re-usability in other scenarios.

Therefore, the functionality that actually orders the transfer of content from camera systems to editing suites needs to be captured at a higher layer, which is the Business Services Layer. This keeps the Adapters workflow agnostic, and enables these services to evolve independently in terms of functionality, without affecting the overall system.

Note however that not all Application layer services are Adapter services. More and more software services which provide intrinsic value and don't just wrap legacy products will be used. Examples are services which perform automatic feature extraction form audiovisual material, or transcoding services, among others.

This separation of functionality through different layers also has its price. It increases the need for inter-service communication, which goes through XML coding and dispatching layers, and therefore introduces additional overhead to the system. Nevertheless, more and more the IT industry is finding that it pays to sacrifice a bit the performance on behavior interfaces, in order to gain on scalability and especially on flexibility to adapt systems to various scenarios.

Therefore, for sure the same conclusions will be drawn in the professional media industry, even because here we are referring to behavior interfaces. The actual heavy weight work of moving around the high bandwidth demanding content is done through different interfaces and protocols, for which MXF over FTP is the most common example.

BUSINESS SERVICE LAYER

It is out of the scope of this document to present an extensive range of usage scenarios for Ingest operations, as well as the full derivation from those of a suitable Service Oriented Architecture for ingest operations. Nevertheless one can say that requirements include the need for hot swap, that is, batch ingest of all material started automatically once the storage media is connected to the system; ability to select material prior to ingest; ability to edit metadata prior to ingest; preview material; automatic scene detection and logging; transfer monitoring, among several others.

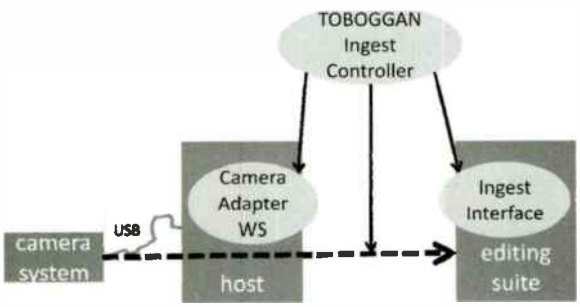


Figure 3 - Business service example

Capturing such logic as Business Services enables such functionality to be re-used when build different kinds of GUIs, and exposing that same functionality to higher level services and even external systems.

Still following the example above, one solution may result into the deployment of a Business service that lists available Clips in connected camera systems, also lists registration of commissioned footage in the exposed editing suite, and order automatic transfer of Clips that match he commissioning. Such Business Service could make use of camera system Adapters and editing suite Adapters.

The same camera system Adapter may also be used in a different scenario, where a Business Service lists the Clips in the camera system Adapter and orders a Transcoding Application service to convert these to another format, which might be more suitable, for example, for internet preview[7].

In both these examples it is obvious that the Application Services are re-usable. However, if we add an additional Business Service on top of it all that enables the transfer to the editing suite of all commissioned material and ensures that only such commissioned material is transcoded out to a lower resolution version, we see that Business Services may re-use other Business services as well. This approach results in a hierarchy whereby services are composed with other services. This matches the "service composibility" [7] principle which is key SOA based systems.

As we move from lower level Application Services to the higher level Business Services, we move from workflow-agnostic to business specific. Therefore, as we move higher in the abstraction layers, the less likely it is that we will find interfaces that can be standardized. Therefore we should expect Adapters and general Application Services to become the first candidates for standardization.

CONCLUSIONS

More and more we see Web Services based systems make their debut in the professional media market. However, in order to get to true Service Oriented Architecture a lot still needs to happen.

Strict top-down approaches are the ideal way to get an elegant SOA design but these may not be cost effective. This bottom-up approach fails to adapt the technology to the Users, ending up trying to adapt Users to the technology instead. An Agile approach is best suited for this market. However, the initial stages of the Agile methodology are bound to be bottom-up design stages, gradually turning into to-down design stages as the technology matures in this field.

Adapter services will be driving the entrance of SOA in this field, with Business Services appearing

[7] Even higher level services at an Orchestration level could signal other business areas at this point (such as Internet Media department) that new content is available for approval and Web publishing. Nevertheless, covering of Orchestration Services layer [7] is outside the scope of this document.

gradually as SOA awareness evolves and designs move away from monolithic systems into distributed systems, where Web Services collaborate to build the full workflow.

The need to automate the interconnection of these services will lead to more demands on standardization of interfaces. Adapter services are bound to become the first candidates for such standardization operations.

REFERENCES

[1] Santos, E., 2007, "Operational Patterns… the MXF Flavors?", NAB 2007 Conference Publication

[2] Santos, E., 2006, "AAF, MXF and XML… Putting it all together", IBC 2006 Conference Publication

[3] OMG, CORBA Specification, http://www.omg.org/technology

[4] W3C, Web Services, http://www.w3.org/2002/ws/

[5] W3C, "SOAP Version 1.2 Part 0: Primer", http://www.w3.org/TR/soap12-part0/

[6] SMPTE, 2007, "SMPTE 434 – MXF - XML Encoding for Metadata and File Structure Information", http://www.smpte.org

[7] Erl, T., 2005, "Service Oriented Architectures: concepts, technology and design", Prentice Hall

MIGRATION TO ALL-IP INFRASTRUCTURES FOR DISTRIBUTION OF BROADCAST SERVICES

Tom Lattie
Harmonic Inc.
Sunnyvale, California

ABSTRACT

Since the mid-1970's, satellite has dominated as the delivery method for primary distribution of pay television services. While technical advances in RF modulation and video compression have enabled satellite distribution to keep pace with the expansion of content during the past 30 years, paradigm shifts in how content is packaged and consumed require a new method for disseminating content moving forward. While the migration of some services to High Definition has accelerated the need to investigate new distribution models, the unprecedented growth of specialized programming and projected explosion in High Definition services demand a fresh look at distribution networks that can scale exponentially.

Content providers no longer rely on a handful of service provider networks to reach consumers. Growth in high-speed connections to Internet and advancements in CE devices have enabled new a new breed of Service Providers with radically different business models As a result, programming can reach the consumer's living room via an ever increasing number of networks and formats. The established model of linear play-out delivered via satellite only serves a subset of service providers in this new market and the approach of building separate systems for each delivery network will not scale. Success and profitability demand Content Providers have the ability to dynamically repackage and disseminate content regardless of the use case.

Technological advancements in video compression and world-wide IP infrastructure allow Content Providers to realize new business models and revenue streams by taking a completely new approach to Primary Distribution. Specifically this paper will discuss:

- Innovations, allowing for reliable delivery of Video across IP networks
- Delivery of "mezzanine quality" material via IP Networks
- Repurposing mezzanine sources at the edge of the network to improve network scalability

INTRODUCTION

For the first 25 years of the pay television market, primary distribution via satellite made a lot of sense. With thousands of cable systems across the country and no interconnection between them, the point to multipoint nature of satellite was ideal. The rapid growth of programming was countered with technical advances in RF modulation and video compression enabling satellite distribution to keep pace.

Looking beyond 2008 it is clear that the status quo distribution model is not scalable enough to survive the tidal wave of High Definition and specialized programming on the horizon. While the past few years have again seen substantial improvements in RF modulation and video compression it is important to re-evaluate the pay television landscape before committing to yet another upgrade of the same old model.

CONSOLIDATION OF PAY TELEVISION DELIVERY FOOD CHAIN

The past ten years have seen simultaneous growth in digital pay television households and consolidation in the number of companies providing services to these households. According to the National Cable & Telecommunications Association, roughly 97 million of the 113 million US Television households are served by a cable or satellite TV provider[1]. Previously, programming flowed through hundreds of cable companies operating thousands of non-connected cable systems but now three-quarters of the nation's pay television customers, 86 million households, can be reached via eight operators. While the number of channels and types of niche programming has increased, during this same period of time the purveyors of the content have decreased. Through the magic of mergers and acquisition, a dozen or so companies originate most of the content, (e.g. Disney owns ABC who owns ESPN).

Creation of high bandwidth IP distribution networks

As important as the relatively small number of U.S. operators is the fact that despite their reach, each operator is heavily interconnected. Two of the operators, Echostar and DIRECTV, distribute national programming to 32 million households from four uplink facilities. The remaining six operators are cable MSOs (Comcast, TimeWarner, Cox, Charter, Cablevision and Bright House) who have spent hundreds of millions[2] of dollars interconnecting each of their respective cable systems. Driven by data and voice initiatives, each MSO has a nation-wide IP backbone network. In July 2006, CED Magazine reported "Comcast, which is using Level 3 as its main provider, has 19,000 miles of dark fiber and 10-GigE connections (with 40 Gig on the roadmap). Today, this backbone covers about 95 percent of Comcast's territories."[3] These massive private networks are capable of supporting IP multicast, which allows the efficient point-to-multipoint distribution of content in the networking world. Unlike point-to-point delivery of video on the Internet today, an operator can reach each cable headend on the network with the same stream. The combination of IP multicast and a nationwide high-speed backbone give operators like Comcast the equivalent of private satellite capacity in the ground. By allocating ten percent of the this 10-GigE backbone to the distribution of video, Comcast could distribute the equivalent of twenty-five QAM 256 transports--seventy-five high-definition channels--to all of their cable headends.

Consolidation results in less Primary Distribution connection points

While not all of the Programmer Networks under a given parent corporation originate in the same location, the number of uplinks sites is quite small. A survey of the nearly 100 HD channels currently available produces a list of roughly 20 origination sites. Generalizing these locations into "Metro Areas" by grouping locations like Denver and Centennial (Colorado) or Culver City, Burbank, Los Angles (California) together, the number of origination locations near ten. When considered in the context of these newly inter-connected mega-service providers, the scope for primary distribution of pay television services is reduced. The majority of the 113 million US Television households can now be served my moving content between a handful of locations. Even with this consolidation amongst programmers and service providers there is still a requirement to connect these sites. The growth of rich media delivery over the Internet has driven the growth of large backbone service providers with nationwide networks to allow high bandwidth inter-connection of programmers and service providers. Fueled by the same market consolidation economics, companies like Level 3 Communications have amassed Internet backbone, content delivery network and satellite teleport properties to create massive media delivery infrastructures. Not surprisingly, a study of Level 3's network "points of presence" correlates nicely to content origination sites and locations of the major service operators.

TECHNOLOGICAL ADVANCES IN PRIMARY DISTRIBUTION: COMPRESSION AND IP

Beyond the physical interconnection of programmers and service providers via a third party, a common and reliable protocol for moving the video is critical. Technological advancements in video compression and transporting video across flexible IP infrastructures enable new distribution methods and business models.

Reliable delivery of Broadcast Video via IP networks

Satellite remains the dominant method for Primary Distribution today because in addition to its point-to-multipoint characteristics, Satellite provides a highly reliable ubiquitous standards-based transport mechanism. Standardization in RF modulation and video compression ensured the video would arrive at its intended destination and could be easily repurposed before final distribution. Fortunately, the same framework exists today in terrestrial networking. TCP/IP, the language of the Internet, provides a common protocol for connecting disparate networks and moving content seamlessly. Based on the premise of two-way communication, IP-based networks were not initially designed for the kind of unidirectional reliability delivery linear broadcast programming requires. With TCP/IP, if the recipient does not receive the entire data set correctly it simply requests a retransmission of the missing data and the sender obliges. When dealing with the real-time delivery of Broadcast video, the buffering, latency and reassembly required for traditional data delivery does not work. Similar to satellite distribution, the video traffic needs to be "wrapped" in such away that it can continue to flow across IP networks in a unidirectional manner most compatible with live Broadcast video. Just as Forward Error Correction (FEC) for RF modulation addressed the problem of packet loss in satellite distribution, FEC can also help in IP based networks. Specifically, SMTPE has standardized an implementation of error correction (SMPTE 2022-2007) ensuring not only reliable delivery but also interoperability between equipment on either side of the network.

Delivery of "mezzanine quality" material Mezzanine most often refers to a level between lower levels in a building or a stadium. The origins of the mezzanine trace back to the Italian word for 'middle' and before that the Latin for 'median'. In the broadcast industry, mezzanine was first used to describe a distribution format for audio, somewhere between the lower level of MPEG1L2 compressed bit-streams and the upper level of the original baseband audio. Mezzanine audio formats like Dolby E attempt to balance the quality versus compression equation in the distribution channel between programmers and service providers, offering an intermediate format that enables a multitude of format options for the final distribution. With an incoming feed of Dolby E, a service provider is relatively unconstrained in what can be provided to the end user. Both 160Kbps stereo and 384kbps 7.1 multi-channels services can be derived from the Dolby E source.

For the most part so-called mezzanine options are not available today for video services. Such a tier would require a 2-3x increase in satellite distribution capacity and would eliminate the ability of the smaller service providers, which today serve ten million households, to cost-effectively pass these services through their networks directly to the end user. As a result, programmers are constrained by the codec and bit rate requirements of downstream networks and service providers. Programmers understand this reality and go to great lengths to find the ideal compromise between video quality, bandwidth and the constraints of the downstream networks.

In the fiercely competitive pay television market, service provides look to differentiate themselves from the competition on the basis of video quality and channel count. As they attempt to fit more programming into limited network spectrum, the ability to achieve better compression efficiency is hindered by the incoming video quality. Each time video is compressed, information is lost, leaving each successive compression step with less information. The higher the compression ratio, the more information removed and the lower the resulting video quality. Based upon their individual business and networks needs, each service provider needs to define the incoming bitrates and video quality from the programmers. The current "one size" of primary distribution does not fit all but with the one-to-many satellite distribution model one size is all that is offered.

By moving to an IP distribution model, programmers can enable cost effective point-to-multipoint and point-to-point models, allowing the service providers to choose what fits best with their business model. DBS operators who are moving to new compression formats and employee a "decode re-encode" model for service aggregation may choose high bit-rate MPEG-2 or MPEG-4 AVC (H.264) feeds. Cable operators who have a large installed base of MPEG-2 decoders and more of a pass through mindset may choose a slightly higher MPEG-2 bit rate than offered today. Given the network topologies of the large operators it is conceivable that a programmer could provide a service provider with an uncompressed version of the programming, enabling the service provider to encode and distribute the programming internal to its network. Technologies like distributed IP-based Statistical Multiplexing allow a DBS operator to physically place encoders at the programmer's facility and multiplex channels at the DBS uplink.

NEW BUSINESS MODELS: SUPPORTING THE ON-DEMAND REVOLUTION

In addition to improving on today's model for distribution of linear pay television services, a migration to IP-based distribution is ideal for non-linear consumption. Delivering a real-time broadcast channel over an IP-based network is far more complex than sending an individual program in file format. Furthermore, using high bandwidth connections, it is possible to send a program faster than real-time, (i.e., transmit the file version of a one hour program in 15 minutes).

With the explosion of video-on-demand offerings and then introduction of the DVR, the consumer's appetite for à la carte consumption of programming on a per episode has grown exponentially. This shift has caused the incumbent service providers to add or expand the service offering to include more file-based offerings and spawned a whole new class of service providers. Apple, Amazon, Vudu and Microsoft all offer à la carte file download services while companies like Joost and Hulu offer on-demand streaming across the Internet.

Similar to the world of linear broadcast, each of these new offerings have specific requirements relating video compression codec and bit-rate. Today the programming is transferred from the programmer to the service provider in tape format. While this offers the service provider a high quality version of the program, it is a very inefficient and unsecured method of dissemination. The combination of a mezzanine level file based asset and an IP-based distribution network would allow the programmer to deliver programs securely and simultaneously to all the required service providers. At the service provider level, having been saved the initial ingest step, large transcoding farms would enable rapid repurposing of each asset.

CONCLUSION: IP DISTRIBUTION OFFERS A PATH FORWARD

Historically pay television has been delivered and consumed in a linear fashion. Service providers simply aggregated dozens of unique program options and bundled them into a packaged offering. In many cases, cable operators in particular, would simply receive the linear broadcast content from the programmer and pass it through to the customer either untouched or with minimal modification complete. Given the finite amount of bandwidth between the consumer and the service provider and the one-to-many nature of satellite distribution between the programmer and the service provider, programmers are forced to package content for the lowest common denominator. In addition to their own RF capacity constraints, programmers have to package their content to pass through a cable network with little or no modification. While this approach is widely deployed today it is fraught with limitations.

Consolidation in the content creation and pay television service provider industries has reduced the number of players such that a radical shift in how programming is disseminated is now possible. Coupled with advancements in both IP infrastructure and protocols for reliably moving video in an IP-based environment, the timing is ideal to consider a new approach for the primary distribution of broadcast television. By migrating to an IP-based method of distributing broadcast television, Programmers cannot only support growth in traditional linear services but are also thrive in the on-demand world.

REFERENCES

[1] National Cable & Telecommunications Association "Industry Statistics" http://www.ncta.com/Statistic/Statistic/Statistics.aspx

[2] CedMagazine.com "To Build or to Buy?"- July 01, 2006

[3] Wall Street Journal "To Ward off New Competitors, Comcast Builds a Mini Internet" October 13, 2005

Forensic Marking for HD VoD and Broadcast Services

Pascal Marie
Thomson
Cesson-Sevigne, France

Introduction

The issue of content security has been a hot topic in the audio-visual market for several years. In more recent times, it has become an increasingly urgent concern. The digital revolution has led to proliferation in digital content creation while the emergence of the Internet and broadband has made it much easier to share illicit copies of content with anyone across the globe.

This paper describes how watermarking technology can introduce content security beyond the set-top box (STB) for Pay-TV and Video on Demand (VoD) services.

The first part of this paper presents some of the global issues related to the future health of the media, entertainment and communication industries. The second part describes digital watermarking technology, highlights some operator use cases and presents, in detail, the integration of the technology in STBs.

CONTENT SECURITY ENVIRONMENT

Increasing piracy risks

Technological environments are making it increasingly easier to copy and redistribute content. Technologies available to consumers are enabling and facilitating piracy:

- Content marketed in HD introduces a risk of circulating very high quality digital pirated copies which can be duplicated indefinitely,
- Software applications are readily available for consumers to process content and to seamlessly edit and compress video,
- Bandwidth is available from broadband Internet connections to circulate content in both directions (upload and download),
- Peer-to-peer (P2P) networks, streaming services and Web 2.0 video file sharing sites originally intended for user generated content (UGC) make it possible to also ''advertise'' pirated copies worldwide all at once.

Each of the consecutive release windows for a given content actually corresponds to a specific instance of piracy. Usual piracy waves for feature films begin with poor quality copies recorded in a theater, and then good SD quality copies obtained from DVD ripping.

For the television industry, the two primary and increasingly concerning piracy issues for premium TV content are the growing illicit redistribution of live TV signal and the posting of TV programs (films, soap, sports, news, etc.) on UGC sites.

The next challenges include the availability of content in HD, from various types of VoD services.

Figure 1 shows the consecutive steps in the content life cycle, pre-release in post-production and distribution, followed by the various release windows as they are currently organized.

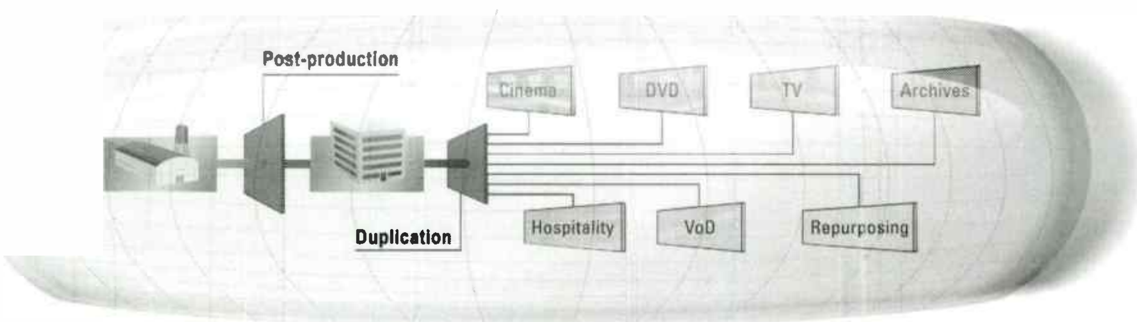


Figure 1: Release windows

Security and VoD release window

There is a variety of diverse VoD services currently available on IPTV and Cable, push-VoD on Satellite, and download or streaming VoD on Internet at large. The availability of content for the first time in VoD is getting closer to the DVD release. Furthermore, premium content is now expected in HD.

As release windows are collapsing and VoD releases become aligned with DVD sales and rentals, another critical source of content theft develops.

More than half of studio revenues for a feature film come from DVD sales. An earlier VoD release window requires extreme caution with regards to content security. The piracy scenario for HD and VoD include content theft using an HD camcorder shooting an HD plasma or LCD display at home or in a hotel room.

Consumer experience

Service providers need to consider the viewer experience among the highest priorities to deliver an attractive and successful service.

For the rental-type VoD model, it will always be necessary to enforce elapsed rights using a DRM system. Furthermore, the agreed usage conditions in a rental situation make it acceptable for the consumer to have access to the content on a single device.

On the other hand, for a download-to-own type VoD model, the consumer expectation is to enjoy content on any device; e.g. play as a DVD or on a portable media player (PMP). As long as the interoperability of DRM systems is not effective, it is important to find alternative ways to make the content available.

An interesting example is the transition of the legal on-line music distribution toward DRM-free operations so as not to lose market opportunities against illegal distribution. What barrier is left if content is made available without any DRM? There is a need to make consumers accountable for the unique copy of content they acquire.

Eliminate the consumer impunity mindset

"Who will ever know that ''I'' made copyrighted content available on the Internet?"

Content owners and service operators need to utilize all available resources to prevent and tackle piracy: technology, legislation, and communication.

Protecting content and managing rights require that piracy also be addressed from bottom to top. This can be done by first monitoring ''content sharing'' networks to locate pirated copies. Once identified - possibly using fingerprinting technology - copyrighted material can be either filtered out or monetized (share revenue from advertising).

The ultimate goal is next to identify the source of piracy, which requires operators to set-up the serialization of the content upon delivery or display. Concurrently, the serialization can be used as a piracy deterrent.

However, technology alone is not enough. Technology requires the support of regulation and communication. Rights owners need proper legislation in place to enforce copyrights and be entitled to organize a ''graduated'' response [1]. The industry also needs to organize targeted communication towards the general public. Educating teenagers and their parents about their liability and the overall impact of piracy is an essential pre-emptive action to prevent piracy.

As we described earlier, an important technological component is the ability to serialize digital content before it can be captured. From there, explaining to consumers that we have the ability to identify the source of piracy and that they face significant penalties, is the best way to decrease the current phenomenon by introducing a deterrent against illicit redistribution.

WHAT SOLUTIONS EXIST BEYOND THE STB?

Forensic marking applications

Watermarking is already widely used by studios and some television content producers prior to content release for the circulation or distribution of preview and master copies. Studios even recently demanded that watermarking be a mandatory component for all digital cinema systems.

A forensic mark can be used to trace illicit copies back to the source of the leak. The rights owner can then organize a graduated response. The main benefit of incorporating this mark is to introduce a deterrent against piracy and reduce illicit redistribution such as posting a feature film or a TV show on an Internet video file sharing site. Lastly, information about watermarking implementation also contributes to the education of consumers and eliminates the general public's impunity mindset.

Major studios are interested especially in the framework of early VoD release windows and the circulation of the feature films in HD format. Indeed, the film industry usually faces two waves of piracy: the first wave, with poor quality copies, - which are camcorded in theaters, - and the second wave, with very good quality copies, from DVD ripping and capture from VoD services. Watermarking solutions are made available for the corresponding formats in digital cinema servers, for Blu-ray discs, and in STB or VoD servers.

Not only is watermarking appropriate for major studios, it also has relevance for all content owners. Television content producers face a growing issue of the global circulation of pirated copies following the first airing of their television program, which might be an exclusive news report, a sports game such as a pay-per-view boxing event, or a popular television episode. The availability of the content from anywhere severely impacts the ability of the rightful owner to negotiate broadcasting rights in other markets.

Digital watermarking technology principles

There is increased recognition within the industry of the important role that watermarking technologies can play in keeping track of audio and visual material. Any watermarking solution must first ensure that a copy of any content can be reliably identified after format changes have occurred (including, in the case of content theft, changes carried out using a camcorder). It must also be certain that it never interferes with the viewer's experience while it carries out this function.

Digital watermarking solutions make individual copies of audio and video content easily and precisely identifiable by embedding specific data, such as the rights holder's or recipient's ID or user number. The watermark, totally invisible to the naked eye, and inaudible in the case of audio watermarking, may be embedded at various points in the content preparation and distribution process.

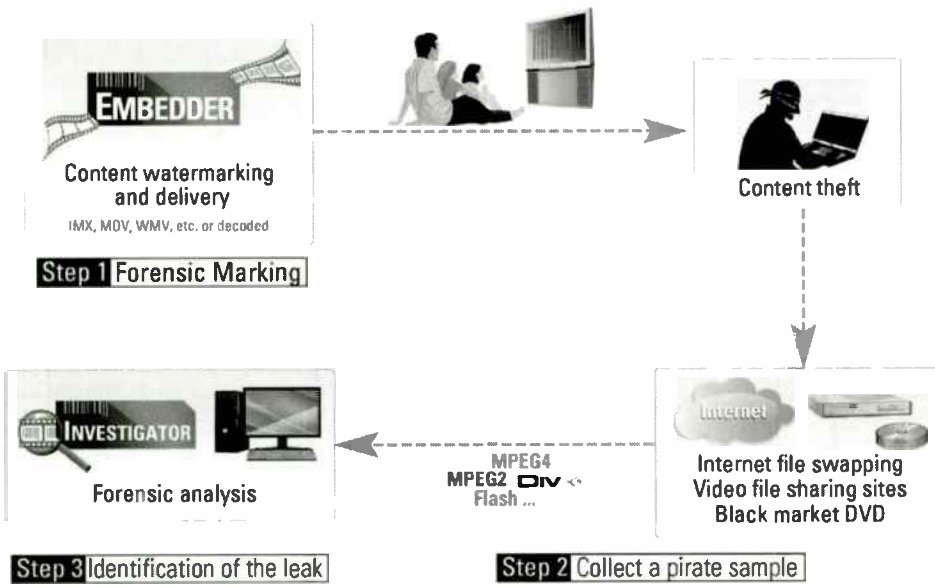


Figure 2: Watermark embedding and investigation

Figure 2 describes how the content is uniquely watermarked at the time the recipient can watch it or capture it (Step 1). In case of content theft, and whatever the video alteration and format change, the analysis of any pirated sample collected [as a DVD from a flea market or as a file or stream from any network (Step 2)] is performed using a specific software which consolidates information along the video duration to detect and retrieve the watermarked data (Step 3).

Watermark detection performance

A video watermark is designed to be invisible to the human eye, or inaudible in case of audio watermarking, but retrievable by a dedicated watermark detection software.

The way the watermark is embedded in the essence of the content is independent from the content format and will survive analog-digital conversions or re-encoding. Accordingly, the detection software is able to accept most usual content file formats as input.

More specifically, for video applications, the watermark information shall be redundant enough to nonetheless be detectable in the cases of:

- resizing (e.g. HD to SD or lower resolution),
- cropping (e.g. selection of a 4/3 area out of a widescreen video),
- frame rate change (e.g. 720p60 to 30 frames per second),
- color filtering, brightness and contrast settings,
- noise reduction,
- digital/analog conversion,
- transcoding, etc. or
- a combination of the above, such as resulting from the camcording of a LCD display.

Watermark detection is a combination of signal processing and statistics. The more the video is degraded, the longer the necessary video sample. For example, a camcorded copy might require a few minutes of video content for the investigation station to provide a result. Watermark detection from a pirate sample is a "blind" process which does not require any additional information nor any reference to the original video.

Forensic marking track records

The use of watermarking increases the responsibility of content recipients because their unique ID is hidden in the content, which means that their identity may be uncovered if an act of piracy occurs. Forensic analysis of pirated video samples has already helped investigative organizations such as the FBI on several occasions [2] to focus their inquiry on a particular copy or recipient and to successfully resolve the piracy case as a result.

The integration of the forensic watermarking solution into digital workflows introduces a piracy deterrent for content rights owners by ensuring the complete traceability of individual copies. Post-production houses, content owners and distributors as well as cable MSO, satellite and IPTV operators can identify the original source of content posted on online video sharing sites.

The purpose of content serialization [3] is to plug the so-called "analog hole", i.e. the ability to capture and redistribute the content. Piracy can be organized with a camcorder shooting a plasma or LCD display; or re-capturing the content on a computer using a video capture card. From there, either individual programs or abstracts can be made available for download, or the TV signal can be redistributed on a continuous basis. Watermarking does not limit the ability to copy or transfer content, but ensures the traceability of individual copies.

OPERATOR USE CASES

Operators' need for content traceability

What benefits can operators expect from content serialization? What is at stake for Satellite, Cable or IPTV operators? There are at least four dimensions to consider:

- liability
- competitive advantage
- acceptability by subscribers
- revenues and market penetration

Overall impact of content theft amounts to billions of dollars [4] and piracy is actively monitored by studios and other content owners. Operators have a certain level of liability vis-à-vis right owners who expect them to implement security measures to protect their assets.

Operators may transform the extra security they put in place into a competitive advantage and gain early access to HD premium content from studios.

Introducing watermarking as a deterrent against piracy is not affecting the viewer experience or limiting his options for personal re-use. The traceability of individual copies will dissuade subscribers from widely circulating copyrighted content, which accordingly extends content security beyond the CAS/DRM of the set-top box.

In the end, forensic marking contributes to the protection of the operator's revenue and market share by limiting local piracy on domestic markets and by having the ability to shut down illicit TV signal redistribution.

Impact for subscribers

Watermarking introduces a security feature beyond the STB or PC client. The CAS or DRM gives access to content as soon as the subscriber has paid for it. From there, it is difficult to control how the subscriber will utilize the content. Operators and rights owners need to meet the expectations of the viewers such as having the ability to create a private copy. The serialization of the content based on watermarking will help to dissuade subscribers from widely redistributing the content.

From a consumer perspective, watermarking delivers undeniable benefits because it is not impacting the viewer experience and the content can still be played on any device. Accordingly, it is possible for consumers to transfer the content they acquired to the various devices they use. Watermarking can contribute to eliminating the issue of non-interoperability introduced by DRM and copy protection techniques, which compel consumers to look for content from illicit sources instead of via on-line content distribution and VoD services. Such flexibility also reduces customer covers live broadcast, replay from PVR, VoD and push-

FROM CAMERA TO THE HOME MANAGING ASPECT RATIO THROUGH THE PRODUCTION AND DISTRIBUTION PROCESS

Larry Thaler, Vice President of Distribution Technology
NBC Universal Inc, New York, NY

INTRODUCTION

As the percentage of high definition televisions in the home continues to grow, networks face increasing pressure to tailor their programming for both the SD and HD audiences. This trend has challenged broadcasters to find innovative ways of producing and distributing their content.

High on this list of challenges is aspect ratio. Programming will continue to originate and will be distributed in both SD and HD formats for years to come. Broadcasters must have a strategy that maximizes production and distribution efficiencies while maintaining the presentation quality for each home viewer, even though their screens may differ greatly.

Adding further complexity, the DTV switchover deadline is looming. What can broadcasters do to ensure consistent delivery of their programming in all markets after February 17, 2009? How could the changing distribution path affect the way programming is displayed at the home? What decisions should you be making now?

This paper will describe steps that NBC Universal has taken to prepare for this transition and makes recommendations that stations can use as the transition approaches. It highlights techniques available to every broadcaster and Production Company that can ensure all programming is optimized for both the SD and HD viewer.

"FRAMING" THE CHALLENGE

At first glance, aspect ratio challenges presented by the HD transition can be simple to understand. It's the age-old puzzle of putting a round peg into a square hole. Or, in this case, putting a 16:9 peg into a 4:3 hole (& vice versa).

When formatting an HD program for the SD audience, the choices are normally Center-Cut (and produced Center-Cut safe) or Letterboxed (giving up roughly 1/3 of the SD screen). Alternatively, when formatting an SD program for the HD audience, the material is normally pillar-boxed (which can sometimes result in home viewers stretching or zooming the content to fill the screen).

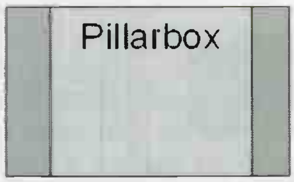


Figure 1: 4:3 Image in 16:9 Frame

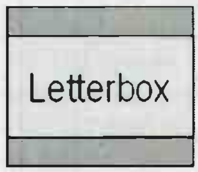


Figure 2: 16:9 Image in 4:3 Frame

There are other options such as shooting material in a compromise 14:9; however building a library in 14:9 may limit its value after the digital transition. Alternatively, material may also be produced and edited twice (in each HD and SD), but production costs make this approach prohibitive.

VARYING PRODUCTION REQUIREMENTS

There are a multitude of options, which may result in an organization attempting to simplify the process by either Center-Cutting everything, or Letterboxing everything. For NBC Universal, and at many other broadcasters, this choice is just not tenable.

NBC Universal's "Entertainment" community like many others express a strong preference for producing all content in 16:9 and down converting the SD in a letter-box format. They argue convincingly that this enables the creative community to tell their story using the entire screen and protects the production investment for the HD future.

Unlike entertainment, NBC Universal's News and Sports divisions expressed a strong desire to produce their programming "Center-Cut safe". Sports communicates concern about the size of the ball in letterbox while News put a high priority on compatibility with 4:3 archived material and format consistency from hard to reach remote productions.

Advertisers each have their own unique opinion on the issues. Regardless of the current preferences, all of this is sure to change as the HD audience increases as well.

DYNAMIC ASPECT RATIO CONTROL

At NBC Universal, it was clear that an adaptive solution was needed. One of the tenants of GE's vaunted Six Sigma design process is to give the decision makers the tools to act on their decisions. Doing this eliminates process steps downstream, avoids mistakes and opportunities for error. A high priority for NBC is to give the program producer creative control as far upstream in the process as possible.

In February of 2005, NBC began a project to convert Saturday Night Live's facilities to High Definition with the following primary requirements:

1. To produce the show in "real HD" (i.e. full 16:9, not center-cut safe)
2. To present the show in letterbox for SD viewers.
3. To avoid increasing number of control rooms needed.
4. To share production facilities with NBC News and Sports.

NBC needed a way to permit a show producer to identify the format of the content upstream, and provide instructions to the down-converter further downstream.

The answer came in the form of AFD (Aspect Format Descriptor). NBC proposed to our vendors that we carry an AFD flag in the VANC (Vertical Ancillary Data Space) of our HD video signals in our upstream production equipment. Downstream down-converters would interpret this flag and automatically switch between 16:9 and 4:3 segments in real-time. This would produce an optimal viewing experience to both SD and HD viewers with one set of production facilities.

AFD was inserted in the 8H control room's embedder, which is used to combine the audio and video signals after the production switcher and audio console. NBC standardized on two main flags ("AFD Full 16:9" and "AFD 16:9 with 4:3 center") to identify material that should be letterboxed or center cut on down conversion.

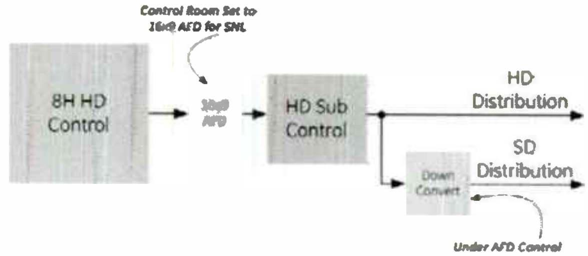


Figure 3: AFD in 8H Control Room

Thanks to support from many vendors, Saturday Night Live went on air with their first HD broadcast on October 1st 2005 with 16:9 programming and 4:3 commercials. The SD viewers saw a perfect letterbox show accompanied by full-frame commercials.

All of NBC's programming has been using AFD technology to produce the proper down-converted aspect ratio since the fall season in 2006. Our audience has been enjoying shows such as Heroes, Scrubs, SNL, Late Night, and Today and Nightly News with Brian Williams with the SD version properly formatted by the production requirements.

In 2007 SMPTE officially adopted an AFD production standard (SMPTE 2016) expanding vendor support for AFD and ensuring interoperability. The ATSC has also included AFD within its transport stream specification (ATSC A/53).

NETWORK ORIGINATION

The NBC Network currently maintains parallel SD and HD distribution from its Genesis Broadcast Operations Center (BOC) at 30 Rock. All programming and commercial playback feeds through BOC to local affiliates for 4 time zones.

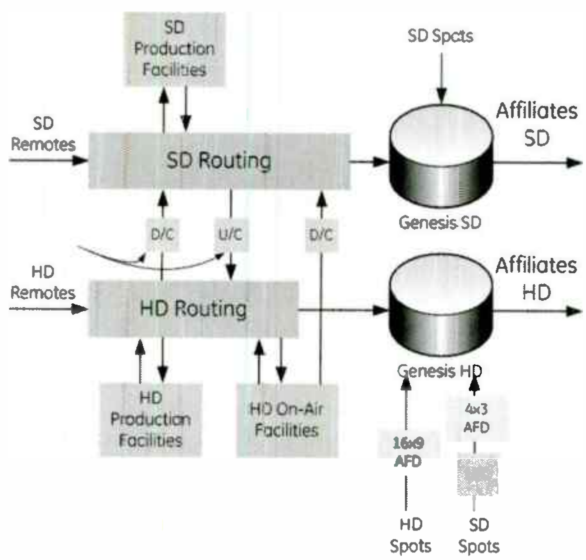


Figure 4: NBC BOC AFD Signal Path

AFD is inserted at this facility when taped material is dubbed into the video servers. With end-to-end support for AFD within Genesis, NBC automatically controls how HD originated content is presented to our SD audience.

Affiliated Stations receive NBC's programming from its Skypath™ satellite distribution system. Both SD and HD feeds are uplinked from "Master" earth stations located in NBC's New York and Burbank facilities. Just as the current Skypath™ system was originally architected to serve NBC's NTSC viewers, a major upgrade will be deployed prior to 2009 Digital Transition.

NBC's new Skypath HD™ system will be the next step in the natural evolution to a fully High Definition network. Once Skypath HD™ is deployed, NBC programming will be distributed to its broadcast affiliates exclusively in HD, simplifying the distribution process.

What makes this all possible is NBC's aforementioned HD/SD infrastructure that has seamlessly incorporated AFD. A local down-converter will generate the main NBC SD feed for its local stations – presented at the proper aspect ratio under dynamic AFD control. With almost 3 years of experience, we have faith that these devices can reliably handle all aspects of the down-conversion accurately.

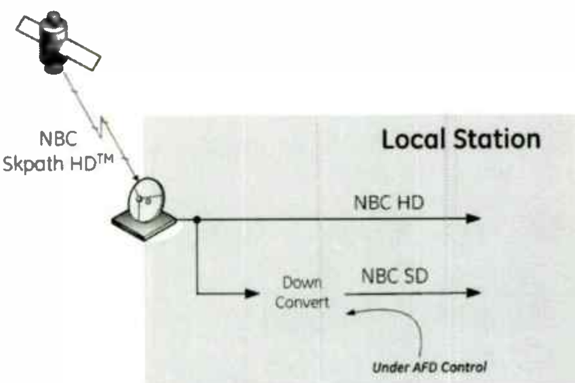


Figure 5: HD Distribution with Down-Conversion

TV STATION DISTRIBUTION PATH

While NBC's new Skypath HD™ system will continue to provide network versions in both HD and SD, many local stations will not have a direct need for the SD signal. Once NBC's signal is fed primarily in HD, the local stations' signal path will also evolve as shown below in the two figures below.

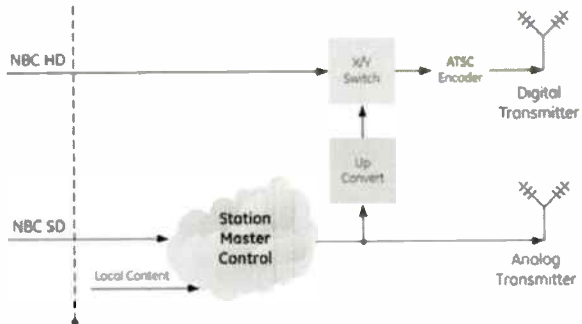


Figure 6: Common Station Signal Path (Today)

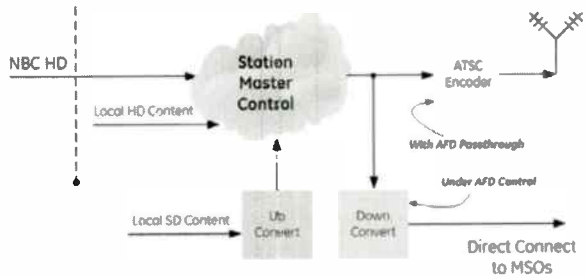


Figure 7: Common Station Signal Path (Post 2009)

Most NBC Stations are currently broadcasting HD programming and NBC Weather Plus on their DTV channels. With increasing pressures to utilize their DTV channel bandwidth for additional programming, mobile DTV, or other data services, NBC SD programming will not be available off-air in most markets.

Considering a large majority of homes on February 17, 2009 will still own standard definition 4:3 sets, NBC is has outlined the following suggested strategy to ensure consistent delivery of programming to its SD viewers.

1. *Direct Fiber Connection.* Where feasible provide direct connection of the local station SD signal to cable head-ends, satellite providers and Telco's.
2. *Propagate End-to-End AFD Support.* Work towards industry adoption of AFD support throughout the distribution path to the home. Ensure ATSC encoder is equipped with update where available to support AFD.

CABLE HEAD-END ARCHITECTURES

While direct fiber connection is the best way to control SD program delivery for cable and satellite viewers, there are large regions of the country where this is not practical. Analyzing typical cable-head architectures will help to illustrate the scenario.

For cable head-ends that redistribute off-air broadcast signals, the diagram below illustrates a typical setup today.

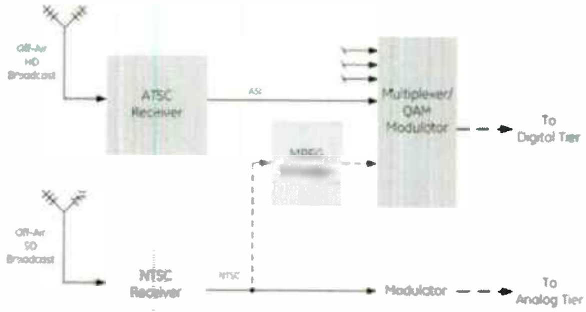


Figure 8: Common Cable-Head Path (Today)

The diagram illustrates the typical parallel redistribution path. Off Air HD Signals are demodulated to ASI and multiplexed into the cable system's digital tier. Off Air NTSC signals received and re-modulated into the cable systems analog tier.

Post February 17, 2009, many cable systems will continue to provide broadcast network signals on their analog tier (or digital SD channels), but will no longer have off-air NTSC signals available to do so.

The best solution is for a new ATSC Receiver device with the ability to down-convert and generate properly formatted SD signals as shown in the diagram below.

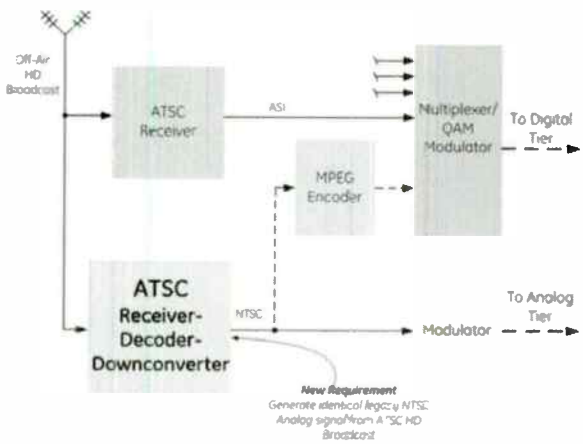


Figure 9: Common Cable-Head Path (Post 2009)

When specifying such a device, it is important to consider aspect ratio control and other elements to ensure that the NTSC signal is properly formatted including the following.

1. Down-conversion under control of AFD
2. EIA-608 Extraction and Conversion of Closed Captions
3. Conversion of XDS Data including V-Chip and Nielsen Ratings
4. Proper Dolby 5.1 Down-mix to LT/RT
5. Proper Choice of Dolby Digital Dynamic Range Profile
6. Carriage of Second Audio Program (SAP)

It is estimated that thousands of these devices will be required in the build out leading up to February 2009. This presents a huge opportunity for ATSC receiver manufacturers that are anxious to meet the industry's needs. It is up to the TV stations to clearly identify these needs when discussing this with their cable operators.

OFF-AIR VIEWERS

It is clear that there will be viewers watching transmitted signals directly over the air. These viewers will not benefit from the down-converters at the television stations or in the cable head-ends. If these viewers have ATSC tuners built into the television sets, they are most likely watching in 16:9 and will not have an aspect ratio issue. Some of the set-top boxes available have the capability to support AFD. As an industry we work towards having AFD support in the government sponsored set top boxes to help ease customers through the Digital Transition.

CONCLUSION

NBC has shown that aspect ratio can be consistently formatted and managed from the content provider, through the program chain, out to affiliate stations and to cable head ends on a day to day basis.

In the short term, we must plan our strategy so that we don't adversely affect the majority of our audience who own SD televisions sets. However, it is clear that over the long-term, more and more of the audience will be upgrading from 4:3 televisions to 16:9 widescreens. Clearly, the ultimate goal will be to maximize the HD viewing experience. This shift of perspective represents an inflection point that we need to prepare for now.

As broadcasters, we have a responsibility to smoothen the DTV transition for our audiences and to create a strategy to handle this inflection point. Full support of AFD through the chain will provide the tools needed to make this choice on our schedule, and on a station-by-station basis. Let's make sure we're ready by laying the groundwork today.

ACKNOWLEDGEMENTS

The presenter wishes to thank Clarence Hau for his vision, project leadership and many contributions to this paper. Also, thanks to Ron Gnidziejko as the first to identify AFD as a potential solution, Jim Starzynski and Glenn Reitmeier for their expertise effort and determination, the entire NBC Media Technology engineering team for their creativity and diligence, Miranda Technologies and our many other vendors for their willingness to push for interoperable solutions.

Thanks also to SMPTE and ATSC for having the foresight to adopt the AFD standard, MSTV for providing an open forum for discussion and to NAB for helping to spread the word.

BRIEF BIOGRAPHY OF AUTHOR

Larry Thaler is Vice President of Distribution Technology at NBC-Universal. During his 25 years with the NBC network, he has led key efforts, which have been at the junction of where traditional broadcasting meets new media. Most recently, he and his team have been responsible for the major overhaul of NBC's 30 Rock headquarters into a state of the art HD production and distribution center, the build-out of the new Cable Network Operations Center in Englewood Cliffs, NJ and the integration of new infrastructure for NBC's new Digital Media delivery platforms.

Larry received a BFA from NYU majoring in Broadcasting with a minor in Computer Science. He has spoken at several industry forums including NAB and SMPTE and is an Emmy award recipient.

THE STATE OF BROADCAST AUTOMATION

Sid Guel
Broadcast Automation Consulting
San Antonio, Texas

INTRODUCTION

In a recent research study performed by "Broadcast Automation Consulting", master control playout automation is surprisingly alive and well considering how many companies exist in the broadcast industry. Latest estimates show 50 broadcast automation companies worldwide. Out of an estimated 50 broadcast automation companies worldwide, 35 are in the United States, and 15 are from the rest of the world.

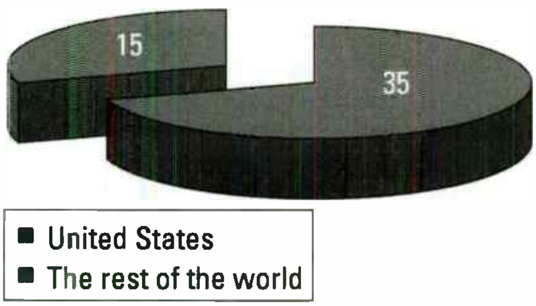


Figure 1. Worldwide Distribution of Broadcast Automation Companies. Graphic provided by Broadcast Engineering magazine.

STATISTICS

When *Broadcast Automation Consulting* conducted research on the age and longevity of automation companies, the results were surprising. We expected to find more new companies, and fewer long-term, veteran companies. This was not the case, in fact, the opposite was true. The majority of companies have been in the business for ten or more years. These veteran companies have been successful in supporting and maintaining customers and keeping their automation technology current to stay competitive and desirable. Plus, some customers are loathe to change and will keep a system well beyond the average five year life span. Many companies have expanded their portfolio of products, adding functionality to sell into other areas of the broadcast facility.

A 2007 study done by the European-based International Association of Broadcasting Manufacturers (IABM) estimated the value of the global broadcast industry for broadcast manufacturers and suppliers at $11 billion. According to IABM, the industry continues to grow at 11% per year with Europe and the Americas keeping the pace. Asia's growth continues to improve, but its gross revenue numbers are much lower. The fastest growth woldwide is the Library Management segment. The value of the key segments is below:

- Automation Segment (4%)
  $430.9 million
- Library Management Segment (1%)
  $106.5 million
- Storage Segment (14%)
  $1,570 million

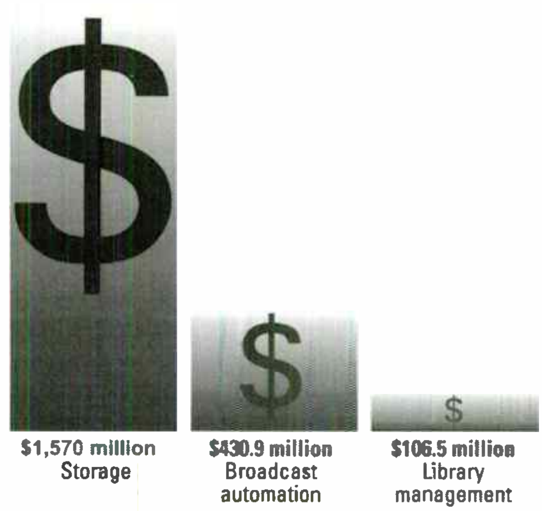
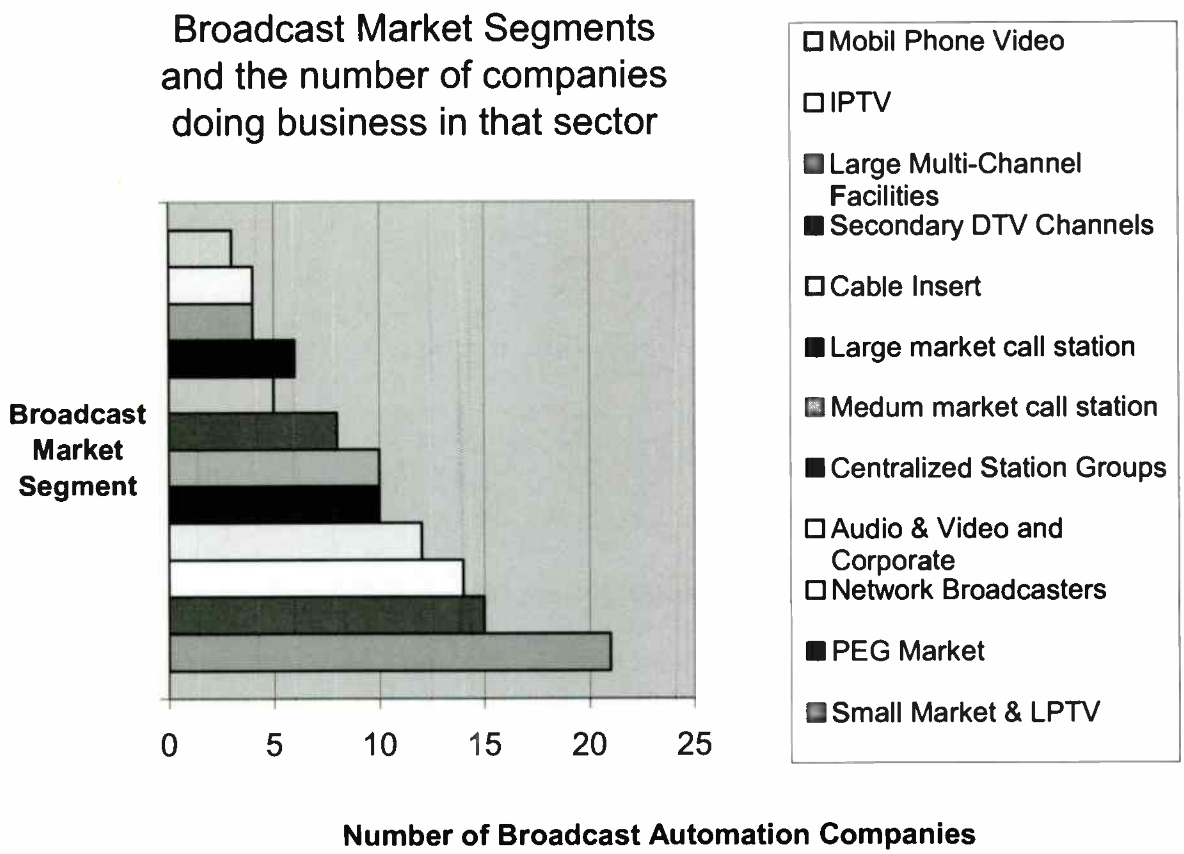


Figure 2. Estimated Value of Broadcast Industry Segments (IABM). Graphic provided by Broadcast Engineering magazine.

Some automation companies now have products that span across more than one segment of the broadcast industry. These overlapping products broaden the scope of offerings allowing automation companies better business stability and security. Some examples of these niche market and new product lines include: News Automation, Media Asset Management, Archive Management, Satellite Antenna and Receiver control, automated ingest, and centralcasting.

SEGMENTING THE MARKET

Broadcast market segments targeted by automation companies range from large centralcasting operations for groups of stations and multi-channel systems, down to Low Power TV (LPTV), PEG channels, and corporate. Few companies are able to offer products across all the market segments, creating niche and specialization opportunities. Even though there are more than fifty broadcast automation companies, most do not compete directly. Each company sells primarily into their own specialized niche of the broadcast market.



TYPES OF AUTOMATION

Technical design, software architecture and platforms vary between the numerous broadcast automation companies. The trend, however, is clear. Certain technologies are now becoming more popular than others. There are three primary types of broadcast automation systems.

- Standard device control.

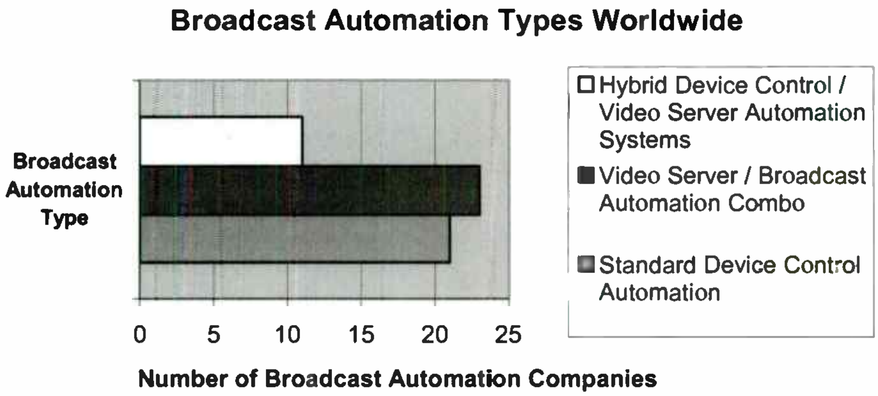
This is the traditional legacy form of automation in which a computer system running advanced software controls a variaty of third-party hardware and can import playlist and export as-run log files to the traffic system.

- Video server and broadcast automation combo.

This is traditional broadcast automation software which that runs internally in a video server.

- Hybrid device control and a video server / broadcast automation combo.

A hybrid system is a combination of all of the other plus somthing extra. Built-in, software only, third-party devices.



TYPES OF AUTOMATION - DEFINED

The trend is moving towards Hybrid systems and a Video server / broadcast automation combo system. These systems have built-in video servers and control more third-party devices in the form of software plug-ins, rather than controlling external hardware boxes.

VIDEO SERVER AND BROADCAST AUTOMATION COMBOS - DEFINED

For years, video server manufacturing companies have joined forces with broadcast automation companies. These partnerships are usually a win-win situation. Each has a better opportunity of selling products, regardless of who was contacted first.

Today, most broadcast automation players now offer video server and broadcast automation combo systems. They are basically broadcast automation software running on video servers with advanced interfaces and device control for third-party equipment. These systems are usually made up of off-the-shelf broadcast-quality video cards and off-the-shelf disk drives for storage. High-end systems use industry video cards, transcoders and RAID sets. There have been a select few broadcast automation companies that have partnered with video server manufacturing companies to include their software in a partners video server. The trend, however, is for the broadcast automaiton companies to develop their own video server system and makert these as one. It's a trend that's here to stay. What's the attraction to these systems? Cost-savings. You get a video server and a broadcast automation system all in one box, a tightly integrated solution, at a lower cost-effective price.

In the U.S., all public spectrum broadcasters are required to switch to DTV within the next two years. Broadcast automation companies can expect good sales in Video server / broadcast automation combo systems, Hybrid device control and video server / broadcast automation systems, and entry-level standard device control broadcast automation systems.

HYBRID DEVICE CONTROL AND VIDEO SERVER / BROADCAST AUTOMATION - DEFINED

The Hybrid automation companies are selling systems with integrated broadcast functionality such as graphics, encoders and decoders, storage, even editing capability. It Gives broadcast automation a smaller footprint because it's software centric and multi-functional. It used to be that manufacturers created boxes and buttons that served a single narrow purpose with the hardware and embedded proprietary software. Now the tables have turned. Harnessing the processing powers of advanced computer chips, automation companies are including this functionality as software, bypassing the traditional boxes and buttons approach.

A single computer can now handle source switching, advanced graphics, playout, video & audio mixing and digital effects. Some broadcast automation companies have mixed their systems with a variety of graphics hardware system capabilities. One broadcast automation company claims it has DTV video muxing capabilities within their automation system. Imagine a video server and broadcast automation combo system with built-in DTV spectrum control mux capabilities. It's a clear trend. The advancement of computer processing power gives broadcast automation companies more capabilities and allow them to offer more options and features with their respective automation systems. This trend will continue.

MS WINDOWS DOMINATES

Microsoft Windows is the most popular operating system on which automation products are based Figure 3 shows the types of operating systems used by broadcast automation companies. The lower cost of Linux-based systems is becoming more attractive to companies as a cost-effective alternative to Windows-based systems.

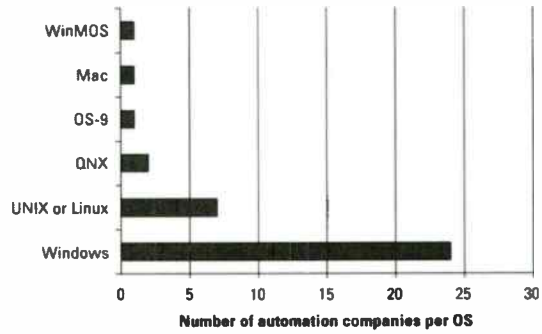


Figure 3. OS used in automation products. Graphic provided by Broadcast Engineering magazine

NEW TECHNOLOGY RISKS

Some industry experts question the reliability of Video Server / Broadcast Automation Combo systems, and Hybrid Broadcast Automation systems. There are many risks with having so much master control functionality in a single system. Industry watchers recommend more built-in redundancy, should be standard in all systems. Like any product, the buyer ultimately decides the level of redundancy they can afford and the level of risk they are comfortable with.

BRANDING AUTOMATION

Graphics and character generator companies are offering their own version of broadcast automation with integrated control of on-screen branding. These systems integrate a graphics playout server, animations, live video, video clips, audio, real-time external data feeds and master control automation functionality. Normally, these systems are serial or API controlled by broadcast automation systems as a third-party device. Today, they are being used as standalone broadcast automation systems for controlling playout on DTV channels.

THE BXF STANDARD

The SMPTE (S22-10) standard now known as Broadcast eXchange Format (BXF), is quickly being accepted by the broadcast automation community for the proper transfer of schedules and as-run logs between traffic and broadcast automation. For years, broadcast automation and traffic companies had to create conversion applications or traffic features to convert schedules and as-run logs for each of the different traffic systems customers were using. The multitudes of proprietary interfaces were difficult to keep up with, especially as traffic upgrades and new features were added requiring another rewrite of the traffic conversion.

BXF is supposed to solve that problem with a common protocol and messaging to exchange data with traffic systems. Not all broadcast automation companies are BXF-compliant, but many vowed at NAB2007 that they would be compatible this year.

After nearly five years of work, will the BXF standard truly promote a single standard way to share metadata, or will it be like many other so-called standards. Which end up in a hundred different implementations that are not compatible – like MPEG2!

Some advertising agencies have become involved in the BXF standard. By involving the advertising agencies, interstitials can have unique identification codes that stay with the metadata throughout the entire workflow – from creation to playout to reconciliation and finally affidavit and billing.

MEDIA ASSET MANAGEMENT

In recent years, media asset management systems have become the center of every media enterprise. From high-end systems for broadcast groups to small systems media asset management systems are in demand. These systems help broadcasters manage the ever-increasing amount of content that broadcaster manage as they play out more channels.

To be successful, a media asset management system should come complete with built-in transcoding features. Transcoding needs to be as transparent and cost-effective as possible. These systems should also include a built-in source-to-destination intelligence management systems to automatically transcode various video formats with ease.

Media asset management companies with more self developed products to offer the broadcasters are usually the best. It's easier for the broadcaster to work with one vendor, rather than several, when issues arise.

There's money to be made by redistributing rich media assets, whether using additional DTV channels, the Internet, IPTV or cell phone video. For the broadcast automation industry, the key is to cater to the different departments in a broadcast facility individually. Once you've built a system for one department, other departments can be added to an organization-wide media asset management system. Centralizing the management of the content assets in a single repository including the resources, networks, and storage is a tantalizing proposition. The infrastructure already exists. You're simply scaling and adding more capabilities to an existing asset management system.

A NEW BUSINESS MODEL - VIRTUAL MASTER CONTROL UPLINK SERVICE PROVIDERS

Hybrid and Combo broadcast automation companies have spurred a new business model. New global broadband service providers manage the broadcasting aspects of distribution and delivery of all forms of digital content. These companies enable content creators and distributors to extend their digital media presence through physical and virtual production, broadcast and infrastructure facilities.

Hybrid and Combo broadcast automation companies provide virtual master control systems without the usual compliment of broadcast hardware found in traditional uplink facilities. These new service providers are targeting media and entertainment organizations worldwide.

EXPANDING SYSTEMS

Some automation companies are expanding their software to integrate both the business and automation side of a broadcast facility. Broadcast automation companies are now providing programming, sales, traffic, master control, asset management and billing applications. Most traffic/automation companies in the United States have separate software systems and databases. Closed loop, tightly integrated traffic interfaces using BXF or based on a common platform is used for bi-directional communication between these applications.

Some automation companies are expanding their products to integrate the engineering side or the news and production side of the broadcast facility. This trend will continue as companies expand their product offerings. Automation companies look to expand because doing so supports and strengthens their business. Broadcasters like it because they get integrated products with centralized databases and automated streamlined workflows.

IPTV SALES OPPORTUNITIES

Broadcast automation players see growth opportunities in the IPTV world as Telcos roll out their own broadband distribution broadcasting services. IPTV may already have solutions for VOD services and pass-through channels, but many Telcos are developing their own network channels. Some of the issues IPTV is trying to solve are problems that have already been solved in the broadcast world through broadcast automation. Even if IPTV is being driven by the IT/computer world, broadcast automation companies existing solutions to offer.

THIRD-PARTY VENDORS

New third-party vendors are developing software plug-ins and add-ons for new business model automation systems. New plug-ins are being developed for services such as content verification, error tracking, fail-over tracking, compliance recording and also legacy third-party hardware products, such as routers, switchers and branding. Since these are software plug-ins, no additional hardware is required, keeping the cost low.

This change is a paradigm shift in how third-party vendors interact with broadcast automation systems. In the past, third-party vendors sold hardware systems that the automation company would control via Serial RS422 or RS232. Now, a third-party device is just a software plug-in. Many of the hybrid and combo broadcast automation companies sell their own third-party device plug-ins, but there is still plenty of room for third-party vendors to develop and sell plug-in options. Other areas could include branding features, switchers, routers and audio servers.

AUTOMATED WORKFLOW MANAGEMENT

There are various types of automation systems within a broadcast facility. Traffic, master control, production and news are examples of areas that use automation systems. Each of these products have advanced to a point that a clear set of business rules can be implemented and configured to create a logical workflow for the successful completion of a goal.

Automated workflow management is either the successful completion of a goal across various departments, or within an automation system. It includes three important levels:

- Day-to-day staff operations
- Operations monitoring, control and performance reporting
- Global reporting and statistics for upper management.

Adhering to and being compatible with SMPTE and industry standards will help tear down the walls these information and workflow silos. Common standards is key goal. Continuing control level software development will also help bridge the gap between various automation systems. With IT networking infrastructures replacing the video engineering in many broadcast facilities, automation companies will need to be more open to interfacing with other broadcast software systems at a horizontal or vertical level, depending on the workflow requirements.

AUTOMATION TOMORROW

In the future, there will be an automated asset management system in which the ingesting and archiving of digital media will be fully automated. Master control will no longer be staffed. It will be monitored by traffic, engineering and production, depending on the situation and needs at the time. Spots and programming will come in from content delivery services in digital form and auto-populate the playout video servers or ingest video server. Metadata with spot and program lengths, segments lengths, titles and identification numbers will be automatically imported into the traffic system.

Hybrid automation systems will be the virtual master control of the future. Instead of various hardware components, there will be software plug-ins. Third-party vendors will already have the software for their hardware systems. Manufacturers would just redesign their product lines without hardware or very little hardware. Instead of a router or switcher hardware tub, they'll provide software plug-ins for the automation system.

Looking further ahead, one day an operator will not be able to make a move unless the computer system gives them the OK. This built-in intelligence, added to monitoring and messaging systems, will decide the best form of action for every request based on pre-defined rules and the current environment. In the broadcast industry, there are already rules-based traffic systems for plotting the highest paid commercials in the log schedule. Broadcasters use these systems to maximize the profit potential of any given time slot. The difference will be who or what is in control. Instead of people controlling processes, these applications, through workflow inboxes, automated intel, etc., will soon be controlling people.

CONTROLLING THE ENTERPRISE

Metadata is the key to everything broadcast. It's required at every level and in every aspect of the broadcast facility. The metadata has to be present, available, accurate and transportable from system to system, department to department, and vendor to vendor. That's our future.

Traffic departments will be in control of the master control room because year after year, the goal is to reduce manpower and to automate processes more. Someday the traffic department will be in full control of everything master control. Since traffic controls the metadata, who better to control the final product? Granted, there will be times when a live person will be needed in master control. This will probably be a dual role with either an engineer or a production person.

THE BOTTOM LINE

It's never easy to predict the future, but there are clear technology trends. Hardware is miniaturizing, and eventually, it will go away completely. Expect more company mergers and acquisitions. Companies with low-cost solutions will do well in the next few years as the final small market level of TV stations switch to DTV.

Smart broadcast automation companies are staying ahead of the curve and developing the next generation of automation products for new business models, such as virtual master control, IPTV, mobile phone, the Internet and DTV business. Expect to hear more about SOA (Service Oriented Architecture) methodology and monitoring and control systems.

Companies are riding the last of the "low hanging fruit" automation business and should soon move toward adding alternative business models. Expanding your product line into new areas in a broadcast facility and working in new business model sectors is critical for increasing market share and staying in business after the big switch to DTV is over.

CREDITS

Along with the support of my friends and family, here are the individuals and companies that assisted in making this report a success.

- Broadcast Automation Consulting
- Broadcast Engineering Magazine –
- IABM – International Association of Broadcast Manufacturers
- John Price – Automation veteran of Odetics, DTG and most recently VCI Automation
- Steve Gottlieb – Formerly with Encoda Systems (Now owned by Harris Corporation)
- Mike Stoker – BIG Picture Concepts
- Carl Amend – TM Television
- Eric Johnson – Redwood Marketing

Audio over IP

Tuesday, April 15, 2008
1:00 PM – 2:30 PM

Chairperson: Talmage Ball
Bonneville International, Salt Lake City, UT

*IP based Audio and Control Distribution over Internet, Satellite, and Wireless Platforms

Johannes Rietschel, Barix AG, Zurich, Switzerland

Rapid Radio Deployment Pack Emergency Edition

Pierre Robidoux, Canadian Broadcasting Corporation / Société Radio-Canada, Montreal, Quebec, Canada

Advanced Tech for IP Remotes

Steve Church, Telos Systems, Cleveland, OH

*Paper not available at the time of publication

Rapid Radio Deployment Pack *Emergency Edition*

PIERRE ROBIDOUX
Canadian Broadcasting Corporation / Société Radio-Canada
Montréal, Québec, Canada

ABSTRACT

This paper describes the RRDP *EE* concept, a portable multiple inputs/outputs IP based modular radio studio including satellite news alert feeds & remote control.
As we all experienced, installing a remote production studio may require a lot of resources and could take quite a long time, which is not suitable in many circumstances. We need a portable studio than can be installed rapidly for special events and emergencies when a phone or an IP codec for a single reporter is not enough and when an OB van is not suitable.
Furthermore, if we have to evacuate the head office, the main network production facility, we need to maintain operations from a remote location; continue basic network programming, continue basic local programming and interrupt all regional stations programming for special news bulletins, when needed.

"DJ" type of Rugged Road Rack Case
which may be used for RRDP *EE*

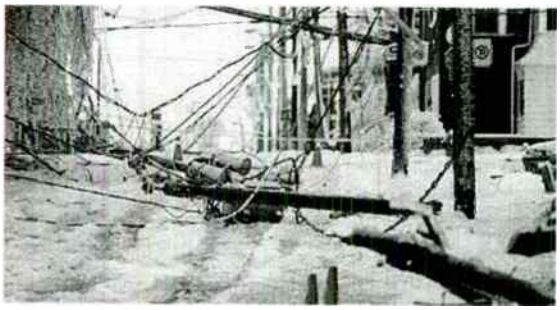
INTRODUCTION

A RRDP *EE* system would have to be very portable, using small vans and/or station wagons for transport. It would also have to be rapidly deplorable; within a few hours, not in several days.
User-friendly and affordability should also be mandatory requirements.

This article presents an overview of the main elements of the project and a brief look at the Internet Protocol, which are:

- **IP Production**, base on protocols such as Livewire™ or Ethersound™ .
- **IP Transport** by audio IP codecs using algorithms such as AAC LD™ or HE apt-X™.
- **IP Remote Control** on UDP/IP and TCP/IP
- **The Internet Protocol (IP)**.

Real-Life Example of Emergency Situation



The Canadian Ice Storm of 1998

After more than 80 hours of freezing rain, in January 1998, 57 communities in Ontario and 200 in Quebec where declared disaster area.
Over 4 million people in the provinces of Ontario, Quebec and New Brunswick lost electrical power.
About 600,000 people had to leave their homes.
28 people died, several of them from hypothermia.
945 people were injured.
About 100,000 people had to go into shelters.
Three weeks after the beginning of the storm, there were still 700,000 people without electrical power.
Among many other material damages, 130 power transmission towers were destroyed and more than 30,000 utility poles fell down.
The total estimated cost of the ice storm was more than $5,400,000,000.
In this type of situation, the challenge for a radio broadcaster is to maintain the services regardless of the situation. A survey from Osaka informs us that :

> "More than 90 percent of community-based FM radio stations nationwide have strengthened preparations for disasters ... 79 percent of the respondents have provided emergency broadcasts during past typhoons and other disasters."
> January 14, 2008 - The Yomiuri Shimbun (Japan)

Are we ready for events of that magnitude ?

If we have to evacuate our main production site …
How can we maintain the service ?
RRDP *EE* may certainly be part of the solution.

IP PRODUCTION

At first sight, IP network packetized transport does not seem to be the easiest solution for audio. IP was not designed to transport real-time data. However. this technology is gaining an increasingly large market share in radio broadcasting, mainly for the following reasons:

- **Routing**: There is no centralized crosspoint audio router. Audio routing is done by standard computer switches.
- **Convergence**: The audio network uses the existing computer infrastructure as the desktop radio workstations, servers and office computers.
- **Cost**: Mainly because of the widespread availability of computer technologies, an IP system is cheaper to install, support and operate.
- **Flexibility**: Compared to a traditional fixed studio installation, a modular IP audio network can be configured, modified and upgraded much more easily.
- **Scalability**: In a multilevel topology, IP is very scalable. The bandwidth would restrict the size of an IP radio site, but Gigabit Ethernet has pushed this limit significantly.

The Time …

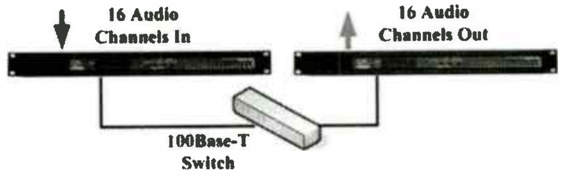
As an RRDP *EE* system would be isolated from the fixed production site, a GPS based time server, using the NTP protocol, will be necessary. In an IP based environment , it will be easy to distribute the signal.

GPS based NTP Time Server

If you use IBOC (HD Radio™) transmission do not forget to compensate for the 8 seconds delay.

A First Step

IP audio technologies offer a wide range of possibilities for audio projects, from a simple Ethernet PC sound card to a mobile production system to a complete multiple-studio radio broadcast site.
As a first example, a 16-pair audio cable is replaced by a single computer cable.



An IP audio project using Rave™ digital audio routers

The first audio IP unit (left) has 16 analog audio inputs. The other audio IP unit (right) has 16 analog outputs. This "Ethernet snake" can use a Cat-5 crossover cable, without a switch, at lengths up to 100 meters (328 feet). With a switch, this maximum doubles. A copper-to-fiber media converter can also be used to significantly extend this maximum length.
In the IP world, the console becomes a control board only. Mixing, processing and other audio functions, such as mix-minus, EQ and auxiliary feeds, are provided by a separate hardware unit called the "core" or "studio engine". The faders, buttons and other controls can be mapped into pages to serve different purposes (assigned to several functions).
One of the IP studio firms offers separate modular (from the studio engine) IP devices for audio signal input and output (analog and digital).

A complete modular (networked) IP-based radio studio

The audio workstation, the audio I/O modules and the control surface are connected to the switch at 100 Mbps. The mixing and processing module and the link to other studio(s) or the master control room are connected to the switch at 1000 Mbps.
The audio modules and audio workstations will be assigned two IP addresses, one for control, the other (a

Class D) for audio streaming. The virtual audio routing among the modules and the audio workstation is done by management and monitoring software.

Using Computers in a RRDP *EE* project

In a mixed client-server/peer-to-peer topology for an IP-based "mobile" radio project, PCs and may multiply very rapidly. An RRDP *EE* site would have DTR client and server software, several control software programs, some administrative software programs, and maybe automation software. Furthermore, clustering and/or redundancy would add on systems. For audio workstations, PCs would be kept in place, mainly for DSP (Digital Signal Processing) functions. For all other types of work, from office to system management, there are solutions for getting jobs done efficiently.

Thin clients and "Blade" server

Here are some items to consider for a reliable mission-critical system and, at the same time, for minimizing the total cost of ownership (TCO) of RRDP *EE* units

- **Blade Servers** are not the solution to every situation, but when there is a need for many servers in a small space, it is one option to consider. A blade enclosure (rack) provides power, cooling, networking, interconnections and system management. Each blade is a complete hot-swappable server (including a hard drive). Blade storage units, blade switches and other accessories are commercially available as well.
- An **Application Server** is a software service (layer) that runs on top of a Windows server and that delivers applications to client computers or thin clients. Microsoft™ uses the Remote Desktop Protocol (RDP) for its graphic terminal services. Citrix™ also offer an application delivery system that runs on a Windows server which uses ICA (Independent Computing Architecture) protocol.
- A **Thin Client** is a graphic terminal that works in a client/server architecture. The server processes all the activities, and the thin client displays them in a Windows environment. Its advantages are mainly lower IT administrative costs, lower hardware costs, the fact that it provides a more secure infrastructure, and easier upgrade capability. Most thin clients come with RDP and ICA protocols.

IP TRANSPORT

Analog leased lines, ISDN switched services, synchronous networks as V.35/X.21 and T1/E1 are coming to an end. Today more and more signal transport is done "over IP". In a near future it will come to an almost "IP only world".

IP availability in disaster situation

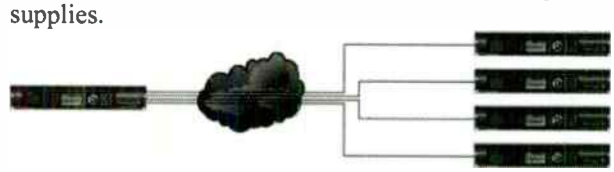
IP Transport is usually done using the corporation private MPLS network. If impossible to access, we may use a virtual IP network over satellite. IP feed over RF (STL) may also be used locally. There are also service providers like satellite BGAN or VSAT, Wideband 3G+ (1xEV-DO or UMTS), WiMAX and the public IP network (Internet).

Via IP audio codecs

Transport via IP service, is done using IP audio codecs.

Multiple-channels IP audio codec

A major installation would need several codecs on the same Ethernet link for 24/7/365 mission-critical reliable service. An IP audio multiplexer or multi-channel codec would be a solution. Some systems offer hot-swappable units, IP redundancy and dual power supplies.



Multiple unicasting

Unicasting describes a peer-to-peer link between two codecs. In a multicast session, a "source" codec has an audio feed available at a single multicast address (Class D address, from 224.0.0.0 to 239.255.255.255). Hosts (other codecs) may "join" in order to receive this audio feed. Because of the risk of network overload multicast is not always permitted and multicast addresses would be blocked. Multiple-unicast is used when a codec has to feed the same audio material to several destinations simultaneously. In this case only one of the addressed codecs can be configured to answer back.

Livewire™ Gateway Audio Codec

RRDP EE is using an IP based topology for production. If it would use the Livewire™ protocol, there is at least one codec firm hat offer a direct gateway (and normal codec analog & AES/EBU I/O) between the Livewire™ production network and the transport network (for distribution or collection) to avoid cascading.

The audio signal

The codecs will packetizise the audio signal (analog or digital) for transport over a regular IP network infrastructure. For the moment, because of the constraints of IP bandwidth availability (and cost), the signal has, most of the time, to be compressed. Linear audio would result in a bit rate of 1.5 Mb/s to 4.5 Mb/s. Beside Linear audio (uncompressed) a radio broadcaster would have to choose between the following coding algorithms: G.722, CELP™, MPEG 4 AAC™ , HE AAC v2™, AAC LD™, AAC ELD™, AEQ LD Extend™, ADPCM, apt-X™, HE apt-X™, BRIC™, MPEG 1/2 Layer 2/3, Music™, Voice™, and probably a many more … An other choice would also have to be made in regards to sample rates; from 8 kHz up to 96 kHz. Most firms offer multi-algorithms audio codecs and extensive settings combinations. The proper choice of a compression algorithm and it's setting will be made according to:

- The IP bandwidth and network service quality from one point to the other of the communication.
- The "acceptable" signal degradation caused by the coding (compression) process.
- The "acceptable" maximum total audio delay (latency).

In spite of some claims, there is not a "best" algorithm and setting to fit all needs, therefore several subjective tests and probably some objective tests would have to be made.
A very first step would be to listen to audio files that where compressed at different rates using different compression algorithms.

The cascading issue

To express the number of time when audio material was copied from one magnetic support to an other we used the term "generation". Generation "0" being the original. When digital audio signal pass through several encode/decode cycles the term generally used is cascading. Some coding schemes are more resilient to multiple cascades. The audio signal degradation will be more severe when several different coding algorithms are used on the same cascade.

Audio Evaluation and Measurements

An IP audio codec is a non linear audio device. Using a fixed bit rate (stream) "masking", bandwidth response and some other specifications will change (continuously) according to the input audio signal. Various encoding methods (compression schemes) remove both redundancy and perceptual irrelevancy in the audio signal so that the bit rate required to encode the signal is significantly reduced. These lossy compression algorithms take into account knowledge of human auditory perception, and typically achieve a reduced bit rate by ignoring audio information that is not likely to be heard by most listeners. The International Telecommunications Union (ITU) describes in detail a standard method for measuring the quality of wide bandwidth compressed audio (ITU Recommendation BS.1387-1). The method is the result of a joint effort among laboratories in Canada , The Netherlands, France, and Germany. The acronym for the measurement model is **PEAQ** (Perceptual Evaluation of Audio Quality).
According to several technical books and papers (from AES EBU, ITU, NAB and others) PEAQ is the only relevant measurement method for audio codecs (using compression).

Overhead and Total IP Data Rate

Packet size	Audio Data Rate	IP Data Rates
128 bytes	384 kb/s	582 kb/s
256 bytes	384 kb/s	483 kb/s
512 bytes	384 kb/s	434 kb/s

Total Audio Payload

The total IP data rate is composed of the audio data rate plus the overhead. The audio data rate is a function of the compression algorithm, the sampling rate and the mode. The overhead can be drastically increased by reducing IP packet size or by the addition of a FEC (Forward Error Correction).
Overhead is usually from 4% up to 25 % but it may also more than the double of the audio data rate.
This very partial table (below) presents the relation between packet size and the total IP data rate.

FEC

To avoid unacceptable audio breaks, these are two schemes may be used on "constrained network". To conceal the effects of lost packets processes as interpolation are implemented. FEC (Forward Error Correction) adds packets to the stream to enable the reconstruction of lost information. As an example FEC50 would add 50 % and FEC100 would add 100% to the audio payload.

SIP

The Session Initiation Protocol (SIP) is an application-layer control (signaling) protocol to create and terminate sessions with an other (remote) codec. This is not offer by all vendors. When SIP is not avaliable, the session has to be create and terminated at each end of the link.

NAT and STUN

To let a codec on a LAN with a private IP address communicate with a codec on the Internet, there needs to be a "translation scheme" between the private and public IP. NAT (Network Address Translation) is a technology that allow multiple devices (such as codecs) with private IP addresses to share a single public IP address on the Internet. STUN (Simple Traversal of UDP through NATs) is a protocol that can assist codecs behind a NAT firewall or router with their packet routing. STUN enables a codec to find out its public IP address and the type of NAT service its sitting behind.

IP Remote Control

Remote control would be an important component of an RRDP EE project.
As it would be the case for a permanent studio installation, a broadcaster would have remote control software for the production and the transport elements of the IP audio system.

IP audio codecs control software

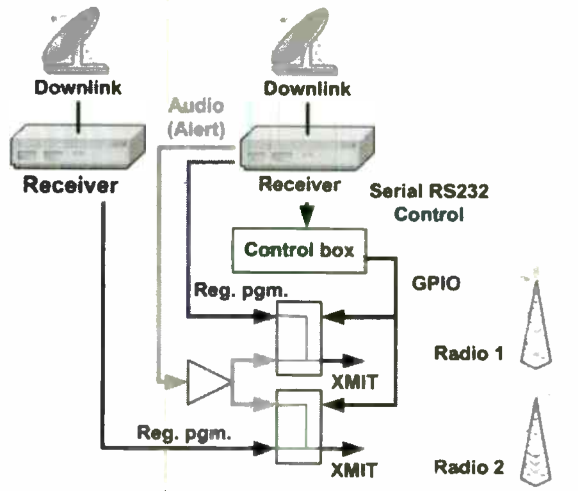
There are basically 4 type of control software for IP audio codecs and IP audio production devices:

- **Web pages** – Each unit having a small web server.
- **Telnet commands** – Useful for maintenance and control by an other software (as automation).
- Individual **Windows Applications** – Each unit is controlled by it's own software.
- **Centralized Control** application – To manage several units.

Theses software may include log files, alerts, and statistics such as codec status, IP throughput, network jitter and latency monitoring.

On the Transmitters side of things

To build a complete RRDP *EE* system, you would also need remote control over the distribution (to the transmitter) feed. The principle is quite simple.
The general idea is to be able to remote control an audio switcher enabling the interruption of the normal programming for an special newscast.
This audio switcher would be feed the regular programming (on default position) and (one or more) alert newscast feeds. Theses emergency feeds may come from an IP audio codec, a satellite link
The figure on the next column presents an example of a remote control (including the audio alert feed) via satellite.

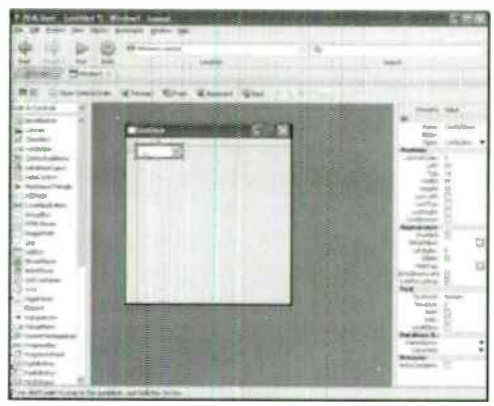


Radio-Canada's News Alert Remote System
(Used by the French Radio Network)

The 2 transmitters, Radio 1 and Radio 2, are receiving the network feeds via 2 satellites. One of the satellites also carries the Alert Audio Signal and the News Alert Control Signal.
Note: The transmission department may ensure the control over which transmitter (main, fixed backup or mobile backup) would be in service.

RAD - Rapid Application Development

To build such a project some inside (or outsourced) development (programming) would be needed.



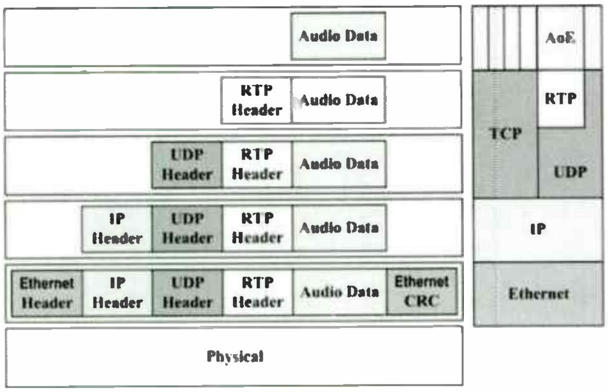
RAD Software for Windows, OSX & Linux

To develop this type of control software the application development software will generate a self-contained executable control application which will not be dependent, on .NET Framework, external DLL or JRE (Java Runtime Environment).

The Internet Protocol (IP)

In-depth computing expertise is not mandatory, but basic IP knowledge is essential for building and supporting the RRDP *EE* IP based system.

The Internet Protocol is layer-structured to provide efficient communication between different kinds of computers. Several versions of this layering model are documented. They have four or five layers, and different names are used for the layers. IP is based on the original DoD (U.S. Department of Defence) ARPANET layered model created in the 1970s. Today, IP is used, among many others, by the largest worldwide public computer network: the Internet.



5-layer concept for IP audio networking

LAYER 1 - Physical (Electrical Interface)
Hardware communication for PCs and IP devices is provided by an NIC (Network Interface Controller, or Network Interface Card). The standard for audio is the Ethernet, using 100Base-T (100 Mbps) and 1000Base-T (1000 Mbps) cables. These cables are made of eight conductors, four UTP (Unshielded Twisted Pairs), and RJ-45 8-pin plugs. Category 5e (enhanced Cat-5) cable may be acceptable, but Cat-6 cable is recommended. Mediatwist^{TM} Cat-6 is often suggested. Cable pinout should be TIA/EIA (Telecommunications Industry Association/Electronic Industries Alliance) T568 B-compliant, a standard commonly used for data applications. 1000BASE-T Gigabit Ethernet (GbE) is also known as the IEEE 802.3ab standard.
LAYER 2 - Ethernet or Data Link (or Switching)
Each NIC has a unique 48-bit ID assigned by manufacturers: the MAC (Media Access Control) address. The Layer 2 header contains this address, which is used for local (intrasite) communications. The MAC address has the following format: 00-11-24-EB-28-DG. The IEEE 802.3 group of standards defines the physical layer (1) and the MAC part of the data link layer (2) (wired Ethernet).
LAYER 3 - IP (Network or Routing)
The task of IP is to get packets of data from the source to the destination. The dotted decimal notation of the 32-bit IP address consists of four octets of eight bits separated by three dots. An example of an IP address is 192.15.121.10. The left-most bits indicate the network address. The others indicate the host (PC or IP device) address. The number of bits assigned to each is determined by the class. For a medium-sized organization, a class B will be needed, but a class C will suffice for a small organization.

Class	Address Range	Max. Hosts
A	1.0.0.0 to 126.0.0.0	167772142
B	128.1.0.0 to 191.254.0.0	65534
C	192.0.1.0 to 223.255.254.0	254
D	224.0.0.0 to 239.255.255.255	Multicast Groups
E	240.0.0.0 to 254.255.255.255	Experimental

IP Addresses classes

Between the Data Link and the network layers is the ARP (Address Resolution Protocol), a service for finding the correspondence between the MAC address and the IP address.
For audio network design, both IP and MAC addresses for each host may be required.
On a PC (using the command line), **ipconfig –all** will provide both the MAC and IP addresses of the PC.
To display the ARP table, enter **arp –a,** using the command line.
The command **ping** is used to test whether a specific system is visible on an IP network. Ping sends an ICMP (Internet Control Message Protocol) "echo request". Example: ping 1.126.23.12
LAYER 4 - UDP/RTP (Transport)
UDP (User Datagram Protocol) is mainly used for applications such as streaming audio and video, where on-time delivery is more important than reliability. UDP uses a "send and forget" strategy with no acknowledgment process, as is the case for TCP. There are four fields in a UDP header. One of them is the source port, another the destination port.
RTP (Real Time Protocol), a layer built on top of UDP, has a timestamp and sequence number fields in its header for synchronization and jitter processes. RTP was originally designed as a multicast protocol for delivering real-time media over an Ethernet network.
LAYER 5 – Application, AoE (Audio-over-Ethernet).
In this layer, we find protocols such as HTTP, FTP and SMTP. In AoE, the protocol used will be, for example, Livewire^{TM}, EtherSound^{TM} and audio streaming. These are real-time audio applications that include some form of control-and-identification dialog,

Version	IHL	ToS	Total Length
Identification		Flags	Frag. Offset
TTL	Protocol		Header Checksum
Source Address			
Destination Address			
IP Options			
Data			

14 fields of an IP packet for IPv4

The specific fields that require special attention for our IP based RRDP *EE* project are:

ToS (Type of Service): This is for assigning different levels of importance to datagrams.
Information on setting this field can be found in the section entitled "Networking" in this paper.
TTL (Time To Live): The counter, set at the source, which decrements at every "hop". If this field reaches the value of zero before the data arrives at its destination, the data will be discarded.
Protocol: Indicates the upper-layer protocol (UDP for audio over IP).
Source Address: The sending host address.
Destination Address: The receiving host address.
Data: Contains upper-layer information (from/to the Transport layers
The **DiffServ**, a Layer 3 ToS (Type of Service), uses one of the IP header fields in IPv4 (the ToS Byte). The IETF (Internet Engineering Task Force) agreed to reuse the ToS byte as the DS field for DiffServ networks. Therefore, DiffServ architecture now supersedes the ToS field.

DS5	DS4	DS3	DS2	DS1	DS0		
1	0	1	1	1	0		

The DiffServ byte

QoS (Quality of Service)

In a IP packet, the **DiffServ** byte, a Layer 3 ToS (Type of Service), uses one of the IP header fields in IPv4 (the ToS Byte). This ToS byte is new reuse as the DS field for DiffServ networks. The six most significant bits of the DiffServ field are called the DSCP (Differentiated Service CodePoint).
The QoS configuration should be set (101110) in Expedited Forwarding (EF PHB) mode to build a low-loss, low-latency, low-jitter assured bandwidth service. PHB stands for "Per-Hop Behaviour".

Four main PHBs classes (and DSCP names/ values)		
BE	Best Effort	Value = 0
CS	Class Selector	7 DSCP values
AF	Assured Fowarding	12 DSCP values
EF	Expedited Fowarding	Value = **46**

Per-Hop Behaviour classes

CONCLUSION

For this type of IP based mobile project, it is not advisable to adopt a "guess and deploy" attitude. Information gathering, discussions with users – mainly emergency procedure team - good planning, proper architecture, and subsequent training are paramount for ensuring a successful state-of-the-art high-tech IP emergency project. Radio maintenance technicians and engineers would have to be knowledgeable about IP networking, in addition to analog and digital audio, because of this new convergence of technologies. Computing (mainly networking) and IP audio manufacturers offer extensive practical documentation, good support and training. Audio-over-IP is no longer a futuristic idea; it is successfully being developed, in radio broadcasting, for fixed and mobile systems.

REFERENCES

1. A Network-Enable Radio Console Architecture, Michael Dosch, Axia Audio.
2. Audio Routing, 2007, Radio Magazine
3. Bridging Between IEEE 802.1Q VLANs, Cisco Systems.
4. Building EtherSound Networks, Digigram.
5. Configuring Cisco Catalyst Switches for Livewire™, Axia Audio.
6. DIFFSERV – The Scalable End-to-End QoS Model, Cisco Systems.
7. Ethernet for Studio Audio Systems, Steve Church, Telos.
8. Experiments with MediaTwist, Belden Web site.
9. Implementing Quality of Service Policies with DSCP, Cisco Systems.
10. Inside TCP/IP third edition, New Riders Publishing.
11. Inter-Switch Link and IEEE 802.1Q Frame Format. Cisco Systems.
    IP-Audio Networks in the Real World, Clark Novak, Axia Audio.
12. IP, QoS and IGMP. Technology Documentation, Cisco Systems.
13. Measuring Delay, Jitter, and Packet Loss with Cisco IOS SAA and RTTMON, Cisco.
14. Networked Audio, 2005, Wheatstone Corp.
15. Overview – An Introduction to the ES-100 Technology, Digigram.
16. Road Ready Cases (Catalog)
17. Quality of Experience for Media over IP, IneoQuest.
18. System Design Reference & Primer, 2004, Axia Audio.
19. Using Category 5/5e/6 for Audio and Video Applications, Belden.
20. Road Ready Cases (Catalog)

THE AUTHOR

Pierre Robidoux has been with CBC/Radio-Canada since 1980. He received a degree in electrical engineering in 1993. Pierre has worked in the VTR, Radio Operations and Radio Maintenance Departments, where he created, among other things, a system for controlling all the French Radio Network transmitters from the Montreal newsroom for the purpose of airing special newscasts. Pierre joined New Broadcast Technologies in 2007. He has worked, among other things, on the evaluation of IP audio codecs and video encoders for digital multimedia broadcasting. Pierre is a member of the AES and the SMPTE.

ADVANCED TECH FOR IP REMOTES

STEVE CHURCH
Telos Systems
Cleveland, OH USA

ABSTRACT

ISDN has served broadcasters well. Indeed, it was a small-scale revolution when it first appeared in the early nineties. For the first time, the dial-up network could be used for high-fidelity remotes. Compared to the equalized analog "broadcast loops" that were the only high-fidelity telephone service before, ISDN was miracle. While ISDN is still a perfectly good technology, it does have some drawbacks. The main one is that usage is billed by the minute. Another is that installation of the line at the remote side usually has a multi-week lead time and has a significant set-up charge. In some parts of the USA and in other countries, ISDN service is being discontinued or has become difficult to get. IP networks are becoming the new way to get broadcast audio to here from there. A broadcast codec taking advantage of new technology and optimized for the real-world conditions on IP networks makes this a practical reality.

IP TO THE RESCUE – BUT...

Years ago, Telcos envisioned ISDN as the way Internet would be delivered to the masses. But DSL now fills that role. Broadcasters remain one of the few users of ISDN basic rate service, but they don't provide enough business for Telcos to justify the expense of maintaining the infrastructure.

So we turn to the ubiquitous Internet and other IP networks for an alternative. At many remote sites, there is an existing IP connection than can be hijacked for an ad-hoc broadcast. High-speed IP links via mobile phone networks offer the chance to connect from almost anywhere within large cities. International connections are no problem. And the per-minute charge is gone.

But IP has its problems. On the Internet, there are no quality-of-service guarantees. That means that the packets may jitter, may be dropped, and may not provide a consistent bandwidth. To avoid audible problems, equipment for audio transmission over IP networks must be designed to cope with these conditions. While older equipment might work over networks that have good QoS, the latest generation IP codecs have been designed from the ground-up to deliver high-quality audio, even when the networks are not.

The first problem to be solved is the delay variation in packet arrival, known as jitter. On the Internet, this ranges from tens to hundreds of milliseconds. In contrast, ISDN has no significant jitter. A long buffer in the receiver, one set to the maximum expected jitter time, can be used to even the flow and deliver consistent audio. But this comes with a cost – audio is delayed. Since codecs at remotes are often used in two-way fashion, this delay can cause trouble for the talent, who find it difficult to speak naturally. And there is always the possibility that a packet will arrive later than the maximum buffer time. Thus, a fixed buffer has only limited utility. Much better would be an adaptive buffer that automatically contracts when the jitter is low, and expands when jitter increases. But this means that audio needs to be squeezed and stretched, not something so easily accomplished in real time.

Dropped packets are a normal condition on the Internet. It was designed so that any router node that becomes overloaded can deliberately drop some percentage of the packets. The TCP (Transmission Control Protocol) part of TCP/IP is intended to deal with this. When it senses that a packet has not reached its destination, it requests a re-transmission. It also lowers its rate of flow to adapt dynamically to network conditions. Retransmission recovers lost packets, but again, at a cost. The receiver must have a buffer long enough to cover the time of the detection and re-transmission procedure. Better would be a system that could deal with lost packets by inaudibly concealing them, rather than requiring re-transmission.

Finally, we must cope with the fact that there is no guarantee for bandwidth. If we set the codec bitrate to some high value to get good fidelity, we might find that over time the network can't provide enough bandwidth to support that rate, causing audio interruptions. So we might then decide to be conservative and set the codec bitrate to some low value to be more confident that there will be no drop-outs. But we now sacrifice audio quality. Wouldn't it be better to have a codec that automatically detects the bitrate that the network can support and then adjust to that rate? Even better would be if this adjustment could be ongoing and dynamic, adapting to network conditions. Until recently, no codec was able to "gearshift" inaudibly, but a new one has been invented which can.

An ideal IP codec system would have the following characteristics:

- Effective, inaudible packet loss concealment
- Adaptive receive buffer, with the necessary time squeeze/stretch capability

- ♦ An efficient codec, to achieve maximum audio fidelity from the least bitrate
- ♦ Adaptive codec bitrate, accommodating dynamically to network conditions

Fortunately, with the latest advances in codec technology, all of these are now possible.

AN IP-OPTIMIZED CODEC SYSTEM

A broadcast codec intended for IP application needs to be optimized for the purpose. An integrated system that pulls together a suite of appropriate pieces will be much more effective than codecs that have not been tuned and optimized specifically for the IP world.

The Codec Core

The MPEG AAC family codecs have always had quite good concealment techniques, but these had been optimized for the bit errors found on non-packet transmission paths. Recent work has expanded the concealment technology so that it can work effectively with packet loss as well. It's a clever technique. The codec keeps an ongoing measure of the spectral shape of the audio. This is easy because the codec already must have a time-to-frequency domain transform as part of its perceptual coding functions. When a packet loss is detected, a synthetic replacement is created by using the spectral values to filter white noise. To the ear, this sounds very much like the original. The amplitude is tuned at each end of the packet to match the preceding and subsequent packets so there is no audible "pop" from the splice. It turns out that this can be very effective, indeed. As much as 20% random packet loss can be inaudibly concealed.

MPEG AAC is an efficient codec, and the Spectral Band Replication (SBR) addition makes it the most efficient within the MPEG family. AAC with SBR is called officially AAC-HE (High Efficiency), but is also known as AAC+. The downside is it has quite long delay – around 150ms. That would mean 300ms for a round-trip, plus yet more for the IP packetization and buffering processes. Too much for interactive two-way conversation. AAC-LD comes to the rescue. It has around 50ms delay, so is much better on that count. But it has 30% less bit-efficiency than vanilla AAC. Since SBR adds approximately the same amount in efficiency, if we could combine that with AAC-LD, we would have a low delay codec with the coding power of plain AAC. And that is just what the new AAC-ELD (Enhanced Low Delay) codec does. It has reasonably good fidelity down to 24kbps and excellent fidelity when used at 64kbps and above. At 128kbps, it is regarded as indistinguishable from the original.

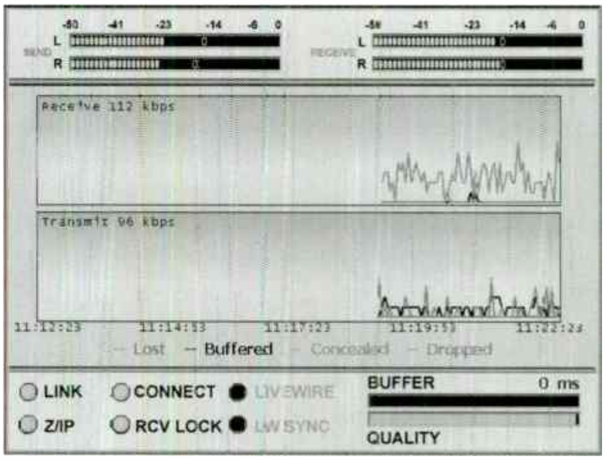
AAC-ELD's wide bitrate range is a good match to the needs of IP networks, since they vary so widely. A mobile phone connection might be limited to perhaps 40kbps, while dedicated links could be sized as desired, supporting codec rates of 256kbps, 384kbps, or more. An international Internet connection might support a 64-96kbps rate.

Adaptive Bitrate

An important feature of the AAC-ELD codec is that it can be made to "gearshift" its bitrate without making audible glitches. Coupled with a tuning algorithm, a broadcast codec can automatically adapt to the available network bandwidth. The algorithm constantly probes the network for the maximum rate that can be carried and sets the codec to this rate. When the network has high capacity, audio is as high-fidelity as possible. When the network has limited bandwidth, the codec adjusts to a low bitrate to ensure that audio gets through.

Adaptive Receive Buffer

Unless we have a guaranteed QoS network, it not possible to predict the jitter. Each packet is subjected to different network load conditions, which affects the transit time. In fact, each packet may take a different route. For uninterrupted audio, a buffer in the receiver must accommodate the longest delay the network presents. If a packet arrives outside of the buffer time, it's as good as lost.



Codec Status Screen Showing Network Performance
Packet Loss, Jitter, and Bandwidth Vary with Time

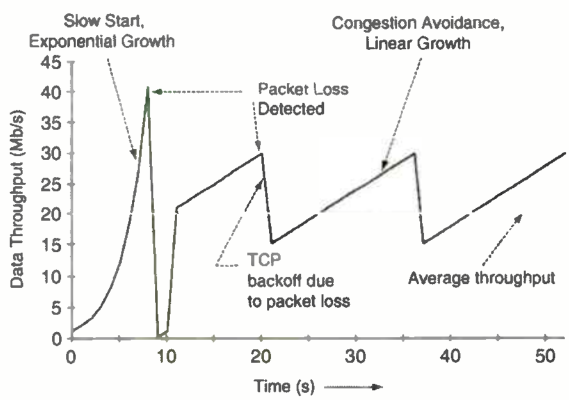
On the other hand, a long buffer translates to a long delay. So we want to optimize the buffer for the conditions that actually exist. And we want this to vary as needed to adapt to changing network conditions. But how do we detect the network condition? Recall that TCP adjusts its flow rate when packet loss is detected, so this is a long established way for attached equipment to respond to the network. TCP is constantly probing the network for the fastest supported speed by increasing the rate until loss is detected, then backing off. We can borrow exactly this idea for our receive buffer adjustment. We start a new connection with an average-length buffer. If packet loss is detected, we expand it. Unless there is an extreme case, the effect of the lost packets are not heard because the codec conceals them.

The adaptive algorithm is constantly pushing to reduce the buffer length. To minimize the number of concealments, the algorithm has a fast-attack/slow-release time characteristic – expanding the buffer quickly when a lost packet is detected, but allowing it to contract only slowly. Just as with TCP this feedback loop causes the buffer to automatically adjust to the optimum length. On networks with low-jitter connections, the buffer is automatically made small so as to minimize delay. But when the jitter is high, the buffer is made long to ensure that there are no audio drop-outs.

The buffer adjustment requires that time be squeezed or stretched, and this must be accomplished inaudibly. Fortunately, this is possible – as many audio editors with this feature demonstrate. We broadcasters are well familiar with profanity delay units that also have stretch/squeeze processing. This concept can be used also in the codec system application.

Transport

As we've seen, TCP solves the lost packet problem via retransmission, but this imposes a delay penalty since the buffer has to be long enough to accommodate the time it takes for the replacement packet to arrive. Streaming audio on the Internet usually uses TCP – and there are usually multiple-second-long buffers in the players. That's why it takes so long for audio to start after you click on the link or play button. When we care about delay, this is not going to do. TCP's flow control algorithms are also a potential problem, since they could needlessly restrict bandwidth.



TCP's Flow Control Causes Available Bandwidth to Vary

Forward Error Correction is another way to deal with packet loss. The principle is simple: both the original and some form of copy of the packets are sent on the network. If one is lost, hopefully the copy was not and the receiver can use it as a replacement. The structure of the original-copy sequence is organized to maximize the chances for successful recovery. You don't want to just put the original and a copy adjacent to each other, since that increases the odds that both will be lost. A minimum 2x2 FEC requires the buffering of four packets, while a more reliable 5x5 FEC would require a 25-packet buffer. The latter has more time spread, so is better able to cover losses. But now, unfortunately, we are back to significant delay – in the 5x5 case, as much as 600ms with the usual packet size. As well, FECs cause streams to take more bandwidth, and a network that is losing packets is one that is probably already near its limit, so adding to the bandwidth requirement is just as likely to create a problem as to solve one. There may be some cases where FECs make sense, but they are generally not useful for audio on the public Internet.

The alternative is User Datagram Protocol (UDP) combined with concealment. Using UDP, we are as close to the underlying network as we can be, so the delay is as low as possible and the bandwidth as high as possible. Responsibility for dealing with packet loss is moved to the "user", which is perfectly OK because we can deal with it in a way specialized for our audio application, rather than accepting the compromise of a general approach that was designed mostly for email, web browsing, and file transfers.

There are clear standards for streaming audio and video over the Internet, written-up in so-called RFCs. (The initials stand for *Request For Comment*, reflecting the open and changing nature of the Internet. But they are, effectively, standards that vendors use to achieve interoperability.) One of these describes the Real Time Protocol (RTP), a scheme for extending UDP to support media streams. This is used for VoIP telephony and is becoming standard for broadcast codecs, as well. RTP adds a sequence number to UDP packets so that the receiver can be sure packets are placed in order at the receiver. A timestamp can be used to synchronize multiple streams, such as audio and video for a television program. Just as TCP/IP are often said in the same breath and considered inseparable, so too RTP/UDP.

To recap, there are only these ways to deal with the inevitable packet loss in most wide area IP networks:

- Retransmission, such by TCP
- Forward Error Correction
- Concealment

Both retransmission and FEC cause an increase in delay that is unacceptable for two-way applications. FEC increases bandwidth and can make the loss problem worse. This leaves concealment as the best solution for interactive applications. Although retransmission may have its place when delay is not an issue, and FEC when retransmission is not possible. RTP/UDP is the Internet standard for media transport and gives us what we need to convey low-delay audio streams over IP networks.

Call Setup: SIP and SDP

Broadcast codecs can simply connect to each other by having the transmitter specify the receiver's IP number. The network passes the packet stream on to the destination device, and that is that. For nailed-up connections over dedicated links, this works perfectly well.

But we might well want our IP codecs to work like their ISDN equivalents, with a dialing function to find and connect to the destination codec. Session Initiation Protocol (SIP) is nearly universally used for VoIP telephone service and is becoming the standard for broadcast codecs. While it is possible to have SIP connect two units with no other component, it is common to use it with a SIP Server that can help get around firewalls, provide "buddy list" features to groups of related users, and supports a relocation service so that a destination can be found regardless of which IP number it is connected to.

SIP serves as a carrier for Session Description Protocol (SDP). This signals between the two ends what codecs are available at each, and allows the system to negotiate the optimum codec among those to be had. Codec users finally have what they have been waiting for – no need to know or set the coding method before attempting a connection. Just "dial" and let the system figure it out.

Delay

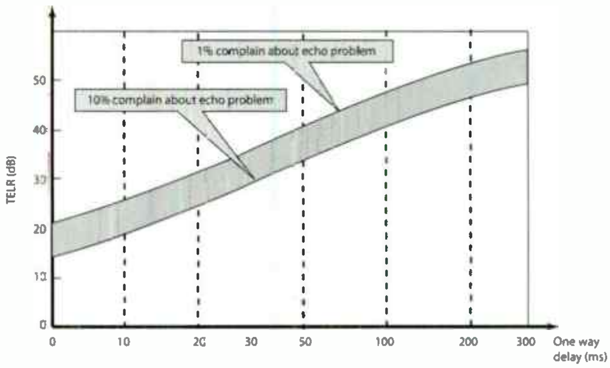
ISDN has almost zero delay, so the delay in ISDN codecs comes from the coding process, not the network. The popular MPEG AAC and Layer 3 (MP3) have around 150ms delay. This is too much for the round-trip, so the usual practice is to use G.722 for the return path, trading-off lower delay for lower fidelity. G.722 has around 20ms delay, so a connection with AAC one way and G.722 the other would have around 170ms delay.

The new AAC-ELD codec has around 60ms delay. Used on both directions, that would result in 120ms delay from the coding process. On a good IP network connection, there could be around 50-100ms delay from packetization and buffering, making the total delay around 170-220ms. That is acceptable for two-way conversation, but is pushing the limit. It's about the same as the delay in mobile phone conversations, which people have become accustomed to, so we can expect talent to be reasonably satisfied.

Echo

With IP we are going to need to give even more thought to echo than with ISDN codecs. In a perfect system, there will be no delayed return at all to the talent's earphones. A local mixer at the remote site sends the talent microphone audio directly to headphones and a mix-minus at the studio blocks the codec feed from returning, entirely avoiding the codec delay. But there are unintended causes of the talent audio making its way back. One is headphones worn by talent in the studio when the volume is high and isolation less than perfect. Another is telephone hybrids that don't have sufficient trans-hybrid loss.

Annoyance from echo is a function of both delay time and amplitude, a phenomenon well-studied in the context of telephony. Thus, anything that can reduce the amplitude is going to help. Don't use open air headphones, for example.



Annoyance from Echo is a Function of Both Amplitude and Delay (from ITU G.131)

NETWORKS

One of the advantages of IP is that there are so many ways to use it. From local networks to satellites, you have many options from which to choose.

The Public Internet

The Internet has the great twin advantages of ubiquity and low cost. One can find an IP connection almost anywhere and simply jack into it without waiting for installation or paying for it. There are no service guarantees, so you take your chances. But with the adaptive codec technology just described, the Internet becomes a reasonable proposition for many broadcast applications. There have already been successful IP remote broadcasts from airplanes, remarkably. International hook-ups are as easy as local ones.

A broadcast codec intended to be used over the Internet can benefit from having an integrated traceroute and network conditions graphing capability, so that you can see the cause of problems when they occur.

Dedicated Links

For studio-to-transmitter links and other high-reliability point-to-point applications, it is often possible to order Telco IP links that have guaranteed performance. Packet loss, jitter, delay, and bandwidth are specified in a contract called a Service Level Agreement.

MPLS Service

This is a Telco IP service that is growing in popularity. It is intended for high-quality VoIP telephony, video conferencing, and the like. Multi-Protocol Label Switching networks analyze traffic at the entry point and attach a label to each packet that describes the path the packet should take within the network. Because routers can see the packets as a stream, reserving a specified bandwidth is possible and usual. MPLS services are attractive to broadcasters since they offer a good cost/performance compromise – more expensive than non-guaranteed public Internet service, but less costly than dedicated links or ISDN.

An EVDO Mobile IP Radio Card

Mobile IP Services

Mobile IP services with fast enough uplink speeds can be used for remotes. In the CDMA world, EVDO Rev A is fast enough for us to use and is widely deployed in the USA. The uplink has a maximum rate of 1.8 Mbit/s, but under normal conditions users experience a rate of approximately 500-700kbps. EVDO Rev B promises yet faster speeds. The bandwidth is shared and there is a possibility of oversubscription and thus packet loss and insufficient bandwidth, but some trials with IP codecs have been successful. In Europe GSM is not generally fast enough in the uplink direction, though a fast service called High-Speed Uplink Packet Access (HSUPA) is just coming online, offering up to 5.76 Mbit/s bandwidth. Both EVDO and HSUPA have QoS capability in their technology specifications, but it is not clear if or to what extent mobile providers will pass this to their customers.

Access to these services is usually via a PC Card-style radio device. These are designed to be used with laptop PCs, but can just as well be plugged into broadcast codecs that include the appropriate connector and software. Or an external box can be used to interface the PC Card to the codec via a wired Ethernet connection.

LANs

Modern local area networks are usually switched Ethernet, but they can sometimes be routed at the IP level. In either case, they will have no packet loss. Recall that packet loss results from a link being overloaded. Wide area links are expensive Telco circuits, so they often have lower capacity than peak demand requires. On LANs, links are fast and cheap, so no packets need be dropped. Thus, there is no need for codecs and the other features needed for WANs. PCM linear coding can be used and no packet loss correction is necessary. While broadcast codecs could be used on these networks, it doesn't make much sense to do so. They are "overqualified" for the job – and simpler, less expensive interfaces are more appropriate. This is the domain of AoIP (Audio over IP) equipment such as Axia's Livewire.

Ethernet Radios

There are plenty of radio systems on the market that can be used for Ethernet links. Most work on the license-free ISM bands at 2.4, 5.2, and 5.6GHz. With high-gain antennas and a line-of-sight path, these can have range up to many tens of miles/kilometers. Bitrates are in the 10s of megabits, so there is much more bandwidth than needed.

WiMax

These are Ethernet radios that conform to a standard, assuring interoperability. Unlike the usual Ethernet radio which is used for point-to-point operation, these permit multiple sites to share a common channel and infrastructure. Since the channels are shared, there would be the possibility of oversubscription and contention for bandwidth. Perhaps WiMax vendors and providers will introduce some form of priority mechanism to offer guaranteed quality of service.

WiFi

IP codecs usually work over WiFi radio links without trouble. These would normally be only a part of the total IP path, perhaps extending an available DSL connection to the required location at a remote site. Again, the bandwidth is in the 10s of megabits, much more than is needed.

Satellites

Many satellite services are now IP-based and can be used for both point-to-point and point-to-multipoint links. While satellites are certainly exotic compared to other IP connection methods, from the perspective of the terminal equipment, they look about the same as any other link.

SERVICE LEVEL AGREEMENTS

With dedicated links and MPLS service, there will generally be a contract with the provider specifying the terms of their obligations with regard to quality of service. These are called Service Level Agreements (SLAs). Typically an SLA will include the following points:

- QoS guarantees: delay, jitter, and packet loss limits
- Non-QoS guarantees such as network availability. For broadcast, this should usually be at least 99.999%.
- The scope of the service. For example, the specific routes involved.
- The traffic profile of the stream sent into the network. This will be the bandwidth required, including any expected burst.
- Monitoring procedures and reporting.
- Support and troubleshooting procedures including response time.
- Administrative and legal aspects.

SIP SERVERS

A system using SIP requires proxy and registrar (also called User Agent) servers to work as a practical service. Although two SIP endpoints can communicate without any other SIP infrastructure, this approach is impractical for a public service. SIP provides a signaling and call set-up protocol for IP-based communications that can support a superset of the call processing functions and features present in the public switched telephone network. But SIP by itself does not define these features. Rather, it has been designed to enable the building of such features using servers placed on the network, which provide functions that permit familiar telephone-like operations: dialing a number, causing a phone to ring, hearing ring-back tones or a busy signal.

Servers can include the following components:

- **Proxy Server**: These are the most common type of server in a SIP environment. When a request is generated, the exact address of the recipient is not known in advance. So the client sends the request to a proxy server. The server on behalf of the client (as if giving a proxy for it) forwards the request to another proxy server or the recipient itself.
- **Redirect Server**: A redirect server redirects the request back to the client indicating that the client needs to try a different route to get to the recipient. This generally happens when a recipient has moved from its original position either temporarily or permanently.
- **Registrar**: As you might have guessed already, one of the prime jobs of the servers is to detect the location of a user in a network. How do they know the location? If you are thinking that users have to register their locations to a Registrar server, you are absolutely right. Users from time to time refresh their locations by re-registering.
- **Location Server**: The addresses registered to a Registrar are stored in a Location Server.
- **Presence Server**: Keeps track of the status of users (such as Available or Do Not Disturb) and makes this available to other users. There is no need to attempt a call to receive the status information.

They also can provide a gateway into the public switched network and other features such as directory and location services.

Calling with SIP

SIP calls can be made using different identification schemes.

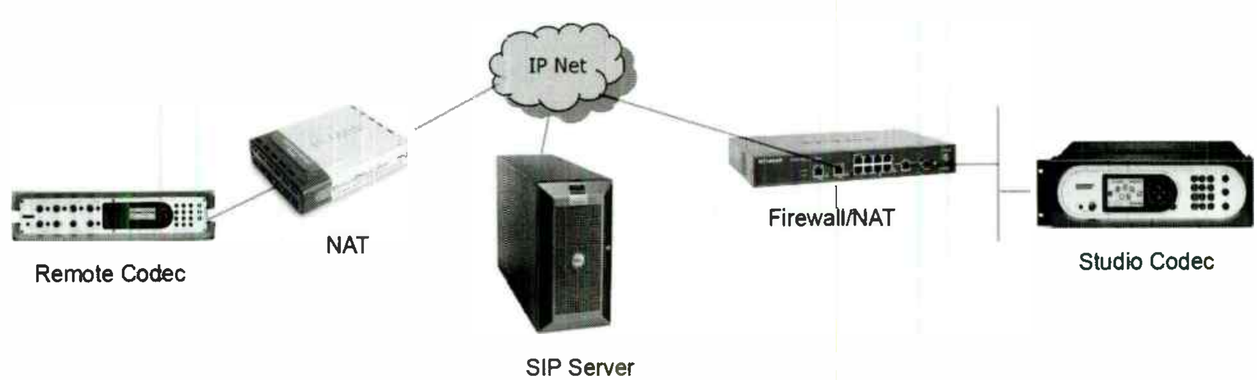
- Directly by IP number (e.g. sip:user@123.5.64.6)
- A form similar to email (e.g. sip:user@host.com)
- A text name processed by a directory server (e.g. WABC ZIP Codec 1)
- Plain old telephone numbers in the so-called E.154 format (e.g. +1 216 241 7225)

When needed, the SIP server translates telephone numbers to IP addresses using a procedure called ENUM (from tElephone NUmber Mapping). (But as VoIP and SIP catch-on, there is not much need for keeping phone numbers around. Wouldn't it be better to just use someone's email address for both text and voice messages? And if you want to call a company, why not pbx@telos-systems.com rather than an obscure 10-digit number string?)

FIREWALLS AND NATS

Network Address Translators are widely used on DSL Internet connections to allow more than one computer on the inside to share a single IP number toward the outside. All connections must originate from a computer on the inside. Since unsolicited incoming traffic can't get through, NATs provide a basic firewall function. This means that any codec inside a NAT would be both invisible and unreachable by another codec on the other side. Firewalls have the same effect. What can we do about this?

A simple solution is to put the studio codec directly on the public Internet, making it visible to codecs at remote sites even when they are behind NAT/firewalls. But it would not be possible to make a call from the studio



Typical Remote-to-Studio Set-Up
The SIP Server, outside firewalls and NATs, allows codecs inside to connect with each other.

to the remote codec, and having anything on the Internet without firewall protection is inviting problems. So we need to think about alternatives.

Consider that web traffic certainly moves both ways past these NAT/firewall devices. This happens because the NAT or firewall is usually "symmetric", meaning that when a packet stream is sent from the inside toward the outside, the NAT/firewall opens a return path for some period of time. Placed outside any firewalls, a SIP server can both receive and make calls to codecs located inside firewalls. Codecs register with the server, which then takes advantage of the open return path through the NAT/firewall to send an acknowledgment. The server sends additional messages periodically to keep the path open. When a codec wants to connect with another, it contacts the server, rather than the other codec directly. The server knows where to find the other codec and has an open path to it, so it can signal that a connection is being requested. Each codec can now send messages and audio streams to the other, thus opening a direct return path the other can use.

With unusually restrictive NATs and firewalls, the server can act as a relay for the set-up messages. In extreme cases, it may even have to relay the audio stream.

The Telos Z/IP Server

The Telos Z/IP server was designed to support the Z/IP codec. It is similar to a SIP server, but is specialized for broadcast codec application. It provides the following functions:

- **Directory Services**: Allows for easy discovery of other devices. The user controls the visibility of their device in the directory. A device may be (1) visible to all, (2) visible only to the group it belongs to, or (3) not visible in the directory. A device always belongs to a group (by default it belongs to the "public" group). The user may create a group at any time as long as the group name is not already in use. By giving others the group password you allow them to add their device to the same group. This also allows them to view devices that are visible only to the group.
- **Presence Services**: Allows a user to view the connection state of their "buddies".
- **NAT Traversal Services**: Allows a device to discover their public address and NAT type, if any. Keeps a connection open to devices that are not usually reachable behind the NAT to allow incoming calls to get through. For more restrictive NAT types, the server relays the signaling information to assist the connection establishment. For the most restrictive NAT types the server can also function as a media relay.
- **Geolocation Services**: Allows a device to find out the geographic location of their device and of the other end as well as the path taken (visual traceroute).
- **QoS Data Collection**: The server keeps track of call QoS data reported by devices. This information can provide us with some insight about the packet drop patterns, bitrates achieved, etc.

SIP servers can be public or private, and the same is possible with the Z/IP server. Telos operates a public server that Z/IP clients can use without having to support their own, or private servers can be installed by Z/IP owners.

With Server Support for Geolocation Services, a Codec can Display a Visual TraceRoute

THE EBU N/ACIP STANDARD

The European Broadcast Union (EBU) has concluded a process that has resulted in a standard for broadcast codecs. Their goal was to ensure that codecs from various vendors have modes that interwork with each other. To be compliant, each manufacturer must support a core set of functional components. A key is the use of SIP for call set-up and SDP for codec description, since this offers automatic negotiation to the "least common denominator" codec. A manufacturer is free to add its own enhancements as additions to the required core, but these might only work when codecs from that manufacturer are talking to each other.

The N/ACIP (Norm/Audio Codec over IP) standard specifies the following codecs:

Required

- G.711 (the standard telephone codec)
- G.722 at 64kbps
- MPEG-1/2 Layer 2 at 32-384kbps
- PCM linear at 12/16/20/24-bits and 32/48kHz

Recommended

- MPEG-4 AAC
- MPEG-4 AAC-LD
- MPEG-1/2 Layer 3 at 32-320kbps

Optional

- MPEG-4 HE-AACv2
- Enhanced APT-X
- Dolby AC-3
- AMR-WB+

Since none of the required codecs have concealment mechanisms and the standard does not require either TCP (or other retransmission) or FEC, it seems the EBU is targeting networks that have guaranteed QoS.

VOIP TELEPHONES FOR BROADCAST

With broadcast codecs migrating to IP and telephony moving to VoIP, seems some kind of convergence is inevitable. Both of these use SIP/SDP, so a particular call can be specified to use either a telephone-grade codec to interwork with the public switched voice network or a high-fidelity one. Broadcast IP codecs that conform to N/ACIP have the G.711 codec that traditional telephony uses, in fact.

Mobile phones might well use VoIP in the future, as some proposals within that industry are suggesting. Already phones like the Nokia E-series and several WiFi enabled mobile phones have SIP clients hardcoded into the firmware. Such clients operate independently of the voice part of the mobile phone network. Some operators actively try to block VoIP traffic from their network and in this case VoIP calls are done over WiFi. Several WiFi only IP hardphones exist, most of them supporting either Skype or SIP.

As a manufacturer of both on-air telephone systems and codecs, we are always thinking about opportunities to make the lives of our clients easier by combining the two systems. As time goes on, they may be entirely merged, though the operator interface would probably remain much as it is today. It probably would be more confusing than simplifying to have codec feeds appearing on telephone call selector button banks.

IP-BASED STUDIO EQUIPMENT

Studio audio distribution and mixing is migrating to IP, as well. Many stations are now using IP systems to distribute pro-grade audio throughout their facilities. With most audio now either originating or being sent to PCs, there is good logic behind using their native Ethernet/IP connections rather than a sequence of a sound card and an analog or AES3 interface to station audio systems. A 24-bit transparent connection is established simply and directly via the low-cost Ethernet link.

We have another advantage from this approach. It allows yet another convergence step as telephone, codec, and studio-grade audio share the same infrastructure. An Ethernet switch or IP router at the core can serve as a distribution router for everything. Control communication can pass on the same network. The mixing console can be tightly integrated with the telephone, codec, and call-screening application. Multiple studios can share "lines". Etc.

A couple of years ago, the *Wall Street Journal* had an article calling IP the "Pac-Man" of protocols, because it seems to be devouring everything in its path. With Ethernet as its LAN transport partner, its advantages are compelling. More IP-based PBX lines are now being installed than the traditional kind. Radio broadcasting technology has always taken cues from telephone world. It now follows the data technology world as well. (Count the number of PCs in your station.) The two are increasingly coming together outside our stations and are becoming ever more wedded within them.

An All-IP Facility Integrates Studio Audio, Telephone, Codec, and Control on a Single Network Infrastructure

Future Broadcast Technologies – A Worldwide Perspective

Tuesday, April 15, 2008
2:30 PM – 4:00 PM

Chairperson: David Wood
EBU, Geneva Switzerland

***Panel**

Bernard Caron, Communications Research Centre, Canada
Klaus Illgner, Director and CEO, IRT, Germany
Alberto Morello, Director, RAI, Italy
K. Tonioka, Director, NHK, Tokyo, Japan
Colin Whitbread, Head of Research and Innovation, BBC , UK

*Paper not available at the time of publication

Next Generation Public Alerting

Tuesday, April 15, 2008
4:00 PM – 5:00 PM

Chairperson: Clay Freinwald
Entercom, Seattle, WA

*Opening Remarks

Derek Poarch, Chief, Public Safety and Homeland Security Bureau. FCC, Washington, DC

*Panel

Edward Czarnecki, SpectraRep, Chantilly, VA
Darryl Parker, TFT Inc., San Jose, CA
Jerry LeBow, Sage Alerting Systems, Tye Brook, NY

*Paper not available at the time of publication

Monitoring and Measurements in the Broadcast Plant - Radio

Wednesday, April 16, 2008
9:00 AM – 12:00 PM

Chairperson: Talmage Ball
Bonneville International, Salt Lake City, UT

*Grounding Systems Why Important & Why Testing is Invalid 95% of the Time
John Howard, Lyncole XIT Grounding, Torrance, CA

VSWR Measurements in Broadcast Transmission Systems
Tim Holt, Director, Bird Technologies Group, Solon, OH

The Role of the Detector in Spectrum Analyzer Measurement of Hybrid Digital Signals
David Maxson, Broadcast Signal Lab, LLP, Medfield, MA

RF Measurement Techniques for Broadcast Engineers
Myron D. Fanton, Electronics Research, Inc., Chandler, IN

RF Signal Performance Measurements of Consumer FM Receivers and Coverage Effects
John Kean, NPR Labs - National Public Radio, Washington, DC

*Paper not available at the time of publication

VSWR Measurements in Broadcast Transmission Systems

Tim Holt, Director Applications and Systems Engineering
Bird Technologies Group
Solon, Ohio

INTRODUCTION

The measurement of Voltage Standing Wave Ratio (VSWR) has long been considered to be the most universal indicator of the health of transmission systems, and the continuous monitoring of transmission system VSWR, from the output of the transmitter to the input to the antenna is an important component in the maintenance of any wireless communication system. While it is possible to use precision reflection measurement instruments such as vector network analyzers to make high quality measurements on inactive systems, the day to day continuous monitoring of system VSWR under operating conditions requires the careful management of specific measurement system components. In this paper, we will explore the best techniques available for VSWR measurement, as well as to outline the most important measurement system parameters that must be considered when making VSWR measurements. While we could spend significant time exploring the definition and the supporting mathematics associated with VSWR measurements, the focus of this paper will be the details associated with performing continuous measurements on a day to day basis.

VSWR Defined - By way of definition, VSWR is a measurement of the ratio of the maximum voltage of a standing wave pattern on a transmission line to the minimum voltage on the line. The standing wave pattern is developed as a result of the interaction of the forward traveling wave with the reflected traveling wave, where the magnitude of the reflected traveling wave is related to the degree to which the load is mismatched to the characteristic impedance of the transmission system. (see Figure 1)

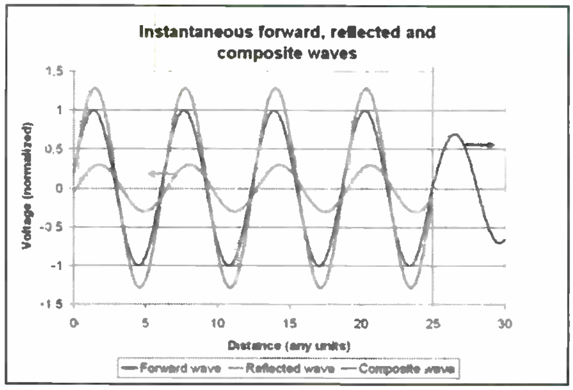
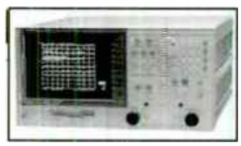


Fig. 1

VSWR Measurement Techniques - Transmission line VSWR may be measured using several techniques, including both scalar and vector network analyzers, and a direct measurement technique using a slotted transmission line equipped with a voltage probe/ detector arrangement has also been used historically.



While the above techniques may provide good results, by far and away the most common technique used today for the ongoing maintenance of broadcast communication networks for routine VSWR measurements is to derive VSWR measurements from forward and reflected power as measured using directional power meters. This technique has been used for over sixty years with good results, particularly in cases where the directional power meters are carefully selected for specific applications according to signal type, frequency, and power range. While this technique is well established, there are several important considerations when making these measurements that will help to insure high quality measurements. Some of these are:

1) **Power Meter Accuracy** - What is the basic accuracy of the directional power meter to be used?
2) **Power Meter Dynamic Range** - Over what range of power measurements will the instrument maintain rated accuracy?

3) **Power Meter Directivity** - What is the directivity of the directional coupler used as a part of the power meter system?
4) **Signal Type** - What are the characteristics of the signals that are being measured?
5) **Location** - Where is the power meter located within the transmission system?

We will deal with each of these measurement considerations along with their impact independently.

Power Meter Accuracy - When using a directional power meter (reflectometer) for VSWR measurement, the VSWR value is based upon the values of forward and reflected power according to the following formula:

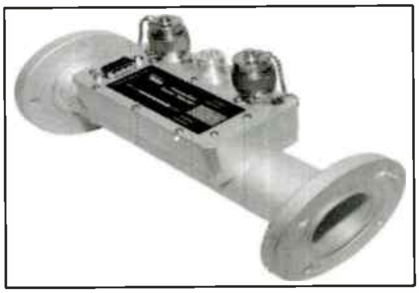
$$VSWR = \frac{1+\sqrt{\frac{Pr}{Pf}}}{1-\sqrt{\frac{Pr}{Pf}}}$$

Where, **PR** and **PF** are values of forward and reflected power as determined by the power meter. Based upon the above formula, it is clear that the basic accuracy of the power meter will have a direct effect upon the VSWR measurement. In addition, the power meter accuracy must be considered over the entire dynamic range of the instrument, rather than simply the accuracy of the power meter at a single point. It is best to use an example to illustrate this point:

The accuracy statements of most directional power meters are based upon either a percentage of the full scale range of the instrument, or as a percentage of the actual reading as determined by the power meter. For example, if a particular power meter is specified to have a **1 kW** full scale forward power range, and the rated accuracy of the instrument is **+/-5%** of full scale, the possible error of the instrument at full scale would be **+/-50W**, or the instrument could provide a reading between **950W and 1050W**, and remain within rated specifications. In this case, with the power meter accuracy specification based upon full scale, if the power meter is operated at **500W**, the error would remain at +/-50W, and the readings could be anywhere between **450W and 500W**, which translates to an error of **+/-10%.** Since the VSWR measurement is comprised of both forward and reflected readings, the accuracy of the reflected power measurement must also be considered. For example, our 1kW power meter mentioned above might use a reflected power range of 100W full scale, again with a rated accuracy of +/-5% of full scale. As in the case of the 1 kW forward channel, the accuracy of the measurement at full scale is +/-5%, but is considerably more than +/-5% for downscale readings.

To illustrate the effect that power meter accuracy has upon VSWR measurement, consider the following example:

Again considering our 1 kW system, if we desire to measure a load **VSWR of 1.1** with the power meter described above, this would mean that the reflected power associated with this VSWR value would be 2.26W if the forward power is 1 kW. If our power meter has a specified accuracy of +/-5% of full scale, the actual VSWR reading will range from **1.09** on the low side, to **1.19** on the high side based upon instrument accuracy alone.



This condition improves considerably with later generation power meters, where the accuracy of the power meter is specified as a percentage of reading, rather than a percentage of full scale. In large part, this has been made possible through the use of square law diode detector techniques, which behave linearly with respect to power.

Power Meter Dynamic Range - The dynamic range of a directional power meter is defined as the range of power values over which the instrument is capable of resolving measurements, while maintaining full rated accuracy. It is important to note that in most cases, power meters used for broadcast applications are configured as full two channel instruments, with separate channels for incident and reflected measurements. Often, the full scale power ranges of each channel are tailored to specific measurement conditions. For the forward channel, the power range is selected based upon 80 to 90% of the transmitter operating power. The reflected channel full scale power is normally based upon the anticipated reflected power as it applies to the expected VSWR of the antenna system. There is however, a limitation to this approach based upon the maximum power handling capability of the directional coupler / detector circuit. For example, if a power meter is to be selected for a 20kW transmitter application and the expected antenna VSWR is 1.1, this would imply that the anticipated reflected power would have a value of approximately 46W. For this reason, it would make sense to set the full scale range of the reflected channel to 100W, so that the expected reflected power could be easily resolved.

Directional Power Meter Directivity - At the heart of every in-line, directional power meter is a directional coupler for the purpose of providing a sample of the transmission line voltage, while discerning between the forward and reflected traveling wave. In most cases, in-line power meters used in broadcast applications are full two channel instruments, with separate directional couplers for forward and reflected measurements.

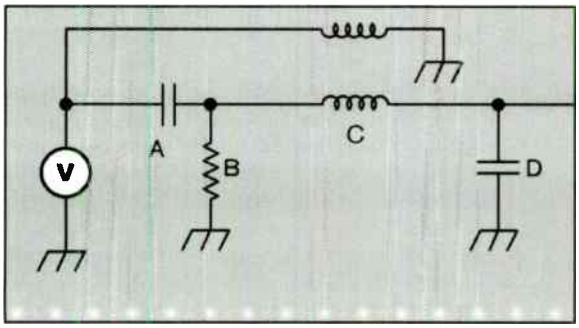


Figure 2- Directional Coupler

The directivity parameter for the couplers is a measure of how well the coupler is capable of discerning between the energy traveling toward the load, and the energy that is being reflected due to mismatch of the characteristic impedance of the transmission system with the impedance of the load. The directional characteristics of these couplers are a result of the fact that the coupler extracts two samples of the energy in the transmission line. One sample is derived from the electric field within the line, and the other is derived from the magnetic field. The directional nature of the magnetic field provides directionality, but the electrical balance between the magnetic field and the electric field samples serves as an indicator of the coupler directivity. The more balanced these samples are, the higher the intrinsic directivity of the coupler. Generally, directional couplers that are not fitted with frequency compensation networks have higher directivity, as the compensation networks are always reactive in nature, and will tend to alter the phase response of the coupler. In **Figure 2**, components A, B, and C form the main coupling elements in a lumped element directional coupler. Note that the magnetic field (current) sample is derived from the mutual coupling that exists between the transmission line center conductor, and the inductor in the coupler assembly. Component D is a frequency compensation capacitor.

As a first order approximation, the directivity value of a coupler may be used to determine the lower limit of the reflected power that a particular directional power meter is able to resolve. This is based upon the fact that although the reflected directional coupler is oriented such that it will respond only to the reflected traveling wave, it will also be affected by the forward traveling wave, in accordance with its directivity. For example, using our 20 kW transmitter example, if we are using a reflected directional coupler with 30dB directivity, this would mean that the reflected coupler would see about 20W of additional energy due to less than perfect directivity.

While the above discussion will serve well as a first order approximation, it is important to understand that the samples of energy collected by a directional coupler will appear as voltages at the coupler sidearm, and thus must be treated as vector quantities. Further, both the magnitude and angle of the load impedance is also important in a complete analysis of the effects of directivity upon directional power measurements. Since the angle of the load impedance is rarely known, the easiest approach to completely understand the effects of directivity is to calculate the change in reflected power appearing at the sidearm of the reflected directional coupler as both the contribution due to directivity, and the actual voltage due to the reflected energy appearing at the coupler sidearm are treated as phasors. Returning to our example:

a) **Calculate Voltage Due to Reflected Power** - This may be determined from the following formula:

Reflected Power (%) = $\left(\frac{VSWR-1}{VSWR+1}\right)^2$ x 100

Reflected Voltage = $\sqrt{P_R / R}$ = $\boxed{V_R}$

b) **Calculate Voltage Contribution Due to Directivity** - First, calculate the directivity power after converting the directivity value to a ratio, and then applying this to the forward power value. Based upon 30dB directivity, and 20 kW forward power, the directivity contribution would be:

Directivity Power Ratio = 10 $^{(\text{Directivity}/10)}$

Directivity Power = $\frac{\text{Forward Power}}{\text{Directivity Ratio}}$ = $\boxed{P_D}$

Directivity Voltage = $\sqrt{P_D / R}$ = $\boxed{V_D}$

c) **Find the Maximum and Minimum Reflected Power Limits** - These are found by simply adding and subtracting the voltages obtained in "a" and "b", and then converting these values back to power.

Maximum Reflected Power = $\frac{(V_D + V_R)^2}{50}$

Minimum Reflected Power = $\frac{(V_D - V_R)^2}{50}$

Based upon the above reflected power values, the VSWR measured at the power meter will range from a low of **1.03** to a high of **1.17**. Keep in mind that this represents the change in VSWR over any condition of load phase angle. It is extremely rare that these conditions would be observed in practice.

Finally, the directivity of a directional coupler will also serve as an indicator of the degree of immunity to line position under conditions of high standing wave. In other words, the higher the coupler directivity, the less immune to line position the coupler will be.

Signal Type - In-Line RF power meters are affected by the nature of the signals that they measure. The extent to which the power meter is affected depends upon the power meter type, as well as the modulation characteristics of the signals being measured. Conventional in line instruments that have been in use since the 1950's use point contact diode detectors, configured as peak detectors. These instruments are typically connected to analog meters calibrated to read average power. Under conditions where modulation is not present (CW), or the crest factor of the signals to be measured is low, these instruments perform very well. As the crest factor of the signal is increased, these instruments tend to follow the envelope of the modulation waveform, resulting in greater inaccuracy. Later generation diode type power meters use detectors operated in the diode square law region, where the diode voltage output is proportional to the square of the input voltage. These detectors will provide detected voltages proportional to true power, as long as the signal applied to the detector is bounded within the diode square law region.

Table 1 below compares the performance of two types of in-line power meters measuring an OFDM signal with a crest factor of 9 dB. The instruments were tested at power levels from 2kW to 12 kW, using a high power RF calorimeter as a primary reference. The data shows that while the square law based instrument closely follows the calorimeter readings, the conventional peak detecting diode power meter is not able to deal with the high crest factor OFDM waveform. One of the most significant aspects of this data is that while it is possible to calibrate a conventional power meter at a particular point for reasonable accuracy, the ability of the instrument to operate over a reasonable dynamic range is severely compromised. This will have a negative impact on the ability of the instrument to be used for VSWR measurements

Table 1- Power Meter Comparison Testing- OFDM

Average Power	Conventional Power Meter	Error (%)	Square-Law Power Meter	Error (%)	
2 kW	2.39 kW	+19.7	1.99 kW	-0.5	RF calorimeter used as reference
4	4.4	+10.2	4.03	+0.8	
6	6.32	+5.3	5.99	-0.2	
8	8.27		7.99	-0.2	
10	9.27	-7.3	9.99	-0.2	
12 kW	10.91	-9.1	11.99	-0.1	

Table 1

In-Line Power Meter Location - The location of the directional power meter within the transmission system may have a significant effect upon the ability of the instrument to resolve antenna VSWR. The prime consideration related to the location of the power meter is that the losses associated with the transmission line, as well as the insertion loss values of any other system components will have the effect of isolating the power meter from the measurement point. The best way to illustrate this point is with an example:

We are operating a 10 kW transmitter, connected to an antenna located **730'** up the tower, using 1-5/8" Heliax transmission line. The VSWR, measured with a directional power meter located at the output of the transmitter is **1.08**. The question is: what is the VSWR at the antenna?

The loss of 1-5/8" Heliax is specified at **0.2 dB per 100'** of length, this means that the one way loss due to the feedline is **1.46 dB**. Since a VSWR or Return Loss measurement is comprised of both incident and reflected parameters, the one way loss must be doubled to arrive at the isolation. A 1.08 VSWR corresponds to a return loss of **28.3 dB**. When the feedline isolation is subtracted from this value, the new return loss, at the antenna is **25.3** dB, which corresponds to a VSWR of **1.12**.

If the path between the in-line power meter and the antenna includes components such as switches, adapters, or other items, the insertion losses of these components will also have the effect of further isolating the power meter from the antenna.

Conclusion - VSWR measurements using in-line power meters are important components in the day to day operation of broadcast transmission systems. Following the above guidelines will help to insure that the highest quality measurements are obtained using this approach.

Tim Holt
Bird Technologies Group
440-519-2324
tholt@bird-technologies.com

THE ROLE OF THE DETECTOR IN SPECTRUM ANALYZER MEASUREMENT OF HYBRID DIGITAL SIGNALS

David Maxson
Broadcast Signal Lab, LLP
Medfield, Massachusetts

ABSTRACT

To obtain consistent, repeatable measurements of hybrid digital signals with a spectrum analyzer requires an understanding of the differences in spectrum analyzer detection methods. Since the initiation of In-Band/On-Channel ("IBOC") hybrid digital broadcasts, also known by the "HD Radio™" brand name of iBiquity Digital Corporation, there has been some variation in the measurements obtained by various methods and instruments. This paper discusses the common differences among spectrum analyzer settings and detectors, and recommends preferred methods.

INTRODUCTION

Hybrid IBOC transmissions present waveforms that are moderately challenging to measure. The term "hybrid" applies to the presence of two types of signal modulation in one communications channel: traditional analog modulation (AM or FM) accompanied by a set of OFDM digital waveforms. Orthogonal Frequency Division Multiplex transmission, from which the acronym OFDM derives, simultaneously modulates hundreds of narrow-bandwidth carriers with a series of digital symbols. In hybrid IBOC transmission, the OFDM waveforms are transmitted at a combined power level that is lower than the accompanying analog signal in the same channel. To measure the OFDM signal power levels, the dynamic range of the measurement instrumentation must accommodate the stronger analog signal while providing a clean enough noise floor to discriminate not only the digital signal components, but also the even lower-level unwanted byproducts of the transmission process. Further complicating such measurements, the manner in which digital signals are measured requires different assumptions than the measurement of analog signals.

Much material has been published on the structure of the hybrid IBOC signals in the AM and FM broadcast bands; this paper does not dwell on these fundamentals. See the National Radio Systems Committee ("NRSC") website for the IBOC transmission standard, which as of this writing is the version called NRSC-5-A. It can be obtained at no charge through www.nrscstandards.org. A detailed explanation of the inner workings of IBOC transmissions is presented in The IBOC Handbook - Understanding HD Radio Technology, written by this author and published by NAB/Focal Press. Further information is available in the NAB Engineering Handbook, 10th edition chapter 4.13 written by Jeff Detweiler of iBiquity Digital Corporation. The IBOC Handbook also contains a bibliography of many papers relating to the technology.

The NRSC-5-A standard contains fairly simple descriptions of the measurement of the digital components of hybrid IBOC signals. The NRSC has been discussing ways to clarify the techniques. Although much of the information in this paper was shared with NRSC, this paper does not represent NRSC determinations or specifications on the subject.

The long experience and sage advice of Joseph Gorin of Agilent Corporation is heartily and gratefully recognized. This paper would not have been possible without his contribution. Mr. Gorin, author of numerous articles and Agilent Application Notes over the years, is intimately familiar with the inner workings of spectrum analyzers, old and new. He generously answered questions and performed bench experiments to test hypotheses.

THE PROBLEM

Experience has shown that two parties can measure the same hybrid IBOC signal and obtain different results. In many cases the measurements are still within a couple of dB of each other, and the differences are not only analyzer-dependent, but also can be attributed to the expected variations in subsequent measurements of a signal: Signal levels vary over time; The setting of a reference level can vary from one time to the next, one operator to the next, one instrument to the next; Repeatability of measurements on one instrument is not perfect, while the use of separate instruments at separate times introduces further uncertainties. Of course, the signal under test might also vary, and the challenge is to identify real variations versus measurement variations. Ultimately the goal is to be certain that a measurement indicating compliance or non-compliance with spectral emissions limits is accurate and repeatable.

Anecdotally, there have been occasional circumstances during the commissioning of a new hybrid IBOC transmission facility where the broadcaster obtains different results than those that were obtained on the same device either at the factory or by the factory technician at the new installation. Also, in some instances two engineers will disagree whether a particular facility is in compliance with the applicable RF power spectral density mask.[1]

The iBiquity reference documents[2] in the NRSC-5-A standard refer rather imprecisely to "averaging the power spectral density of the signal in a 1 kHz bandwidth over a 30-second segment of time." [300 Hz for AM] Subsequent recommendations by iBiquity also suggest a minimum average of 100 swept traces of a spectrum analyzer. No other instrument settings are specified.

TRACTABLE VERSUS OPERATIONAL DESCRIPTIONS

The iBiquity description on making a power spectral density measurement is an operational description for the most part. In other words, it describes some fundamental settings that one would apply to a spectrum analyzer, i.e. 30-second averaging, 1 kHz resolution bandwidth, 100 sweeps. However it leaves out details on specific settings or instrument characteristics. Mr. Gorin explained to this author the difference between such operational specifications and a tractable specification this way:

> A tractable specification is one that is defined in a way that its measurement can be made without dependence on the particular characteristics of a measuring device. This goal can be contrasted with an "operational definition." An operational definition might say "configure an analyzer as follows, and ensure that the results are within this range." A tractable specification would say, "compute the sum of the power, regardless of the statistical distribution of that power (noise-like vs. CW-like statistics), across a passband with a width of x Hz subject to a fitting within a mask described in figure y." A tractable standard can be measured with older spectrum analyzers combined with computers, or newer spectrum analyzers, or dedicated hardware.

Perhaps the only clue we have to a tractable specification is that the power spectral density per kHz [300 Hz for AM] should remain below the applicable spectrum mask. The devil is in the details. Spectrum analyzer operators may not realize that differences in detection, filtering and computation among analyzers can contribute to differences in results. In his example of a tractable specification, Mr. Gorin adds the caveat that the measurement technique should accurately gauge the power per kHz regardless of the statistical distribution of the waveform's energy. This concern is explained further below.

POWER SPECTRAL DENSITY ON ANALYZERS

Spectrum analyzers have a variety of methods for "averaging" their measurements to obtain a result to approximate average power. The spectral masks specify how much energy in each slice of the radio spectrum may be emitted by the hybrid IBOC transmission. To evaluate the digital components of these emissions, one must start with a reference power level. With an unmodulated AM or FM carrier, the spectrum analyzer can be set to a narrow resolution bandwidth ("RBW") and the reference power level can be established.[3] This is the only time that the peak detector can be employed in hybrid IBOC measurements.

The Detectors

Spectrum analyzers have two layers of "detectors." The first form of detection is analogous to that which occurs in radio receivers. In a receiver, this stage demodulates the radio frequency waveform to recover the information imposed upon it. The radio receiver detects the baseband signal (audio plus any additional information such as stereo components) from the incoming modulated RF signal. Spectrum analyzers also detect the RF signal, not to reproduce the intelligence on the baseband but to sample the distribution of energy across the spectrum of interest. It is the equivalent of an AM detector, which converts the incoming filtered RF energy to a voltage. It detects the envelope of the incoming signal.

---

[1] NRSC-5-A Reference Documents: HD Radio™ FM Transmission System Specifications and HD Radio™ AM Transmission System Specifications

[2] ibid

[3] It is assumed that the operator has practiced good RF "hygiene" by ensuring that there are not significant signal levels from undesired sources that could be overloading the front end of the analyzer.

However, a spectrum analyzer sweeps across a selected spectrum at a selected RBW, so it is not detecting the baseband of the incoming signal the way a fixed-frequency AM receiver does. Instead, it is RF-detecting a continuously moving window on the spectrum.

In a classic analog spectrum analyzer, the display trace is swept across the screen at a rate that matches the frequency sweep of the local oscillator. The result is a continuous trace of the voltage of the RF envelope across a span of frequencies (the span) over a period of time (the sweep time).

The current state of the art in spectrum analyzers includes a digital display. Such digital displays have become the norm over the past two decades of analyzer evolution. Instead of a continuous analog sweep of the trace across the display, the trace is divided into discrete pixels representing "bins" (also known as "buckets") at specific points on the display. Digital displays commonly have between about 500 and 1000 such points horizontally. An inexpensive analyzer may have as few as 100, and highly specialized units may have several thousand.

As with the analog swept analyzer, the digital analyzer also employs a continuously swept local oscillator. The digital analyzer employs an analog-to-digital ("A-D") converter that collects a series of samples during the time that the local oscillator sweeps across each bin. For the duration of each bin, numerous voltage samples are collected.

The level of sophistication of the spectrum analyzer determines what can be done with the A-D sample in the bin. Early digital analyzers had barely enough processing horsepower to offer a choice of one of three features in each bin. The analyzer could identify (a) the highest value in each bin (the peak) and place that value on the display. Instead, (b) a minimum value (the pit or the min) could be displayed. Alternating peak and pit samples on the display produces a digital simulation of an analog display (often called a max/min display). The third way to select a value for a display bin was (c) to take the "sample" point in the bin. Commonly, the last data point collected by the A-D converter in each bin would be presented. Since there is usually no correlation between the sweep rate of the analyzer and the rate of modulation of the signal, the act of grabbing the last data point in each bin is a random sampling of the modulation in the bin, although it occurs repeatedly at a fixed frequency in the bin. This is result of "sample detection."

These three means for collecting information from a bin – peak, pit, and sample detection– are the second layer of detection found in a spectrum analyzer, often referred to at the display detector. Analog analyzers could perform a form of averaging by employing a video filter on the trace, reducing the vertical rate of change on the trace to filter out a noisy sweep. Peak detection would employ circuitry that caused the trace to ride the high side of the detected envelope. Now display-detection and filtering can occur in the digital domain, enabling more precise forms of filtering and detection.

Peak Detector

Returning to the concept of using an unmodulated carrier as a reference level source, this reference signal could be measured with a traditional peak detector. The detector would indicate the power of the unmodulated carrier by converting the voltage to an inferred power value based on the input impedance and the assumption that the measured signal is sinusoidal. As the unmodulated carrier is sinusoidal and spectrally falls entirely within a bin, no matter how narrow the RBW, the peak-detected reading of that signal's envelope as presented in units of power should be equal to the power of the signal.

Since, with a perfect continuous wave source, there is no variation in the carrier envelope during the time the analyzer spends in the center-frequency bin, the peak, pit, and sample will have the same amplitude. Thus, each of the three traditional display detectors should indicate the same level on an unmodulated carrier. This is one way to establish a reference level on a spectrum analyzer.

The same would not necessarily be true for a modulated signal. An AM broadcast signal has sidebands as well as energy on the carrier frequency. Averaged over time, the power specifically on the carrier frequency should closely approximate the average power of the signal, however a single sweep with a spectrum analyzer and peak detector would not achieve a successful reference level measurement with a modulated AM signal.

With an FM signal, a narrow RBW would not capture the reference level. However, since the modulated FM signal is transmitted at a constant power, a resolution bandwidth wide enough to capture the total power of the signal will provide the peak detector with the nearly sinusoidal frequency modulated signal to establish a power reference. However, since there is some amplitude modulation of any FM signal, the peak detector will err in favor of slightly overstating the power reference level, while application of the min detector will similarly understate it. (Typically there is so little amplitude modulation that the differences are inconsequential.) To obtain a more precise reference value of the modulated FM carrier power, the sample detector could be employed with averaging a series of sweeps.

Sample Detector

With a suitable reference level established, the sample detector can be employed to obtain power spectral density measurements. Since the sample detector acts as a random sample of the energy in each bin, a single sweep of a modulated waveform will not be particularly meaningful. The sample in one bin could be a peak, a pit, or anything between. The capability of the sample detector derives from the averaging of successive traces. Once the reference level is set, the hybrid IBOC signal is sample-detected, consistent with the instructions to be "averaging the power spectral density of the signal in a 1 kHz bandwidth…" Over the period of numerous trace sweeps, and with the continuously varying modulation of the OFDM waveform, a series of sample-detected data points in a single bin will provide a data set for obtaining an average level for that frequency bin.

With sample detection and the ability to accumulate numerous traces in memory and average them (or to statistically compute a rolling n-sample average), simpler (i.e. older) digital spectrum analyzers can be employed to generate an apparent average power spectral density of incoming signal. On a hybrid IBOC signal, the resulting indication of the OFDM power spectral density can be pretty close to the real values. Nevertheless, this method falls short on accuracy because of the nature of the averaging and the detection of digital waveforms.

Traditional Detector Averaging Error

Densely modulated digital waveforms behave in a white-noise-like fashion, which affects certain detector/averaging arrangements. Using the sample detector, a series of the samples from successive traces is averaged, and then converted to an equivalent power. This is not a pure power computation. It is essentially the square of the average rather than the average of the squares. A series of sample-detected samples of the white-noise-like digital waveform can be expected to be representative of the distribution of voltages in the waveform. However, an average of the values of these points is an average voltage. Converting the result to an equivalent power on a linear scale results on an understatement of the actual power by 1.05 dB.[4] Using the logarithmic scale the understatement is 2.51 dB. In other words, using a sample detector on noise-like digital waveform and averaging the results from a series of sweeps produces a power reading in each frequency bin that is 2.51 dB lower than the actual power level when viewed on a logarithmic display.

Older digital spectrum analyzers may only have sample detection plus trace averaging available to estimate power spectral density levels. By understanding how the instrument detects and averages the result, a more accurate interpretation of RF mask compliance can be made.

Is the IBOC digital waveform sufficiently white-noise-like to fit this model? After all, in the FM hybrid waveform, the individual OFDM carriers are 363.4 Hz apart and slowly modulated at a 344.5 Hz rate. Is it possible that the waveform within each frequency bin, (1 kHz on FM and 300 Hz on AM hybrid transmissions) consisting of about two or three OFDM carriers per bin, would behave more in a sinusoidal fashion than in a fully Gaussian (noise-like) fashion?

To answer this question, Mr. Gorin performed a digital simulation of an OFDM waveform similar to that of the FM IBOC signal and detected it using software similar to that in a digital spectrum analyzer. The result showed that the actual power understatement for sample detection of the FM IBOC OFDM at a 1 kHz RBW with trace averaging was 2.46 dB, which is very close to the 2.51 dB expected of a white-noise-like digital waveform.

The AM OFDM carriers are spaced at 181.7 Hz – half the FM OFDM frequency spacing – and measured at 300 Hz RBW – three-tenths the specified FM RBW. With this scaling in frequency and RBW for AM OFDM carriers, it is likely that these OFDM carriers will behave in a similarly noise-like fashion on a spectrum analyzer. Thus, the same averaging errors can be expected on hybrid AM measurements.

Since the reference level for hybrid IBOC power spectral density measurements is the analog carrier, if the sample detector/trace average method accurately measures the power of the continuous wave carrier as the reference and the same method understates the power of the OFDM waveform, then the relative power level of the OFDM sidebands will be understated. In

[4] Agilent Application Note AN 1303, Spectrum Analyzer Measurements and Noise

other words, a properly operating hybrid signal will appear to have slightly low OFDM carrier levels on traditional digital spectrum analysis.

STATE OF THE ART DETECTION AND AVERAGING

As spectrum analyzers have evolved, their processing power has followed the evolution of the digital signal processing ("DSP") industry. Processor speed, sophistication, and memory have increased, while cost, size, and power consumption have reduced. Many currently offered spectrum analyzers employ not only digital display capability but also digital detection and averaging. Processing in the digital domain permits analyzers to provide results that are less dependent on the mechanical and electrical limitations of physical filters and detectors.

The Average Detector

With the ability to store and quickly process all the data points collected in one bin, the fully digital spectrum analyzer does not need to rely on the single-data-point-per-bin approach of the sample-detection scheme. Instead, the analyzer can process the full set of data points in the bin to provide an average value. Some instruments call this the average detector, which can be set to a voltage average mode or a power average mode. Other instruments differentiate between the two average detector modes by calling one the average detector and the other the RMS detector. In either case, the power or RMS detector performs the same function of providing an accurate power calculation of the energy in each bin.

An average detector should not be confused with the trace averaging function, in which a series of traces is averaged to obtain time diversity on the measurement in each bin. Average detectors can be employed with trace averaging. Trace averaging a series of average-detected sweeps is not likely to cause the same error as trace averaging with a sample detector. This is due to the fact that each average-detected bin value is an accurate power value, having smoothed out the variations in level for the duration of the bin. The variations between successive sweeps across the same bin are likely to be substantially less than the white-noise-like distribution of a series of single samples, one from each bin.

Alternatively, by running one very slow sweep there can be enough dwell time in each bin for the average detector to accumulate a reasonable long term average per bin, without resorting to the uncertainty of trace averaging. In this manner, utilizing the 30-second averaging specification, an instrument with an average detector could be employed with a single 30-second sweep. However, there is something reassuring in the act of sweeping more quickly and returning to each bin numerous times and averaging the resulting set of traces. The average detector still insures the best average is obtained in each bin for each sweep. The trace averaging provides the opportunity to average the activity in each bin from non-contiguous modulation symbols, offering a time-diverse data set for each bin. There may be no significant difference in the result from spending n milliseconds in the bin once, versus spending $n/100$ milliseconds in the bin 100 times and averaging the resulting 100 values.

The best accuracy from a measurement perspective would be to use the average detector and avoid the trace averaging, unless it is certain what method the analyzer uses to trace average. Some instruments accumulate the underlying bin data to properly average detect all the data aggregated from a series of sweeps. Others rely on the more inaccurate (mildly so) method of averaging the resulting display traces, which can contribute to the trace averaging error discussed in this paper. It may be difficult to get a correct answer from the manufacturer regarding these subtle distinctions in analyzer algorithms, as they may not be fully described in the documentation.

Channel Power Utility

The channel power ("CP") and adjacent channel power ("ACP") utilities offered on spectrum analyzers can be employed to obtain power data on specified bandwidths other than the standard RBW settings. For instance, with a hybrid FM IBOC signal operating in basic MP1 mode, one could set the ACP utility to measure the 69.4 kHz spectra from the 129.2 to 198.6 kHz offsets, which contains the Primary Main ("PM") OFDM spectrum on each side of the analog signal. This will provide the power of each entire PM sideband, which can be compared to the analog power reference. Such measurements can corroborate the power spectral density plots of the 1 kHz and 300 Hz RBW measurements.

In the case of the hybrid FM signal, the total power in the PM sideband would measure –23 dBc/69 kHz while the target for the 1 kHz RBW measurements is proportionally lower (about -41.4 dBc/kHz). The channel power utilities perform computations on the sampled data in the background and do not rely on the trace measurement bins. Channel power measurements are reliable and accurate on digital waveforms, but may not be dependable on analog signals because of the interaction between the modulation, the sweep rate and the resolution bandwidth.

MASK SLOPES

When measuring the slopes of the edges of the hybrid IBOC OFDM waveform, the filter employed in the analyzer can influence the outcome. An RBW filter with relatively wide skirts can create the appearance of a slope that does not exist on the spectrum.

Consider the case of the AM hybrid IBOC signal, whose last OFDM subcarrier exists at 14.717 kHz offset from center frequency. The spectral mask wraps around this last subcarrier and contains a slope that leads down and out from the subcarrier. (Mask segment appears in Figure 2.)

Mr. Gorin experimented with the effects of two types of RBW filters. One is the typical shape of the passband of a four-pole filter commonly employed in spectrum analyzers. The other is typical of a Gaussian filter employed in the digital domain in fully digital spectrum analyzers. Figure 1 shows the approximate skirt shapes of the four-pole and Gaussian filters at 300 Hz RBW. Figure 2 shows the NRSC-5-A hybrid AM IBOC spectrum mask above two curves. The two curves below the mask are created by sweeping each filter past the end of a signal "cliff" at 14.7 kHz. Actual OFDM carriers were not simulated in this model, so there is no phase noise or intermodulation energy outside the edge of the virtual signal ending at the 14.7 kHz offset.

Figure 1 illustrates how effectively a Gaussian filter cuts off energy leakage from adjacent spectrum. Where the four-pole skirt is 29 dB down at 720 Hz off center, the Gaussian filter is 69 dB down.

In Figure 2, the two filters are swept past an energy cliff at 14.7 kHz offset from center frequency. This depicts the steepest possible slope that could be obtained when measuring the hybrid AM IBOC OFDM subcarrier group. No matter how sharp is the cutoff of energy above 14.7 kHz, the spectrum analyzer filters will display a slope in the spectrum that is related to their resolution bandwidth and their shape. (The energy below 14.7 MHz is simulated on the figure as crosshatching, for illustration only.)

As expected, the skirt of the 4-pole filter lets in more energy from the edge of the occupied spectrum as it is swept above 14.7 kHz. The resulting slope closely parallels the spectrum mask, with about 250 Hz horizontally and 12 dB vertically between the filter

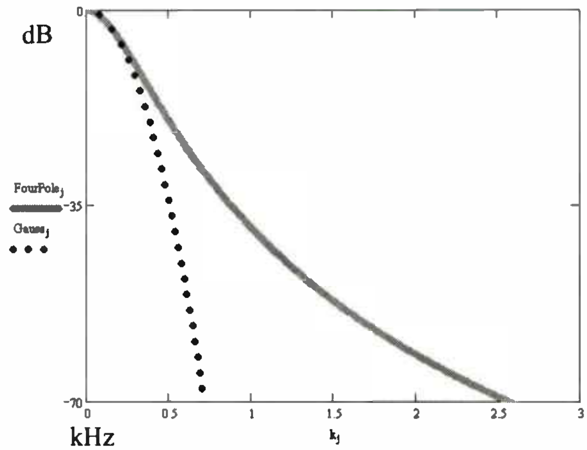


Figure 1

Response Curves of Typical 300 HZ RBW Four-Pole Filter (solid line) and Gaussian Filter (dotted line) in Spectrum Analyzers (based on algorithms provided by J. Gorin)

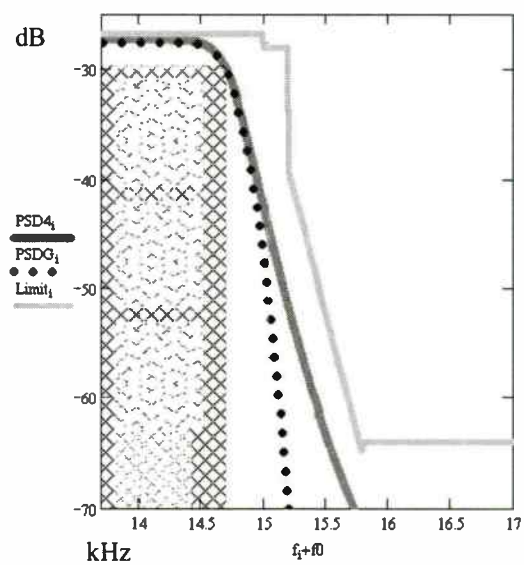


Figure 2

Idealized Power Spectral Density with 300 Hz RBW Filters Swept Up-Frequency Past a Signal Cliff at 14.7 kHz Showing Gaussian (dotted line) and Four-Pole (Solid Curve) Response in Relation to Spectrum Mask (Concatenated Line Segments). (based on algorithms provided by J. Gorin)

response and the mask. This does not appear to be an accident. If the NRSC-5-A spectrum mask were tighter than it is, having a steeper slope by the 14.7 kHz cliff, then a spectrum analyzer with conventional filtering could show a failing response with even a perfect signal. The spectrum mask must not have a slope that is steeper than the ability of a spectrum analyzer to resolve.

The Gaussian filter provides more breathing room between filter response and mask (in other words, better selectivity).

If there were intermodulation and phase noise products in the slope area, their power would be added to the baseline filter response shown in Figure 2 and could potentially create the appearance of non-compliance with the mask. The operator of the spectrum analyzer would be equipped to address such a condition by being aware of the actual filter response on his instrument. A simple sweep of a continuous wave test signal will create a snapshot of the instrument's filter response.

Hybrid FM measurements are not as constrained by filter slopes as hybrid AM. The slope of the hybrid AM mask is much steeper than that of the hybrid FM mask. As can be gleaned from Figure 2, as well as the NRSC-5-A AM Transmission System Specifications reference document, the pitch from the 15.2 to 15.8 kHz offset is 43.3 dB per kHz. The hybrid FM pitch specified in the NRSC-5-A FM Transmission System Specifications reference document is a comparatively shallow 0.9 dB per kHz from the 200 to 215 kHz offset. Even allowing for the expansion to a 1 kHz RBW filter for FM mask measurements, the four-pole filter response is substantially steeper than the mask, ensuring the analyzer filter will not be a significant factor in the event that a measurement indicates non-compliance on the slope. (iBiquity has since recommended the FM slope be further shallowed by extending it from 200 kHz to 250 kHz instead of to 215 kHz, resulting in a 0.26 dB per kHz slope.)

SWEEP RATE

The optimum sweep rate is determined by the relationship between the RBW and the span. Spectrum analyzers provide an autocouple mode in which the appropriate sweep rate is automatically selected for the chosen span and RBW.

A rule of thumb can be applied to estimate the optimum RBW for a given sweep time and span.[5]

$$R_o \approx \sqrt{S/T}$$

[5] Engelson, M., Modern Spectrum Analyzer Theory and Application, p. 213, Artech House, 1984

Selecting a 600 kHz span for a look at the hybrid FM spectrum with a 1 kHz RBW, and solving for sweep time, the optimum rate would be about 0.6 seconds per sweep. Similarly, with a 50 kHz span and 300 Hz RBW, the sweep time would be about 0.6 seconds.

In practice, spectrum analyzers appear to be conservative with this computation, by about a factor of two. It is common to see a sweep rate of about 1.2 seconds per span for 600 kHz span and 1 kHz RBW. Similarly, at 50 kHz span and 300 Hz RBW, the automatic sweep rate also typically sets to a rate of about 1.2 seconds per span.

This performance represents less than 30 sweeps in 30 seconds. Hence, the 30-second sweep specification in NRSC5-A reference documents is incompatible with the subsequent 100-sweep recommendation. A tractable specification would perhaps be most useful if it were to specify the total measurement time that each 1 kHz of spectrum should accumulate to ensure a representative sampling of the spectral power within that bin, regardless whether the total time is accumulated in one pass of the sweep, or in multiple passes. Based on the 100-sweep recommendation, and an assumption that the sweeps were 1.2 seconds in duration, that would yield 120/600,000 = 0.2 milliseconds per kHz on the FM signal. The 600 kHz span permits the operator to view sidebands that are 100 kHz outside each side of the occupied bandwidth. On the AM measurement, the 50,000 kHz span permits the operator to view 10 kHz on each side of the occupied bandwidth. With 100 sweeps at 1.2 seconds each, the dwell time per kHz would be 2.4 milliseconds per kHz.

The calculated AM and FM capture times per kHz differ by an order of magnitude. The FM symbol rate is about 344.5 Hz, with duration of 2.9 ms per symbol. The AM rate is half that, with a duration of 4.8 ms per symbol. Clearly, with capture times per kHz being quite a bit less than the symbol duration, the modulation is sufficiently noise-like that it does not require the capture of one or more complete symbols to evaluate the power spectral density. Nevertheless, while selecting a capture time per kHz may be the most effective way to write a tractable specification, the optimum capture times are not obvious. More analysis is necessary to determine whether there are structural reasons to perform trace averaging to obtain time diversity sampling of the modulated OFDM waveforms.

Capture time for an average detector is literally the time spent per kHz per sweep. In contrast, the capture time for the sample detector is only one data point per sweep. Therefore, it may take a sample detector with trace averaging more time to achieve optimum averaging than an average detector with or without trace averaging.

CONCLUSIONS

The specifications for measuring hybrid IBOC spectral occupancy presently lack tractability and are not described with enough detail operationally. Without tractability or at least a reference operational description on reference equipment, measurement disagreements and borderline non-compliance conditions will be challenging to resolve. The NRSC has taken up this issue and may release further clarification in the near future.

Spectrum analyzer operators should become fully aware of the detection and averaging algorithms employed by their devices.

The spectrum analyzer operator is best served by employing an analyzer that has Gaussian digital filtering and an average power (a.k.a. RMS) detector.

If a less sophisticated digital spectrum analyzer is employed, the sample detector may suffice with trace averaging. However, the operator should verify that at least 2.5 dB appears between the measured trace and the limit prescribed in the mask. This margin accounts for the understatement made in power spectral density when sample detection and trace averaging are employed in a log display.

Peak detection might indicate compliance with the mask, but it is not reliable for indicating non-compliance.

Channel power utilities are helpful for obtaining accurate power measurements of entire OFDM subcarrier sideband groups as a quick check on the relative power of the analog with the hybrid digital signals.

Care should be taken with setting the reference level of the analog carrier prior to taking the mask compliance measurement.

RF MEASUREMENT TECHNIQUES FOR BROADCAST ENGINEERS

Myron D. Fanton, PE
Electronics Research, Inc.

ABSTRACT

Broadcast engineers must be competent in many areas of discipline: IT, DSP, Facility and Power Engineering. In many cases Radio Frequency (RF) technology may be a low priority for a station engineer. Modern methods and tools for RF measurement are discussed.

Passive RF measurements of transmission equipment are reviewed: impedance, return-loss, insertion-loss, VSWR, isolation, directivity, and coupling. The discussion details the measuring and troubleshooting of filters, combiners, hybrids, transmission lines, and antennas. Also, the Vector Network Analyzer equipment used to perform these measurements is described, outlining modern techniques for system measurements.

Other RF measurements facing the broadcast engineer are explained, including the measurement of coverage, transmitter power output, and inter-modulation products. The specialized equipment and techniques involved are discussed, and common mistakes are revealed.

INTRODUCTION

The introduction of RF circuit principals occurs when the features of a component become a significant fraction of a wavelength. Compute the wavelength by,

$$v = f\lambda \qquad (1),$$

where v is the velocity of propagation, f is the frequency, and λ is the wavelength. The velocity of electromagnetic waves is the speed of light within the media, 11,803 million inches per second in a vacuum. For transmission lines, the velocity is specified as a factor of the speed of light in a vacuum.

At the point that systems and components become large with respect to wavelength, the relations of voltages and currents depart from the conventional lumped regime and are described by distributed models. Ohm's law and ohmmeters are replaced by Maxwell's equations and Vector Network Analyzers (VNA).

RF Equipment in the Broadcast Facility

The Broadcast Facility incorporates an RF transmission system delivering information to a number of receivers via a wireless channel: antenna, transmission line, filters, RF components, and high-power amplifier. Measurements associated with the transmitter have been regularly discussed [1] and are regulated [2]. Measurements of the passive RF equipment in the broadcast plant are presently discussed.

Fundamental RF Parameters

Given the distributed nature of RF structures, wave interactions must be described. It cannot be assumed, for example, that the voltage applied to a circuit is instantaneously applied to a load. Instead, the voltage along a circuit, $V(z)$ is the sum of incident, V_+, and reflected, V_-, voltages propagating as waves in the steady state, according to,

$$V(z) = V_+ e^{-\gamma z} + V_- e^{\gamma z} \qquad (2),$$

where the propagation constant, $\gamma = \alpha + j\beta$, and z is the position. α and β are respectively the attenuation and phase constants. At some positions, these interfering waves add constructively and at others destructively. The ratio of the resulting maxima and minima is the voltage standing wave ratio, *VSWR*. The ratio of the forward and reflected waves is the reflection coefficient, Γ.

$$\Gamma = \frac{V_-}{V_+} = \frac{Z_L - Z_O}{Z_L + Z_O} \qquad (3)$$

$$VSWR = \frac{1+|\Gamma|}{1-|\Gamma|} \qquad (4)$$

The voltages in Equations (2) – (3) may be complex numbers, meaning both magnitude and phase must be taken into account. Typically the magnitude of the *VSWR* is noted (the relative position of the standing wave nulls is neglected) or the logarithmic expression of the reflection coefficient, return loss,

$$RL = -20\log|\Gamma| \qquad (5).$$

Assumed in the above is that the waves are propagating along some uniform media, perhaps coaxial transmission line or free space, that may be described by its constituent material to have a wave impedance, Z_O, that is the ratio of the voltage and current at a given position. This transmission media is terminated with a discrete load impedance, Z_L, being an antenna or dummy load in the broadcast plant.

Vector Network Analyzer

The VNA is test equipment used to measure the RF network parameters of circuits. Network parameters are generally the scattering parameters, possibly the entire scattering matrix. The reflection coefficient of Equation (3) is one of the scattering parameters, and another is the insertion loss or coupling between ports. The scattering parameters are converted to other familiar RF parameters: VSWR, Return Loss, etc.

The output power of the VNA is much smaller than the typical broadcast system power at approximately 1W maximum, meaning that devices are characterized at low power. This in no way renders the measurements inaccurate; the VNA is sensitive to very low powers. In fact, care must be taken at broadcast sites to reduce interference and prevent damage to the sensitive receivers.

The VNA typically transmits an un-modulated sine wave, slowly swept over a frequency range, and is equipped with phase-locked receivers that measure both amplitude and phase. At minimum a VNA has one port that delivers the incident wave to the component under test while measuring both the incident and reflected waves with a dual-directional coupler. With multiple ports, the waves passing to and from multiple test points may be measured simultaneously.

Thus equipped, the VNA measures and displays the complex reflection coefficient, the quantity expressed in Equation (3). All of the other relevant parameters are computed from it: Return Loss, VSWR, phase, group delay, etc. The complex quantities can be displayed in polar plots, and the scalar components (magnitude, phase, group delay, etc.) may be displayed as a function of the frequency. Functions vital to modeling or tuning of a component or systems are also available, such as introducing time delay or time-domain transforms and gating. Because the input connectors (typically type-N, 7mm or 3.5mm) for a VNA are much smaller than typical broadcast transmission lines, transitions from type-N are required to execute the measurement.

Slotted Line

The slotted line has some application in broadcast equipment, and the measurement is therefore described. The specification and measurement of many RF components and system reflection is VSWR, the meaning of which follows directly and intuitively from the slotted line. From the sum of transmitted and reflected voltages on a uniform transmission line the expression for the measured standing wave voltage may be derived:

$$\frac{|V^2|}{|V_o{}^2|} = 1 + |\Gamma_L|^2 + 2|\Gamma_L|\cos(2\beta x) \qquad (6),$$

where Γ_L is the reflection coefficient of the impedance terminating the line (see Figure 1), β is the phase constant in the transmission line, and x is the distance from the load reference plane.

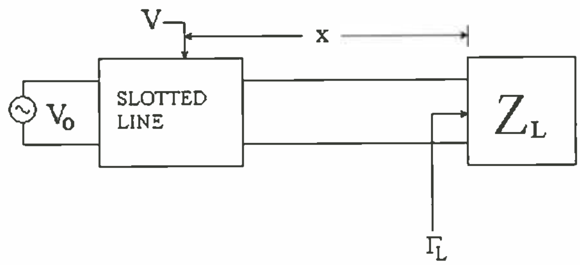


Figure 1: Slotted Line Circuit

Typically, the probe is moved along the slot while recording the voltage of the standing wave. At some positions the voltage maxima occurs, and the minima occurs at others. The ratio of the maxima to the minima is the VSWR. The magnitude of the reflection coefficient is then computed from the VSWR. In addition, the distance from the desired reference plane to a voltage minimum, x_m, is measured, and the phase of the reflection coefficient may then be computed as follows [3]:

$$\begin{aligned} \theta &= 2\beta x_m \pm \pi \\ |\Gamma| &= \frac{VSWR-1}{VSWR+1} \\ \Gamma &= |\Gamma| e^{j\theta} \end{aligned} \qquad (7)$$

TRANSMISSION SYSTEMS

Because the Broadcast Engineer is typically faced with measurements of a complete, assembled system, we consider this problem first and turn to each component later. Facilities for the transmission of FM and TV broadcast are systems comprised of antennas, coaxial or waveguide transmission line, connectors, elbows, combiners, diplexers, and filters [4]. Coaxial transmission line contains support insulators and bullet interconnections as well. Modeling this assembly of components appeals first to two-port network theory.

System Model

The network description of two components separated by uniform transmission line is shown in Figure 2. The total reflection, Γ_t, is related to the reflection of component B, Γ_B, and the S-parameters of component S.

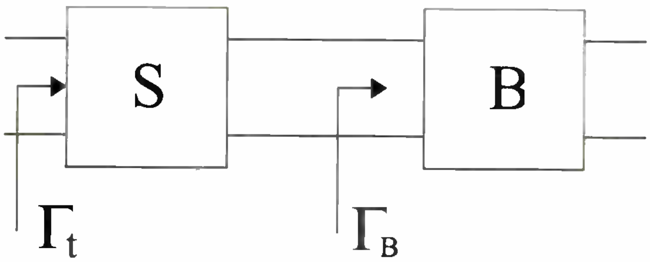


Figure 2: Network Model.

The total reflection is given by

$$\Gamma_t = S_{11} + \frac{S_{12}S_{21}\Gamma_B}{1 - S_{22}\Gamma_B} \qquad (8)$$

Assuming that the through losses and level of reflections are much smaller than unity (S_{22}<<1, S_{12}=1, S_{21}=1), a reasonable assumption in the vast majority of system components, the total reflection becomes

$$\Gamma_t \approx S_{11} + \Gamma_B \qquad (9)$$

which is the addition of two complex numbers. As the two components are separated by a distance, *d*, the total is

$$\Gamma_t \approx S_{11} + \Gamma_B e^{-2j\beta d} \qquad (10)$$

where the phase constant $\beta = 2\pi/\lambda$.

A depiction of this complex addition in terms of two vectors is shown in Figure 3. The magnitude of the total reflection varies according to the relative phase of the two component reflections. As Equation (10) indicates, the relative phase of the two reflections depends on the distance between the two components and the frequency. At some frequencies, they add constructively (long red arrow, Figure 3), and at some frequencies, they add destructively (short, blue arrow).

Extending this to the total reflection of the entire RF system yields the summation of the complex voltage reflection coefficients of each component. With this model, the cardinal system design principle has been established: minimizing the component reflections reduces the total system reflection.

Considering only the magnitude of each component, the maximum total reflection may be computed. Component values adding to a system VSWR of 1.10 are tabulated in Table 1. The excessive number of decimal places is included to allow the computations to be verified.

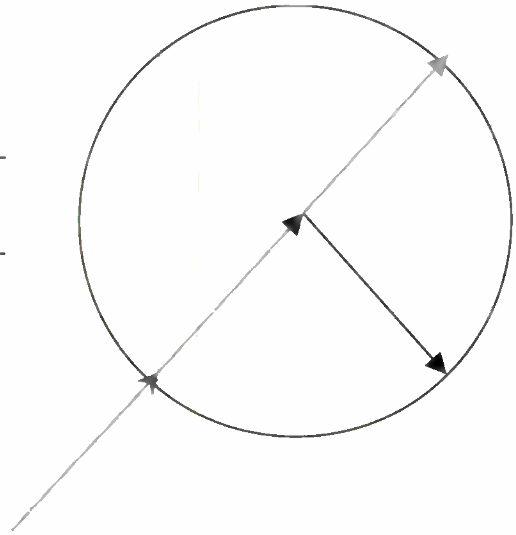


Figure 3: Vector Addition of Reflections.

# Components	Reflection Coef.	VSWR	Return Loss
2	0.0238	1.0488	32.46
5	0.0095	1.0192	40.42
10	0.0048	1.0096	46.44
20	0.0024	1.0048	52.46
30	0.0016	1.0032	55.99
50	0.0010	1.0019	60.42

Table 1. Component VSWR for 1.10 System

Using Equation (10), many typical measurement scenarios may be analyzed. The input of a transmission line system contributes the first of many reflections, often the large transition from the VNA connector to the large transmission line. Though this may seem undesirable or otherwise annoying, it provides a convenient reference with which to measure the distance to other reflections.

Naturally, an extended run of transmission line precedes the antenna; therefore the reflection due to the antenna occurs a long distance from the system input. An example of the return loss of such an antenna system is shown in Figure 4. The two reflection coefficients were assumed to be 0.01, located 500 inches apart.

This measurement scenario may be generalized and applied to any system composed of two reflections separated by a significant distance. The "rippled" characteristic of the return loss may be used to deduce the distance, *d*, between the contributing reflections with the following relation,

$$d = \frac{(n-1)cv_f}{2(f_2 - f_1)} \qquad (11),$$

where n is the number of nulls across the band, f_1 and f_2 are the respective frequencies at which the first and last null occur, c is the speed of light, and v_f is the velocity factor in the transmission line. Equation (11) simply follows from Equations (9) and (10). For convenience, the equation for the velocity factor is given as well:

$$v_f = \sqrt{1 - \left(\frac{f_c}{f}\right)^2} \qquad (12).$$

f_c and f are the cutoff and operation frequencies respectively. Equation (12) is used in waveguide systems while the specified velocity factor is used in coax systems.

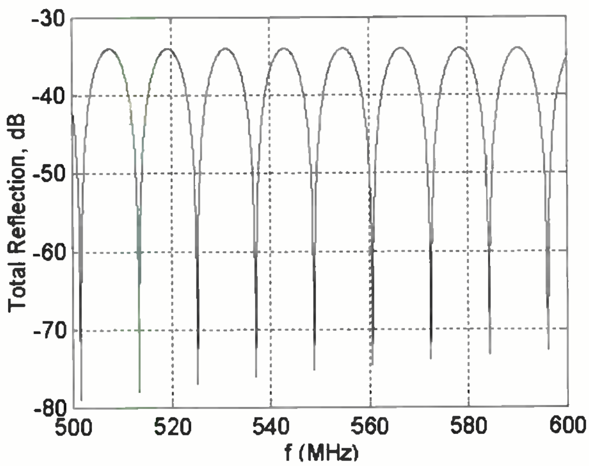


Figure 4: The Sum of Two Reflections

It must be noted that digital sampling at discrete frequencies across the band may produce a "ripple" in the measured data that may be interpreted erroneously as a reflection. Increasing the number of measurement points and decreasing the measured bandwidth will eliminate this error and possibly reveal a contributing reflection from a great distance away.

Similar techniques may be applied to RF system measurements to pinpoint reflection problems. Though the plots become more complicated, the location of multiple reflection points can be deduced by noting the nulls and applying Equation (11). The minima with a small separation in frequency correspond to a reflection that is far away while those with a large separation are closer together. Minding the characteristics of the nulls and employing Equations (11) and (12), the locations of the multiple reflections may be computed. The return loss predicted by Equation (10) is shown in Figure 5 for a three-component system: reflection 1: 0ft., reflection 2: 194.7ft., and reflection 3: 1220.8ft.

The concepts outlined above aid in the practical location of reflections such as poor connectors, transmission line obstructions, elbows, and damaged components. Improvement of such reflections must be carried out by improving the deficient component with tuners. Extending this technique to its ultimate end leads to the time-domain tools in the VNA.

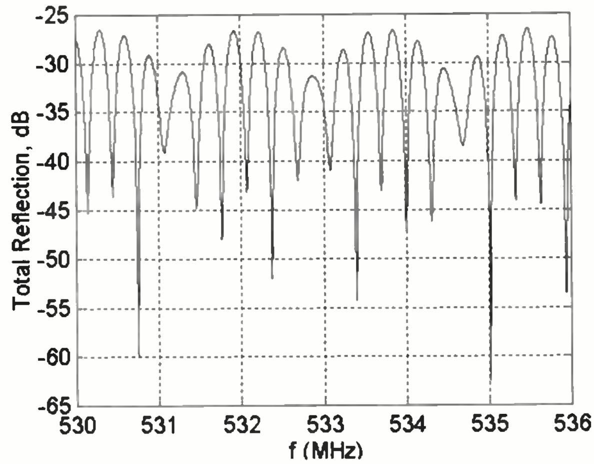


Figure 5: Predicted Return Loss of the System

Component reflections are complex quantities. A design criteria relating to the magnitude has been developed above, and one concerning the phase may be developed. Because the magnitude of the total reflection given in Equation (10) depends on the distance between components, the spacing of components may be designed to minimize the total reflection. For a narrow band application, like a single channel system, the phase relationship reflections separated by large distances may be exploited to tune the system.

In a transmission line system, this requires that the section lengths be varied along the total length of the line according to the patented method used in ERI WIDELine™ transmission line [5]. Essentially the spacing of support insulators and bullet interconnects is varied by a half-wavelength along the full transmission line length, making the last stick a half-wavelength shorter than the first.

Many 6 and 12 MHz TV broadcast systems as well as Single or Multi-Channel FM broadcast systems are successfully designed with the application of a "fine-tuner" section. These sections are a short length of transmission line with a number of capacitive probes spaced along its length. The addition of these tuner sections allows the narrowband optimization of the total system reflection. Similar narrowband tuning is accomplished by the addition of capacitive rings or "slugs" at select locations along the line. Neither of these methods will result in a broadband minimization of reflection because the reactance of a shunt capacitor is directly proportional to frequency.

Instrumentation in transmitter and control equipment that monitors VSWR typically computes VSWR from power measurements from a directional coupler. The physical limitation of the transmitter output is the reflected power, and the measurement is typically performed with a scalar, average power detector. In a scheme that average power varies with modulation (NTSC for example), the VSWR displayed by these instruments will subsequently vary. If only reflected power is measured, then the displayed VSWR is only correct for the output power and modulation state at which it was calibrated.

TDT Technique

The design principles established in the previous sections have highlighted the important contribution of small reflections to the total system reflection. Designing components with Return Loss performance better than 60 dB presents a challenge to conventional laboratory techniques. Through the use of a VNA and time-domain transform (TDT) analysis, the components may be accurately characterized.

Components in a broadcast transmission system require transitions from large coaxial transmission lines to small standard connectors on the VNA (typically type-N or 7 mm connectors). Such transitions are often narrowband components tuned by using narrowband techniques discussed above. As such, the reflections caused by the transitions are typically much higher than the 60 dB reflections demanded from the broadband components. Overcoming this limitation requires the use of TDT analysis.

The reflection measurements of the VNA are performed in the frequency domain. The VNA contains vector voltage detectors and a synthesized. swept frequency source. Frequency domain data are transformed to the time domain using a Chirp-Z transform, a numerically efficient generalization of the discrete time Fourier transform (DTFT). The data may be digitally filtered in the time domain by using "gating" and inverse-transformed back to the frequency domain. Such gates may be positioned in the time domain to remove the troublesome reflections from the test adaptors. Gating may also be applied to filter the TDT data to eliminate all reflections in a system but those of interest, focusing the measurement on a particular component.

BW, MHz	Np	Range, ft	Resolution, ft
6	1601	131136	157
6	801	65568	157
12	1601	65568	79
25	1601	31473	38
100	1601	7868	9
300	1601	2623	3
1000	1601	787	1

Table 2. TDT Range and Resolution

The swept bandwidth, or frequency span, affects the range and resolution in the time domain. The range (in seconds), being the largest value of time computed in the transform, is related to the number of data points and the swept bandwidth by the following equation:

$$range = \frac{N_p - 1}{BW} \tag{13}$$

where BW is the swept bandwidth in Hz, and N_p is the number of points. The time range may be converted to distance by multiplying by the speed of light (approx. 1 ft/ns). This range is the round-trip time of the reflected wave and must be halved to relate to the physical distance.

The resolution, in seconds, of the TDT data is given by

$$resolution = \frac{1.92}{BW} \tag{14}$$

where the 1.92 constant represents a band-pass transform by using a normal window and ranges from 0.45 to 2.88, depending on the VNA window and transform settings. Note that the resolution is not dependent on the number of points and may be converted to distance in the same manner as the range.

In the laboratory, the range is typically not a problem and the bandwidth is set solely based on the desired resolution. The range may become a factor in field measurements of transmission systems mounted atop tall towers. Measuring the reflections from an antenna atop a 2000 ft tower requires a range of 4 μs, which requires a bandwidth of 196.7 MHz with 801 points. Sweeping greater than this bandwidth will reduce the range and not allow TDT data manipulation of items outside the range of the transform. The corresponding resolution at this bandwidth is 4.8 ft, which can only be improved by increasing the bandwidth. By increasing the number of points to 1601 (at the expense of measurement time), the range may be doubled to 4000 ft. Alternatively, the bandwidth may be doubled, maintaining the 2000 ft range and improving the resolution to 2.4 ft. Table 2 contains a number of useful range and resolution relations.

FILTERS, COMBINERS, RF COMPONENTS

The components in the transmitter room are the most available for the broadcast engineer to measure. The transmitter often requires a mask filter to eliminate off-channel harmonics and spurious emissions, and notch filters are included when a specific modulation product must be eliminated. Filters are specified and measured by their insertion-loss and return-loss over the channel.

Channel combiners also involve filters, and have multiple narrow-band inputs and often a single broadband output. The filters reduce the cross-coupling from one channel to another, and this isolation along with the on-channel return-loss are specified. The insertion-loss of each channel input to the output is also a key parameter of the combiner.

Power combiners, the scheme of combining the output of multiple transmitters, are essentially the application of a RF Hybrid circuit. The hybrid has, at minimum four ports, and in the case of power combining, two ports are inputs, one the combined output, and one the isolated port containing a dummy-load. The transmitters are phased appropriately so that the two are combined in phase at the output of the hybrid circuit. The channel return-loss and coupling to the isolated port are key parameters in a hybrid circuit.

Other components that route and switch the RF transmission through the transmitter building are specified by their return-loss.

In high-power operation, if a high VSWR exists in a section of transmission line it is possible to feel hot-spots at points along the component. These will occur every half-wavelength along the component until the VSWR is reduced with a tuning device. It is possible to have a short length of transmission line with high VSWR having low VSWR sections immediately before and after.

TRANSMISSION LINES

Transmission lines may be rigid coax, coaxial cable, or single-conductor waveguide. Transmission lines are by nature, distributed and broadband devices. However, the periodic insulators in rigid line limit the useable frequencies because the reflections of hundreds of insulators add constructively at some frequencies. This fact gives rise to the various standard lengths of rigid line used for different channels.

The key RF parameter for transmission line is the return loss, and typically the insertion loss is computed from attenuation data. The return loss may be easily measured after installation, and the insertion loss is not. TDT techniques may be employed to isolate the transmission line reflections from the other components.

Some components in transmission line systems are not broadband, particularly the elbows, gas barriers, and fine matchers. Reflections from these items may be isolated by TDT gating.

Another common measurement that must be performed at the transmitter site is the relative phase length of multiple transmission lines feeding an antenna. For power capacity or system redundancy motives, an antenna may be fed with two or more transmission lines run from the ground. The relative phase length of these lines is critical in creating the antenna radiation pattern, and their length must be measured and adjusted upon installation. The electrical length may be measured by placing a terminating short (or open) circuit at the junction with the antenna input. The reflection over a significant bandwidth must be measured, computing the electrical length from the phase of the reflection for each line. The line lengths are adjusted to set the necessary phase difference.

ANTENNAS

The RF parameters of an antenna may be measured, and are often monitored with the transmitter instrumentation in conjunction with the other components in the transmission system. The antenna manufacturer measures the VSWR, patterns, and gain of the antenna alone. Once installed, characterization of the antenna VSWR is difficult to separate from the other system components. This may be accomplished using time-domain transform techniques, discussed above.

The azimuthal antenna radiation pattern may be measured using an aircraft to fly around the antenna, recording position and received signal level, or field strength. Though the aerial measurements are not likely to be performed by station personnel, radiation patterns may also be investigated by measuring the received field strength at various locations in the service area, on the ground. Specialized test equipment is used for field strength measurement, such as the Patomic Instruments FIM-71 with standard antennas [2]. To limited accuracy, the antenna gain may also be confirmed from the measured field strength values.

Measuring the antenna VSWR after installation must be performed with a VNA. The time-domain techniques described above may be used to isolate the antenna from other system components. Antennas terminate the RF system, and care must be taken to set the measurement bandwidth and number of points so that the range and resolution of the time-domain data include the antenna. Often transmission line elbows are positioned very near the antenna and cannot be gated-out of the reflection measurement.

REFERENCES

[1] Holt, Tim. "RF Power Measurement Techniques for Digital Broadcast Systems," 56^{th} Annual NAB Broadcast Engineering Conference Proceedings, 2002.

[2] E.A. Williams, ed., NAB Engineering Handbook, 10th Edition, Focal Press, New York, 2007.

[3] Montgomery, Carol G., ed., Techniques of Microwave Measurements, Boston Tech. Lith., MA, 1963, pp. 473-514.

[4] G.W. Collins, *Fundamentals of Digital Television Transmission*, John Wiley & Sons, Inc., New York, 2001.

[5] E.L. Ostertag, "Apparatus for Reducing VSWR In Rigid Transmission Lines," US Patent 5,999,071, Dec. 1999. Andrew Corporation.

RF SIGNAL PERFORMANCE MEASUREMENTS OF CONSUMER FM RECEIVERS AND COVERAGE EFFECTS

John Kean
NPR Labs, National Public Radio
Washington, DC, USA
jkean@npr.org

ABSTRACT

This paper describes a comprehensive study of consumer FM and hybrid digital ("HD Radio®") receiver performance, which required the development of several innovative laboratory techniques. The testing led to the development of an interference model for HD Radio reception. The impact of these measurements on consumer broadcast radio coverage is discussed.

INTRODUCTION

Our study of analog and digital radio coverage required receiver performance data for a variety of consumer receiver types: automotive, indoor and portables. Also, our study required us to predict reception performance under actual conditions, which include interference susceptibility and the effects of environmental RF noise. Measurements of this complexity for consumer receivers are mostly incomplete or entirely unavailable.

Predicting actual reception performance requires that one establish a definition of service quality. For analog radios, this is usually expressed in terms of a signal-to-noise ratio. For digital radios, their audio cutoff at threshold (sometimes called the "cliff effect") is more straightforward. For this study we chose a special meter called a "psophometer" that is intended to match the human hearing system's reaction to low-level noise and interference.[i] While psophometers are relatively unknown in the U.S. this meter is defined by ITU Recommendation 468 and is widely used overseas for audio circuit noise measurements.[ii] To help reference these weighted quasi-peak signal-to-noise ratio (WQPSNR) measurements to consumer perception we conducted a 30-listener test to rate audio samples impaired with specific levels of noise and to determine at what level of impairment consumers would turn off the radio. A full report of this study is a separate paper presented at the NAB 2008 Engineering Conference.[iii]

Measurement Objectives

Our objective was to collect performance data for analog FM in both monophonic and stereophonic modes. Measurements were collected with both analog-to-analog and hybrid-to-analog interference at three signal powers: -50 dBm, -60 dBm and -70 dBm, representing strong, medium and weak receiving conditions, respectively. Three SCA subcarrier receivers of the type used by radio reading services for the blind were also tested.

HD Radio receivers were tested similarly to the analog receivers, but were tested both as analog-only as well as digital receivers. Using threshold digital receive failure as the criteria, the measurement objectives of the receivers were:

- Unimpaired sensitivity and with Additive White Gaussian Noise
- Co-channel interference ratio
- 1st-adjacent channel interference ratio (single and dual)
- 2nd-adjacent channel interference ratio (single and dual)
- 3rd-adjacent channel interference ratio
- Tests of the above with Raleigh fading on either the desired channel, interfering channel, or both

TEST BED DESIGN AND OPERATION

To determine the interference ratios affecting HD Radio reception, as well as other potential impairments, an RF Test Bed was built to measure consumer receivers.

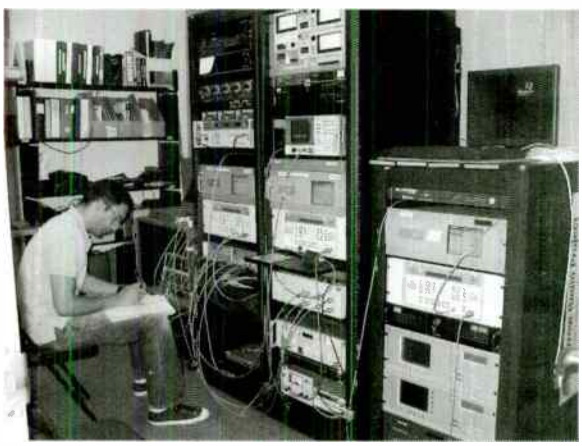


Figure 1- Lab setup for Test Bed

For the hybrid signal generation systems a Harris Dexstar HD Radio exciter was combined with a Hewlett Packard 8647A FM generator. The Dexstar and the FM generator operate at a constant +3 dBm output; with a 20 dB attenuator connected to the Dexstar, the combined output level of the system was approximately 0 dBm. Three of these systems were used; one for the Desired Channel and two for Undesired Channels #1 and #2. All of the HD Radio exciters were connected to a central GPS antenna on the roof of NPR Labs. A 10 MHz GPS-derived signal was used to synchronize all of the other RF equipment.

As shown in the simplified diagram of Figure 2, the signal generator systems were connected to a Hewlett Packard 11759C Channel Simulator, which was used to introduce Rayleigh fading. A two-channel RF attenuator system by Aeroflex/Weischel, Inc. provided control of the RF levels over a 0 to 103 dB range in 1 dB steps.

System levels for the interfering signals were increased by two high-performance RF amplifiers made by Mini-Circuits® so that the maximum input level to the Receiver Under Test was equivalent to 50-ohm dipole field strength of 109 dBuV, producing a D/U ratio of up to -62 dBm. Higher input powers to the receiver would be necessary to cause 2nd- and 3rd-adjacent reception failure with the best receivers, such as automotive units, but the dynamic range of the system was a determined to be a good tradeoff between noise floor and high-level receive conditions.

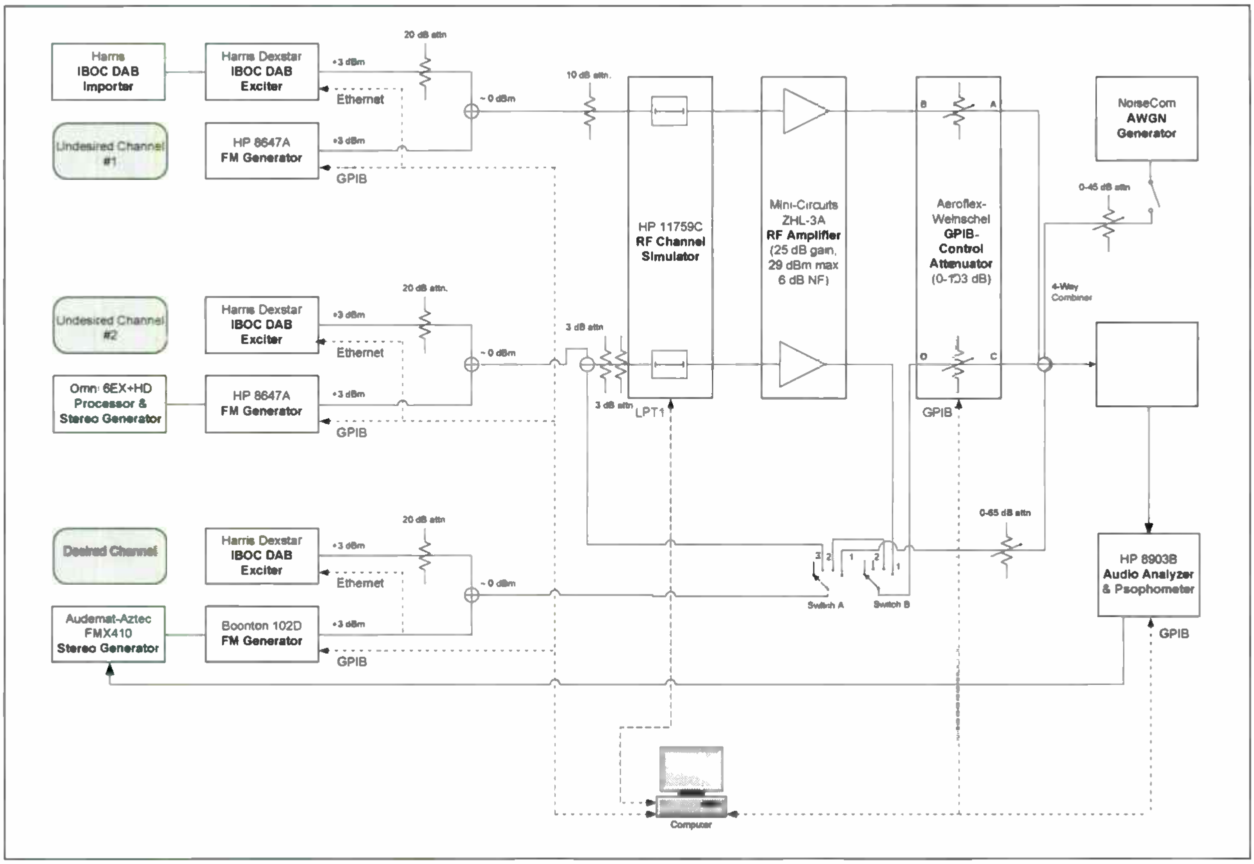


Figure 2- Test Bed diagram

A photo of the RF Test Bed is shown in Figure 1. Not visible is the computer workstation that was networked to the instrumentation. This computer runs MATLAB software to automate the measurements. (It was decided early in the project that the number of tests would be cumbersome to collect manually. As we learned later, as the number of measurement points climbed, automation was a fortunate decision.) An Ethernet switch is connected to Dexstar exciters to control emissions and frequencies and to an Audemat-Aztec FMX480 stereo generator on the Desired Channel generator system. A GPIB network was set up to communicate to the other instrumentation. A Hewlett Packard 8903B audio analyzer was included in the test bed for audio SNR testing and generation of audio test signals under MATLAB control.

To the left of the tall racks, but just visible in the photo beyond our lab intern's knee, is a Ramsey Electronics STE5000 shielded test enclosure. With double-shielded coaxial cables, this cabinet shields consumer receivers from off-air signal leakage and avoids the need for a large screen room.

An Additive White Gaussian Noise Generator, using a NoiseCom NC6110 noise source was available for AWGN impairment tests. Noise levels of 30,000°K and 300,000°K (degrees Kelvin) were chosen for sensitivity impairment tests. These levels were proposed by iBiquity Digital, which commissioned two consulting firms in 1999 and 2000 to determine the noise level caused by co-channel analog interference. The "medium level" background was determined to be 30,000°K, or a field strength of 15 dBuV.[iv] While this noise was intended as a test bed additive for analog compatibility testing of IBOC DAB by the NRSC, AWGN seems a good model for a large number of environments in which to test HD Radio receivers. For comparison, a level 10 dB higher (300,000°K) was included.

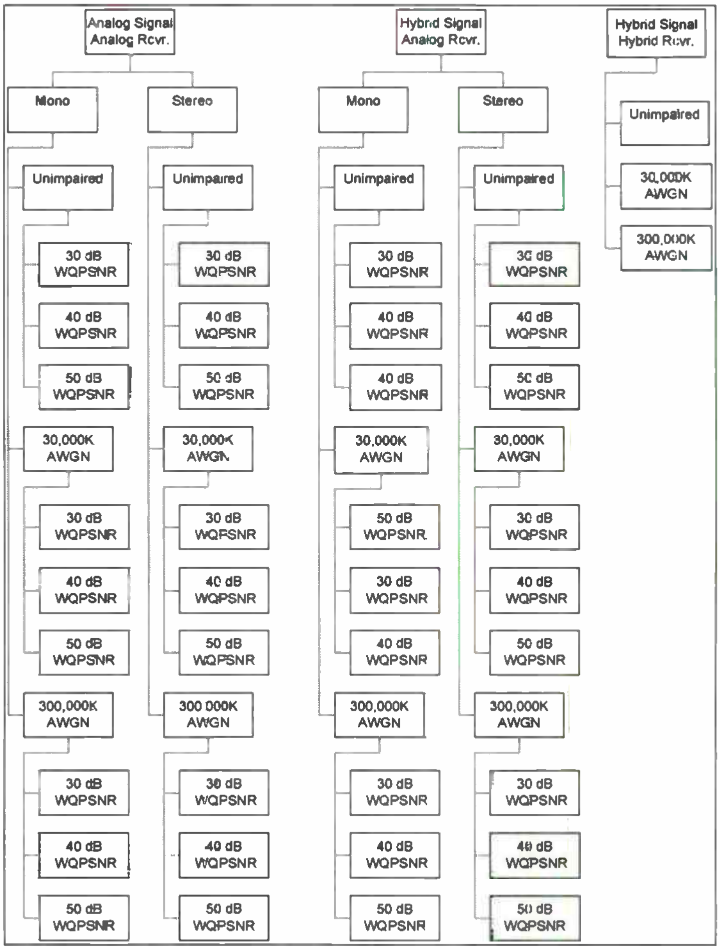


Figure 3- Sensitivity measurement chart for analog and hybrid receivers

The process to collect sensitivity measurements for the analog and hybrid receivers is shown in Figure 3. Three paths are shown. On the left, the analog FM receiver is tested in both monophonic and stereophonic modes, first with unimpaired (no added noise) and then with 30,000°K and with 300,000°K AWGN. The automated Test Bed targeted 30 dB, 40 dB and 50 dB WQPSNR for each impairment condition, automatically zeroing in on the RF level for each target SNR. Identical measurements were collected with the analog receivers using hybrid signals on the desired channel. Measurements are recorded in dBm, referenced to 50 ohms at the antenna input to the receiver. We did not attempt to measure the actual complex impedance of the receivers' input.

Automated sensitivity tests, shown on the right side of Figure 3, are also performed with unimpaired conditions, 30,000°K and with 300,000°K AWGN using the receiver's audio turn-on threshold.

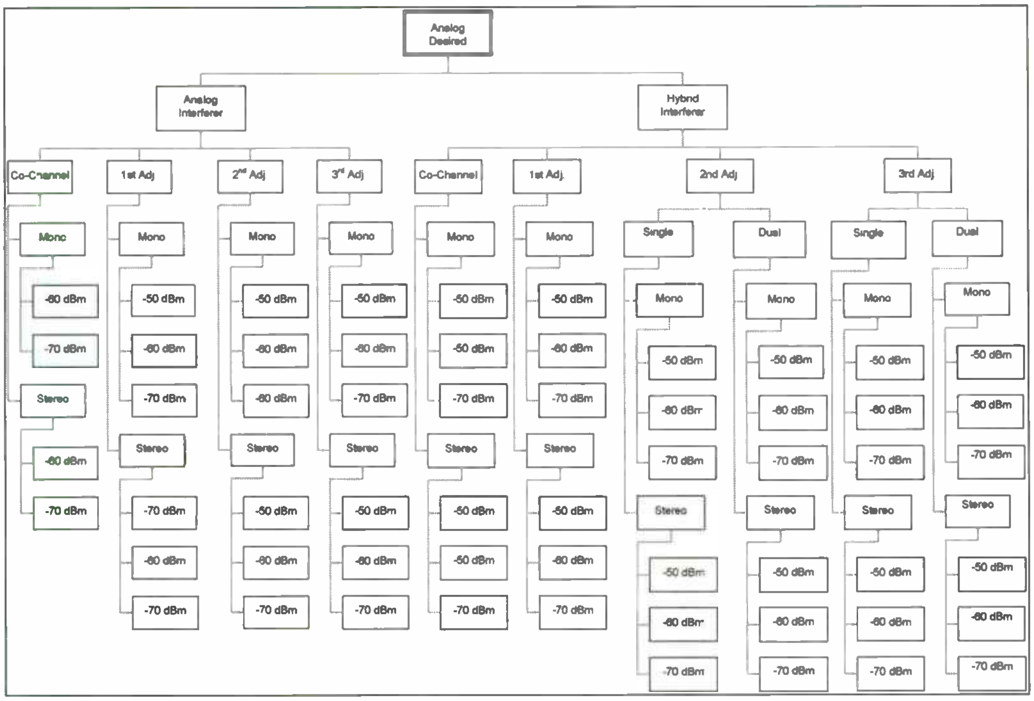


Figure 4- Analog-to-analog and hybrid-to-analog interference measurement chart

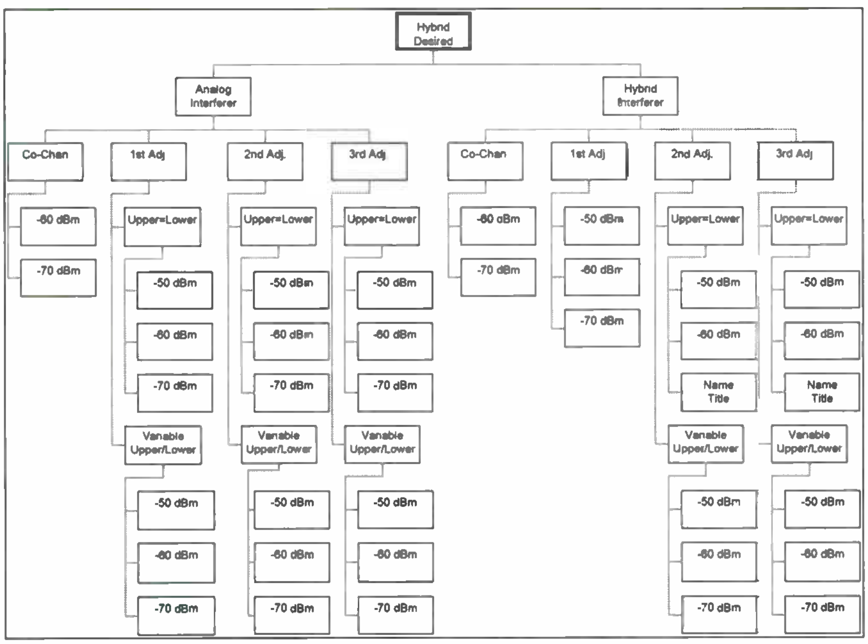


Figure 5- Hybrid receiver interference chart

Figure 4 shows the interference measurement process for analog desired signals. Analog-to-analog and hybrid-to-analog test are similar, except that cochannel analog interference was tested at the two lower signal powers (it would not be probable to receive cochannel interference at a high desired signal power), and hybrid 2nd- and 3rd-adjacent interference was tested with both single and dual carriers.

The process for testing hybrid receivers with a hybrid desired signal is shown in Figure 5. Both analog and hybrid interferers are tested identically. The 1st-, 2nd- and 3rd-adjacent channel tests are also performed with a combination of varying levels on the alternate upper and lower interferers. The reasons for this are explained in the next section. All of the measurements were automatically stored in separate Excel® spreadsheets with 11 graphs to display each receiver's performance data.

Resolving Receiver Behavior

Interference measurements for some of the receivers were surprisingly erratic. Figure 6 shows a typical 1^{st}-adjacent interference behavior from an analog home stereo. The monophonic (upper) curves indicate a gradual decline in WQPSNR as the D/U ratio declines, to the left. The stereo curves also show a gradual decline in WQPSNR until 10 dB D/U; below which the WQPSNR actually increases as the radio switches from stereo to monophonic reception. This and other noise suppression tactics, such as stereo blending, soft muting, and high-frequency roll off may result in two (or more) D/U values having the same WQPSNR value. This was a challenge for our original convergence algorithm, which attempted to "zero in" on a D/U value for a target WQPSNR.

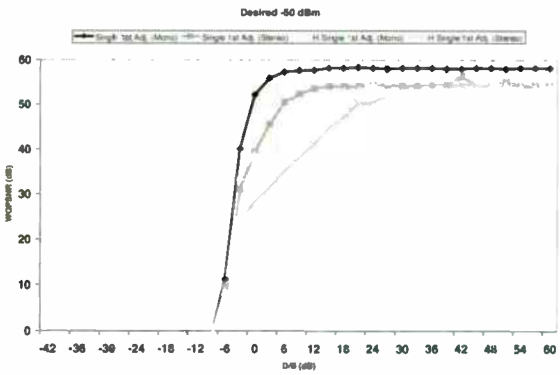


Figure 6- Routine mode-switching in cochannel D/U

A case of stranger behavior is shown in Figure 7, in which the WQPSNR values decline to 0 a D/U of around -10 dB, but then fluctuates upward at lower D/U ratios, eventually shutting down the audio output. We observed this on several HD Radio units operating as analog receivers, which led us to conclude that the digital signal processing in these radios was causing these effects. To produce accurate readings we rewrote the MATLAB® software to take readings every 3 dB across the entire 103 dB D/U range, rather than converge of a target WQPSNR value. A side-benefit of this was graphs of interference behavior collected for every receiver.

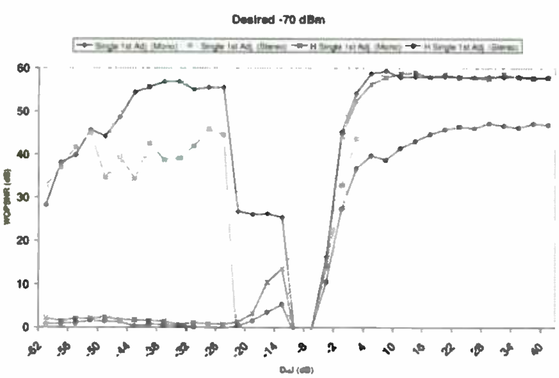


Figure 7- Irregular 1st-adjacent D/U behavior

Receivers Tested

The 25 receivers listed in Table 2 were supplied by the Consumer Electronics Association; the rest were purchased by NPR Labs, including the HD Radio receivers in Table 1. Not shown are several units with anomalous performance, which were disqualified from the tests.

Table 1- HD Radio receivers tested

Receiver Category	Brand	Description
auto adapter	AGT/Visteon	HD Zoom HDZ300
auto adapter	Directed Electronics	DMHD-1000
Car in-dash CD	JVC	KD-SHX900J
Car in-dash CD	Kenwood	KTC-HR100TR
Car in-dash CD	Panasonic	CQ-CB9900U
Car in-dash CD	Sony	XDR-S3HD
Auto/home transport.	AGT/Visteon	HD Jump HDP250
component tuner	Sangean	HDT-1
tabletop	Polk	I-Sonic
tabletop	Radiosophy	HD100

Table 2 - Analog Receiver Tested

Receiver Category	Brand	Description
Home stereo	Sony	STRDE697
Home stereo	Yamaha	HTR-5740
Home stereo	Denon	DRA-295
Home stereo	Denon	TU-680NAB
Home stereo	Pioneer	VSX-D814K
Shelf/mini system	Panasonic	SC-EN7
Shelf/mini system	Sony	CMTNE3
Shelf/mini system	Bose	Wave
Shelf/mini system	RCA	RS23035
Portable CD boombox	Panasonic	RXFS430A
Portable CD boombox	Aiwa	JAX-S77
Portable CD boombox	Grundig	S350
Portable CD boombox	GE	SuperRadio III
Portable CD boombox	CCRadio Plus	CCRadio Plus
Portable CD boombox	RCA	RCD147
Car in-dash CD	Pioneer	DEH-P6600
Car in-dash CD	Kenwood	KDC-3025
Car in-dash cassette	JVC	KS-FX490
OEM auto	Chevrolet	1995 Camaro
OEM auto	Chevrolet	2000 Tahoe
OEM auto	Chevrolet	2002 Suburban
OEM auto	Ford	2002 Mustang
OEM auto	Honda	2002 Accord
OEM auto	Honda	1997 Civic
Clock	Audiovox	CE256
Shelf/mini system	Panasonic	SA-PM19

Measurement Results for Analog Receivers

The consumer ratings of audio noise impairment, reported separately, indicated that a 35 dB weighted quasi-peak signal-to-noise ratio is needed to provide a quality listening experience and to maintain listeners in a quiet indoor environment such as a living room. Table 3, Table 4 and Table 5 list the required D/U ratios required for mobile, indoor and portable receivers, respectively, at a 40 dB WQPSNR. Due to the performance of some receivers and constraints on the minimum D/U ratio that could be generated by the test bed (approximately -55 dB), some 2nd- and 3rd-adjacent channel measurements are estimated.

Table 3- D/U ratios for mobile receivers

40 dB Stereo WQPSNR	Analog -60 dBm	Analog -70 dBm	1% IBOC -60 dBm	1% IBOC -70 dBm
Cochannel	34	31	34	31
1st -Adj.	-9	-8	12	10
2nd -Adj.	-51*	-57	-50*	-57
3rd -Adj.	-51*	-60	-50*	-57

Table 4- D/U ratios for indoor receivers

40 dB Stereo WQPSNR	Analog -60 dBm	Analog -70 dBm	1% IBOC -60 dBm	1% IBOC -70 dBm
Cochannel	34	34	34	34
1st -Adj.	1	0	13	12
2nd -Adj.	-45	-46	-32	-36
3rd -Adj.	-47	-48	-38	-42

Table 5- D/U ratios for portable receivers

40 dB Mono WQPSNR	Analog -60 dBm	Analog -70 dBm	1% IBOC -60 dBm	1% IBOC -70 dBm
Cochannel	21	21*	18	16
1st -Adj.	4	-1	0	-1
2nd -Adj.	-38	-42	-22	-22
3rd -Adj.	-42	-46	-35	-41

While a lower audio SNR may be acceptable for mobile and portable reception conditions, it should be noted that the effects of joint, independent fading on the desired and interfering channels will result in large sudden increases in noise (large drops in audio SNR). Thus, a higher standard of audio SNR is appropriate for predictions of interference-limited coverage, based on median (50th percentile) location variability.

Measurement Results for Hybrid Receivers

To date, a total of 17 HD Radio receivers have been measured on the Test Bed, including those listed in Table 1. The types represented are: after-market car radios, home hi-fi tuners, professional tuner/monitors, and table radios. (Portable HD Radio receivers are not yet available.) Our project grant contract restricts us from publicly releasing all the receiver test data until after review of the final report, which follows the deadline for this paper. For this report, we present the major results of the Kenwood KTC-HR100MC after-market car radio, one of the best-performing receivers and the basis of the NPR HD Radio Logger measurements, discussed below. The results were obtained with hybrid (HD Radio) interferers operating service mode MP3 (including two additional Extended Hybrid Partitions).

Figure 8 plots the ratio of co-channel desired and undesired signals at which digital receive failure occurs. It is apparent that for unfaded (steady-state) signals the D/U ratio remains relatively close to 4 dB over a wide range of desired signal powers. With a "Trimmed Urban Fast" Rayleigh fading profile (60 km/hr) used by the NRSC, and produced by the Hewlett Packard RF Channel Simulator shown in Figure 2, the cochannel measurements required an increase of approximately 3 dB in the D/U protection ratios, increasing gradually at lower signal powers. This increase is due to the inclusion of 30,000°K additive white Gaussian noise in the test bed,[v] which raised the required signal margin for a given data error rate.

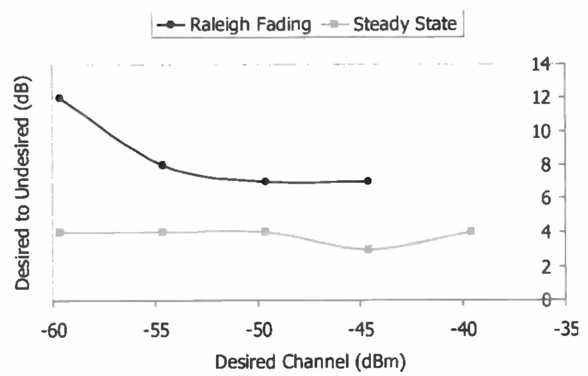


Figure 8- Cochannel interference ratio for steady and Raleigh fading signal.

Figure 9 plots the ratios of 1st-adjacent interference necessary to cause digital receive failure with the Kenwood tuner. The upper (89.9 MHz) and lower (89.5 MHz) adjacent channels are shown for a Desired Channel frequency was 89.7 MHz with steady signals. The D/U ratios remain relatively constant, around -14 dB ±2 dB, over a very wide signal range. This receiver in test also showed good symmetry in upper and lower channel interference ratios, which was a valuable factor later in choosing the receiver for field measurements with dual (upper and lower) interferers.

For reference, the FCC's allocation rules require a D/U ratio of at least 6 dB for 1st-adjacent interfering stations at the 60 dBu service contour of a protected station. This particular aspect of IBOC receiver performance is approximately 20 dB more tolerant of interference than the FCC's standard for analog FM. It is also slightly better than the -9 dB D/U ratio measured for vehicular receivers, as reported in Table 3.

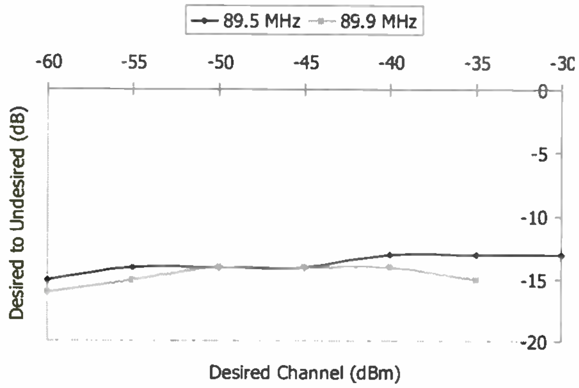


Figure 9- Single 1st-adjacent channel interference ratios for upper and lower channels.

Broadcast engineers are aware that HD Radio transmits data simultaneously in upper and lower data carrier groups, which extend from approximately 129 to 199 kHz on each side of the FM carrier frequency (for the P1 mode of transmission). Interference to either carrier group alone by an adjacent channel interferer has a minor effect due to selective combining of the data from the two carriers.

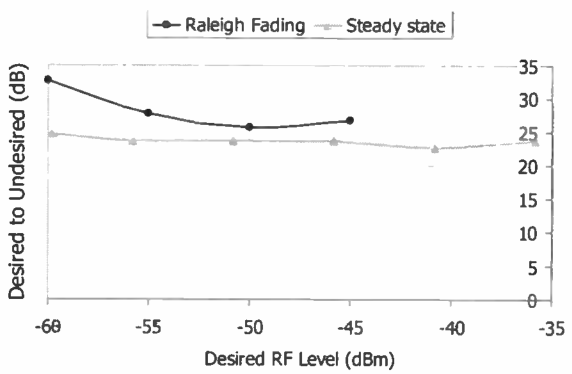


Figure 10- Ratio of dual 1st-adjacent channel interference at equal level, relative to a Desired signal.

The IBOC receiver's performance changes markedly with interference occurs to both upper and lower carrier groups. Figure 10 plots the results of interference from dual adjacent channel analog interferers (one above, one below the desired channel) having equal power. The Desired Channel signal was constant for the Raleigh fading test.

The frequency of an adjacent-channel FM carrier does not fall on the frequencies occupied by the IBOC carriers of the desired channel. However, the modulation sidebands of the interfering station do cross through the desired IBOC carriers. Consequently, the form and extent of the modulation will affect the level of adjacent-channel IBOC interference. We experimented with normal broadcast audio, processed by a Telos Omnia 6EX-HD, to determine a stable and repeatable test signal for interference ratio testing. For these tests we used a 1 kHz sinusoidal modulation on the adjacent-channel generators that was set to provide interference ratios that correlated with the processed audio results.

The graph of Figure 10 is similar to Figure 8, except that the D/U scale is shifted from around -15 dB to nearly +25, or approximately 30 dB higher protection ratios than with a single adjacent interferer. This test is more to demonstrate a point of dual-adjacent interference, as the likelihood of identical levels is remote.

A more probable condition for interference is illustrated in Figure 11, in which threshold interference ratios are plotted as a function of both adjacent interferers. Considering both on a two-dimensional plot requires a unique presentation: the graph lines show the signal ratio of the Desired Channel to the *stronger* adjacent channel interferer on the x-axis *versus* the ratio of the *weaker* to stronger adjacent channel interferer on the y-axis. Tests were performed at Desired Channel signal levels of -60 dBm and -50 dBm (medium and strong signals) and show good agreement.

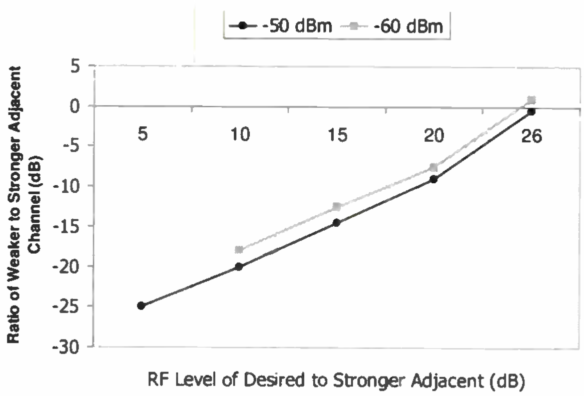


Figure 11- Dual 1st-adjacent channel interference ratios with desired channel at -50 dBm and -60 dBm.

The data shown here was incorporated into a technical model for prediction of HD Radio coverage and interference. The model has been extensively tested at 10 public radio stations specially selected to represent a range of facility sizes and topography. The results from thousands of miles of drive-test data provided a confirmation of the model, which is reported in a separate paper.[vi] Due to the model's simplicity and accuracy a provisional patent for the model was filed with the U.S. patent office in 2007.[vii] As part of the project for the Corporation for Public Broadcasting, which funded much of this research, NPR Labs is now completing coverage maps of more than 850 CPB-qualified radio stations in the United States.

ACKNOWLEDGMENTS

Thanks to our interns Sam Goldman for developing the original MATLAB software to automate the Test Bed, and Babak Monajemi for optimizing software and conducting many of the receiver measurements. Our thanks to Harris Broadcast for loan of some test instrumentation, and our appreciation to the Corporation for Public Broadcasting for the vision to fund this *Digital Radio Coverage and Interference Assessment* project.

REFERENCES

---

[i] *The Assessment of Noise in Audio-Frequency Circuits*, BBC Research Report EL-17, 1968.

[ii] *ITU-R 468* is a standard relating to Noise measurement currently maintained by the European Broadcasting Union, which took it over from the CCIR. It was previously CCIR Recommendation 468-4.

[iii] *Consumer Ratings Of Impaired Audio At Varying Signal/Noise Ratios*, Ellyn Sheffield, John Kean, Dan Schwab, NAB 2008 Engineering Conference.

[iv] *The FM Broadcast Band: Service Area Noise Floors in the US*, iBiquity Digital Corporation, (presented to NRSC TPWG 11/09/2000).

[v] The use of 30,000 °K AWGN to represent noise pollution in the FM band, primarily due to cochannel interference, was advanced by iBiquity Digital Corp. in *The FM Broadcast Band: Service Area Noise Floors in the US*, presented to the NRSC TPWG on Nov. 9, 2000.

[vi] *An Improved Coverage Prediction Method for HD Radio*, to be presented by the author at the NAB 2008 Engineering Conference.

[vii] *A Method for Determining the Usable Signal Coverage of In-Band On-Channel Digital Audio Broadcast Transmissions*, UPTO # 60984421, Nov. 1, 2007.

Monitoring and Measurements in the Broadcast Plant - TV

Wednesday, April 16, 2008
9:00 AM – 11:00 AM

Chairperson: John Merrill
CBS 5 KPHO-TV, Phoenix, AZ

Optimizing the QC Process in a File-based Workflow Facility
Rob Zwiebel, Harris Corporation, Pottstown, PA

System Wide Video Quality Assurance
Ralph Bachofen, Triveni Digital, Princeton Junction, NJ

Controlling and Measuring Loudness for Digital Television Broadcast
Michael Babbitt, Dolby Laboratories, San Francisco, CA

***Reducing the Effects of Bit Errors in Serial-Digital Interface Links**
Paul Briscoe, Harris Broadcast Communications, Mason, OH

*Paper not available at the time of publication

OPTIMIZING THE QC PROCESS IN A FILE-BASED WORKFLOW FACILITY

Robert C. Zwiebel
Harris Corporation
Pottstown, PA

ABSTRACT

Ensuring content quality in a multi-channel explosion has become increasingly challenging for broadcasters as compressed digital content may be comprised of differing standards, formats. and resolutions. Automated media analysis server technology provides broadcasters with a more efficient, consistent, and cost-effective method of verifying content of all popular formats and parameters. The need to support many file formats (MXF, QuickTime etc...) and compression standards, including but not limited to MPEG2/IMX 50, MPEG-2, DV25, DVCPRO 50, DVCPRO 100, VC1 (WM9) and MPEG4-H.264, necessitates a content verification tool that functions irrespective of ingest format.

While traditional stream monitors test for compression and syntax errors, a robust media analysis server checks visual (pixel-by-pixel) and aural (sample-by-sample) quality assurance (QA) parameters, including video and audio levels, video color space compatibility, audio phase, low audio, and letter and pillar box limits. By automating the quality assurance process at rates faster than real time, Quality Control(QC) productivity increases providing broadcaster's confidence in the quality of their digital assets.

THE QC WORKFLOW CHALLENGE

One of the biggest challenges facing today's technologists is developing an effective workflow that incorporates both existing and new technology into a streamline process designed for efficiency and complete fulfillment of internal business objectives. Emerging technology and mandated transitions, such as the move to all digital television, forces the broadcaster to redesign infrastructures, and discern over a workflow that remains viable for the present as well as the future. Choosing the right format and associated equipment can be a daunting task. Equipment interoperability issues, such as differing file structures and file formats, forces compromises to be made; specifically, leading-edge features may not be available throughout the entire data path. In some cases, ingested files must be converted throughout the QC process specifically to match equipment input capabilities. This process increases the amount of time necessary to complete the validation of the content and introduces additional failure points throughout the data chain.

The QC process must be established and proven valuable in assessing the overall quality of the product. Essentially this objective of producing quality programming has not changed as the supporting technology evolves. It further burdens the broadcaster to adapt to continue to deliver the highest level of quality possible to the end user. In order to accomplish this goal, QC operators must be trained in the process, the process must be consistently executed, and various factors producing inconsistent results such as human fatigue must be eliminated.

Executing an effective QC process involves time and resources. Additional complications arise as computer dominance and inexpensive storage allow content to reside within the digital domain as apposed to exclusively in tape-based storage. Storage being at a premium. content must be compressed to save space.

Currently, content is immediately ingested and stored as files and remain in file format throughout its life cycle. Unless some process or action is enacted on the file, there is no existing means to determine the quality of data found within the file. At a minimum, the file would have to be played out using any readily available means, and verified by completely watching the encapsulated content, in real-time, for the entire duration of the segment. This is a major challenge to the QC workflow, assessing the quality of a file or files that exist as digital information.

This new process is completely file-based. Media has now been transformed into a format best suited for the broadcaster's methodology for distributing audio and video content. This content, existing in containers used to store machine-readable information, is certainly multifaceted. A broadcaster's fundamental function is to distribute audio and video, the simplest form of a file in its entirety, would contain only raw essence. Essence in this case is defined as broadcast quality audio or video.

Raw essence may be the entire file or only part of a file. Multi-essence may be grouped to form logical associations as the file complexity increases. As an example, a wrapped file may include compressed video and multi-tracks of audio. For example, Material Exchange Format (MXF), as defined by SMPTE 377M-2004, is a container format used to associate essence with Metadata and can contain single or ganged packages with various item complexities. QuickTime, developed by Apple Inc., is a framework designed specifically to handle various media formats and images. MPEG-4 Part 14, is another multimedia container format.

The file wrapper functions as a means to bind information. The information found in the wrapper contains content as well as Metadata. Metadata provides supplemental information about information contained in the file.

The eliminatory steps in the QC process remain unchanged. At a minimum, assets must be received into the system logged or associated by some means as to be identified for the future; or, when required, it can be called up for some later use. It must move through some quality assessment. After all, why bother to store it in the system if the actual content is unusable? Finally, it must be archived for later retrieval.

However, the complexity level of the process increases dramatically as files now represent complex entities necessitated by the designed workflow. The outcome of the QC process must be that a file is processed through a system that accurately and effectively ascertains the quality of the file against a user defined criteria. A versatile process can be developed that works equally well for all combinations of files and essence.

CONTENT ASSESSMENT, A MORE EFFECTIVE MEANS

Imagine that the quality of a file was known prior to an action taken place, whether that action was playing a file, transferring a file, or archiving a file. Media could be confidently distributed without concern.

In reality, content can arrive in any state ranging from being utterly unusable to completely acceptable in accordance to the established QC criteria.

To revisit the existing QC process, content is received, and the quality of the file and associated content must still be ascertained, but now the QC operator is aided with a Media Analysis Sever (MAS). The MAS is a tool developed specifically to augment the QC process by providing a computer-based analysis of file-based content. The QC operator task still centers on determining the overall quality score of the content, but with the addition of the MAS, the system would scrutinize all aspects of the file, layer by layer, generating customizable metrics in user-defined outputs.

A general overview of the MAS's functionality can be broken down into a four basic stages:

- Process all formats (Handle all file types)
- Break the file into elements (Audio/Video/Data)
- Analyze each element
- Report Findings

When it comes down to it, the format that was chosen for the actual essence is not important. Each of the potentials have their merits. It is important that, regardless of the selection, the QC process can be carried forward with confidence that the quality of the media can be validated.

This certainly applies to the QC process as a versatile process and must work equally as well for all combinations of files and essence. In effect the QC process should be one that can be easily and repeatedly executed, always achieving the same results while utilizing minimal resources.

How does the QC operator asses the quality of the package or wrapper, which in this case includes not only the content but also the delivery mechanism for the content? After all, the file wrapper is nothing more than a binder associating content and Metadata.

At the file level, the Media Analysis Sever must interpret the file as a whole by providing feedback to the end user, in this case the QC operator, as to status of the file. If the file is defective or incorrectly generated, the content could be inaccessible. So if files can be received in various wrapper formats such as (MXF, QuickTime etc...), the tool must provide a mechanism to inform the user as to what type of files are found and what contents are contained within file.

If it is expected that the media must be received as a QuickTime reference format and that there should be one video track and two associated audio tracks, it can be readily determined that the file was correctly constructed through a reporting mechanism. Extracted from the wrapper, we find the individual components that make up the file. Depending on what is contained within the file, the MAS must identify and process these elements. Essence breaks down into data contained within the file.

A compression standard, MPEG2/IMX 50, MPEG-2, DV25…, for HD or SD content has been applied prior to encapsulating the content into the file. It is equally important to verify that the video can be uncompressed to verify that the compression settings and parameters were setup correctly during the compression process. Once the video is uncompressed, the quality of that content still has to be determined.

Transmission aspects of the essence (how will the video look and how will the audio sound after decode) should be done at a baseband level. For video this equates to the pixel, and for audio this equates to the sample. For a QC operator to access the quality of video/audio using these basic elements is impossible.

In active video a 1080i/60 video format has well over two million pixels. SMPTE defines standards that can either be directly converted or derived into analysis criteria that are used against the data to deterministically conclude if the standard is in violation.

Using a piece of test and measurement equipment, the QC operator would have to monitor rasterized signals at display extremes in order to detect errors.
It is nearly invisible with the naked eye, using common test and measurement equipment, to visually detect chroma or luma excursions especially at any distance away from the display.

This unwarranted consumption of time is absorbed by the robust tool which checks visual (pixel-by-pixel) and aural (sample-by-sample) quality assurance parameters at faster than real-time rates.

The Media Analysis Server is transformed into a valuable test and measurement device allowing actual decoded content to be completely verified. This includes details such as video and audio levels, video color space compatibility, audio phase, low audio, and letter and pillar box limits.

Subject screening analysis can also be determined. This includes field dominance, blockiness, blurriness, and freeze detect. The summation of all results would be effectively communicated via any currently available means (a Pass/Fail assessment of the quality of the file).

The results of the analysis can be included as Metadata, if support by the chosen container.
Standardizing Metadata fields would allow any program to immediately determine the quality of the file and the associated date of assessment. The results may also be exported in a report format for historical bookkeeping.

With the technical side of the QC equation confidently determined, the QC operator is then free to concentrate on evaluating content by watching and listening, as an end product. A file labeled Mayberry R.F.D., Season 1, Episode 2, has content with associated audio. An ISCI code is used to partially verify the contents of the tape. If labeled correctly there is a means to identify the associated content. In file-format, the file name could easily be changed as it moves throughout the workflow, this leads to the problem of associating a file name with actual content. The machine can verify the technical side of the equation, but not the intention behind file.

This does not eliminate the human element. Instead, it compliments it, because certain elements of the QC process must be performed by humans and a certain percentage of the process must be performed by machine allowing greater accuracy and repeatability of the process. As QC productivity increases the broadcaster's confidence, the quality of their digital assets increases. The QC process is now represented in Figure – 1.

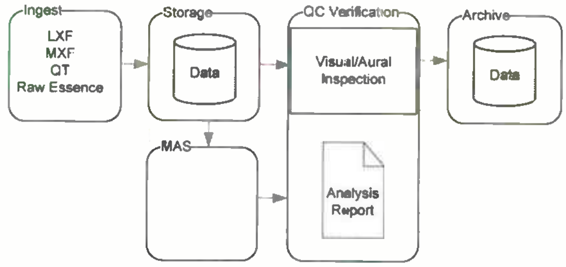


Figure - 1

As a side note, if the MAS has the capability of detecting problems, it would only seem logically advantageous to the process if there was a means to instantiate some form of correction for the detected errors. Valuable time can be saved if the tool performs failure recovery in-house. There is no lost time sending content back to the originator for rework. Even if there are edit suites in the facility, it still consumes time to schedule the rework and if the time to air deadline is drawing near, this will cause undue anxiety and unwarranted stress.

For example, working backwards from the video frame, the tool would combine video legalizers with video proc amps to audio limiters to compensate or eliminate detected excursions. Audio clicks and pops could be eliminated. Loudness could uniformly be set across audio tracks. Encoding headers may be corrected to match actual video content. The enhanced MAS with correction details is shown in Figure – 2.

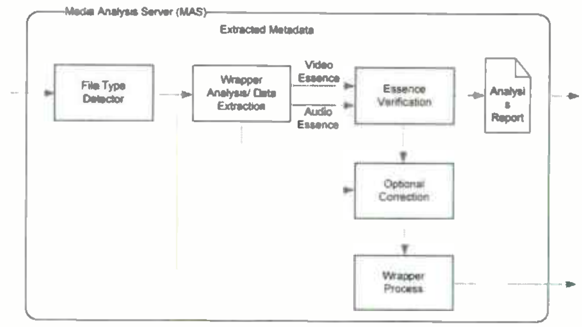


Figure - 2

AUTOMATING THE PROCESS

The MAS has the capabilities of analyzing content off-line allowing QC operators to begin their daily task with a clear understanding as to the QC issues that must be addressed. If there are no obvious failures, such as loss of audio, content may be spot checked thus completing the QC process. This again is dependent on the established QC process. As previously stated, there must be some means to verify that the actual content matches the expected content. The amount of time that the QC operator spends reviewing content diminishes, but cannot completely be eliminated.

Digital Asset Management (DAS) software can be used with the MAS to automate the process. As content is introduced into the system, the QC process can be triggered automatically from the automation system. With a powerful interface the automation system can even vary the parameter of the QC process, allowing content specific detailing to be introduced at the beginning of the cycle. After the completion of the analysis, the automation system can redirect the file into a quarantine container. The file must be further scrutinized by performing a detailed analysis or moved into a passed location, one where the rest of the QC process can be completed knowing that the technical side of the analysis has been completed without issues.

Through effective communications, SNMP, or Email, the user remains informed as to the status of the file.

CONCLUSION

To increase the productivity of the file-based QC process, a certain percentage of the task must be off loaded to a highly specialized tool set which aids the QC operator in ascertaining the ultimate quality of the product. The right tool can identify flaws in content previously undetectable in a manual process. This nets a result of reducing man hours necessary to complete the process in ascertaining the overall quality of the digital asset.

With the use of automation, workflow can be tailored allowing valuable time to be spent on verifying content of prescreened passed files as opposed to spending little or no time on files with apparent defects. Process wise, anything that we can do to improve the time involved in QCing when time is critical is not just helpful it is necessary.

SYSTEM WIDE VIDEO QUALITY ASSURANCE

RALPH BACHOFEN
RICH CHERNOCK
Triveni Digital
Princeton Junction, New Jersey

ABSTRACT:

Digital television broadcast has become quite complex. Ensuring the quality of the broadcast is necessary for viewer satisfaction and retention. Additionally, business pressures are forcing consolidation of operations and expertise for broadcast station groups. A consequence of this consolidation is the need to be able to assure video quality for an entire enterprise from centralized locations – either because operations or expertise have become centralized.

This paper introduces an architecture for centralized system wide quality assurance. Components of this architecture include:

- Monitoring: Unattended comparison of stream parameters against pre-established rules
- Troubleshooting: Active analysis of stream conditions to determine causes of impairments
- Consolidation: Collection of system-wide conditions into a usable form
- Communication: Generation of reports and identification of targeted alarm responses
- Configuration: Centralized configuration of all monitoring and analysis systems

With a system wide video quality assurance strategy, impairments can be found and corrected more quickly. The overall effect of having an integrated view of monitoring will be a higher quality service, with a lower cost of operation.

INTRODUCTION

What matters most to a television viewer is the ability to enjoy their preferred programming or service without glitches. Customers don't really care about technical details, such as MPEG or 8-VSB. The only way to satisfy today's television viewers is to provide solutions which 'Just Plain Work.' When viewers encounter difficulties such as audio lip sync, blocking, or black screens, they turn to other channels. Viewer satisfaction suffers when customers find issues with transmission before engineers at the studio do. Therefore, it is imperative that television engineers find and fix network, encoding, and transmission problems before their customers become aware of them. One of the top goals for a DTV broadcast installation is to provide a defect free, uninterrupted viewing experience to the customers of the service. The discussion below outlines some of the key considerations for ensuring a quality experience for the viewer.

Most of the approaches to video quality to date have been of the reactive variety. Problem detection is driven by viewer calls, which then kick off the steps of fault localization, fault analysis and isolation, and fault remediation. Many broadcast stations across the country use set top boxes and video monitors to confirm that they are 'on air.' With an analog signal this might have been sufficient; but digital broadcasting introduces another element to the monitoring equation - software. Every digital set top box (STB) has software running on it. Depending on the implementation of the software in a specific STB, that receiver may react differently to a specific non-compliance in the bitstream. Problems that affect users of one type of STB, may not be visible to users of another brand of set top, or even a later model. This approach to monitoring is clearly not recommended.

The ability exists today to operate in a proactive mode – identify problem areas before viewers do and then take actions to cure the faults before calls arrive. A comprehensive monitoring architecture, coupled with comprehensive troubleshooting/analysis tools, allows this shift to a proactive mode. At the same time, broadcast operations have become more complex, often growing beyond the single station/transmitter model to distributed transmission and centralized control. Distributed monitoring solutions have evolved to address this additional complexity, retaining the ability to determine system health from a centralized location.

SERVICE CONCERNS E.G. WHAT CAN IMPACT YOUR VIDEO DELIVERY IN YOUR NETWORK

Common examples of audio and video aberrations that can result from bitstream issues include video tiling, audio lip sync errors, and intermittent tuning. Without stream monitoring, it is not easy to identify the specific cause of a problem. For example, a given tuning issue may be due to dropped packets, or due to metadata errors in PSIP and MPEG tables. Another prevalent problem is known as "Tiling" (macroblocking or pixelization). Tiling can come from a number of root causes, such as PCR Jitter, Video buffer under/overflow and under provisioning – created by any of the MPEG affecting devices in the broadcast chain, for example a multiplexer. In order to understand which device is causing the problem, it is often necessary to examine the MPEG layer comprehensively and in real-time. Basic or off-line MPEG analysis may not be able to isolate the fault.

Without a monitoring system, the engineer's only tool is trial and error. Monitors help engineers find out where a problem originates in their network by tracing viewer problems to a specific deviation from the applicable standards. Once the bitstream monitor identifies a problem in the stream, the monitor can be used to find out which device in the network introduced the fault. Monitors, unlike STBs, can be equipped with multiple input types. Engineers can monitor the RF using 8-VSB and then examine the output of a MUX using SMPTE 310 or ASI.

CENTRALIZED ARCHITECTURE OVERVIEW (WHY IS THE SERVICE VIEW SO IMPORTANT)

Digital Television broadcast architectures span a wide range of complexity from single stations through station groups with centralized operations to large station groups that have both centralized and regionalized operations. As seen in Fig 1, below, the content flow within a single station is relatively simple. Strategies for deploying monitoring equipment aimed at fault detection and localization are equally simple.

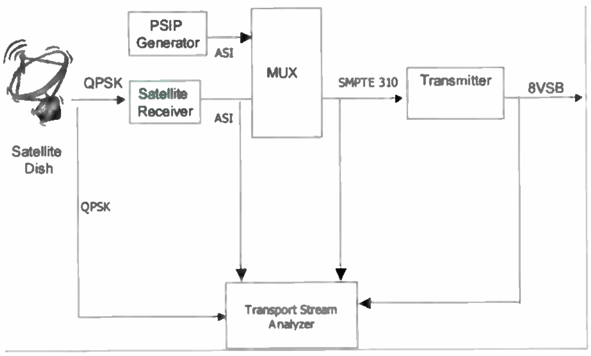


Figure 1 Single Broadcast Station

Figures 2 & 3 illustrate more complex examples. Figure 2 shows station groups with centralized operations: transmitter sites geographically scattered; but with a single, centralized "Network Operations Center" (NOC).

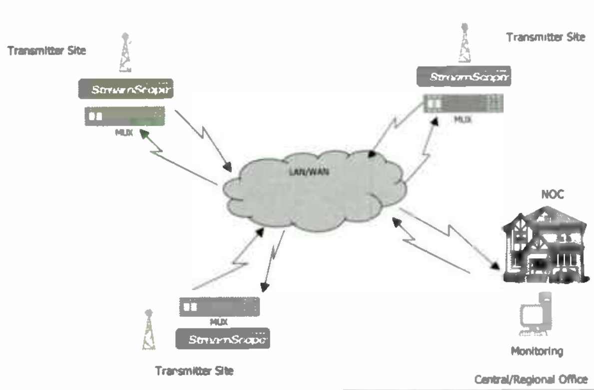


Figure 2 Centralized Station Group

Figure 3 shows an extension of this concept to a single centralized NOC with a system-wide view; coupled with a number of regional NOC's, each responsible for operations within their own region.

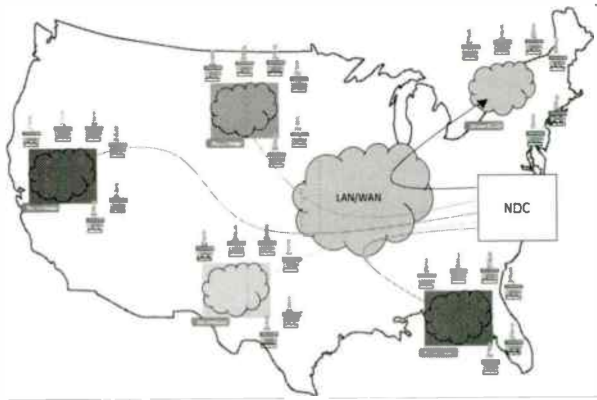


Figure 3 Centralized/Regionalized Station Group

For either of these situations, in addition to the ability to obtain monitoring information at the broadcast locations, there's a need to be able to view the status at a centralized NOC and regional NOCs. Increasingly, because of budget restrictions, personnel with the ability to understand and deal with service impairments are being consolidated at the operations centers, rather than being located at the edges. A consequence of this is that to be most useful, the information must be moved to these operations centers and consolidated in a way that allows quick understanding of the health of the overall system and rapid localization of any trouble points.

STRATEGIC VIDEO MONITORING POINTS

Real-time comprehensive monitoring is the main approach to enabling proactive fault detection. The goal of this approach is to be able to determine when faults occur within the distribution system, localize problem source to candidate equipment, analyze and isolate faults, and initiate a remediation process. With properly placed monitoring units, which are examining the right parameters and comparing against rules that make sense, proactive problem detection can take place.

While an ideal monitoring strategy might involve placing monitors at all points that manipulate the digital signal, the costs of doing so would be prohibitive. An alternative would be to the place permanent monitoring equipment at strategic points and use portable equipment to uncover stream impairments at tactical points (discussed below), as needed. In defining which points would be considered strategic, there are basically two goals: A) To be able to determine that an actionable impairment has taken place and B) To be able to localize the source of that impairment.

Figure 4, below shows a single broadcast station and suggests strategic points for monitoring (marked with "*"). These points have been selected to allow quality determination at locations that will allow impairments to be localized. The rationale for these points is as follows:

- 8-VSB represents the final output transmission signal from the station and is the one directly

impacting the viewers. Monitoring this point allows determination of the quality of the broadcast reaching the consumers.
- The mux output represents the output signal from the studio to the transmitter.
- The input from the satellite dish (if used) represents compressed signals coming into the plant.

Normal engineering practices will allow problem localization. For example, if problems are seen in the broadcast emission and similar problems are seen in the satellite feed, then it is clear that the root cause is not within the plant and that the originator of the content should be contacted. Similarly, if problems were seen in the broadcast emission, but not at the mux output, then the STL or exciter would be suspect.

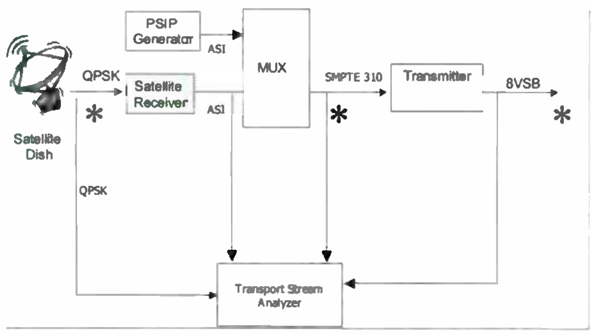


Figure 4 Strategic Monitoring Points for Single Station

Figure 5 (below) shows a more complex DTV broadcast distribution system and suggests strategic points for monitoring. As above, "*" are used to indicate recommended strategic monitoring points. By monitoring the output of each transmitter site, as well as the input to that broadcast station, it is possible to quickly isolate faults to a reasonable subset of systems to troubleshoot.

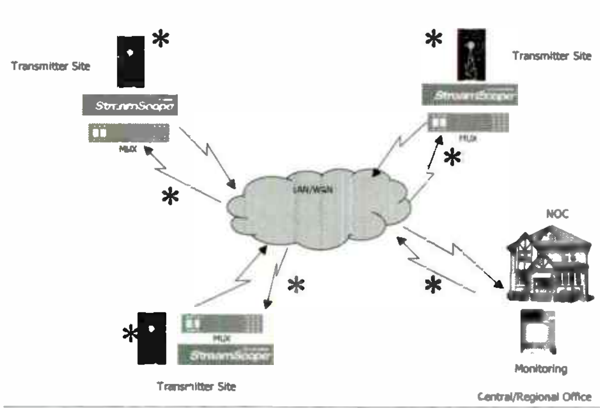


Figure 5 – Strategic Monitoring Points for Centralized Station Group

TACTICAL VIDEO MONITORING POINTS

The discussion above covers points which would be considered "strategic" for monitoring – i.e. points that should be permanently monitored. Other components within the broadcast system can (and do) introduce impairments – but it isn't practical (or more specifically, cost-effective) to place monitoring equipment there. A method to address this would be to utilize the strategic monitoring approach initially. If problems are found, then portable analysis equipment can be put into place to further isolate the problem to a single component. Equivalently, if a particular area is found to be trouble-prone, monitoring equipment can be temporarily place there.

The tactical points can also be defined by location such a distance and accessibility within the broadcast infrastructure. While a satellite down/up link or transmitter site might be better served with a remote monitoring unit, areas in the local TV station could be served with a portable analysis solution.

It is imperative however, to bring these two models together for a unified and consistent point of view from a reporting perspective. This is especially true, if the focus of the monitoring and analyses function is based on the end-to-end service.

SERVICE BASED CONFIGURATION OF THE VIDEO ASSURANCE NETWORK

There are (at least) two ways to view quality in a delivery network. The most common is based on location – allowing questions to be answered such as: "Are there any particular problems at location X?" or "Which of the broadcast stations in the system require attention?" Within a distributed system, there is another way to collect and view the information – service based. As an example - a large, distributed station group may carry different network feeds. There is often the need to view the health of a particular network across all of the different broadcast locations, which allows a different form of question to be answered: "What is the overall health of the XYZ network across the entire business?"

By incorporating a view of the flow topology for the distribution system, it is possible to offer either of these views from the same monitoring architecture. Figure 6, below, shows a consolidated view of the system health based on the topology. Figure 7 shows an alternate view, based on the actual services flowing through the system. The service-based view may become more important when a broadcaster transmits services with a higher viewer demographic and therefore a larger revenue is at stake

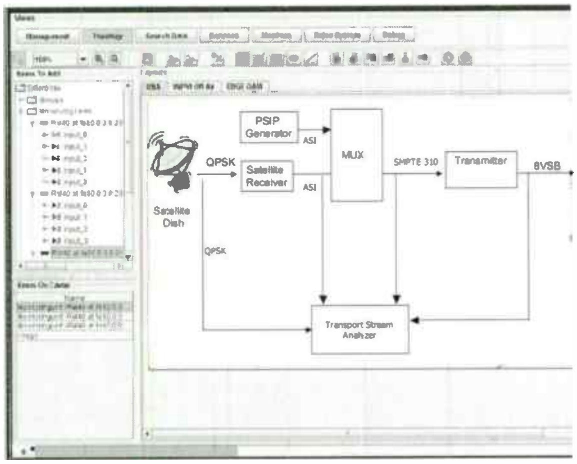


Figure 6 Topology View

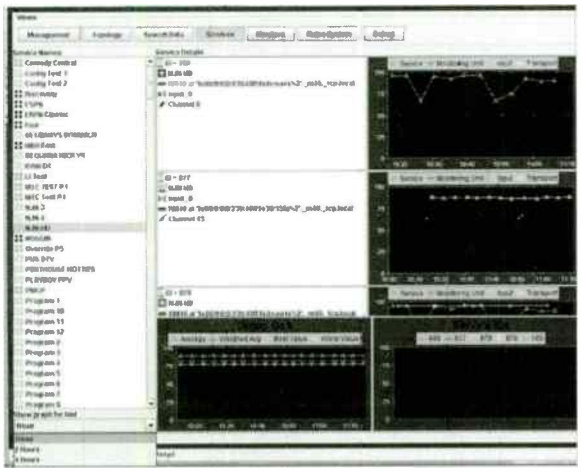


Figure 7 Service Based View

CONSOLIDATED TROUBLESHOOTING REPORTS FOR SERVICES AND THEIR INTERPRETATION

Unless implemented appropriately, a monitoring system with many monitoring devices has the possibility of introducing an information overload condition. Establishing an overall consolation functionality for this system, allows pinpointing to the actual location or area of the problem.

Monitoring strictly against standards compliance (for example, utilizing ESTI ETR101-290 [1]) can create numerous alarms for conditions that, while being strictly out of compliance, effectively do not make a visible difference. The ATSC has introduced a recommended practice (A/78 [2]) that allows error severity to be taken into account. When these concepts are incorporated into the design of a monitoring system, it is possible to have alerts directly routed to operations personnel for serious errors (missing PAT/PMT, missing video elementary streams…) and less intrusive actions for QOS or lesser errors (such as table timings slightly too slow or slight buffer overflows – these conditions can be logged for later attention during routine maintenance).

Once a problem is found, an analysis/troubleshooting phase must take place to determine the nature and location of the problem and the appropriate fix. For most video quality problems, comprehensive examination of the MPEG layer is necessary to determine what the problem is and the appropriate fix. The ability to perform the troubleshooting in real-time (i.e. no capture/offline analysis delays) is critical to a rapid solution. Since MPEG is still a somewhat foreign language to many within the broadcast plant, tools that lead the user by the hand directly to the problem area and explains the details of the problem is quite important.

As above, the need exists to be able to directly analyze the layer responsible for causing the impairments – in many cases the MPEG layer. Because of the lack of one to one correspondence between cause and effect, it is necessary to be able to analyze the MPEG layer completely – examine all potential causes of defects.

Another set of monitoring rules addresses the business intent. The monitor can learn or build a template based on a transport stream known to include all the necessary elements, including the number of audio and video components, the number of PSIP tables in the stream, and the bandwidth of audio and video components. From that point on, the stream is validated against the known default.

Stations that make more frequent changes often use monitors that download station configuration information from the station's PSIP generator. A template downloaded from the PSIP generator can provide information such as the number of audio channels to expect in the transport stream. The monitor can compare the actual stream with the expected stream and deliver an alert in the case of any inconsistency. This type of monitoring closes the loop linking station policy, business priorities and execution, allowing broadcasters to meet both FCC regulations and their own business strategies.

TREND ANALYSIS

Often, impairments don't simply appear out of nowhere – they are the result of a gradual degradation in performance of one or more components. By observing the monitoring results over a period of time, this degradation can be discovered – often before the parameters drift out of compliance. Incorporating historical trend analysis into the monitoring architecture allows a further extension of proactive response, allowing faults to be discovered and repaired before any viewers are made aware of them.

An important enhancement to trend analysis would be the capability for historical replay, being able to recreate stream conditions at a time in the past. Even with a proactive approach, there are times when impairments will not be discovered. Historical replay capabilities would allow correlation of an upturn in trouble calls in the recent past with what was actually taking place in the broadcast stream at that time. This, in turn, would allow a level of localization to the device (or area) that was causing the impairments. Once the area was located, further testing could either identify the failing device (which would then be repaired) – or monitoring equipment could be temporarily placed to uncover repetition of the problem with finer resolution, ultimately leading to correction of the problem.

SUMMARY

The more information the monitoring solution provides about the DTV signal, the better the broadcaster can respond to issues within the stream. When presented logically and simply through an easy-to-use interface, information provided by such an analytical tool can make the transition to digital broadcasting and subsequent management of the DTV signal a straightforward operation. What's more, with a comprehensive system wide quality assurance solution in place, broadcasters can maintain the integrity of their digital broadcasts without investing a great deal of time or resources in becoming familiar with the complex details of broadcast and signal standards. Remote analysis thus eliminates the need for an in-in-person visit, allows for more efficient use of technical staff and reduces the time and cost involved in solving transport stream issues.

Though the ability to ensure the integrity of digital television services may seem like a future goal, it's an important part of broadcasters' operations today. Increasingly viewers are investing in new television sets and tuning in to digital broadcasts. As stations across the United States continue to launch digital transmission and prepare for the analog shut-down, effective technologies for delivering healthy transport streams are playing a critical role in the success of those broadcasts.

REFERENCES

[1] ETSI TR 101 290 v1.2.1 (2001-05) Digital Video Broadcasting (DVB) Measurement guidelines for DVB systems.

[2] ATSC Recommended Practice A/78A – Transport Stream Verification

CONTROLLING AND MEASURING LOUDNESS FOR DIGITAL TELEVISION BROADCAST

MICHAEL BABBITT
Dolby Laboratories
San Francisco, CA

ABSTRACT

Loudness discrepancies have been a major source of headaches for the broadcast professional, and only seem to be getting worse in digital television broadcasts. Before corrective action can be taken, program loudness must be measured and the behavior of the professional integrated receiver/decoder (IRD) understood. A broadcaster's best efforts at controlling loudness and reducing the number and frequency of viewer complaints can be undermined by just a few outlier programs or a delivery path that has not been appropriately managed. Rather than approaching the task piecemeal, a systemic approach is necessary to achieve the best results and reduce viewer complaints.

WHY IS LOUDNESS SUCH A PROBLEM IN DIGITAL TELEVISION?

Current practices in NTSC broadcast television provision the audio signal such that it approaches the highest permissible level.[1] These legacy practices have provided, as an unintended byproduct, a measure of program loudness normalization within a specific channel's broadcasts, as will be explained later in this paper. Inter-channel differences have been apparent (i.e. one channel may be louder or quieter than another channel), but inter-program loudness (i.e. different programs broadcast on a specific channel) has generally maintained a relatively consistent level due to these legacy practices.

Digital television has the capability of providing a cinematic experience to the viewer, complete with stunning visuals and dynamic, multichannel surround sound. Maintaining wide dynamic range audio within digital programming is a key and desirable feature of digital television, but wide dynamic range audio can present challenges to equipment, broadcasters and viewers alike.

Figure 1 shows three day's worth of the exact same primetime, high value programming presented and measured simultaneously on both the analog and the digital channels of a single terrestrial network affiliate. While the analog signals (the higher cluster of data points) clearly appear to be much louder than the digital signals (the lower, and more varied cluster of data points), the analog signals are clearly more consistent, program to program. Hence, when a viewer tunes to a specific channel, in many cases a single volume adjustment will be sufficient to "calibrate" their listening environment to their desired acoustic volume level, and most programs on that channel will not require a volume shift to maintain the viewer's desired target playback level. The digital signals (the lower and more varied collection of data points) show the wide dynamic range of high production value programming, and there is clearly a wide variance from program to program. Fortunately, systems and processes within most broadcasters' existing equipment and signal paths can help to maintain consistent levels within and between both analog and digital broadcast signals at no additional expense when understood and properly utilized.

UNDERSTANDING RECEIVER BEHAVIOR

At the local station or cable headend, the behavior of an integrated receiver/decoder is often overlooked, as long as downstream equipment appears to receive the appropriate signals and the appropriate program is on air with both sound and picture. Engineering resources are generally reserved for those systems that are reporting error conditions, and loudness complaints from viewers have historically not been accompanied by an error message displayed on encoding or processing equipment. Unbeknownst to many station and headend engineers is the fact that appropriate settings in the audio decoding subsystem within these IRDs can provide a major benefit to reducing viewer complaints and presenting an appropriate and compelling viewing experience in the home.

Digitally Tiered Analog Services

For those local stations and cable headends that are receiving a pre-encoded Dolby Digital (AC3) data stream and decoding this signal within the IRD (i.e. not using an external, professional reference decoder) for transmission on an analog tiered service or for other purposes (bug insertion, time shifting, etc.), inappropriate settings within the IRD audio decoding subsystems are a primary source of viewer complaints. The lack of understanding of these settings and the manufacturer's default conditions of the IRD are contributing to this issue. In order to best match existing analog programming on these delivery systems, a specific mode must be invoked within the Dolby Digital (AC3) decoding subsystem in order to maintain a match between programs and neighboring channels.

This decoding mode is known as "RF Mode" in Dolby parlance, but IRD manufacturers may use another term such as "heavy" or "narrow" in their units. The behavior of this mode will be discussed in detail in the following section. It is important to understand that this mode will only affect the decoded analog or digital output of the IRD. Any selected decoding mode will not affect the Dolby Digital (AC3) bitstream output of the IRD, and those systems and signal paths conveying AC3 via a digital pass through architecture will maintain the capability of conveying full dynamic range audio to their viewers.

The Effects of Audio Metadata

Dolby Digital (AC3) data streams contain audio metadata, which provides necessary information to the receiver/decoder. A main component of this audio metadata is the dialog level parameter (aka dialnorm), which is intended to convey to the decoder the average A-weighted (or K-weighted) level of conversational speech within the audio program. Studies have shown that conversational speech and dialog elements within programming are the anchors by which viewers adjust their playback volume to a desired level.[2] When the dialnorm parameter is set such that it accurately conveys the average A (or K)-weighted level of conversational speech, the average speech level in the decoded output of a consumer IRD will be maintained around -31dBFS. This provides enough headroom for a wide dynamic range, cinematic experience in the home via a digital television broadcast.

However, the dialog leveling subsystem via the dialnorm metadata parameter does not live in a vacuum. Other important processes within the AC3 decoding process depend on an accurately set dialnorm parameter. In addition to the most commonly understood function of provisioning the average level of speech to a specific target level and therefore "calibrating" the viewer's listening environment, the dialnorm parameter also provides the basis for dynamic range processing (aka dynamic range control, or DRC) within the Dolby Digital (AC3) decoder (Figure 2).

The dialnorm parameter, when properly reflecting the average level of conversational speech within the program, provides the center point to a "null band" where (when invoked manually by the viewer or as defined by the manufacturer to reflect the capabilities of the playback equipment) no DRC processing occurs. When used correctly, these two processes prevent inappropriate processing from occurring on conversational dialog within the program in addition to normalizing the level of speech to the target level of -31dBFS.

In the United States, cable systems typically receive content via both digital and analog methods, and relay this content often via both digital and analog systems to their subscribers. This combination of dual delivery methods requires the ability to provision digital services to match analog services with respect to level. Often, broadcast engineers are unaware that existing equipment likely contains all the necessary processing to match levels within their digitally tiered services without any additional expense, provided that the content itself is provisioned properly prior to delivery to the broadcaster. Methods and processes to properly provision content for delivery will be discussed later in this paper.

As most consumer digital IRDs are capable of receiving varied input signals (AC3, MPEG-1 Layer 2, QAM, DVB, ATSC, etc.) from multiple sources and provide the audio program within those received input signals at several different output connections, the gain structure within these IRDs is provisioned with the assumption that the dialnorm parameter within a received Dolby Digital (AC3) data stream is set correctly when tuned to a digital service.[1] When tuned to an analog or other non-Dolby Digital service, the gain structure is provisioned with the assumption that the average level of conversational speech is about -17dB below 100% modulation.[1] Legacy NTSC signals are generally normalized via audio processing in an effort to maximize the permitted broadcast level. This has the effect of providing a "brute force" method of controlling loudness differences by minimizing dynamic range and making all audio elements (speech, music, sound effects, etc.) within a program near to the same level. The Dolby Digital (AC3) system contains specific processes that, when properly utilized, can help to match levels within digital programming to existing analog broadcasts while simultaneously helping to retain dynamic range to create a more transparent and compelling listening and viewing experience.

The decoded target level of speech at -31dB is too quiet when compared to standard NTSC broadcasts. In order to match the average level of speech and the wide dynamic range signals that are commonly found in digital broadcasts to existing analog programming, a specific dynamic range control (DRC) profile should be invoked to better match the narrow dynamic range of programs broadcast on analog services. When the dialnorm parameter accurately reflects the average A (or K)-weighted level of conversational speech in the encoded program stream, RF Mode DRC appropriately scales high level dynamics downwards and low level dynamics upwards thus "compressing" the dynamic range of the digital program (Figure 2). In addition, the Dolby Digital (AC3) decoder invokes an 11dB level shift upwards when RF Mode DRC is enabled to better match the expected level of speech within the gain structure of RF demodulators. When the dialnorm parameter accurately conveys the average level of conversational speech within the program, the 11dB shift above the normalized level of -31dB invoked by the RF Mode DRC setting places the average level of speech at -20dB for stereo programming, and when summed to mono, matches the expected target level within NTSC broadcasts of -17dB. In addition, RF Mode provides 6dB more headroom than typical NTSC programming, providing, "...a similar dynamic range to NTSC without the need for the violent non-linear processing commonly employed to avoid over-modulation."[1]

The keys to the predictable behavior of a receiver and an appropriate and compelling audio experience in the home are the correct setting of these audio metadata parameters within the received program audio stream and the correct setting of the Dolby Digital (AC3) decoding subsystem within the professional IRD.

The Consequences of Improperly Set Audio Metadata Parameters within Delivered Signals

When audio metadata within a delivered Dolby Digital (AC3) data stream is correct, the licensed decoder within a receiver processes the audio prior to delivery to the output connections in an appropriate fashion based upon settings within the delivered audio metadata set. The main processes include leveling (dialnorm or the dialog level parameter), dynamic range compression (when invoked), and downmixing (summing multichannel audio appropriately for presentation on a stereo or mono playback system).

In all cases within licensed consumer decoders, and in most cases within professional licensed decoders (note that in some select professional receivers, special modes are available that allow for the disabling of the dialnorm subsystem; these special modes must be used with care), dialnorm processing is a mandatory process and is invoked on the decoded output at all times. When considering consumer and professional broadcast receivers, the decoded output of these units in most cases will be stereo signals, and when receiving multichannel programming, a cascade of processing occurs within the Dolby Digital (AC3) decoder. Specifically, 1) the audio is "leveled" according to the dialnorm parameter setting, 2) the dynamic range of the audio is controlled, and 3) the multichannel program is intelligently combined (downmixed) to create an appropriate stereo or mono signal.

It is important to understand that these processes work together. The dialnorm parameter identifies an area within the audio program that will not be processed by the dynamic range control subsystem. This "null band" provides a safe area for conversational speech. When the dynamic range control subsystem is invoked either through manual selection by the operator or, in the case of some consumer systems, by the manufacturer as a standard condition in order to protect small speakers from distortion and/or damage, audio elements that occur above the null band as identified by the dialnorm parameter are lowered in level and elements that occur below the null band are raised in level (Figure 2). The amount of dynamic range control is dictated by the mode selected by the operator and calculations made by the Dolby Digital (AC3) encoder that are embedded within the incoming data stream. Lastly, in the case of downmixing a multichannel signal to a stereo or mono presentation, the DRC subsystem, which in the case of downmixing is invoked automatically, helps to prevent clipping when these signals are combined.

If the dialnorm parameter is set such that it does not match the average level of conversational speech within the program, this cascade of processes can create audible artifacts as processing will take place on critical signals. Consider a program with the measured, average conversational speech level at -20dBFS but the dialnorm parameter is set to -31. In this case, the null band that is intended to protect critical conversational speech elements from inappropriate processing will be ineffective, and these critical speech elements will be pushed into areas of active dynamic range processing, putting the program at risk of audible compression artifacts and impacting viewer satisfaction.

A SYSTEMIC APPROACH

While the proper provisioning of broadcast reception and delivery equipment is a component of controlling loudness, the content itself must either contain the correct audio metadata or be provisioned such that the average level of speech within the content is either known or standardized to a consistent target level. One or both strategies can be implemented depending on the activity taking place and the equipment within the broadcast infrastructure.

Considering the varied delivery systems present in the typical broadcast infrastructure, it should be clear that program loudness cannot be successfully managed with a piecemeal approach. If locally produced content is properly managed but network signals, advertisements, and syndicated programming is not, level differences between programs will be objectionable, possibly resulting in viewer complaints. Additionally, when an analog feed is originated from a digital service, the incorrect selection of processing modes within the Dolby Digital (AC3) decoder may result in inappropriate audible artifacts in addition to level differences, also generating viewer complaints. Clearly what is necessary to achieve viewer satisfaction is a systemic approach which brings to bear equipment and workflow processes to achieve the desired result.

Live Programming

Controlling the audio program from acquisition to transmission provides the best opportunity to manage program loudness and eliminate unnecessary and inappropriate processing. For stereo or multichannel content that will be delivered using Dolby Digital (AC3), networks and local stations alike must decide upon a specific and consistent target level for conversational speech, and during the content creation process, use appropriate loudness metering during periods of conversational speech to ensure that the standardized average level of speech is maintained throughout the program. Often, the target level of speech is defined within a delivery specifications document. For stereo and multichannel programs, the target level of average conversational speech should match the specified level in the delivery specifications document. Again, when the dialnorm value within the Dolby Digital (AC3) data stream correctly matches the average level of speech, RF Mode DRC (aka narrow or heavy) will ensure that average speech levels will match existing NTSC programming, and the analog broadcast viewer will experience a seamless transition when channel surfing.

With respect to local stations, a consistent speech level is likely already being maintained, as there are fewer operators and the content is chiefly local news and special interest programs, both of which lend themselves to consistent procedures and practices. Empirical evidence suggests that production crews produce consistent results over time (Figure 3). As the data in Figure 3 suggests, any level differences in locally produced content for some stations may simply be the result of a shift change, and even that difference is likely to be slight. A medium-to-long term measurement strategy that includes proper metering will best diagnose any differences between production crews and provide the local station with a pathway to improve consistency.

For special events, the audio mixer must have a broadcast loudness meter in the audio booth, and throughout the event, must periodically monitor the measured result of the speech content, maintaining an artistic balance between program elements while also endeavoring to maintain a relatively consistent level with respect to speech. With the measured value of speech conveyed to the mixer via a clear, numeric display, consistent audio levels can be maintained while the mixer creates an appropriate balance between the audio elements, irrespective of operator fatigue.

Program Ingest

Content created for broadcast must meet specific standards, as described in the network or local delivery specifications. It is important that delivery specifications include an average level of speech along with a clear methodology to attain it so that a broadcaster can have some assurance that content received from outside providers for broadcast matches content created in house or from other outside vendors. Creating and publishing such a specification significantly reduces the chance of an externally produced program being rejected.

Here is an example of a loudness specification within a generic program delivery specification document:
The average level of dialog within the program on all non-compressed (i.e. non-Dolby E encoded) content (whether multichannel or stereo programs), evaluated using a Dolby LM100 Broadcast Loudness Meter or equivalent Leq(A) meter, shall return an average level of dialog of -27dBFS, +/-2dB measured over the entire length of the program. The LM100 should be set to "Infinite Mode" with "Dialog Intelligence" enabled. The measurement shall begin at the start of the program and stopped at the end of the program. The resulting Infinite Term measurement value shall be -27dBFS, +/-2dB.

Checking for adherence to published delivery specifications generally occurs at the point of ingest.

When the delivery specifications contain a standardized average level of conversational speech and a clear methodology by which the content provider can attain it, these quality assurance (QA) processes must also include a check for the same specification using the same methodology. Similar to the Live Programming discussion earlier, the ingest operator should have an average loudness meter which is sensitive to speech in their toolkit, and understand how to use it. Additionally, the methodology outlined in the delivery specifications should be adhered to for the submitted content to stand the best chance of being approved for air. A detailed outline of measurement modes, settings and procedures will benefit both the content producer and the broadcaster.

Commercials and Ad Insertion

Television commercials have historically been blamed for most of the loudness problems within the medium. As Figure 4 suggests however, commercials in general, while loud, may appear to be quite consistent with respect to level. Considering the fact that commercial production techniques require that the "story" be told in a very short time period, this observed consistency with respect to level would seem to be a natural byproduct of the production process. Certainly they can be loud, in the case of Figure 4, averaging around -22dBFS, but the perceived loudness of any program will be judged in the home by the content preceding it. Herein lays the issue: If the program preceding the commercial break is perceived as louder than the commercials, the outlier will be the program content, while the opposite will be true if the program content is perceived as quieter than the commercials. The key is to either a) have each program segment (scripted programming, interstitials, news, commercials, etc.) contain a dialnorm value that matches the average speech level within it, or b) leverage the observed consistency of commercial content to match the level of the program content to the commercials.

Adjusting the level of commercial content may be a controversial task, as advertisers are paying for air time during a particular program. There are two strategies for tackling the issue of commercial loudness, and each may prove effective provided the appropriate content is measured prior to invoking a specific solution.

The first strategy involves either a) changing the level of the commercial content such that it matches a selected target level, or b) encoding the commercial content with the correct audio metadata. As commercials are often received locally via a fiber or satellite service and stored on servers prior to air, these programs can be provisioned automatically using non-real-time, file-based analysis and encoding products. These commercial files are handled by an automation or asset management system, and during their stay within a local station's IT network, can be analyzed and corrected through the use of faster-than-real-time, file-based processing engines, like the Dolby DP600 Program Optimizer. While this strategy processes commercial content and thus changes the level within, it is also automated and transparent, creating an attractive and cost effective solution to controlling program level.

These processing engines can also encode the file-based content prior to air and create the appropriate audio metadata that reflects the needs of the audio program within. In this fashion, the audio essence of the commercial is untouched thus removing the possibly controversial aspect of level adjustment, and the correct audio metadata is conveyed to the home where the Dolby Digital (AC3) decoder can properly provision it based upon the viewer's preference or the equipment's capabilities.

The second strategy is, in effect, the opposite of the previous one. In order to better match level differences between scripted programming and commercial content, the general consistency of commercial content observed in Figure 4 can be leveraged to provide an average standardized level for scripted programming. (*Note: Each facility should measure and analyze a representative sample of its own commercial content rather than using the Figure 4 example as a basis for a level correction or processing strategy.*) In this case, scripted programming would be required to match a level specified by the network or local station that matches the average level of commercial content. The burden of matching levels is placed on the content provider, but must be checked by the network or local station during ingest as described previously to ensure that it meets the mandated delivery requirements with respect to the average level of conversational speech. This strategy relies upon the assumption that the average level of every commercial will reside within a few dB of a selected target level. Because of the high probability of the level of some commercial content falling outside of this target acceptance window, there will likely be some instances of wider differentials between programs (i.e. average speech levels outside the target boundaries), and a higher incidence of viewer complaints, than if the first strategy is used exclusively.

Syndicated Programming

Like commercials, syndicated content is often received via a subscriber-based fiber or satellite delivery systems and stored on a server prior to air. Also like commercials, this delay in transmission can be used to the benefit of the broadcaster by allowing the content to be analyzed and corrected prior to air using non-real-time, file-based processing equipment.

Unlike commercials and ad content, syndicated programming does not benefit from a set of common production techniques. Syndicated programming can consist of chat shows, confrontational and reality programming, news and scripted programming. While chat shows and confrontational programming may use similar production techniques and audio processing equipment, scripted programming and even some reality-based programming generally involve a higher production value which in turn will result in a wider dynamic range.

Because of the wide variety of content available and the different production techniques used in syndicated content, it will not provide a generally consistent level to which other content can be matched. A local carrier can implement a manual leveling process involving a costly and burdensome series of checks and adjustments through an additional ingest path, but this is not a realistic solution for the modern broadcaster. In order to assure that syndicated content either a) carries an accurate dialnorm value that reflects the average level of conversational speech within the program, or b) the content itself matches a chosen target level for broadcast, the only realistic solution for managing program loudness on this type of programming is via non-real-time analysis, correction, and processing tools. These systems can perform leveling, encoding and trans-coding functions on various file types received by the local station, and do them transparently as an additional series of processes within the local station's normal IT network reception-to-transmission path.

Video on Demand Services

Premium programming can also benefit from non-real-time tools. When ingested onto the IT network, the program content can be analyzed and processed to provision it for air. In these instances, the content's audio essence can remain unchanged (i.e. not adjusted for level or limited in dynamic range) thus satisfying the viewers. program producers and any contractual obligations that may be in force. These non-real-time tools would be instructed to analyze the content and use the results of that analysis to create metadata values that will accurately reflect the content itself. Using these metadata values, the processes can then encode the content into Dolby Digital (AC3) and be ready for air, or create a broadcast wave file containing a metadata "chunk" which may be read by subsequent processing engines.

As with syndicated programming and commercial content, when a program resides within an IT network as a file, file-based analysis, correction and processing tools provide a "one stop shop" for managing loudness and provisioning content for air within a modern broadcast infrastructure.

CONCLUSION

Studies show that viewers will adjust the volume of their playback system in an effort to make the level of the speech element within a program consistent. Through the dialnorm parameter within audio metadata, Dolby Digital (AC3) provides a method to convey the measured average level of conversational speech within a program to the home, and when the dialnorm parameter is accurately set, dynamic range control subsystems within the Dolby Digital (AC3) decoder can appropriately provision the audio for the viewer's playback equipment or acoustic environment. When properly understood and utilized, the dynamic range subsystem within a Dolby Digital (AC3) decoder can be leveraged to provide an appropriate level match to existing analog content within a professional broadcast environment.

Loudness complaints from viewers are the result of perceived differences between program content. As such, even if 90% of a broadcaster's programming is level matched with respect to speech, viewer complaints will still occur due to the remaining 10% of unmatched content. For this reason, comprehensive loudness management and control requires a systemic approach that engages each program path within a broadcast infrastructure.

For live-to-air programming and content creation (news, sports, program interstitials, news teasers, etc.), identification of a target loudness level and placement of a broadcast loudness meter in the audio suite will provide real-time control as the content is being created. For program content that exists on a server (commercials, syndicated programming, video-on-demand programming, etc.), file-based, non-real-time tools provide the best method to both analyze and correct audio level issues as well as provision the audio content for air.

If networks, local stations, MSOs and content providers undertake a systemic approach to loudness control (Figure 5), viewer complaints can be reduced while simultaneously conveying wide dynamic range, compelling and engaging content which fulfills the promise of high definition television.

REFERENCES

1) EIA/CEA-CEB11, "NTSC/ATSC Loudness Matching," 2002.

2) Riedmiller, Lyman and, Robinson, "Intelligent Program Loudness Measurement and Control: What Satisfies Listeners?" presented at the 115th AES Convention, New York, 2003.

.

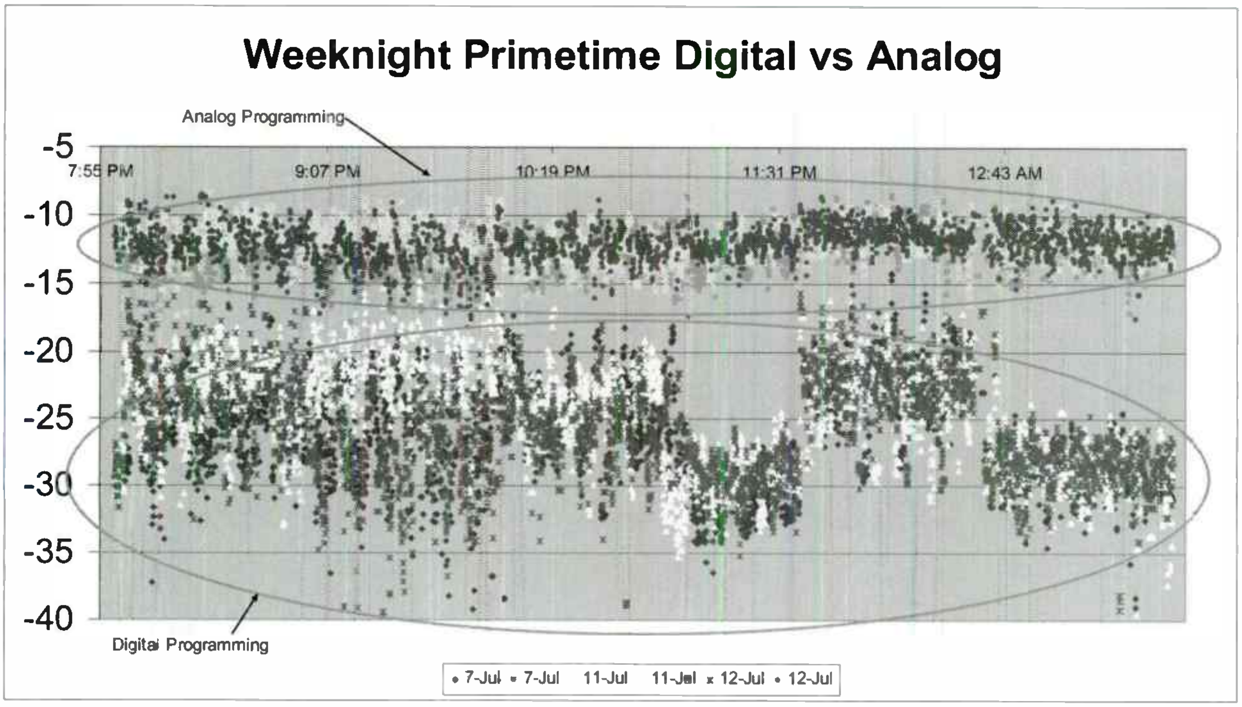


Figure 1

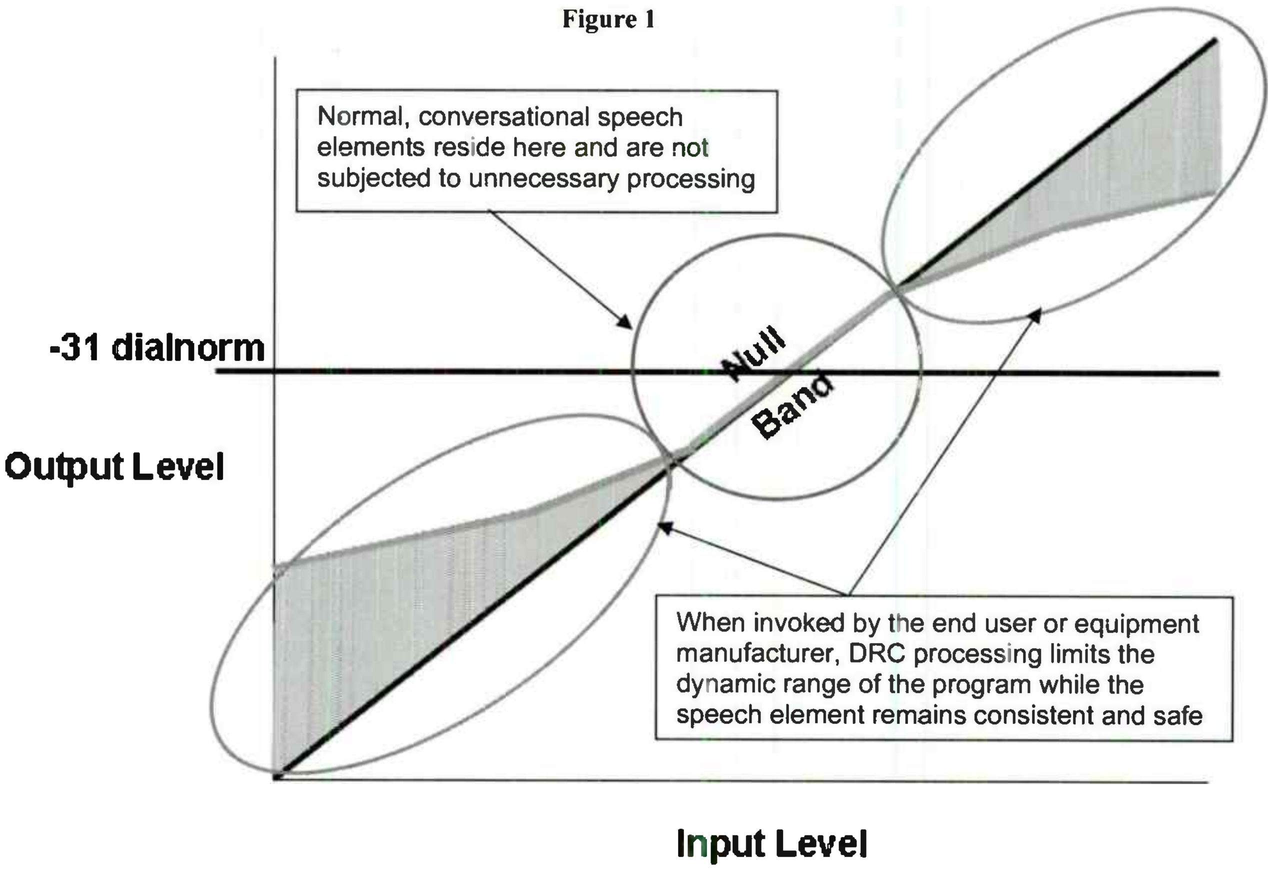


Figure 2

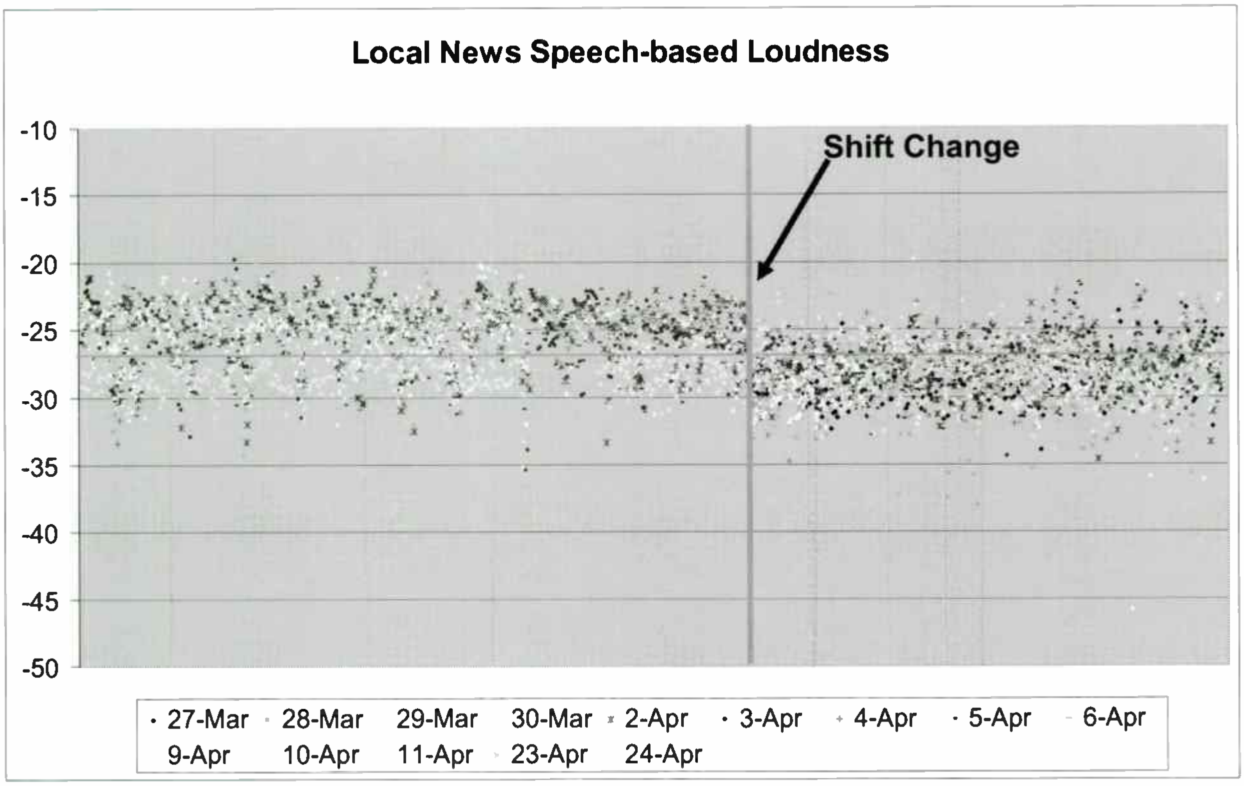


Figure 3

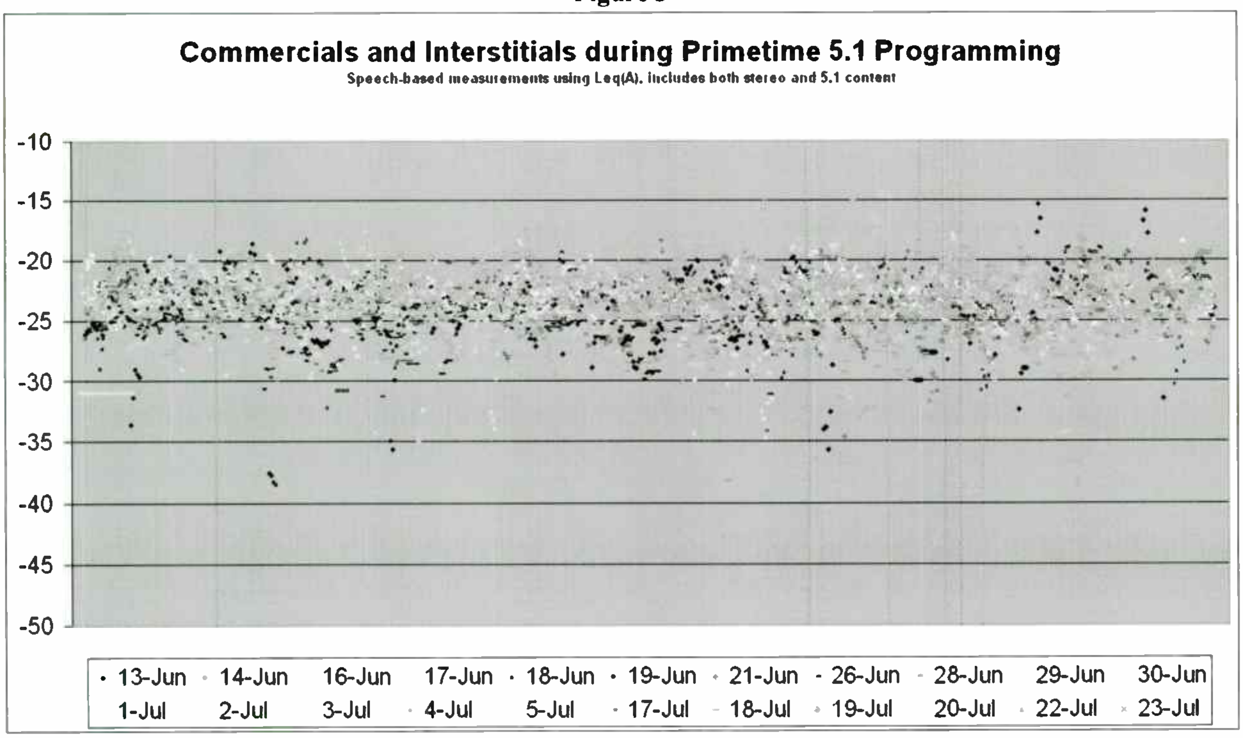


Figure 4

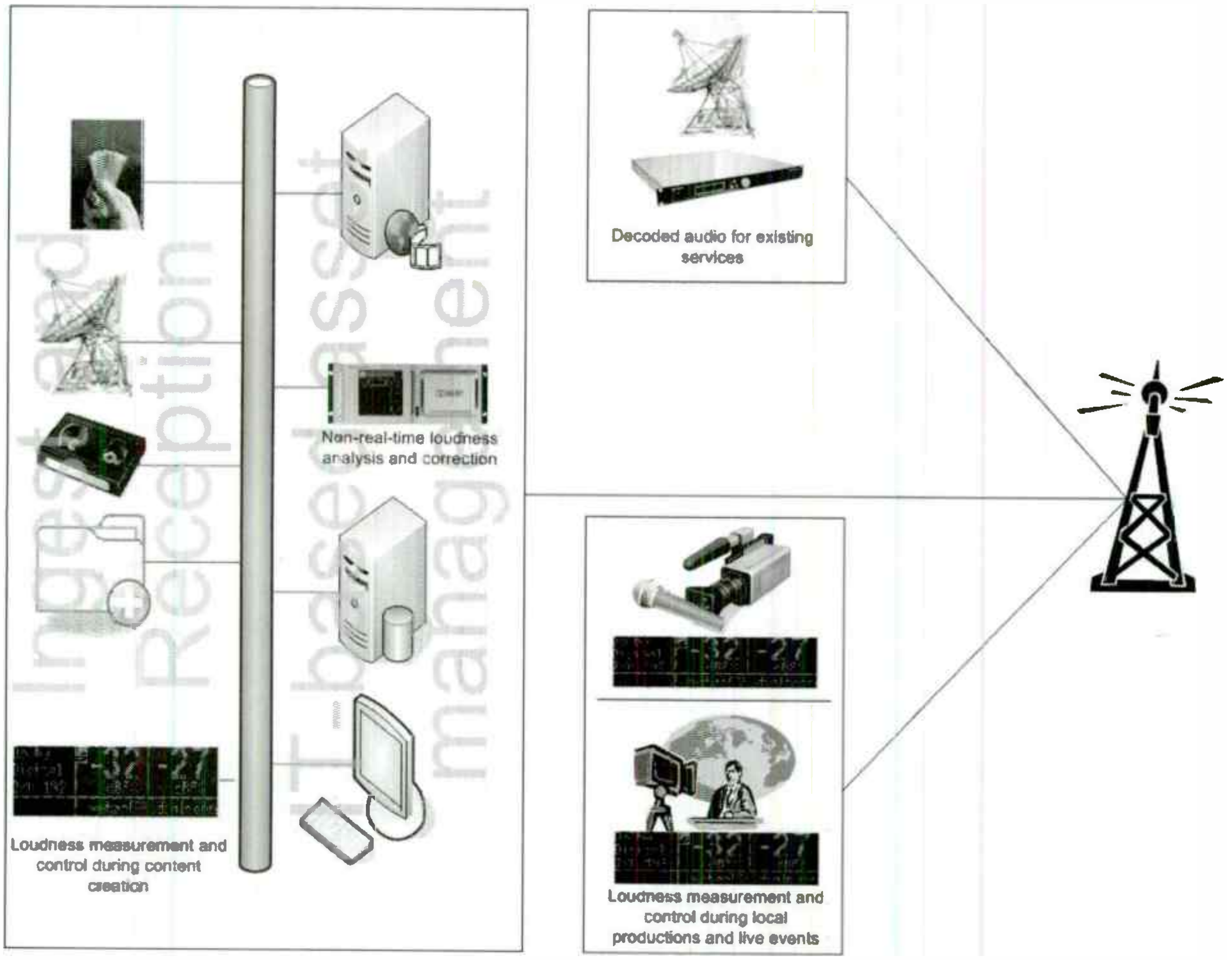

Ingest and Reception
IT-based asset management
Non-real-time loudness analysis and correction
Loudness measurement and control during content creation
Decoded audio for existing services
Loudness measurement and control during local productions and live events

Figure 5

Television RF & Transmission Systems

Wednesday, April 16, 2008
11:00 AM – 4:30 PM

Chairperson: Louis Libin
BroadComm, Inc., Woodmere, NY
Chairperson: Victor Tawil
Association for Maximum Service Television, Washington, DC

***Implementing an 8-Transmitter Distributed Transmission Network**
S. Merrill Weiss, Merrill Weiss Group LLC, Metuchen, NJ

Antennas for Distributed Transmission and Single Frequency Networks
Myron D. Fanton, Electronics Research, Inc., Chandler, IN

Evaluation of Buildings Penetration Loss for 100 Buildings in Belgium
David Plets, Ghent University, Department of Information Technology, Ghent, Belgium

***Vertical Polarization for UHF DTV**
Kerry Cozad, Dielectric Communications, Raymond, ME

RF Coverage and Tower Motion
Myron D. Fanton, Electronics Research, Inc., Chandler, IN

***Field Test of the Distributed Translators in Korea**
Young-Woo Suh, Korean Broadcasting System, Seoul, Korea

*Paper not available at the time of publication

ANTENNAS FOR DISTRIBUTED TRANSMISSION AND SINGLE FREQUENCY NETWORKS

Myron D. Fanton, PE
Electronics Research, Inc.

ABSTRACT

Much like traditional broadcasting, antennas deployed for distributed transmission systems are key to the success of the system. Coverage and interference are directly related to antenna parameters, with the additional issue of self-interference. Antenna arrays tailored to the network topology maximize coverage and minimize self-interference.

Antenna array theory is brought to bear on the self-interference problem. Regions of sharp "roll-off " in the azimuth patterns are discussed. Techniques are given for creating a rapid drop in amplitude over a short distance, and the pitfalls are revealed that will be encountered without measurement and validation of the array.

INTRODUCTION

Having appropriately equipped receivers, broadcasters now employ Single Frequency Networks (SFN) to fulfill their obligations to provide service to their communities. ATSC receivers contain adaptive equalizer circuits to recover signal information in a multipath channel, imposing practical and achievable restrictions on multiple transmitters. The various benefits of SFN broadcasting in relation to conventional, single antenna systems, are known [1]: redundancy, improved fault tolerance, improved reliability, improved coverage, improved interference protection, improved spectral efficiency, reduced interference generation, reduced power requirement. Basic design methods have been outlined, and standards and regulatory issues have been formulated [2].

At this writing, several ATSC SFN systems have been successfully implemented. Most of these are large-cell, shadow-filling schemes in which the regions blocked by terrain from the traditional, high-power transmitter are served with low-power distributed transmitters. The unique role that antenna equipment plays in the SFN is discussed presently.

ANTENNA ARRAYS

Broadcast frequencies and power levels mandate the use of the array antenna. Generally the arrays take the form of a linear array of similar elements, and can thus be separated into the element pattern and array factor. The patterns of the complete antenna result from the product of the element and array patterns. The element, one bay or level, typically defines the antennas azimuth pattern while the geometry and feeding scheme defines its elevation pattern.

Azimuth Pattern

Antenna azimuth patterns characterize the radiation in the horizontal plane. An antenna with a large aperture will have a very narrow beam, like a microwave antenna. Antennas with broad beams that change greatly in amplitude over a small angular region also require large apertures. Apertures of this size are not generally available at broadcast frequencies, and typically the objective is to cover a broad angular region.

For slot antennas, the radiation from a single narrow slot on a cylinder is described by [4],

$$E(\phi)=\sum_{n=-\infty}^{\infty}\frac{j^{n}C_{n}\cos(n\phi)}{H_{n}^{(2)'}(ka)},\begin{cases}C_{n}=1,n=0\\C_{n}=2,n\neq 0\end{cases}\quad(1),$$

where ϕ is the azimuth angle, k is the propagation constant, a is the radius of the cylinder, and $H_n^{(2)}$ is a Hankel function of the second kind, order n. With this as the element pattern, a circular array of slots is computed with the array factor [3],

$$AF(\phi)=\sum\alpha_{i}e^{j(ka\cos(\phi-\phi_{i})+\beta_{i})}\quad(2),$$

where ϕ_i is the azimuth angle of each slot and α_i, β_i are respectively the amplitude and phase of each radiating element. Patterns resulting from Equations (1)-(2) are plotted in Figure 1a.

In a SFN service region, the topography often shadows some important areas, providing isolation between transmitters. In fact, these shadows are the most natural introduction of the SFN into the one, high-power transmitter model of broadcasting. The very geographic conditions that necessitate the use of SFN are often critical in their implementation. Without topographic isolation between regions, antennas with angular areas of rapid signal roll-off are necessary to reduce interference to adjacent coverage areas as well as the self-interference within the SFN. This requirement imposes if not new, non-traditional limitations and specifications on the azimuth patterns, namely front-to-back ratio, beam-width, and roll-off. Front-back ratio is a measure of the power over an angular region behind the antenna relative to the power in the main beam peak, and this ratio speaks directly to the geographic isolation of SFN cells. The antenna azimuth 3dB beam-width describes the coverage area in terms of azimuth angle, and the roll-off describes how rapidly the amplitude changes over a given angle. The roll-off is plotted as differential amplitude in Figure 1.

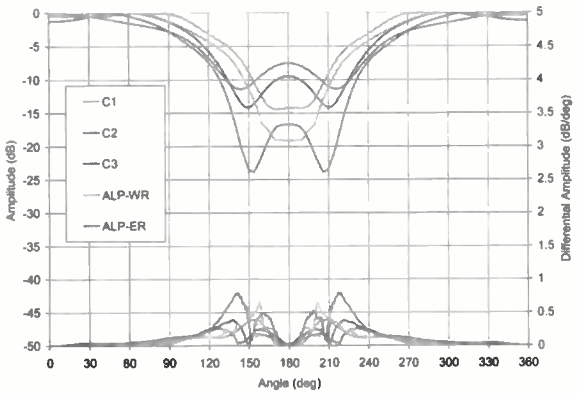


Figure 1a: Slot Antenna Az Pattern and Differential Amplitude

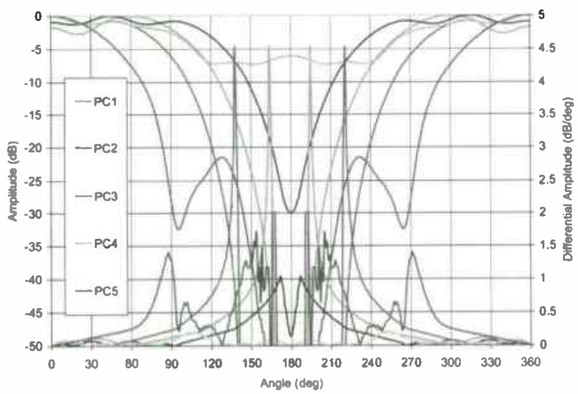


Figure 1b: Panel Antenna Az Pattern and Differential Amplitude

The patterns resulting from slots on cylindrical conductors generally change slowly with azimuth angle and do not provide high front-back ratios. These pattern characteristics along with the inherent mechanical efficiency give slot arrays their prominent place in broadcast applications. Typical slot antenna patterns are plotted in Figure 1a.

Integrating reflector systems into slot arrays can increase the roll-off and front-back ratio at the expense of increasing the aperture size. A minimum limit to the nulls in an azimuth pattern of 15dB has been mandated at times, and a 20dB front-back ratio and roll-off less than 5dB over 10° is conservatively assumed in planning [5].

Antenna arrays composed of panel antennas generally consume an aperture larger than that of a slot antenna. This allows a larger roll-off and front-back specification. The azimuth pattern of a panel antenna is also governed by the array factor of Equation (2), the element pattern being narrower than that produced by a slot on a cylinder. Typical panel array patterns are plotted in Figure 1b, along with their differential amplitude.

The application of slant linear, or circular polarized antennas to SFN has the potential to impact the interference problem. Cross-polar isolation improves the front-back and roll-off of sector antennas, improving the isolation between cells.

Elevation Pattern

The radiated fields of an antenna are related to the distribution of fields at the antenna aperture. For arrays of discrete elements, the aperture distribution is related to the fields at each element. Given the amplitude and phase of each radiating element, α_i, β_i respectively, and the known (linear) array geometry, d_i, the array factor may be computed by the following equation [3],

$$AF(\theta) = \sum \alpha_i e^{j(kd_i \cos\theta + \beta_i)} \qquad (3),$$

where k is the propagation constant and θ is the elevation angle. If the array consists of similar elements, the product of the array factor and the pattern of one element produce the elevation pattern. Note that the quantity computed by Equation (3) is a complex number, providing both amplitude and phase of the elevation pattern.

The main beam is synthesized from the constructive addition of waves emanating from all array elements. The aperture distributions of a typical slot array antenna are shown in Figure 2 and result in pattern characteristics ideal for a broadcast station: high side-lobes and heavy null fill.

Elements fed with uniform amplitude and phase will result in antennas with higher gain than those fed otherwise. Antenna arrays with greater numbers of elements, longer arrays, will also generally have higher gain. With increasing gain, more energy is concentrated in the main beam, and the slope is greater of the main beam edges.

SFN typically require antennas mounted much lower than traditional, high-power broadcast antennas. At the same time, the power levels are much lower, and the available aperture is often smaller. Mounting at lower heights has a number of benefits, one of which is increasing the order of signal decay due to propagation. The FCC propagation curves already account for this order, and the propagation loss they predict is greater than the inverse-square loss of free-space. For example, the curves yield the 6dB per octave free-space loss for tall towers and close range and 12dB per octave for short towers. The propagation situation for the low-height antennas in SFN is often the latter, 12dB per octave [6].

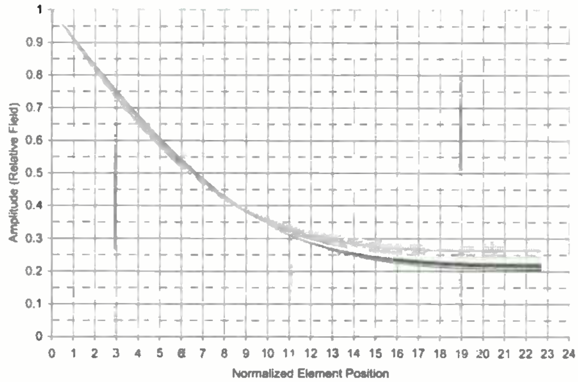


Figure 2a: Amplitude Aperture Distribution of End-Fed Array

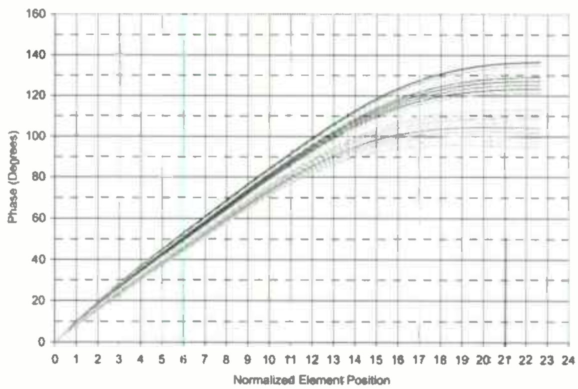


Figure 2b: Phase Aperture Distribution of End-Fed Array

The elevation pattern parameters affecting the performance of antennas within SFN are gain, beam-tilt, side-lobe level, and null-fill. As such, no unusual specifications are required for SFN than conventional broadcast antennas. The lower mounting height typically dictates lower down-tilt. However, increased tilt can increase the propagation loss (adding the loss of the ERP above the main-beam), improving isolation between cells.

Tower height reduction may necessitate a reduction in side-lobe level, as the side-lobes that cover the area within a mile of the tower may produce hazardous non-ionizing energy density or a signal level that saturates receivers. Another side-lobe constraint occurs with energy radiated above the main beam or above the horizon. These side-lobes may intersect tall buildings or propagate over the horizon.

The gain and shape of the antenna main-beam is determined by the interaction of each element in the antenna array. The level and shape of the side lobe structure is, to a large extent, determined by the elements at each edge of the array. Side-lobes are the result of the vector interference of the waves emanating from these end-most elements, adding to in-phase at some angles as peaks and out of phase at other angles as nulls. If the amplitude of the end-most elements are different, like those in Figure 2a, the destructive interference results in heavily-filled nulls. Conversely, if the amplitudes are equal, nulls are very deep. Aperture illumination that peaks at the middle of the array will generally have lower side-lobe peaks than those with peaks at the ends or edges of the aperture. These characteristics may be exploited to design arrays of slots or panel elements for optimum SFN performance.

CONCLUSION

Single frequency networks require antennas with specifications somewhat alien to broadcast applications. This requirement imposes limitations and specifications on the azimuth patterns, namely front-to-back ratio, beam-width, and roll-off. Antennas with slant-linear or circular polarization can increase the front-to-back and roll-off performance of an antenna array with impact on the self-interference of SFN cells. SFN antennas with higher beam-tilts and lower side-lobe levels have the benefit of reducing interference and radiation hazards for lower mounting heights.

REFERENCES

[1] Mattsson, A., "Single Frequency Networks in DTV," IEEE Trans. Broadcasting, vol. 51, pp. 413-422, Dec. 2005.

[2] E.A. Williams, ed., NAB Engineering Handbook, 10th Edition, Focal Press, New York, 2007.

[3] C.A. Balanis, *Antenna Theory – Analysis and Design*, Harper & Row, New York, 1982.

[4] E.A. Wolf, *Antenna Analysis,* Artech House, Norwood, MA, 1988.

[5] R. Beutler, *Frequency Assignment and Network Planning for Digital Terrestrial Broadcasting Systems,* Kluwer Academic Publishers, Boston, 2004.

[6] J.D. Parsons, *The Mobile Radio Propagation Channel,* John Wiley & Sons, Inc., New York, 2000.

Evaluation Of Building Penetration Loss For 100 Buildings in Belgium

DAVID PLETS, WOUT JOSEPH, LEEN VERLOOCK, EMMERIC TANGHE, LUC MARTENS
Ghent University, Dept. of Information Technology / Interdisciplinary institute for BroadBand Technology (IBBT); B-9050 Ghent, Belgium [david.plets@intec.ugent.be]

ETIENNE DEVENTER, HUGO GAUDERIS
Flemish Radio and Television Network (VRT); B-1043 Brussels, Belgium

ABSTRACT

Building penetration loss measurements of a DVB-H signal at 602 MHz have been performed in 100 buildings in Ghent, Belgium. Buildings are categorized in different types (office buildings, apartments, villas, mansions, terraced houses, stations) and the rooms are classified according to the number of outside radiated walls. The cumulative distribution function of the penetration loss is determined and lognormality is investigated. Models are developed to calculate the penetration loss as a function of the number of radiated walls. This research enables the calculation of indoor coverage probability for wireless networks.

INTRODUCTION

The digital broadcasting standard DVB-H (Digital Video Broadcasting - Handheld) enables a high data rate broadcast access for hand-held terminals (e.g., portable, pocket-size battery-operated phones). It is based on the specifications and guidelines of ETSI [1–4]. The broadband downstream channel features a useful data rate of up to several Mbps and may be used for audio and video streaming applications, file downloads, and many other kinds of services.

There is only limited literature available about penetration loss at UHF frequencies [4, 7, 8]. Mostly, only a limited number of buildings is investigated (e.g., 4 in [7]).

In this paper, the penetration loss of 100 buildings in Ghent will be investigated at 602 MHz (248 measured rooms in total). Buildings will be categorized in different types (office buildings, apartments, villas, mansions, terraced houses, stations) and models will be provided dependent on the number of radiated walls. This research will enable the broadcasters to calculate the indoor coverage probability for wireless networks, especially DVB-H networks.

The outline of this paper is as follows. First, the transmitting network and the measurement method will be described, followed by a discussion of the results and the proposed models. Finally, the conclusions are presented.

METHOD

Transmitting Network

The transmitting network is located in a suburban environment in Ghent, Belgium. The SFN (Single-Frequency Network) contains three base station antennas (BS). All transmitting antennas Tx are omnidirectional. The heights of these Tx are $h_{Tx} = 57$ m, $h_{Tx} = 64$ m, and $h_{Tx} = 63$ m, respectively. The EIRP (Equivalent Radiated Power) used for these Tx is 36.62 dBW, 39.93 dBW, and 40.90 dBW, respectively. The constellation used for the tests is 16-QAM 1/2 with an MPE-FEC rate of 7/8, corresponding with a useful bit rate of 9.68 Mbps [5,6].

Measurement Method

The penetration loss of a building is determined as the ratio of the average electric-field strength outside a building E_{out} on the ground floor and the average electric-field strength inside that building E_{in} (possibly on different floors):

$$\text{PenL [dB]} = E_{out}\ [\text{dB}\mu\text{V/m}] - E_{in}\ [\text{dB}\mu\text{V/m}], \qquad (1)$$

For detached houses, the average field strength outside the building is determined by measuring the field strength along a route around the building. For buildings attached to other buildings on one or more sides next to it (e.g., terraced houses), the field strength is measured by collecting samples on as many sides of the building as possible (in front of the building, garden or courtyard, etc.). To process the field strengths inside the buildings, each room is classified on the number of radiated sides (RS). A side of a room is radiated if it is part of an outer wall of the building.

For this measurement campaign, 100 buildings in Ghent have been investigated. These buildings are divided into 6 categories:

- office buildings (large building with multiple stories, rather large rooms, many large windows): 9 investigated buildings
  - coated (office building with coated windows): 2 investigated buildings
  - non-coated (office building without coated windows): 7 investigated buildings

- apartments (high building with lots of different (rather small) rooms): 7 investigated buildings
- mansions (large house, multiple stories, high ceilings, wooden floors, houses on both sides): 15 investigated buildings
- detached houses or villas (house with no building on any side of the house): 17 investigated buildings
- terraced houses (smaller house with buildings on both sides): 51 investigated buildings
    - shops (terraced house with a large front window): 5 investigated buildings
    - bank office (terraced house with a thick front window): 2 investigated buildings
    - private house (regular terraced house): 44 investigated buildings
- station (train station): 1 investigated building

This selection of the categories is representative for cities in Belgium.

The measurements are performed with a DVB-H tool implemented on a PCMCIA (Personal Computer Memory Card International Association) card with a small receiver antenna Rx (gain -5 dBi). The electric-field values are measured with this tool. The PCMCIA card is plugged into a laptop, which is used to collect and process the measurement data [5,6].

RESULTS

General Results

Table 1 shows the average penetration losses μ and standard deviations σ for the different building types and for a different number of radiated sides. The number of investigated rooms is shown between brackets for each combination of building type and number of radiated sides. Table 1 shows that the penetration loss decreases for rooms with more radiated sides: on average 9.24 dB for 0 RS, 8.53 dB for 1 RS, 7.57 dB for 2 RS, and 3.58 for 3 RS. The average penetration loss is smaller for villas (4.67 dB) than for mansions (8.19 dB), because villas mostly have more radiated sides than mansions. Private houses have an even larger penetration loss (9.76 dB), due to even less radiated sides, but a lower one than coated office buildings (21.94 dB). Coated office buildings have of course a larger loss than non-coated office buildings (5.30 dB), due to the metallized coating on the windows. The standard deviation varies from 2.37 dB (villas) to 9.37 dB (shops). In [4], a median value of 11 dB is assumed, with a standard deviation of 6 dB. Here, an average value for the penetration loss of 8.10 dB and a median value of 7.28 dB was obtained. The standard deviation is 6.23 dB. In the Flemish DVB-H trial, the network is designed for indoor coverage of terraced houses with 1 RS, i.e. μ = 11.19 dB and σ = 5.73 dB. These values agree reasonably well with those of [4], but the average value of 8.10 dB of course depends on the weight given to each building category. Although the houses are classified on the number of radiated sides, there is still a difference in penetration loss between the types of houses, even for rooms with the same number of radiated sides. The difference is caused by the non-radiated sides. The more rooms between a non-radiated wall and outside, the higher the expected penetration loss will be. For detached houses, the number of rooms with non-radiated walls is likely to be low as all four sides of the building are radiated. Terraced houses are likely to have (at least) two non-radiated walls with a larger number of rooms next to it, because other houses are present on either side of the house. The loss for terraced houses in Table 1 is indeed higher than for detached houses with the same number of radiated sides (1 RS: 11.19 dB vs. 6.22 dB, 2 RS: 8.78 dB vs. 4.74 dB).

Only mansions in between other buildings have been selected. The investigated mansions are thus terraced houses, but with different characteristics: older buildings, larger rooms, higher ceilings, wooden ceilings,... Their behavior is in between that of terraced houses and detached houses. Penetration losses equal 7.77 dB for 1 RS and 6.40 dB for 2 RS.

Cumulative Distribution Function

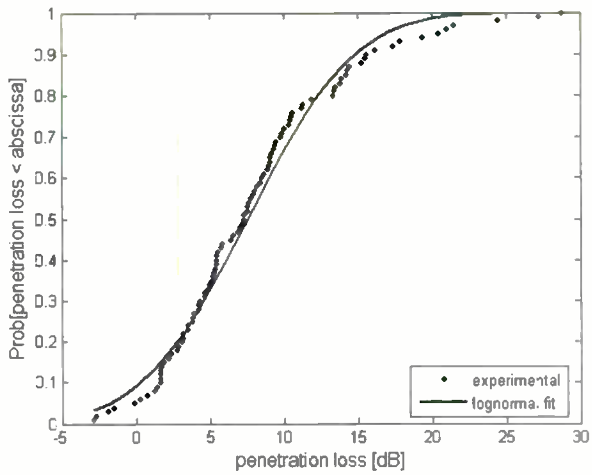


Figure 1: Experimental cumulative distribution function of the penetration losses of the 100 investigated buildings and lognormal fit.

Fig. 1 shows the cumulative distribution function of the penetration losses of the 100 investigated buildings. This cdf is then being fit using a cdf of a normal distribution (in dB). The average value of the experimental penetration losses is 8.10 dB, the standard deviation 6.23 dB. The lognormal fit has an average value of 7.57 dB, and a standard deviation of 5.68 dB.

The respective average values and standard deviations agree fairly well with the experimental values (Table 1, row 'Total'). Moreover, the data (in dB) also passed a Kolmogorov-Smirnov (K-S) test for normality at significance level α = 5%, indicating that the penetration losses (in dB) are normally distributed. In Fig. 2 the cumulative distribution functions are compared for 17 villas, 44 private houses, 15 mansions, and all 100 buildings together. It shows that villas have the lowest average penetration loss (4.67 dB), and the lowest standard deviation (2.37 dB, steepest slope of the cdf). Mansions have higher average penetration losses (8.19 dB), and the standard deviation is higher (3.70 dB). Private houses have an average penetration loss of 9.76 dB, with a standard deviation of 6.04 dB. The cdf of all 100 buildings lies in between these three curves (average value 8.10 dB, standard deviation 6.23 dB).

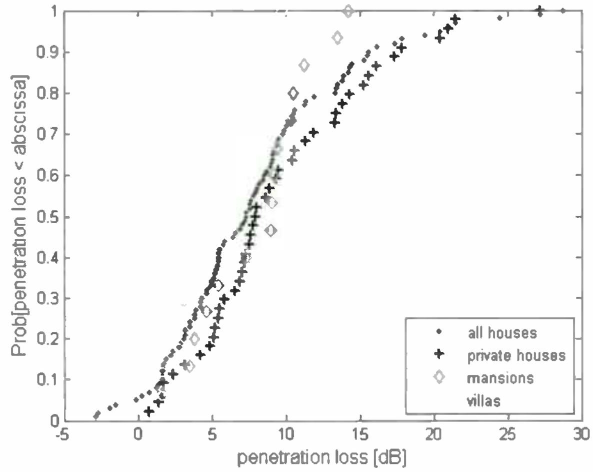


Figure 2: Experimental cumulative distribution function of the penetration losses of villas, private houses, mansions, and all 100 investigated buildings

Modeling Of Penetration Loss As A Function Of The Number Of Radiated Sides

In this section the penetration loss is modeled as a function of the number of radiated sides for the different categories of buildings. Fig. 3 shows the average penetration loss for different numbers of radiated sides and for three different building types (villas, mansions, private houses). The penetration loss

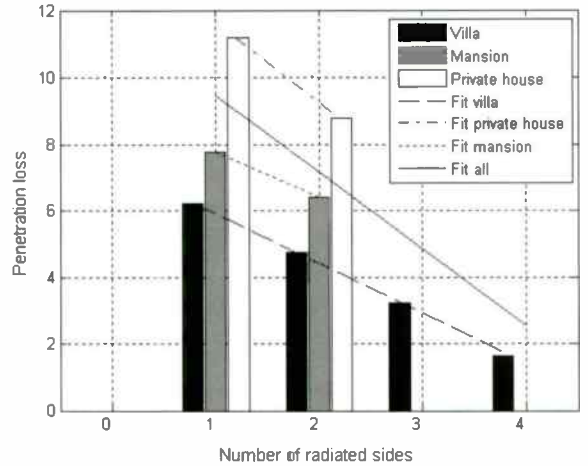


Figure 3: Average penetration loss for a different number of radiated sides for villas, mansions, and private houses and fit with one parameter (p) for each of the building types and for all three categories together.

is lower for villas than for mansions and private houses. The penetration loss also decreases for rooms with more radiated sides. The experimental values for the different number of radiated sides #RS have been fit for the different building categories (private house, mansion, villa) to the following model:

$$\text{PenL [dB]} = \text{PenL}_1 \text{ [dB]} - (\#\text{RS} - 1) \cdot \text{p [dB]} + \chi, \quad (2)$$

where PenL_1 [dB] is the penetration loss for one radiated side and p [dB] is a decrease factor (called the penetration loss decrease here), χ is the statistical variation on the model and has a standard deviation σ. Two types of fits will be investigated, one with two parameters (PenL_1 [dB] and p, noted as model 1), and one with one parameter (p, noted as model 2).

# Radiated sides		0 RS		1 RS		2 RS		3 RS		Total	
Type		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Office building	Non-coated	8.06 (5)	7.90	8.64 (3)	4.25	0.34 (3)	8.87	1.57 (1)	0	5.30 (7)	6.08
	Coated	25.05 (2)	1e-3	18.23 (2)	6.30	-	-	-	-	21.94 (2)	3.57
Apartment		2.49 (7)	5.56	-1.05 (5)	4.06	9.73 (3)	7.05	7.02 (1)	0	1.20 (7)	4.12
Villa		-	-	6.22 (4)	3.08	4.74 (13)	2.58	3.23 (4)	2.00	4.67 (17)	2.37
Mansion		-	-	7.77 (10)	3.36	6.40 (7)	5.32	-	-	8.19 (15)	3.70
Terraced house	Bank office	-	-	12.11 (1)	0	1.93 (1)	0	-	-	7.92 (2)	7.76
	Shop	30.66 (1)	0	7.57 (3)	2.28	21.15 (2)	7.97	-	-	13.39 (5)	9.37
	Private house	-	-	11.19 (17)	5.73	8.78 (32)	5.81	-	-	9.76 (44)	6.04
Total (average)		9.24 (15)	11.08	8.53 (46)	6.04	7.57 (61)	6.26	3.58 (6)	2.38	8.10 (100)	6.23

Table 1: Average penetration loss μ and standard deviations σ for different building types and different number of radiated sides with the number of investigated rooms between brackets.

Fit with one parameter (p)						
	μ_{dev} [dB]	σ [dB]	μ_{fit} [dB]	σ_{fit} [dB]	$PenL_1$ [dB]	p [dB]
Private	1.1e-3	5.72	-4.3e-5	5.53	11.19	2.41
Mansion	2.5e-3	4.12	-6.3e-2	4.46	7.77	1.37
Villa	7.3e-3	2.39	-0.28	2.42	6.22	1.50
All three	0.09	5.10	-0.43	4.59	9.45	2.30
Fit with two parameters (p and $PenL_1$)						
	μ_{dev} [dB]	σ [dB]	μ_{fit} [dB]	σ_{fit} [dB]	$PenL_1$ [dB]	p [dB]
Private	1.1e-5	5.72	-0.53	5.40	11.19	2.41
Mansion	3.2e-5	4.12	-0.06	4.46	7.77	1.37
Villa	1.5e-5	2.39	-0.22	2.44	6.24	1.52
All three	-2.3e-5	5.10	-0.56	4.54	9.68	2.49

Table 2: Values for μ_{dev}, σ, μ_{fit}, σ_{fit}, $PenL_1$, and penetration loss decrease p.

The root-mean-square (RMS) deviation of the measurement points was minimized with a linear regression fit, where $PenL_1$ [dB] and p were adjusted. First, a fit with two parameters ($PenL_1$ [dB] and p, model 1) was performed. Table 2 shows the resulting values for $PenL_1$ [dB] and p. The parameters $PenL_1$ [dB] equal 11.19 dB, 7.77 dB, and 6.24 dB and the penetration loss decreases p = 2.41 dB, 1.37 dB, 1.52 dB for private houses, mansions, and villas, respectively. The standard deviations σ are 5.72 dB, 4.12 dB, and 2.39 dB for private houses, mansions, and villas, respectively. Secondly, we perform a fit where only the parameter p of equation (2) is adjusted (model 2) and $PenL_1$ [dB] is defined as the average value of the penetration loss values for one radiated side (Table 1). The standard deviations σ are 5.72 dB, 4.12 dB, and 2.39 dB for private houses, mansions, and villas, respectively.

Table 2 shows that the values for $PenL_1$ [dB] and the penetration loss decrease p agree excellently for both types of fit, showing that the model with 1 parameter is sufficient.

To investigate the lognormality, the cdf of the difference between the experimental data and the model is analyzed. This cdf is then being fit using a cdf of a normal distribution (in dB). The RMS deviation is minimized with a linear regression fit, where the mean value and the standard deviation are adjusted. The resulting mean value is noted as μ_{fit} and the standard deviation as σ_{fit} and is compared with μ_{dev} and σ obtained from the difference of the experimental data and the model. Table 2 shows the values for μ_{fit}, σ_{fit}, μ_{dev}, and σ. Excellent agreement is obtained between μ_{fit} and μ_{dev}, and σ_{fit} and σ, respectively, indicating that the measured values are lognormally distributed around the model. The data also passed a Kolmogorov-Smirnov (K-S) test for normality at significance level $\alpha = 5\%$ for all four investigated models (private houses, mansions, villas, and all three categories together). Fig. 3 shows the average penetration loss for different numbers of radiated sides and for three different building types (villas, mansions, private houses) and also the model (fit with one parameter, p) for each of the building types and for all three types together. This figure shows that the respective models correspond excellently with the average penetration loss values for the different number of radiated sides for a certain building type.

CONCLUSIONS

In this paper penetration loss measurements for 100 buildings at 602 MHz in a DVB-H system are analyzed and discussed. A classification has been made on building type and on number of outer walls of the rooms (or number of radiated sides). The average penetration loss of 100 buildings equals 8.10 dB, the standard deviation is 6.23 dB. These values agree well with the ETSI values. Penetration losses have been modeled as a function of the number of radiated sides for terraced houses, mansions, and villas. Lognormal fits and K-S tests have demonstrated that the values are lognormally distributed around the different models. Rooms with less radiated sides have higher penetration losses (from 9.24 dB for 0 radiated sides to 3.58 dB for 3 radiated sides). Terraced houses have higher penetration losses than mansions and villas.

ACKNOWLEDGMENT

This work was supported by the IBBT−MADUF project, co-funded by the IBBT (Interdisciplinary institute for BroadBand Technology), a research institute founded by the Flemish Government in 2004, and the involved companies and institutions. W. Joseph is a Post-Doctoral Fellow of the FWO-V (Research Foundation-Flanders).

REFERENCES

[1] ETSI, EN 302 304 v1.1.1, "Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H)," Oct. 2004.

[2] ETSI, EN 300 744 v1.5.1, "Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television," Nov. 2004.

[3] ETSI, TR 102 401 v1.1.1, "Digital Video Broadcasting (DVB); Transmission to Handheld Terminals (DVB-H); Validation Task Force Report," May 2005.

[4] ETSI, TR 102 377 v1.1.1, "Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines," Feb. 2005.

[5] D. Plets, W. Joseph, L. Martens, E. Deventer, and H. Gauderis, "Evaluation and Validation of the Performance of a DVB-H Network," in 2007 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, Orlando, Florida, USA, March 2007.

[6] D. Plets, W. Joseph, L. Verloock, E. Tanghe, L. Martens, E. Deventer, and H. Gauderis, "Influence of Reception Condition, MPE-FEC rate and Modulation Scheme on Performance of DVB-H Networks," IEEE Transactions on Broadcasting (Special Issue on Quality Issues in Multimedia Broadcasting), submitted.

[7] E. F. T. Martijn and M. H. A. J. Herben, "Characterization of Radio Wave Propagation Into Buildings at 1800 MHz", IEEE Antennas and Wireless Propagation Letters, vol. 2, pp. 122-125, 2003.

[8] A. M. D. Turkmani and A. F. de Toledo, "Modelling of radio transmissions into and within multistory buildings at 900, 1800 and 2300 MHz", IEE Proceedings-I, Vol. 140, No. 6, pp. 462-470, Dec. 1993.

RF COVERAGE AND TOWER MOTION

Myron D. Fanton, PE
Electronics Research, Inc.

ABSTRACT

The Structural Standard for Antenna Support Structures and Antennas, TIA-222-G [1], include descriptions and guidelines for controlling microwave antenna sway. Though unregulated by the code, the same phenomena occur in broadcast antennas. With operational wind pressure, the antenna structures atop broadcast towers deflect, altering the location of the main-beam, side-lobes, and nulls of the radiation pattern. This in turn changes the coverage and received signal strength in the broadcast population areas.

The important parameters of the problem are defined, including sway, joint rotation, deflection, and beam-tilt. The relationship between the antenna electrical parameters and the structural parameters of the tower and antenna system is derived. In many broadcast structures used for TV and FM transmission, a half-degree change in the down-tilt of the main beam degrades the received signal over 10dB, potentially losing signal reception over large population areas.

The focus of the additions to the structural standards is to define and place engineering parameters on the design of new antennas and tower structures. Improving the design of antenna and tower equipment will result in improved coverage and more reliable transmission systems.

INTRODUCTION

The broadcast tower is an antenna support structure and is casually regarded a static structure. Towers and buildings supporting antennas move, and the movements affect the antennas and ultimately the signal coverage of the transmission system. The types and magnitudes of these movements are specified and controlled in the tower design process, and in other contexts are governed by the TIA-222-G structural standard [1].

The system dynamics of the antenna and tower structure are discussed herein, including the impact on broadcast coverage. Several previous revisions of TIA-222 contained a generic, minimal standard of 10dB RF signal degradation for any antenna [2]. This minimal standard, liberal as it is, has been abandoned in the recent TIA-222-G revision, which accounts only for microwave dish antennas. This relegates the integrity of the broadcast transmission system to conscientious tower designers and ultimately to the broadcast facility engineer.

TOWER AND ANTENNA GEOMETRY

The system and definitions are shown in Figure 1. Twist, ϕ, is a horizontal plane angular rotation of the antenna beam, and sway, θ, is an angular rotation in the vertical plane. Displacement, D, is defined as a horizontal translation of the tower. The total sway of the tower considers horizontal displacement of tower over its total height, and joint rotation, θ_1, considers only the rotation of the top of the tower in the vertical plane. The sway of a cantilevered antenna and a side-mounted antenna are shown as θ_2 and θ_3 respectively.

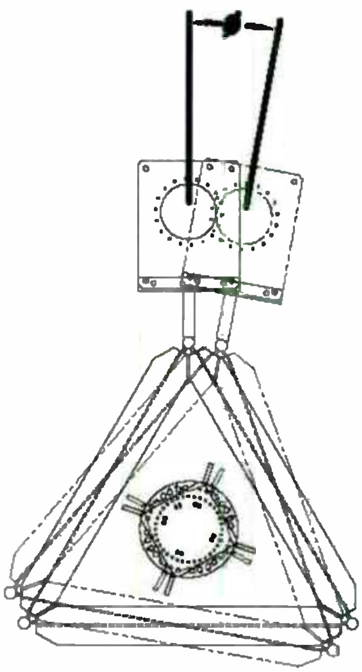


Figure 1a: Antenna/Tower Twist

The twist in the horizontal plane has little effect on coverage in a non-directional azimuth pattern and can be very important for a directional antenna, especially an antenna with regions of rapid roll-off. Maximum allowable twist is 5°. Coverage will definitely be affected by sway and joint rotation because the elevation pattern has regions where amplitude varies greatly. A typical value for tower-top joint rotation is 1.5°. Some antenna suppliers consider

1-2° antenna sway acceptable, and the maximum allowable sway is 4°. ERI high-gain antennas are designed with sways much smaller than 0.5° because of the important role this plays in system coverage.

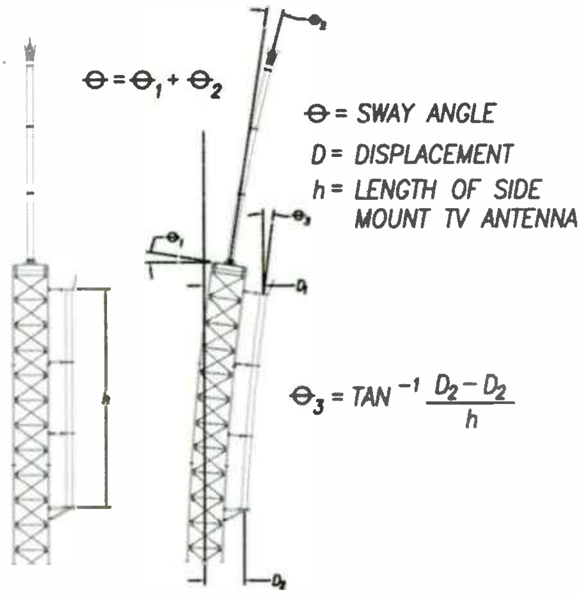


Figure 1b: Antenna/Tower Sway

ANTENNA PATTERNS

The antenna patterns in terms of the relevant structure geometry are the azimuth and elevation patterns. As antennas deflect and move in the dynamic mounting system, the patterns are rotated, and in some cases the patterns actually change. With significant changes in patterns, the antenna gain also changes.

Azimuth Pattern

Antenna azimuth pattern characterize the radiation in the horizontal plane. An antenna with a large aperture will have a very narrow beam, like a microwave antenna. Antennas with broad beams that change greatly in amplitude over a small angular region, also require large apertures. Apertures of this size are not generally available at broadcast frequencies, and the objective is to cover a broad angular region. Therefore most azimuth patterns are insensitive to typical twist magnitudes. In municipal coverage and interference scenarios, twisting antennas with sharp roll-off regions affects a large population and becomes an area of concern.

The antenna pattern parameter of interest is the change in relative field with respect to a change in angle. In most pattern plots, this is the slope of the radiation pattern. In many cases, the azimuth pattern data is plotted on polar coordinates to preserve the connection with map geometry. A polar plot will mask the field changes. Azimuth patterns for slot antennas and panel antennas are shown in Figure 2a and Figure 2b respectively. The differential amplitude for slot antennas can reach 0.8dB/°, and for panel antennas it can reach 4.5dB/°. An application with 5° of twist will produce 4dB of signal variation in a slot antenna and over 20dB in a panel antenna with steep roll-off.

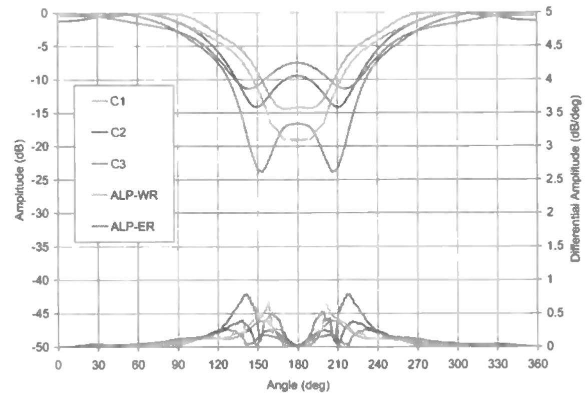


Figure 2a: Slot Antenna Az Pattern and Differential Amplitude

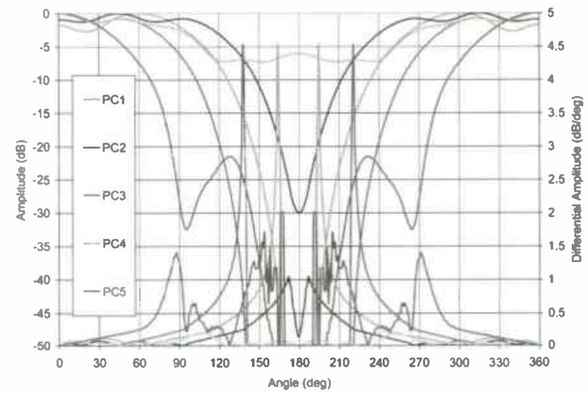


Figure 2b: Panel Antenna Az Pattern and Differential Amplitude

Elevation Pattern

The radiated fields of an antenna are related to the distribution of fields at the antenna aperture. For arrays of discrete elements, the aperture distribution is related to the fields at each element. Given the amplitude and phase of each radiating element, α_i, β_i respectively, and the known (linear) array geometry, d_i, the array factor may be computed by the following equation [3],

$$AF(\theta) = \sum \alpha_i e^{j(kd_i \cos\theta + \beta_i)} \quad (1),$$

where k is the propagation constant and θ is the elevation angle. If the array consists of similar elements, the product of the array factor and the pattern of one element produce the elevation pattern. Note that the quantity computed by Equation (1) is a complex number, providing both amplitude and phase of the elevation pattern.

The shape of the main beam of the antenna array is most pertinent to the discussion of antenna and tower sway, and the main beam is synthesized from the constructive addition of all array elements. The aperture distributions of a typical slot array antenna are shown in Figure 3 and result in pattern characteristics ideal for a broadcast station: high side-lobes and heavy null fill.

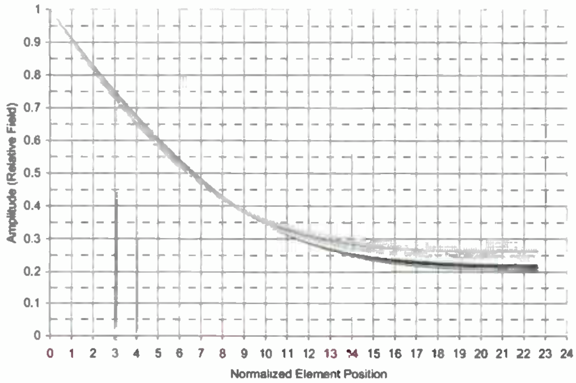


Figure 3a: Amplitude Aperture Distribution of End-Fed Array

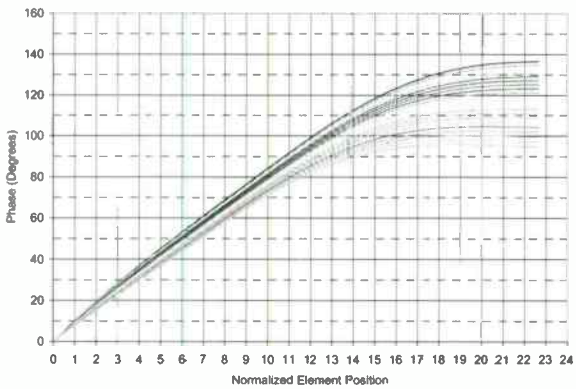


Figure 3b: Phase Aperture Distribution of End-Fed Array

Elements fed with uniform amplitude and phase will result in antennas with higher gain than those fed otherwise. Antenna arrays with greater numbers of elements, longer arrays, will also generally have higher gain. With increasing gain, more energy is concentrated in the main beam, and the slope is greater of the main beam edges. This is shown in the radiation pattern plots of Figure 4, the patterns of full-wave spaced, non-uniformly fed elements. The patterns have 20% 1st null-fill and have numeric directivity (with respect to a half-wave dipole) values very near their number of bays.

The differential amplitudes for the elevation patterns range from 2 to 15 dB/° and reach their maximum at the edge of the main beam. These areas of high differential amplitude are most susceptible to large sway dynamics, representing the majority of the coverage area.

The deflection of an antenna structure, be it cantilever or side mounted, introduces a phase taper across the antenna aperture that closely follows a parabola. The phase of the waves emanating from top element is advanced in the direction opposite of prevailing winds, causing a change in beam-tilt, null-fill, and antenna gain. As seen in Figure 5, the dominant change is that of beam-tilt, and it may be approximated in large UHF arrays by the value of the sway.

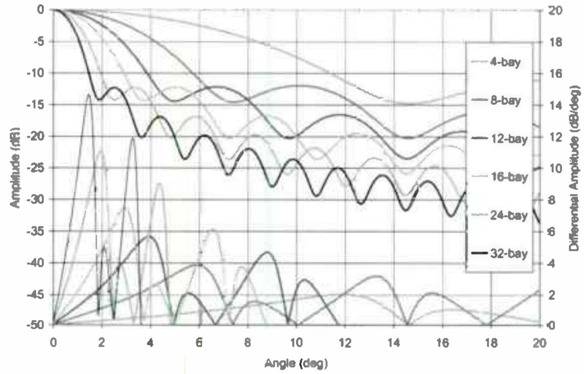


Figure 4: Elevation Patterns and Differential Amplitude

A deflection of the top element by one-quarter wavelength causes 90° of phase taper across the antenna aperture, causing a change in beam-tilt of 0.4° in a 32-element array. The sway for this array in the UHF-band is 0.4°, justifying the equivalence of the sway and change in beam-tilt. The parabolic displacement of each radiator reverses some of the null-fill, making the first null about 5dB deeper in our 32-bay example in Figure 5. As expected, a difference in amplitude occurs in the sides of the main beam.

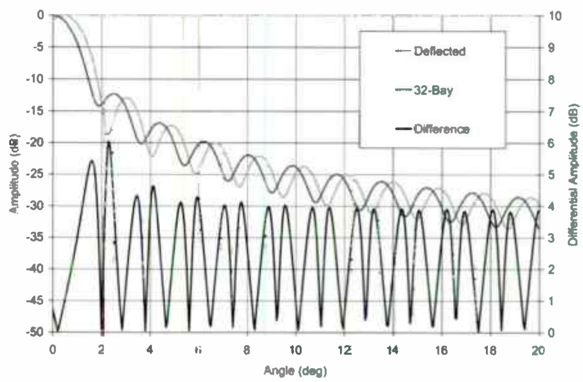


Figure 5: Parabolic Deflection in a 32-Bay Array

Bays/Gain	3dB Half-Beam Width (°)	10dB Half-Beam Width (°)	Differential Amplitude (dB/°)
4	6.65	11.40	2.0
8	3.30	5.75	4.0
12	2.20	3.85	5.5
16	1.65	2.90	7.5
24	1.10	1.95	11.0
32	0.85	1.45	14.5

Table 1: Antenna Beam Parameters

A summary of key antenna parameters is shown in Table 1: the antenna gain, 3dB and 10dB half-beam width of the elevation pattern, and the maximum differential amplitude from Figure 4. The beam width indicates how much sway would induce a 3dB or 10dB drop in signal level, and the differential amplitude indicates the worst signal degradation with one degree of sway.

BEAM SWAY SPECIFICATION

The sway of top-mounted antenna structures includes the joint rotation of the top-plate and the contribution of the antenna. Considering 500ft tall self-supporting towers as well as 1000-1500ft guyed towers, the joint rotation at the top plate is 1.6° for structures built to the limits of TIA-222-G. This combined with antenna sway, 0.4° in our 32-bay case, yields 2.0° total sway. This will move the 32-bay antenna beam from its peak down to the first side-lobe, about 15dB change in signal level.

Clearly the sway must be reduced in the top-mount situation. The antenna induced sway of 0.4° is within reasonable fabrication methods, and though this critical antenna parameter is not often specified directly, it may be assumed that when comparing, a heavier antenna has less sway. Also note that various designers define these parameters differently, so explicit, comparable specifications may be difficult to obtain. The sway of the tower top is a parameter that can be constrained to a value much smaller than 1.6°: 0.2° is achievable, 0.8° is typical. This reduces the total sway of the above example to 0.6°, smaller than the 3dB beam-width and relating to less than 9dB of signal change in the worst spot in the pattern.

The deflection of the tower top in the tower study produces a 1-1.5° sway in a side-mounted antenna, being worse for the self-supporting tower. For a 32-bay antenna, this is a change of 10dB (reference Table 1). The tower sway of the side-mount case is difficult to reduce in the tower design.

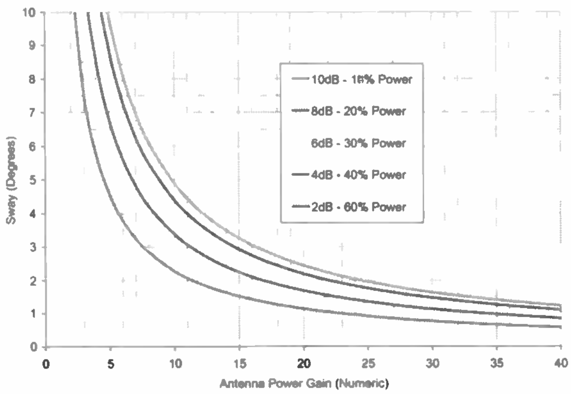


Figure 6: Allowable Sway vs. Antenna Gain

The case for a high-gain antenna is the most critical: it requires the largest structure and imposes the strictest constraints on the sway. The relationship of allowable sway and antenna gain is provided in Figure 6. The curves represent various antenna main beam-widths, meaning that sway of a given magnitude will reduce the signal in the main-beam areas by the value of the curve. For example, a sway of 1.45° will cause a 10dB change for a 32-bay antenna and a 2dB change for a 16-bay antenna. The curves of Figure 6 may be used to specify constraints on the tower and antenna structure and to specify the maximum gain of the antenna.

CONCLUSION

Tower and antenna twist and sway have been considered and analyzed. Tower twist affects the azimuth radiation pattern and coverage area, having greater impact on directional panel antennas with large differential amplitudes or areas of steep pattern roll-off. Combined sway of antenna and tower structures affects the elevation pattern, essentially changing the antenna beam-tilt the amount of the total sway. A high-gain UHF antenna contributes about as much sway as the typical tower structure, amounting to about 0.5° for each. This decreases the tilt in toward the wind and increases the tilt away from the wind, possibly changing the coverage by 10dB in the main beam region. The design curves of Figure 6 may be used to constrain the system sway and specify antenna gain to limit the coverage degradation due to tower dynamics.

REFERENCES

[1] ANSI/TIA 222-G-2005, "Structural Standards for Steel Antenna Towers and Antenna Supporting Structures," Aug. 2005.

[2] ANSI/TIA 222-F-1996 (R2003), "Structural Standards for Steel Antenna Towers and Antenna Supporting Structures," Mar. 1996.

[3] C.A. Balanis, *Antenna Theory – Analysis and Design,* Harper & Row, New York, 1982.

Radio RF & Transmission Systems

Wednesday, April 16, 2008
2:00 PM – 5:30 PM

Chairperson: Gary Kline
Cumulus Media, Atlanta, GA

***Save That Tower!**
Anne Gabriel, *Current*, Forest Lake, MN

AM Co-location: Money on the Table?
Lawrence Behr, LBA Group, Inc., Greenville, NC

Radio Transmitter Maintenance "Back to Basics"
Paul Shulins, Greater Media, Boston, MA

Implications of IBOC Injection Levels above -20dB
Gary Liebisch, Nautel, Bangor, ME

FM IBOC Building Penetration Tests at Elevated Digital Subcarrier Levels
E. Glynn Walden, CBS Radio, New York, NY

***Linear Effects of AM Narrow Band Antenna Systems: Characterization by Direct Measurement and Transmitter Based Equalization**
Tim Hardy, Nautel, Hacketts Cove, NS, Canada
Benjamin Dawson, Hatfield and Dawson, Seattle, WA

Free Software Tools for Design of AM Antennas
Van Richards-Smith, RadioTAB Network, Brisbane, Australia

AM Co-location: Money on the Table?

LAWRENCE BEHR
LBA Group, Inc.
Greenville, NC

INTRODUCTION

AM radio broadcast towers are rapidly becoming the new "hot topic" in the wireless communications industry. The name of the game today for wireless carriers and tower owners alike is "co-location, co-location, co-location" on existing structures. To date, this has not necessarily included attachment to an existing AM radio broadcast tower, even though there are an estimated 10,000 AM towers in existence in the United States. In many locales, new site opportunities are becoming stressed, leaving AM towers as strategic, or often, the only possible locations for new site opportunities. Even where open sites exist for new towers, local zoning and planning authorities often require that all collocation options be exhausted before "Greenfield" towers are permitted.

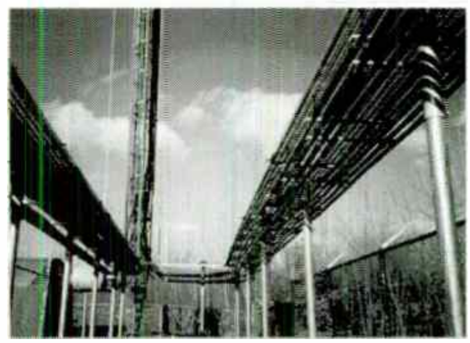
This places AM tower owners directly in the path of economic opportunity. How much opportunity? In many locations, cellular and PCS antenna locations rent for about $2500 per month. If four tenants can be attracted to an AM tower, that's potentially a $10,000 per month revenue stream. Even better, the wireless carriers typically pick up the costs of tower modification and on-going maintenance!

To tap this pot of gold, it's important to understand how AM colocation works from a carrier perspective, and what you, the AM broadcaster, must consider to present a viable site opportunity to the wireless industry.

After a bit of orientation, some specific operational and technical concerns will be presented.

WHY NOT AM CO-LOCATION?

Historically, the wireless industry has been warned to stay clear of AM radio stations at all costs. This not only had to do with attaching to them, but also with reference to locating nearby and interfering with their broadcast patterns. Indeed, wireless carriers must prove to the Federal Communications Commission that they have considered and corrected all such problems when constructing or modifying any tower within three kilometers of an AM station. This has left underserved "holes" around many AM's, often in prime suburban areas.



A Folded Unipole AM tower handles dozens of antennas and cable runs for cellular and PCS systems

In the past, AM towers were considered unsuitable for antenna attachment by most cellular and PCS wireless carriers due to presumed grounding difficulties, interference and safety considerations. Coordinating construction between the vastly different AM and wireless cultures was frequently a slow and painful process. From an engineering perspective the process of integration and demonstrating license compliance to the FCC often required tinkering, delays and costs unacceptable to wireless carriers and broadcasters alike. AM station owners have also wanted assurance that the co-location methods proposed by the wireless operator were reliable, proven, and acceptable to the FCC and would not harm their signal coverage pattern. In the past, these outcomes could not be readily or easily assured. Many wireless system designers, and AM tower owners themselves, are yet unaware that new technologies are now available to solve these problems and efficiently integrate wireless and AM systems at reasonable costs.

In the AM band, the tower itself is the radiating element without a need for attached antennas. However, wireless antennas and coaxial cables are self-contained systems that merely attach to their support structures. Achieving compatibility with the AM tower through electrical integration or isolation of wireless antennas is a challenging engineering exercise. Since AM broadcasting is a specialized field, many wireless system designers and constructors, not being conversant with lower frequency technology, have been unaware of the techniques available to make wireless compatible with AM. Few broadcast engineers and consultants have experience in co-location, except for the odd STL or FM standby antenna. Thus, many potential AM tower co-locations have been avoided as technically impossible or prohibitively expensive. Sometimes large additional costs have even been incurred for detuning the resulting new tower near the avoided AM station!

The solution to these problems is actually straightforward in most cases, and can be readily implemented at reasonable cost by using a qualified consultant and the latest hardware solutions.

LBA Technology ColoCoil™ isolates multiple wireless coaxial cables

CURRENT TECHNOLOGY FOR AM CO-LOCATION

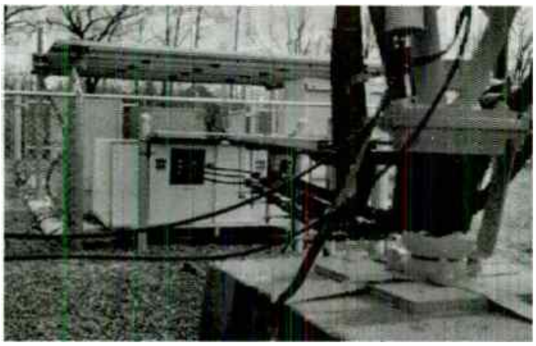
Today, equipment is available from several AM antenna system manufacturers which will provide the ability to co-locate communications transmission systems on both single tower and multiple tower AM antenna systems. Companies such as LBA Technology Inc. and Kintronics Labs have equipment which once installed, insures that the wireless antenna and coaxial cable installations have virtually no effect on the host AM tower(s), and the AM signal has no effect on the wireless antenna. Moreover, additional antennas and transmission lines can be added to the tower in the future without reengineering the AM isolation. This means the owner can lease additional space to other wireless carriers, limited only by the tower structural capability.

On non-directional towers, an advanced folded unipole isolation system is typically used. This results in direct grounding of the AM tower. Wireless antennas and transmission lines are mounted and bonded directly onto the structure. The folded unipole uses a unique wire cage impedance transformer. In the LBA Technology implementation, lower portions of the cage are heavily insulated and spaced away from the tower to allow ready operational access to the wireless antenna system by tower climbers, even while "hot." Such folded unipole co-location systems benefit the AM station with improved efficiency, "air sound", and lightning protection, thus enhancing the co-location experience for the station. Conversion to an advanced folded unipole type co-location system may entail replacing the station's ATU (antenna tuning unit) and some "on tower" construction.

Directional stations use multiple towers to form an FCC licensed radiation pattern crucial to protecting other stations from interference. This licensed pattern may not be disrupted by co-location. The cost-effective approach to this end is to employ specially designed isolation systems between the base station equipment and the AM tower. The manufacturers mentioned earlier, along with a qualified Consulting Engineer, can assist in this design.

AM CO-LOCATION PLANNING

Planning for AM co-location begins with an analysis of the station facility. While all AM stations may theoretically be used for wireless co-locations, practical factors may make some facilities economically or technically unattractive to develop. Where multiple towers exist, the most favorable of the towers must be chosen. A wrong choice may add thousands of dollars to project costs. AM operations are at times complex, with different towers or even different sites being used for day or night transmissions, at several power levels. This may impact costs and operational aspects of the project. For instance, selection of a tower used only at night could be a benefit to daytime construction and maintenance activities.



LBA ColoCoils™ in use at directional AM station with fence isolation from hot tower

Normal site factors, such as access and construction convenience, must be evaluated. Structural suitability of the tower and any required augmentation must be considered and viewed in terms of AM system parameter impacts. Further, each AM tower has beneath it a radial ground network of miles of copper wire. This is essential to proper AM operation and is mandated by the FCC. Special planning and construction precautions are needed to protect ground system integrity to avoid disruption and expensive replacement of the system, but benefits the wireless installation with superior lightning protection!

Because AM towers operate "hot" at high RF voltages, proper selection of candidate towers is very important to cost effective and operationally supportable co-location. There are significant safety and operational issues, which must be carefully dealt with in installation and maintenance of wireless equipment near AM towers. Fortunately, these RF concerns can be managed. For instance, it is not true that AM stations must always be shut down for installation and maintenance of co-located antenna equipment. Both the FCC and OSHA permit work on "hot" AM towers with proper power levels and precautions. Alternate operating modes and temporarily deployable AM towers similar to the familiar wireless "COW" can be employed should shutdown be needed. With a high level of expertise employed in the planning phase, these safety and operational concerns can be addressed.

The location of the wireless equipment shelter or pad must also be carefully chosen to minimize AM interactions, and appropriate shielding, carrier equipment grounding, and filtering may be needed. Electromagnetic field modeling techniques allow experienced designers to specify exact locations and outfitting for equipment packages to minimize interactions.

It is important to understand that the wireless carrier will essentially drive the installation and will demand the highest quality hardware, detailed planning, and full documentation. These are not always well understood or appreciated by AM operators and engineers.

It is important that the AM co-location integrator be involved in the process at the site acquisition stage. There are many subtleties to negotiation of a satisfactory lease or acquisition agreement for which expert input is mandatory. Advance screening of all site candidates can also reveal possible AM co-location candidates and avoid unneeded detuning situations.

SUCCESSFUL OUTCOMES

In summary, professionally managed AM co-location is not only possible, it has been successfully accomplished throughout the country for tower service providers and for wireless carriers such as Sprint Nextel Corp., AT&T Mobility and Verizon Wireless. With careful planning, competent project management, and the use of quality hardware integrated into the overall AM co-location site development process, success is certain. Furthermore, professional interaction with the AM host makes the station a willing and positive partner in the long-term co-location relationship.

APPENDIX A – PLANNING AND OPERATIONAL ISSUES

• WIRELESS CARRIER CULTURE

AM broadcasting and wireless carrier operations are two very different cultures. Neither typically knows or understands much about the other. Part of the art of collocation is successfully blending the two.

Carriers are highly organized and follow rigid procedures on planning, construction, and regulatory matters. They use the best materials, and are highly time and cost driven on projects. They design for 24/7 reliability, and have the technical depth to back it up. Usually, they will insist on strict control of the entire colocation project.

AM stations need to cover themselves with a fair, but protective lease with appropriate indemnities. AM colocation is first a business and legal project. Once the project is engineered and agreed upon, the AM local technical resources typically have only a minor role in the carrier's execution. The AM must ensure that the carrier has engineering resources who understand AM to oversee its successful completion. The use of specialist broadcast business attorneys in these negotiations is recommended.

• ZONING AND PERMITTING APPLIES TO YOU

AM antenna towers are attractive to wireless carriers for several reasons. High among them are often favorable zoning and permitting biases that reduce red tape and community opposition to new cell antennas. The carriers have entire departments and legal staff dedicated to zoning and permitting, and so will vet these issues early in the negotiation process. Broadcasters tend to have a "do it yourself" approach to such things, and more than one co-location project has ended up on the rocks because of broadcaster end-run attempts on zoning. It is important to have a specialist site advisor in carrier zoning issues on your side. This is something that must be done formally and by the book!

• TOWER REPLACEMENT OR REINFORCEMENT

Often AM towers are inadequate to support cellular antenna systems and lines. Frequently, if the tower is already zoned and permitted for additional antennas, the cellular company or tower company that leases the AM tower will structurally upgrade the tower, or even replace it. While this is good, the plans must be carefully drawn with AM RF issues in mind. More than one replacement tower has gone up without insulators!

• PROTECT ONGOING OPERATIONS

An ongoing operational agreement should be put in place. It should include a protocol for coordination between parties on maintenance with procedures and designated authorities on each behalf to ensure compliance with RF safety plans and to protect operational integrity of both carrier and AM facilities.

The AM must recognize that there will be considerable access traffic to the tower and compound. Not just tower crews, but electronics maintenance techs, generator techs, generator fuel trucks, and yard maintenance crews, for instance. Almost never is an AM tower visited, in the experience of most stations. This means that the AM may have to beef up general site security and inspect its transmitter site more often.

• PROVIDE FOR ADEQUATE CARRIER SPACE

Successful colocation sites require adequate carrier ground space at the tower base, and proper access for roads and utilities. Carriers spend a good deal of A&E (architect and engineer) money to design the tower base environment, to say nothing of building it out. Expect the AM tower base and ATU to be in a separate fenced compound. The colocation devices should be located at that boundary so the "hot" part is inside the fence, and carrier techs can safely access coax ports on their side for testing. Tower mounted isolation devices can be a safety hazard, and, as such, are not as much favored.

Carrier equipment may be housed in shelter buildings or stand alone environmental cabinets, all of which must be specially shielded, or properly situated to avoid RF ingress. Unless your AM tower is very close to a building, there is little chance that building space will be required. Of course, 365/24/7 access is a given, imposing special challenges upon the use of a tower in swampy or inhospitable locations.

• RF HAZARDS NEED ATTENTION

Contrary to general belief, it is possible to work on "hot" AM towers. FCC guidelines permit low power operation without RF hazard strictures at antenna power ranging from about 3000 watts at 540 kHz to 300 watts at 1700 kHz. Appropriate engineering investigation can guide tower selection and/or work practices to facilitate carrier tenant maintenance.

Generally speaking, the emissions from most cell carrier installations are insignificant at ground level and won't be a problem for AM stations. However, a proper RF safety plan is essential to ensure safe and FCC compliant on-tower work. Particular care in evaluation must be exercised if the tower is to also accommodate higher power services such as FM, terrestrial satellite repeaters or mobile TV systems. Furthermore, many local authorities are now requiring formal proof of compliance with FCC RF hazard rules.

- **HOW MUCH AM POWER IS TOO MUCH?**

As power on a tower goes up, potential carrier colocation issues increase. The feasibility of colocation is a matter of budget and operational convenience for both the AM and carrier. The choice of carrier isolation technology is an important factor in the analysis. In a directional array, the important factor is the power in the tower under consideration, not the licensed power. A thorough engineering investigation is necessary in all but the lowest power situations. It has been our experience that successful collocates can be done on 10,000 watt towers, but usually "lower is better"!

- **AM DIRECTIONAL COMPATIBILITY**

Collocation on one or more towers of a directional array is often quite practical. In fact, in terms of maintenance flexibility, such installations can offer advantages over non-directional collocations. A good directional array co-location requires a balance between carrier operational logistics and AM technical factors.

The number one factor in tower selection from a carrier perspective is convenience for fast and safe maintenance and/or system modifications. This suggests the choice of a tower that is inactive in one mode (preferably daytime) and that uses low power. At the same time, that tower should be low impedance and stable for minimum perturbation of array operation. Part of the equation is accessibility for roads and underground utilities without unreasonable disruption of AM transmission line and ground system elements.

For the protection of both the AM and the carrier, it is very important that coordination of RF grounding, equipment placement, interim operation and RF safety plans be accomplished and memorialized before construction begins.

Construction will inevitably result in some AM array disruption, and plans should be in place for appropriate FAA notices, FCC STA's, power reduction, pattern changes, or other actions to facilitate the construction phase. Of course, associated costs should be factored into the lease agreement.

- **COLLOCATION ON MULTIPLEXED AM SYSTEMS**

It is possible to collocate on a multiplexed AM tower. On a folded unipole system, no special provisions need be made. If multiple isocouplers are employed, they will generally not be AM frequency sensitive, but the combined shunt capacity may have significant impacts on the individual AM antenna impedances, which may compromise the multiplexer tuning. Devices like the LBA ColoCoil™ are AM frequency sensitive, but can be fabricated for multiplexer AM frequency operation with high isolation and minimal multiplexer impact. In each case, a careful engineering review of present and future system requirements is essential before selecting an isolation technology.

APPENDIX B – TECHNICAL ISSUES

- **FCC AND FAA MATTERS**

In general, no prior FCC or FCC authority is required, of the station, to add collocation antennas to an AM tower, so long as the height is not increased. On completion of construction, FCC Form 302 can be filed to recognize any changes beyond current license limits.

- **IMPACTS ON RADIATION PATTERNS**

In most cases, adding carrier antennas to AM towers will not significantly change the vertical radiation pattern or efficiency. There will be somewhat greater changes to the antenna impedance because the electrical length of the tower is made a bit longer as the effective radius of the tower is increased by antennas and lines. However, this can be good news because the "Q" of the tower will often be decreased, improving bandwidth. Particularly in a critical directional array, these matters should receive engineering evaluation to verify acceptability

- **IMPEDANCE IMPACT OF COAXIAL LINES**

Many broadcasters are familiar with isolation involving a single coaxial cable as for an FM or STL. In carrier operations, multiple antennas are employed. It is not unusual to have 12, 24, or more cables and antennas on the tower. Not only can the physical presence of this hardware affect the impedance of the tower at AM frequencies, but the isolation devices may have large cumulative impacts on impedance. For instance, if 24 isocouplers with 20 pf capacity each to ground are arrayed, that is 480 pf total, or about 300 ohms shunt reactance at 1 MHz. That will transform an antenna impedance of 89 -j120 to about 45 –j96, a value perhaps inconvenient for the installed ATU!

- **GROUNDING AND BONDING PRACTICES**

Carrier engineering departments have developed standardized grounding practices which are quite different and often conflicting with the mostly ad hoc practices employed in AM radio. Further, carrier practices are lightning, and not RF, oriented. Integration of these approaches is a significant issue to resolve that requires engineering knowledge of both. The protection, repair, or replacement of the AM buried radial ground system needs to be factored into early planning activities.

- **PROVIDING FOR COAX REPLACEMENT, ADDITIONS**

Rarely is a collocation tower dressed with lines and antennas, then forgotten. Rapid technology changes in carrier systems result in almost routine changes or additions of antennas and lines. The isolation method and infrastructure deployment must give maximum flexibility to the carrier, and keep tower crew workloads to a minimum. The old quarter wave stub isolation method, used for years for its simplicity in isolating a single coax, fails seriously when confronted by modern cable installation requirements, and is rarely, if ever, employed today.

- **ISOLATION DEVICE CONSIDERATIONS**

Modern carrier systems often employ tower top amplifiers and antenna positioning devises powered over the signal coaxial cable. Devices such as isocouplers do not pass DC or AC, and thus have limited utility. The isolation system should be designed for end-to-end DC and AC connectivity. for which devices like folded unipoles and ColoCoils™ are well suited.

Because of the wide range of frequencies employed by carrier systems (700 – 3600 MHz), and stringent VSWR specifications on the system, it is preferable to have no cable discontinuities. Devices, such as isocouplers, that are tuned to a pass frequency seriously limit broadband carrier installations. The most desirable installation is on a folded unipole tower where no interruptions in the coaxial cable are required at all.

- **FOLDED UNIPOLES CAN IMPROVE BANDWIDTH**

Folded unipoles are time tested in their ability to both match an AM tower, and make it "cold" at the same time. Towers with folded unipoles installed may be ready for colocation with little, if any work. In many cases, the folded unipole will need to be replaced with one that is built specifically for collocation. These systems employ wires that are spaced and insulated in such a way as to allow safe tower climber access. They are also built to accommodate clearance to multiple coaxial lines and ice bridges.

Non-directional series-fed towers can often be converted to folded unipoles with only replacement of the antenna tuning unit. The complexity and economics are a bit different with towers in a directional array, however. Since the folded unipole significantly alters the impedance and phase matching conditions at the base, you will likely need a new ATU to accommodate these parameters. This will usually require redesign by your consulting engineer, retuning of the array, and a partial proof of performance.

The upside of a folded unipole is not only isolation for unlimited coaxial lines, but better bandwidth and lightning protection. It also eliminates the need for tower lighting and sample loop isolation.

ACKNOWLEDGEMENTS:

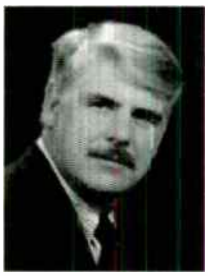
Thanks to the following people that have provided valuable insights to the subject matter of the paper:

Chuck Martin, VP Site Services
Lawrence Behr Associates, Inc.
cmartin@lbagroup.com

Richard Connor Riley, Esq.
Riley Broadcast Siting Services
rcriley@siting-advisors.com

Erwin G. Krasnow, Esq.
ekrasnow@gsblaw.com
Henry A. Solomom, Esq.
hsolomon@gsblaw.com
Garvey Schubert Barer
Broadcasting Business Attorneys

THE AUTHOR



Lawrence Behr is an RF systems expert with over 40 years in AM broadcast and wireless communications. He is CEO of LBA Group, which provides consultations and hardware for AM Collocation. He may be contacted at LBA Group, Inc., 3400 Tupper Drive, Greenville, NC 27834. 252-757-0279. lbwireless@lbagroup.com.

Radio Transmitter Maintenance "Back to Basics"

PAUL SHULINS
Greater Media
Boston, Massachusetts

ABSTRACT

Now more than ever broadcasters need to make sure their transmitter facilities are in good shape. While it is true that over the years there has been a dramatic improvement in the efficiency, ease of service, and reliability of RF Transmission equipment, the pool of experienced RF Engineers has seen a substantial decrease in population. Some of this has been due to many of the older engineers reaching retirement age, but a more significant reason has been the recent trend of young men and women entering the field to be more geared toward the Information Technology (or computer) side of radio engineering, and why not? The line between computers and transmitters is becoming less defined each day. Most radio transmitters manufactured during the previous two decades contain some kind of micro processor or in some cases are full blown personal computers! With the advent of HD Radio this has become particularly true. In this age of computerized broadcasting, it is vitally important that we not loose sight of some of the basic engineering principles and standards that have always ensured high quality and reliable service to our listeners. This paper will review some of the more important areas of transmitter site engineering and maintenance that have been long established, but sometimes forgotten.

CLEAN IS GOOD

Obviously a clean transmitter site is a good transmitter site. Keeping a site clean and organized has many benefits. It is usually beneficial to employ a redundant closed loop HVAC System. A closed loop system keeps all types of dirt, pollen, and humidity out of your shelter. Transmitters that operate with high voltages will especially benefit from the clean environment because high voltage tends to attract dirt and can lead to unwanted arcing. Proper air flow in transmitters, audio processing gear, and computers is also critical, and dirt and pollen can quickly clog air filters leading to overheating of components and eventual failure of critical systems. In addition, having redundant HVAC systems helps to assure that un-attended transmitter sites will be able to continue to stay cool if one of the air conditioners were to fail, until repairs can be facilitated. It has become a common practice to install a tertiary ventilation system that will simply bring in filtered outside air and draw it across the room to cool the area in the unlikely event that both HVAC Units were to fail. It goes without saying that a system as important as HVAC requires regular service and a contractor should be maintained to perform this service as required (usually at least two times annually).

PRIMARY AND BACKUP VERSES MAIN AND ALTERNATE

An ideal transmitter facility will have redundancy for most all the equipment. In many cases the costs associated with buying "two of everything" are impractical. If you are able to duplicate equipment one strategy that has proven to increase overall reliability is to have a main and an alternate "chain". In other words building a facility where each transmitter is fed by its own set of processing and STL path, helps to insure that in the event of a problem (i.e. RF off the air, audio off the air, or audio problems like distortion etc.) you have the very best chance of bypassing the problem as quickly as possible until the problem an be addressed by simply switching to the alternate transmitter chain, including the RF amplifier.

A "main/alternate" configuration implies that both systems are close to being equal in terms of reliability and quality. An ideal main/alternate configuration will be run on a regular changeover schedule. Depending on your individual circumstances the times will vary, but a popular method has been to alternate systems on a quarterly basis. This helps prove to yourself that both systems are working properly, and in the event of a failure on any one system, you be confident that you can reliably switch to the alternate system and be sure your business will not be compromised.

If you have a primary/backup system, your backup transmitter may not be as modern and therefore may not offer you the efficiency and reliability of your main transmitter. This backup may also not be a full power transmitter. So while you may not want to spend much time with this backup on the air, it is extremely important to test it regularly and maintain it as well as your primary… after all if you need to go to this transmitter, you may be relying on it as your last resort until you repair the primary.

SAFETY FOR PEOPLE AND EQUIPMENT

Often times preventative or required maintenance is performed on transmission gear with lethal voltages present inside. While all manufactures of modern broadcast equipment take precautions to interlock their equipment when a door or access panel is removed, occasionally it is necessary to bypass these safety measures in order to troubleshoot the problem. When this happens the engineer opens himself or herself up to added risks. The consequences of making a careless mistake can be deadly. Compounding the problem is the fact that many of these maintenance sessions are performed late at night at a time when the engineer may be fatigued, and could carelessly make a mistake. The issue is a serious one, and requires one to carefully manage the risks associated with transmitter maintenance. Some of the tools available to the engineer to lower this risk include using checklists, making sure the shorting stick is used before contacting any potential high voltage points, and never working alone. Having a second person double checking your safety measures can be a life saving decision.

The equipment itself needs to also be protected. If you employ an air cooled dummy load make sure that it is interlocked to remove the RF source from whatever transmitter is connected to it, should the cooling fan in the load fail to operate. The same is even more relevant with a water cooled dummy load. If sufficient water flow is not maintained the load element will be destroyed in a matter of seconds. A properly adjusted, high quality external water flow switch is usually required to protect a water cooled dummy load.

Just as the load needs to be protected, those who use a manual or motor driven RF transfer switch need to make sure that these switches are interlocked to any transmitters they are connected to. Usually moving an RF contactor with RF present can cause serious damage to the switch.

Probably the most important and unfortunately one of the widest overlooked interlocks needed at transmitter site are the VSWR Protection Interlocks. Even though modern transmitters provide internal VSWR Fold-back protection, it is absolutely necessary to have an external VSWR meter capable of interlocking the transmitter(s) installed at an appropriate place in the transmission line system, and tested on a regular basis. The cost of installing and maintaining such a device is a small insurance premium to pay for a whole lot of protection afforded to your expensive transmission line and antenna system.

Another vitally important consideration for those using transmission lines with an air dielectric is proper pressurization of the transmission line system with either dry air, or an inert gas such as nitrogen. The object here is to keep the voltages between the inner and outer conductors of the transmission lines from arcing over and destroying the line. Using a dehydrator to dry the air to a very low dew point, guarantees that the moisture content in the line will be low enough to avoid arcing. The other popular method that is very effective is to infuse the line with nitrogen gas (either from a bottle or a nitrogen generator). This effectively removes any moisture in the line and helps to prevent flash over. Keeping the transmission lines full of dry air at several pounds per square inch is just about the best thing you can do to help protect your antenna and transmission line system. It is generally agreed that having an electric pressure switch to monitor the transmission line's gas pressure (and relay the status to the operator on duty) is imperative so that if an air leak develops, it does not go un-noticed.

Finally for air cooled tube type transmitters a routine testing program for the air flow safety systems is important. Even though all tube type air cooled transmitters employ a safety air flow switch, unless they are tested on a regular basis, it is unknown whether they are actually working to protect your equipment. Some manufacturers provide a set of instructions for testing the air flow switch, and those should be followed if they exist. One popular method is to slowly restrict the air flow to the intake filter on the transmitter under test (with a piece of cardboard), until the air interlock is no longer satisfied, and verify that the filament gets turned off. Having this airflow switch operational when a blower failure occurs may save the transmitter from burning up!

EXTENDING TUBE LIFE

Most tube manufacturers furnish a data sheet with the tube that along with specifications for tube operating parameters also depicts recommended filament voltage settings. Typically it is advantageous for a brand new tube to be run at the rated filament voltage for the first several hundred hours and then at a reduced voltage for the life of the tube. Eventually when the emission drops off the voltage can be increased to a certain point to further extend the useful life of the tube. Since tube characteristics vary widely, the tube manufacturer is the best source of advice on the optimum program for your tube.

GEOGRAPHICAL SITE DIVERSITY

Have a backup or alternate transmitter site is the ultimate in terms of added security for your signal. Such an arrangement is often very expensive, but in many high stakes markets where loss of air time can have huge consequences, many broadcasters have seen the wisdom in providing a backup site. Events as simple and routine as tower maintenance all they way up to actual catastrophes when a tower, antenna or transmission line is damaged by vandalism, earthquake or weather, have resulted in broadcasters being forced off the air for long periods of time. Having another place to go is indeed a saving grace.

Transmitter Maintenance Logs

Running your transmitter facility efficiently and with a minimum amount of down time requires good record keeping. Most engineers accomplish this with maintenance logs. These records help keep an accurate record of service to equipment, as well as regular records of calibrations. Perhaps the most useful aspect is the fact that engineers can refer to a record of what normal parameters are, and thereby easily understand how trends in parameters are affecting the operation. In addition, a well thought out transmitter maintenance log acts as a type of check list, requiring the engineer to go through all the parameters at the facility and record them individually on a routine basis. This of course allows for a complete comparison to the previous entries to determine if any values are drifting. These drifting values can help predict maintenance that will be required in the near future to help avoid unexpected failures.

Normally a paper log template is custom designed by the engineer maintaining the facility. This log acts a check list and a record for each significant visit to the site where maintenance is performed or a log is recorded. These records then are usually kept on a clip board or a loose leaf notebook for easy access. Looking up the previous log can be very enlightening to an engineer when something looks wrong. A record of what is normal compared to today's readings can help pinpoint where a problem may be.

With the availability of smaller computers and PDA's today it is now practical to electronically record and analyze maintenance log data. This has a few advantages over the paper method.

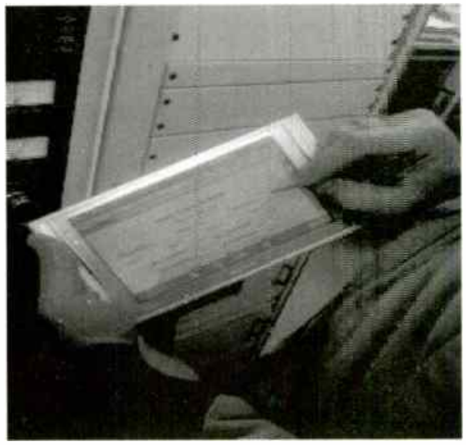


Figure 1. Portable Sony Location Free TV Web Browser being used for data entry into the electronic maintenance log

For example an engineer using a portable electronic device such as a laptop computer, PDA, or other hand held web browser today can directly input information read from the equipment meters for accurate storage, and later retrieval. Information handled in this way has several advantages over the paper method:

1. The data is stored in a logically formatted structure and is always easy to read.
2. The data can be manipulated easily and sorted by field
3. It becomes apparent when trends in readings start to occur and an engineer can more quickly take action to address any equipment shifts in performance.
4. Maintenance Logs can be printed out in a uniform way with color coding and different size fonts to highlight changes the engineer might want to be aware of.
5. Maintenance Logs can also be posted to an internal or external web page so that it is possible to access the data from many different locations.
6. Log fields can be individually sorted by value so the person analyzing the data can see the highest or lowest value in a field.
7. Comment fields in logs can be sorted by keyword. For example search the comments field for the last time a tube was changed.
8. Mathematical calculations can be routinely done to calculate a running time tally on items such as the time between

tube changes in a particular transmitter, or elapsed run time of a dehydrator. These calculations are of course possible to do manually, but to have them done routinely and automatically each time a log is completed can be invaluable especially when an engineer would not have the time to run through each calculation every week. Here the statistics present themselves for the engineer to evaluate and take action if required.

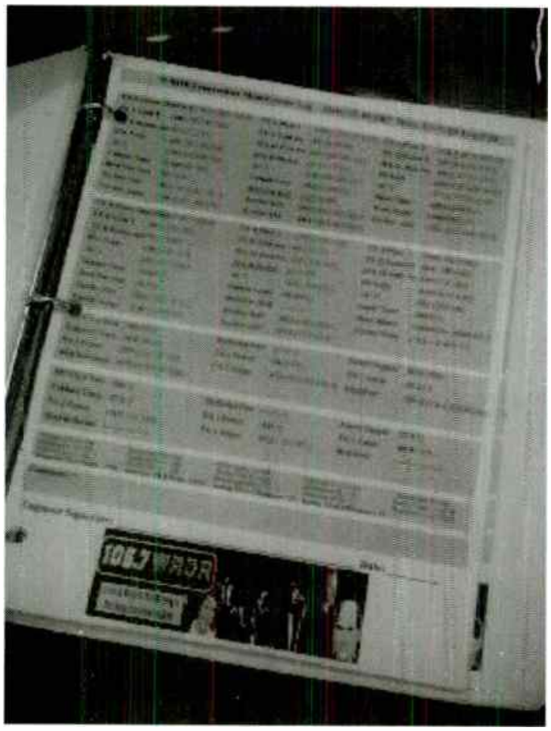


Figure 2. Completed maintenance log in loose leaf notebook at transmitter facility

WEB ACCESS

Even when you are not at the transmitter site, if you write these logs to a web page, it gives you the flexibility to access maintenance data from anywhere. This allows an engineer to more thoroughly analyze the historical data in a cleaner, quieter, and often more relaxed atmosphere (like the office or home) where he or she can spend more time looking at trends and understanding the changes that might be occurring at the transmitter site.

WMJX 106.7 Mhz Transmitter Log

WMJX Transmitter Maintenance Log from the P

	Date	Time	DOM	DAY	AIRLOAD	Pwr	Plt I	Plt KV
1	06-04-2007	17:05:22	31.00	THU	V5	100	1.61	9.20
2	06-05-2007	17:04:52	31.00	THU	V5	100	1.61	9.20
3	06-05-2007	17:03:41	31.00	THU	V5	100	1.61	9.20
4	06-05-2007	17:01:00	31.00	THU	V5	100	1.61	9.20
5	06-07-2007	22:16:18	7.00	THU	V5	100	1.65	9.10
6	06-07-2007	22:26:50	7.00	THU	V5	100	1.65	9.10
7	06-07-2007	22:48:04	7.00	THU	V5	100	2.11	9.22
8	06-07-2007	23:29:08	7.00	THU	ON AIR	100	2.11	9.22
9	06-07-2007	23:46:39	7.00	THU	ON AIR	100	2.11	9.22
10	06-08-2007	00:27:30	7.00	THU	ON AIR	100	2.11	9.22
11	06-08-2007	00:32:14	7.00	THU	ON AIR	100	1.65	9.10
12	06-08-2007	00:47:59	7.00	THU	ON AIR	100	2.11	9.22
13	06-08-2007	23:10:29	8.00	FRI	ON AIR	100	2.30	9.22
14	06-13-2007	12:43:31	13.00	WED	ON AIR	100	1.65	9.20

Figure 3. Web Page with Log

In figure three, the web page shown uses a java applet to enable the user to sort the individual fields by value. For example, it is possible to show what the entire log looked like on the day when the room temperature was highest, or when the transmitter plate voltage was lowest. Having these tools to analyze the entire plant can sometimes be a helpful diagnostic tool.

CLEAN POWER

It is now becoming more important than ever to supply good clean and un-interruptible AC Power to the equipment at your transmitter site. This is especially true since it is not un-common for transmitter sites to be located in remote areas where utility power is not the most reliable, and physical access to the site can be restricted by weather conditions or other environmental factors. Since almost all the gear used these days is CPU based, a power glitch will end up resetting most equipment used at a transmitter. Most equipment does not reset instantly, and will normally take a minute or two to come back to life. This always seems like an eternity when you are off the air and listeners are dropping like flies! Having a clean source of power available to your mission critical equipment is essential these days. Often it is more economical to provide a single yet redundant UPS System that will be able to provide clean power to all your racks. If your plant has a reliable generator, the battery backup run time on the UPS only needs to be a few minutes in order to allow the generator to come up to full power. Many UPS Units have settings that may need to be fine tuned in order for the UPS to see the "less perfect" generator power as a truly good source so the UPS will not run off batteries during the entire time that the generator is running.

Of course the UPS needs to be maintained from time to time, and during those times the UPS may need to be isolated from the load. In order to do this without going off the air, a properly installed make before break bypass switch needs to be installed to allow the UPS to be bypassed for maintenance or due to a failure. In many cases this type of switch can be purchased through the UPS manufacturer.

In general it is advisable to go to a centralized UPS system if possible. This eliminates the need to monitor and maintain a large number of smaller UPS Units.

TOWER LIGHT MONITORING

The FCC and FAA have recently modernized the rules regarding tower sites. Make sure that the base of the tower and guy wires are secure, and that the tower registration number is displayed prominently not only at the tower base, but at a point that can be easily seen from the street (or a public point) leading to the tower.

Tower light monitoring remains one of the more important rules that the FCC is enforcing. A system must be in place to allow your operators to immediately detect a full or partial failure of the tower lighting system. If your tower lighting system is not operating as specified in the station's authorization, notification to the FAA (through a special toll free number) is required within 30 minutes. Notification is again required when the problem has been resolved and the lights are operating in accordance with the station's license. If you are a tenant on a tower that you do not own, it is generally a good idea to make sure you still have access to tower light telemetry and coordinate with the tower owner or other tenants to make sure there is a plan for notifying the FAA of lighting problems.

CONCLUSIONS

As the broadcasting industry changes both the type of engineer and the type of equipment at the transmitter sites will change. The successful and efficient operator will need to use a combination of traditional good sense and modern tools like those described above to keep ahead of any issues that can be predicted or prevented.

Acknowledgements

The author wishes to thank his colleagues at Greater Media, Eimac, and ERI Corporation for their assistance.

Implications of IBOC Injection Levels Above -20 dB

GARY LIEBISCH
Nautel
Bangor, Maine

ABSTRACT

Over the past year, there has been much discussion among broadcasters about the prospect of raising the permissible digital carrier levels of HD Radio™ hybrid FM signals. The current established level in Standard NRSC-5A (September, 2005) is 20 dB (1%) below the reference analog carrier power. But with the number of stations deploying HD2 or HD3 multicast channels increasing, the integrity of program streams offered as SPS (Supplementary Program Services) has taken on new significance. This paper does not attempt to debate the interference implications of modifying the standard. Rather it addresses broadcaster curiosity about what a changed standard implies in terms of implementation methods at their respective stations. For stations still planning conversion for the future, this paper provides some initial guidance on implementation that will help smooth the transition to higher digital power, should it become a reality.

DIGITAL COVERAGE

Most FM stations who have deployed HD Radio have found that reliable digital coverage can be expected out to at least the 70 dBu (3.16 mV) city grade contour. In relatively flat, unobstructed terrain, consistent coverage out to 60 dBu and further is possible. But given the amount of real world obstructions a broadcast signal must overcome on its way to the receiver, coupled with the natural disinclination of the average person to use an external antenna, it may turn out that more "work" needs to be done on the signal before it leaves the transmitter.

Of course, the impact of elevated digital carriers must be exhaustively studied in the field to determine its effect on adjacent channels. To that end, several broadcasters have undertaken STA operations to study the effects, and inevitably the NRSC will need to get involved. It could easily be years before we see a change in the standard level, if any, and there is no way to predict what that final level will be.

It is not the purpose of this paper to debate the viability or adjacent channel interference impact of -10 dB carriers. What I would like to address, though, is the natural curiosity that broadcasters have about protecting their existing HD Radio investment in the -20 dB standard, if the permissible level were to be raised. Vendors of HD Radio transmission equipment are now being queried regularly about avoiding obsolescence if they invest in the technology today. Since there are multiple implementation methods available to FM stations, is one upgrade path better than another with respect to reuse? And for stations not yet converted, is there an upgrade path that will not obsolete what is purchased today?

DIGITAL AMPLIFICATION

To begin to understand the impact on implementation method of raising digital levels, we must first understand a few things about the nature of digital signals, how they pass through a transmitter, and how we calculate and rate transmitter power output.

The first thing to understand is that IBOC signals, unlike FM, are varying amplitude signals. Conventional FM signals are, of course, constant amplitude signals, where only the varying frequency or phase conveys information. HD Radio transmission employs Orthogonal Frequency Division Multiplexing (OFDM) modulation, where symbols, representing data, are decoded from their instantaneous (and rapidly changing) phase and amplitude. As many as 512 different vectors may represent a frame of data.

Amplifiers used for FM transmission are typically biased into Class C mode, which provide high efficiency, but which distort (discard) amplitude variations because their conduction angle is so short. When a digital signal must pass through the same amplifier, amplitude excursions are intentionally created, and must be preserved. To do this, the amplifier is "re-biased" to conduct for a greater time during each cycle, at the expense of efficiency, but with higher linearity. While we could go to a very linear Class A, the efficiency would be unacceptably low. So we typically bias instead at class AB for greater efficiency, at the expense of some linearity and resulting intermodulation products. This re-biasing is one of the reasons why all FM transmitters are "de-rated" when operated in HD modes. Thus, power output rating, therefore, becomes a balancing act between efficiency, linearity, and mask compliance.

We now know that a typical 10 kW FM solid-state transmitter typically is derated to 8 kW when operated in hybrid mode, and 3 to 3.5 kW when operated in

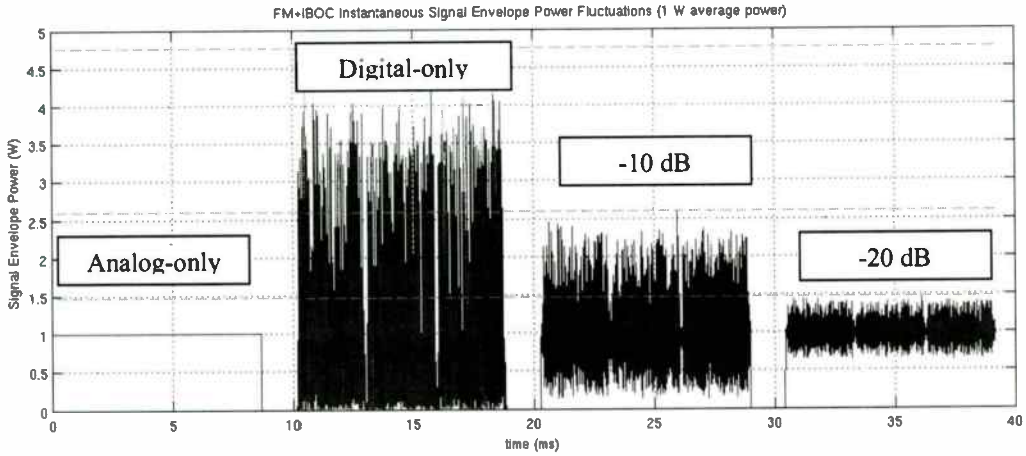


Figure 1: Analog, digital only, -10 dB, and -20 dB envelope signals

digital-only mode. Aside from the re-biasing for linearization that takes place in the amplifier, it must also be recognized that the digital signal now has a peak and an average component. While we are accustomed to thinking only of the average digital power when sizing an FM transmitter, the peak to average (PAR) ratio plays an important role in determining what the maximum power output capability of a transmitter can be. Typical PAR for a digital-only HD Radio signal would theoretically be 7 to 8 dB. But in actual practice, about 2.5 dB of OFDM-compression is tolerated, yielding an effective PAR of about 5 to 6 dB. This is not to be confused with amplifier compression point, discussed below. To accommodate this high PAR, an operating point is selected that is typically 5 dB below (35%) its FM nameplate rating. This is the "backoff" for digital-only mode. Similarly, for common amplification, where the digital component is only a portion of the total signal, the backoff requirement is about 1 to 1.2 dB, resulting in a de-rated TPO that is 75-80% of nameplate rating. Figure 1 is a side-by-side plot of four baseband envelope signals. The signals have been normalized to each produce one watt of *average* power. The analog FM-only baseband (far left) is compared to baseband signals for digital only, -10 dB, and -20db signals. For the purpose of this illustration, the additional 1% average power for the -20 dB signal, and the additional 10% average power for the -10 dB signal are not accounted for, --we are simply trying to illustrate how the peak-to-average changes with a constant average power. Note the increased dynamic range in each step that must be accounted for in amplifier design and analog backoff. To take the phase of the digital carriers into account, Figure 2 is a vector representation of a hybrid (FM+HD) signal operating at a 20 dB ratio, alongside a vector representation of an analog-only FM signal. The accompanying scale shows a new normalized analog component power that is 0.75 times the analog-only power. But note that at purely random times (because analog and digital carriers are not correlated), the digital and FM vectors sum together, and the amplifier nudges saturation. Saturation is defined as the point where instantaneous increases in input amplitude result in no further increase in output, and clipping occurs. Just below saturation, however, the transmitter goes into a compression knee and relies on pre-correction to "linearize" its transfer function. Because the non-linear transfer function is predictable, precorrection, and especially adaptive precorrection, is a valuable tool in extending the linear dynamic range of the amplifier. The result of this linearization process is a slightly less conservative backoff.

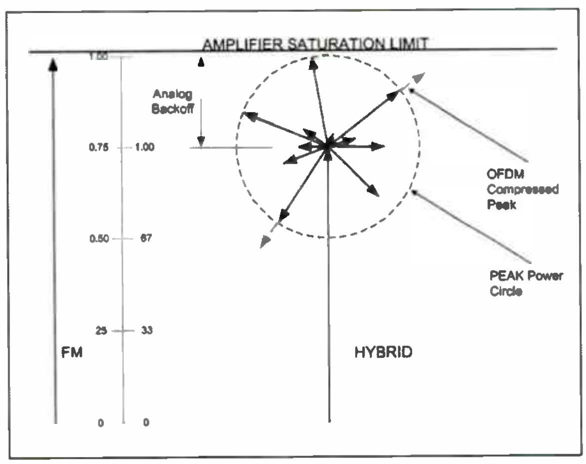


Figure 2: Vector Representation of Hybrid Amplification

HYBRID TRANSMITTERS AT -10 DB

Armed with that knowledge of amplifier behavior, it now becomes apparent why maintaining the required peak-to-average ratio while avoiding excessive amplifier compression forces a further backoff of FM power output capability of the transmitter. Moreover, if we are to maintain the same noise levels at >200 kHz relative to the analog carrier that we have for adjacent channel protection in the current standard, then we are asking the amplifier for 10 dB BETTER mask noise performance. And the price for that performance is an even further backoff of the total power output. How much? Very preliminary test results indicate a backoff of 4 to 5 dB may be necessary. To put that in perspective, your 10 kW solid state transmitter that used to do 10 kW of pure FM was first derated to 8 kW in (-20 dB) hybrid mode, and would be further derated to just 3 to 4 kW of analog power (plus 10% of average power digital) in -10 dB hybrid mode. And your tube rig that started out as a 35 kW FM analog transmitter, will likely top out in the 11 to 13 kW range. In the best case, to maintain hybrid operation, you would be at least *doubling* the number of amplifiers, as a minimum.

For non-hybrid systems (high level, space combined, etc.) that employ "digital only" mode transmitters, there will be no change in efficiency or output rating of the digital transmitter, but you still face the challenge of getting 10 dB more RF to the antenna, either through a different combining scheme, or larger/more amplification. Transmitter output will be mainly mask-limited, and not thermally limited.

With the prospect of declining efficiency, higher operating cost, and expanded footprint to accommodate additional hybrid amplifiers, it becomes readily apparent that in many cases, it is worth taking another look at all the combining methods on the table. Methods dismissed in the initial implementation might be more cost effective if the standard is modified.

SPACE COMBINED SOLUTIONS

Space combining has always enjoyed the distinction of having the lowest operating cost of all the combining methods. Stations that chose this method for their initial implementation will have to purchase a larger digital transmitter, but will continue to be able to operate most efficiently. Space combined digital transmitters that are currently in the 350 watt (average) or less range (1% of a 35 kW FM transmitter), will need a new box the size of a traditional 10 kW FM transmitter, which will yield up to 3.5 kW of digital-only power. The power handling capability of the antenna and transmission line will need to be reviewed as well, but fortunately many stations have these components already rated in the kilowatts range, since they are licensed as auxiliary (FM) antennas.

Stations that have implemented using high level 10 dB combining can also consider a conversion to space combining, because that extra 10 dB of RF is already being produced—and sent to a reject load. While it may not be necessary to purchase a new transmitter, they will require the vertical real estate for the second antenna and transmission line, as well as a ferrite isolator in many cases. This, of course, will be easier where stations own the tower and have the space available, more difficult where the space is leased or limited. In those cases, replacement with an interleaved or dual-input antenna is another solution. When you remove a 10 dB coupler, however, and redirect all its power to a second antenna, keep in mind that the out-of-band noise floor will come up by 10 dB also. So make sure you have the mask margin to do it before considering this solution. In some cases, a mask filter may be required.

Space combining requires adequate isolation between antennas to avoid intermodulation products. Typical minimum value would be 40 dB. This can be achieved by adequate antenna spacing, keeping in mind the FCC requirement that the "aux" antenna used for digital must be at least 70% of the HAAT of the main antenna. The taller the tower is, the easier it is to meet the FCC requirement and still have adequately spaced isolation. Half-wave spaced antennas offer better suppression of vertical radiation, and so improve on the isolation figure. Where space isolation is not adequate, and you many not know until you actually install the system, the use of ferrite isolators can improve performance. These are more common in dual feed or interleaved antennas where the self isolation tends to be significantly less than widely spaced (vertically) antennas. Where isolators are used, their power handling capability must include the SUM of the forward and coupled (reverse) power. If the digital power is raised by 10 dB, it imposes a higher spec on the power handling capability of the digital transmitter's isolator, if used. In the reciprocal direction, digital to analog, it is quite possible that an isolator will be required on the *analog* transmitter to suppress the (higher) digital power now appearing at the analog transmitter's output. At Greater Media's WCSX in Detroit, operating under an STA to test the -10 dB concept, a rather large isolator was added to the analog transmitter feedline. See Fig 3.

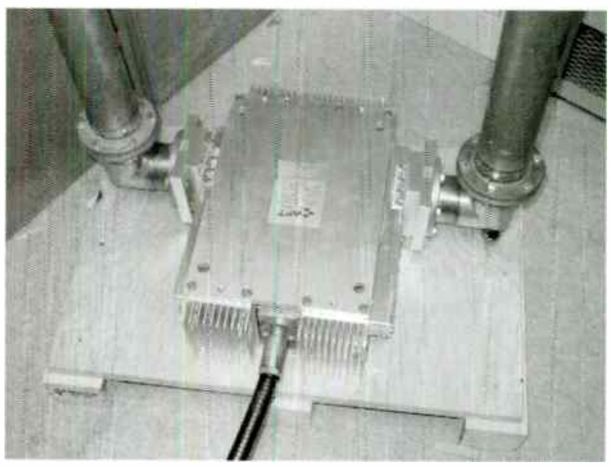


Figure 3: 20 kW ferrite isolator at WCSX

The station uses an ERI dual input antenna, supplemented by AFT isolators on both the digital and analog transmitters. It is worth an inquiry to the manufacturer(s) of both the analog and digital transmitters as to what the turnaround loss spec is for that model. A higher number (20 dB is considered good) decreases the likelihood of undesirable intermodulation products. Solid state transmitters generally have better turnaround loss than tube transmitters.

As has been noted in many articles on implementation, space combining is not without its own caveats. Close attention must be given to producing, as near as possible, identical radiation patterns. When geometrically different antennas are used for digital and analog, particularly with different bay configurations, the vertical null patterns produced can destroy analog or digital coverage close-in to the antenna. This effect can be minimized somewhat with identical gain antennas, and even more so with interleaved and dual input antennas.

VARYING THE COUPLING RATIO

What about high-level combining? Is it still a viable solution with more digital power in the equation? As the desired output ratio of digital to analog increases, it becomes immediately apparent that the 10 dB coupler-injector is no longer an effective way to combine the signals. A 10 kW FM station, for example, would have to produce 10 kW of digital power, just to get 1 kW out at the antenna! An analysis of different coupling ratios concludes that 5 to 6 dB of coupling yields the best AC to RF efficiency for the entire system when an injection ratio of -10 dB is the goal. See Table 1.

Coupling Ratio	10 dB	9 dB	8 dB	7 dB	6 dB	5 dB	4 dB
Analog Tx	11.1 kW	11.4 kW	11.8 kW	12.4 kW	13.1 kW	14.1 kW	15.6 kW
Digital Tx	10. kW	8.0 kW	6.3 kW	5.0 kW	4.0 kW	3.1 kW	2.5 kW
Reject Power	10.1 kW	8.4 kW	7.1 kW	6.4 kW	6.1 kW	6.2 kW	7.1 kW

Table 1

So far, we have considered only the implications of an increase of 10 dB in digital carrier levels. But it is likely that if authorized, intermediate increases in digital power would also be allowed. Staying put at -20 dB should also be an option. So if a station could not cost effectively implement a 10 dB increase, they may instead elect to implement whatever increase may be most cost effective for their configuration, UP TO 10 dB. Figure 4 shows optimum coupling ratios for other injection ratios between -10 db and -20 dB. For a high level combined station, by replacing a 10 dB injector-coupler with a 6 dB version, for example, 25% of the digital transmitter's power now goes to the antenna, compared to 10% before, an increase of 4 dB without replacing the digital transmitter. Of course, it also imposes a 25% headroom requirement on the analog side! A cost effective approach might be to look first at the available headroom in the analog transmitter and then select a new coupling ratio based on the most efficient configuration to arrive at some intermediate boost in digital signal power. Using this approach, no new transmitter is purchased, as the new interim digital

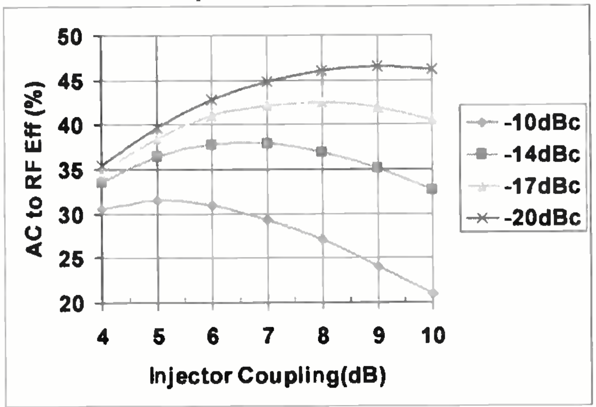


Figure 4: Optimum Coupling Ratios

power is arrived at through replacement of only the injector, and possibly the reject load. In Table 2, we show the metrics for this analysis. All powers are shown as a % of the licensed analog TPO. The irony of this solution is that we actually get more digital power to the antenna by increasing the *analog transmitter* power!

Injectors for ratios other than 10 dB are already quite common. Shively Labs, Inc. manufactures injectors in 1 dB increments from 6 to 10 dB and can do custom injection levels as well. Dielectric's Dibrid™ coupler also allows for varying coupling ratios. 9 dB injectors are sometimes used even today where headroom exists in the analog transmitter but the digital transmitter is running "wide open". Some might be led to believe that substituting injectors of less than 10 dB compromises the isolation between the two transmitters. But this is not the case. Isolation is purely a function of having matched terminations at the antenna and reject ports. 10 dB was just a convenient and efficient coupling ratio selected to keep the analog headroom requirement to a modest (10%) level. But with any change in injector ratio comes an increase in the overhead requirement of the analog transmitter, and all of that overhead goes to the reject load as wasted heat. At some point, high level combining becomes less attractive.

ATTENTION TO SPECIFICATIONS

Having looked at space combined and alternate injector solutions, some stations may still be left with only the common amp hybrid transmitter alternative. And as noted earlier, the backoff numbers will be changing if the new ratio is implemented. This will require extra

Coupling Ratio	Lic. TPO	P_{Analog}	$P_{Digital}$	New Digital Power	Boost (dB)	P_{Reject}
10	100.00	111.00	10.00	1.00	0	20.00
9	100.00	114.17	10.00	1.26	1	22.92
8	100.00	118.36	10.00	1.58	2	26.78
7	100.00	123.93	10.00	2.00	3	31.94
6	100.00	131.43	10.00	2.51	4	38.92
5	100.00	141.62	10.00	3.16	5	48.46
4	100.00	155.66	10.00	3.98	6	61.68

Table 2: Utilizing Existing Headroom in Analog Tx for more digital power
All power shown as % of licensed TPO

attention is paid to how power output specs are stated and compared. To put power output comparisons in the proper context and compare apples to apples, it should always be stated under a specific VSWR condition, such as 1.2:1 or 1.5:1. Also, power output without regard to mask performance margin can be meaningless. Under digital operating modes, a transmitter can go out of mask spec long before its components are thermally stressed, and may happily operate in that condition indefinitely. So the fact that the transmitter can get to 110% or 120% of its rated power may not be of much value. Ambient operating temperature and altitude should also be stated and considered as part of the power output rating.

CONCLUSIONS

It is still too early to predict the exact number for further derating of hybrid transmitters, but a doubling of the number of amplifiers required to achieve a given TPO is a rough estimation. Clearly, the scales of implementation method will tilt more toward less lossy solutions, such as space combining, interleaved antennas, and dual input antennas. Each facility will need to approach any upgrade with an open book, considering and reconsidering all the implementation methods that may have been rejected before. Intermediate steps may be the most cost effective approach for many stations. Large groups may inevitably swap transmitters around to get the most bang for the buck. And above all. the mask must rule in any implementation, if we are to maintain required protection of adjacent channel neighbors.

ACKNOWLEDGEMENTS

Philip Schmid, Nautel Ltd., Halifax, N.S.
Bob Surette, Shively Labs, Bridgeton, ME
Mike Kernan, Greater Media, Detroit, MI
Matt Leland, Dielectric Communications, Raymond, ME

FM IBOC BUILDING PENETRATION TESTS AT ELEVATED DIGITAL SUBCARRIER LEVELS

Glynn Walden, Senior VP Radio Engineering
CBS Radio, New York, NY

OVERVIEW

This report documents the performance of indoor reception of conventional FM analog signals and the digital component of iBiquity's hybrid IBOC FM system. Primarily, these tests document improvements in indoor reception with the IBOC radio system operating with the digital subcarrier power level increased to -10 dBc (overall) vs. the authorized -20 dBc power level.

THE IBOC SIGNAL

The authorized FM IBOC hybrid transmission mode consists of IBOC digital subcarriers inserted on both sides of the host analog FM signal. On each sideband there are 191 subcarriers, each at a power level of -41.4 dBc, referenced to a 1 kHz bandwidth. The total power of the 382 digital subcarriers is 20 dBc below the total power in the analog host. Building penetration tests were conducted in the Los Angeles metropolitan area on KROQ FM with the digital signal operating at the authorized power of -41.4 dBc per subcarrier level (total digital power -20 dBc) and with the digital power elevated to –31.4 dBc per subcarrier (total digital power -10 dBc).

KROQ-FM TRANSMISSION FACILITY

The tests of elevated digital IBOC power tests were conducted using KROQ-FM (FCC ID 28622), licensed to CBS Radio, Pasadena California. KROQ-FM, Class B FM commercial radio station has been continuously operating as an IBOC station on 106.7 MHz since 09/27/02. The transmitter is located at 34-11-49.0 N latitude 118-15-30.0 W longitude. KROQ broadcasts with 5.6 kW Effective Radiated Power ("ERP") at a Height Above Average Terrain ("HAAT") of 423 m. The KROQ transmitter operates at 4.3 kW of Transmitter Power Output ("TPO") to produce its licensed ERP.

The IBOC -20 dBc and -10 dBc IBOC signals were generated using a Harris FlexStar exciter amplified by a Harris Z-16 Plus transmitter employing common amplification. The transmitter was connected to a 3-bay antenna resulting in 5.6 kW of analog ERP and a digital ERP of either 56 W at -20 dBc, or 560 W at -10 dBc. Figure 1 depicts a spectral representation of the FM hybrid mode operating with total digital power of -20 dBc and Figure 2 shows the same spectrum with the total digital power elevated to -10 dBc. The rectangular areas contain the digital subcarriers, and the triangular area represents the analog host FM signal.

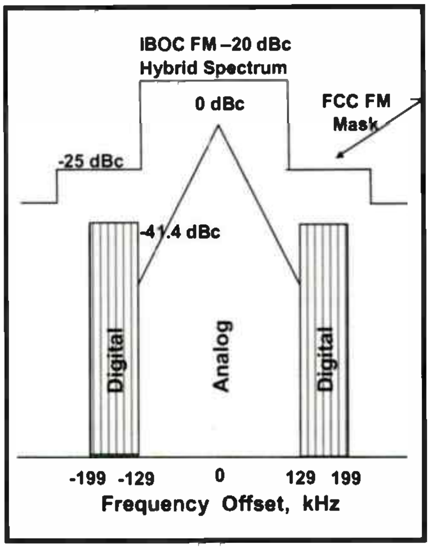


Figure 1 - FM Hybrid -20 dBc Spectrum

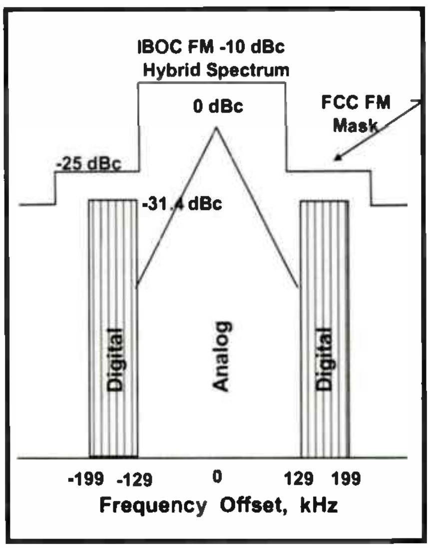


Figure 2 - FM Hybrid-10 dBc Spectrum

Table 1 – Buildings Used For Testing

SITE NUMBER	DESCRIPTION	ADDRESS	HEIGHT, STORIES	FLOOR MEASURED
00	KROQ Main Transmitter Site	1250 Beaudry Blvd., Glendale	N/A	N/A
01	KROQ Studio/Office Conference Room	5901 Venice Blvd., Los Angeles	2	1
02	Chubb Insurance Building	801 South Figueroa St., Los Angeles	31	17
03	8-Unit apartment building	4127 East 1st St., Long Beach	2	2
04	5670 Wilshire high-rise office building	5670 Wilshire Blvd., Los Angeles	30	2
05	"SBS" medium-rise office building	10281 W. Pico Blvd., Los Angeles	4	1
06	"Hollywood/Highland" entertainment complex	6801 Hollywood Blvd., Los Angeles	3	3
07	single-family residence	1313 35th St. San Pedro	1	1
08	32-unit apartment complex	4616 Cahuenga Blvd., Toluca Lake	2	2
09	Underground Parking "Studio City Place"	11239 Ventura Blvd., Studio City	2	2 below ground
10	Underground Parking Ralphs Shopping Center	211 N. Glendale Ave., Glendale	2	2 below ground

STATION CONFIGURATION

KROQ-FM's transmission setup is shown in Figure 3. As part of these tests a new Harris FlexStar Analog / Digital signal generator was installed and programmed to generate a combined analog and digital signal. In addition KROQ-FM's Harris Z-16 HD transmitter was upgraded to a Z16-HD-Plus w/ options transmitter. The upgraded transmitter provides for a higher degree of linearity and enables the transmitter to operate with 10 dB higher digital power level while maintaining the Out-of-Band Emissions ("OBE") within the iBiquity-suggested mask for operation at the -20 dBc level.

OPERATING POWER

Pursuant to KROQ-FM's license and Experimental Authorization, granted by a letter from the FCC dated Feb. 16, 2007, KROQ-FM was operated at its licensed analog power with total digital ERP of - 20 dBc, 56 watts, or -10 dBc, 560 watts.

BUILDING SELECTION

The CBS Engineering team identified 10 buildings differing in construction type and usage. The buildings ranged from a residential dwelling to an office building in downtown Los Angeles. The sites are further described in Table 1 and the locations noted on the map shown in Figure 4.

TEST PROCEDURE

CBS Radio conducted tests to determine the ability of the analog and digital signals to penetrate buildings varying in construction type, use and size. The testing procedures are intended to document the improvements in building penetration resulting from increasing the power of the IBOC radio systems digital carriers from -20 dBc to 10 dBc and provide some information as to the attenuation losses found in these buildings.

As depicted in Figure 5, the tests were conducted with a common FM whip antenna feeding a -3dB splitter connected to the Anritsu spectrum analyzer and the Boston Acoustics Recepter HD receiver.

At each location the following measurements were made:

1. At an outside location the antenna was placed in a position where a visual observation indicated the signal was being received with minimal multipath[1]. Spectrum analyzer data and signal level were recorded to establish the unattended RF signal level being received from KROQ-FM.
2. The antenna, receiver and spectrum analyzer were moved into an indoor location near a window where the engineer attempted to receive the digital broadcast while transmitting with the IBOC carriers operating at -20 dBc.[2] Spectrum analyzer data and signal level were recorded to establish the indoor signal level being received from KROQ-FM.
3. The antenna was moved to a more interior point within the building in an attempt to find a point where the digital signal failed, Point of Failure (POF).[3] The distance in meters from the

[1] The antenna was moved around until a location could be found where the IBOC digital sidebands were symmetrical in level.
[2] In many locations it was not possible to receive the digital signal at the -20 dBc transmission level.
[3] -20 dBc POF location.

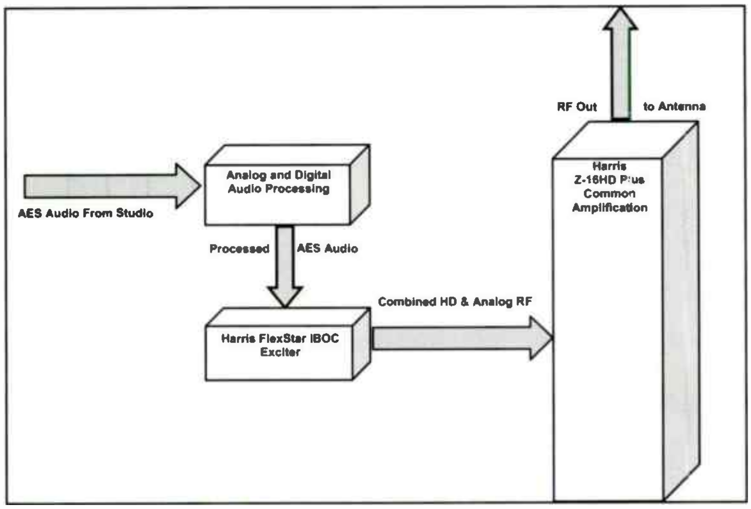


Figure 3 - Diagram of KROQ-FM IBOC Common Amplification FM Transmitter Setup

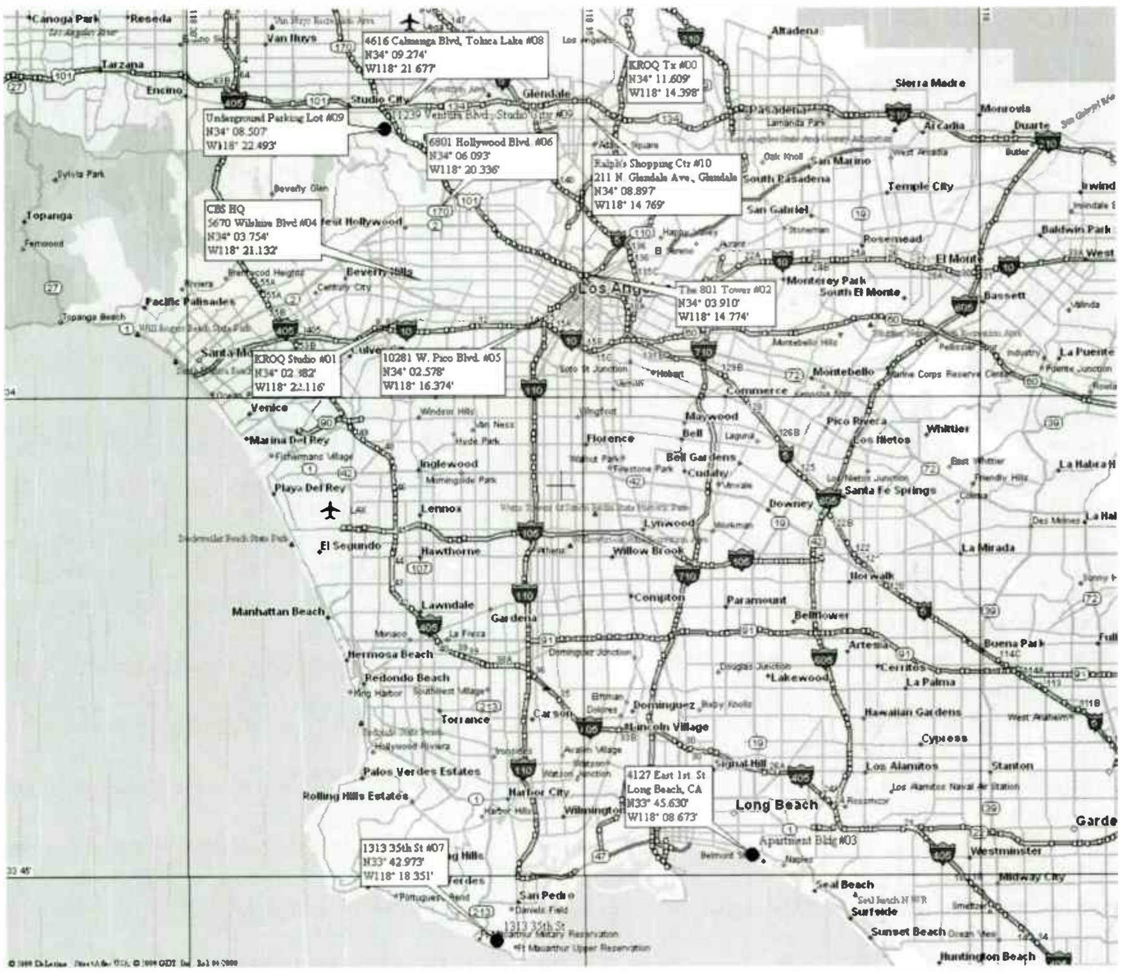


Figure 4 Map Showing Test Locations

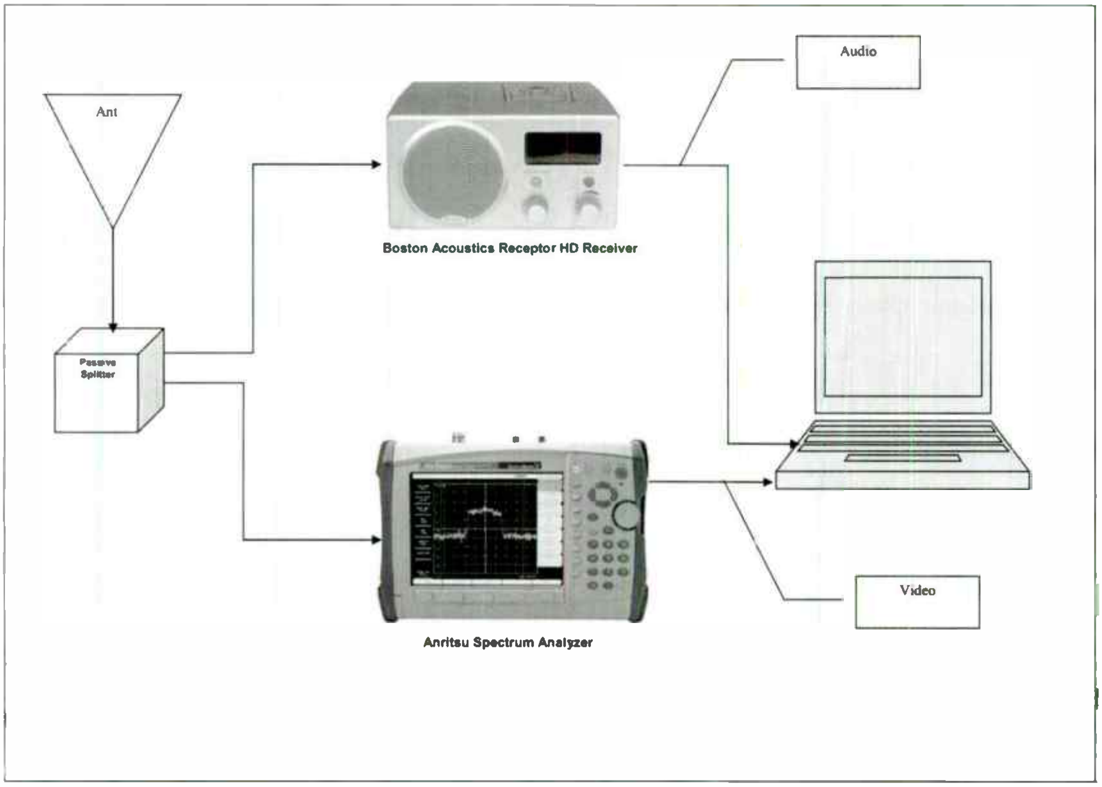


Figure 5 - Diagram of Field Measurements and Recording Setup

Table 2 – Test Results Summary

Note that measurements made outside location 04 (CBS Radio Hdq) were made at ground level therefore the inside signal levels on the 3rd floor are higher, likewise the measurement made outside location 08 were higher than inside. The negative attenuation numbers from these two buildings are not included in the attenuation calculations.

IBOC Building Penetration Test Results, KROQ, Los Angeles, CA															
#	Description	Dwelling Type	Longitude (DMS)	Latitude (DMS)	Distance From Tx (km)	Predicted dBµ	Rx Ht. m	dBµ Out	dBµ In	Bld. Atten. dB	dBµ @ -20 dBc POF	dBµ @ -10 dBc POF	Analog Reception	IBOC @ -20 dBc	IBOC @ -10 dBc
00	KROQ Transmitter	Concrete Block	118-14-23.9 W	34-11-36.5 N	KROQ Transmitter Site -- Calibration Only										
01	KROQ Studios	Concrete Block	118-22-07.0 W	34-02-22.9 N	20.19	63	2	65.7	47.3	18.4			Noisy	None	Solid
02	801 Tower	Concrete & Steel	118-14-46.4 W	34-03-54.6 N	14.66	67	17	70.7	58.9	11.8			Good w/ Multipat	Intermittent	Solid
03	Long Beach Apts.	Wood Frame Stucco	118-08-40.4 W	34-45-37.8 N	45.59	56	2	46.3	43.1	3.2		41.6	Poor	None	Intermittent
04	CBS Radio Hdq..	Concrete & Steel	118-21-07.9 W	34-03-45.2 N	17.24	44	2	62.0	66.7	-4.7		44.0	None	None	Solid
05	SBS Office Building	Concrete & Steel	118-16-22.4 W	34-02-34.7 N	17.13	64	2	61.8	33.1	28.7			None	None	Solid
06	Hollywood Highland	Concrete & Steel	118-20-20.2 W	34-06-05.6 N	12.93	53	3	63.8	51.0	12.8			Poor	None	Solid
07	San Pedro Residence	Wood Frame Stucco	118-18-21.1 W	34-42-58.4 N	53.5	34	2	44.4	41.6	2.8		41.6	Poor	None	Intermittent
08	Toluca Lake Apts.	Wood Frame Stucco	118-21-40.6 W	34-09-16.4 N	10.59	69	2	68.1	73.6	-5.5			Good	Solid	Solid
09	Studio City Pk. Garage	Concrete & Steel	118-22-29.6 W	34-08-30.4 N	12.37	68	2	82.6	55.7	26.9		39.3	Point of Failure Tests Only		
10	Ralph's Pk. Garage	Concrete & Steel	118-14-46.1 W	34-08-52.7 N	5.55	75	2	68.2	49.4	18.9	49.4	41.2	Point of Failure Tests Only		

Table 3 – Test Results Summary

Structure	Attenuation (dB)
Concrete Block	18.4
Concrete & Steel	19.8
Wood Frame	3.0
All Structures (Avg.)	17.5

"window" location was recorded along with the analog audio and the spectrum analyzer measurements.

4. The digital power at the KROQ transmitter was increased and the antenna moved further into the interior of the building to a point where the digital signal failed, POF, -10 dBc. As above the distance in meters form the "window" location was recorded along with the analog audio and the spectrum analyzer measurements.
5. The building attenuation was calculated by subtracting the inside measurement from the outside measurement.[4]

RESULTS

The results of the tests are summarized in Table 2,. In 75% of the buildings tested the analog reception was described as non-existent, noisy or poor and there was no digital reception when the IBOC power levels were operating at the -20 dBc level.

With the digital subcarriers elevated to -10 dBc, the digital radio signal could be received reliably in 75% of the building and at selected locations in the remaining 25% of the buildings. The goals of the test plan were not entirely met as the ability to receive the digital part of the broadcast when operating at an IBOC power level of -20 dBc was only possible in one building and when operating at -10 dBc the signal was available uniformly throughout the structure except in two of the buildings. While the tests did not yield the desired quantitative numbers the qualitative information shows that increasing the digital power by 10 dB results in very robust indoor reception. It is important to note that the improvements in digital reception did not change the quality or reception of the host analog signal.

Testing was conducted in two underground parking garages to determine the POF when the receiver is located in very obstructed and attenuated spaces. In this test, CBS Radio engineers drove a test van outfitted with the Boston Acoustics Receptor HD receiver into underground parking garages to demonstrate how much deeper into a building that IBOC signals could be received when the power was increased to the -10 dBc digital power level. In building 9, for -20 dBc, the digital signal failed at the bottom of the ramp even though the analog reception was good. With the power increased to -10 dBc digital reception was available at most of the locations in the building while analog reception in these same locations was reported to be noisy. In building 10, the digital reception with the IBOC carriers at –20 dBc was available to about 30 feet into the building. With the IBOC carriers elevated to -10 dBc the digital signal was received to 165 feet within the building. At the point of digital failure, the analog signal had deteriorated to the point that would not be acceptable to most listeners. The results are summarized in Tables 2 and 3 ("Test Results") with a detailed description of the results following in Appendixes 00 -10.

[4] The absolute field intensity read off the spectrum analyzer in dBm was converted to dBu by adding 116dB. 3dB was added to the resultant number to account for the loss in the attenuation of the splitter.

CONCLUSIONS

When operating at a digital power of -20 dBc, the digital portion of hybrid FM IBOC signals are difficult to receive inside most structures except when they are located close to the transmission site. This report shows that by elevating the IBOC digital subcarrier levels to -10 dBc that the hybrid IBOC system can provide equivalent indoor digital reception with better quality than analog while providing additional digital channels and services. The power increase will help to promote the widespread adoption of digital radio.

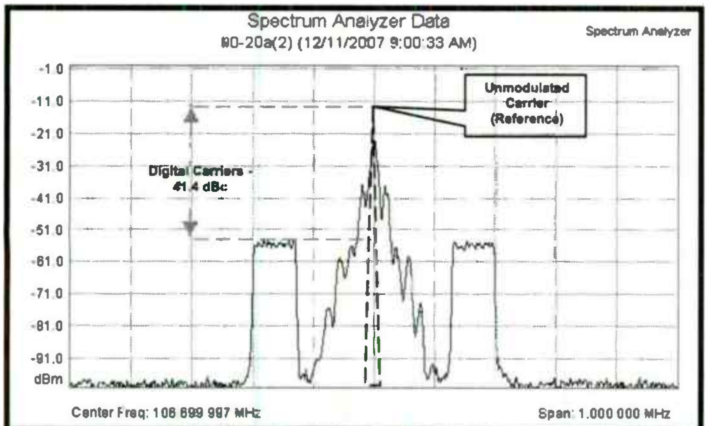
APPENDIXES

Site	Location
00	KROQ Transmitter Site / 1250 Beaudry Blvd, Glendale, CA
01	KROQ Studios / 5901 Venice Blvd., Los Angeles, CA
02	The 801 Tower / 801 South Figueroa St., Los Angeles, CA
03	8 Unit Apt. Building / 4127 East 1st St., Long Beach, CA
04	CBS Radio / 5670 Wilshire Blvd, Los Angeles, CA
05	SBS Office Building /10281 W. Pico Blvd, Los Angeles, CA
06	Hollywood Highland Entertainment Complex / 6801 Hollywood Blvd, Los Angeles, CA
07	Single Residence / 1313 35th St, San Pedro, CA
08	36 Unit Apt. Complex / 4646 Cahuenga Blvd, Toluca Lake
09	Parking Garage / 11239 Ventura Blvd., Studio City, CA
10	Parking Garage / 211 N. Glendale Ave., Glendale, CA
	Table of Audio Cuts

SITE 00 - KROQ TRANSMITTER SITE / 1250 BEAUDRY BLVD, GLENDALE, CA

(N34° 11.609' / W118° 14.398')

The spectrum analyzer traces identified as 00-20 and 00-10 were captured with the analyzer fed from a transmitter RF sample port through a 30 dB attenuator, solely for the purpose of checking the analyzer and establishing benchmark measurements. This was done at the KROQ main transmitter site in the Verdugo Mountains. We adhered to this trace numbering convention as closely as possible throughout this report.

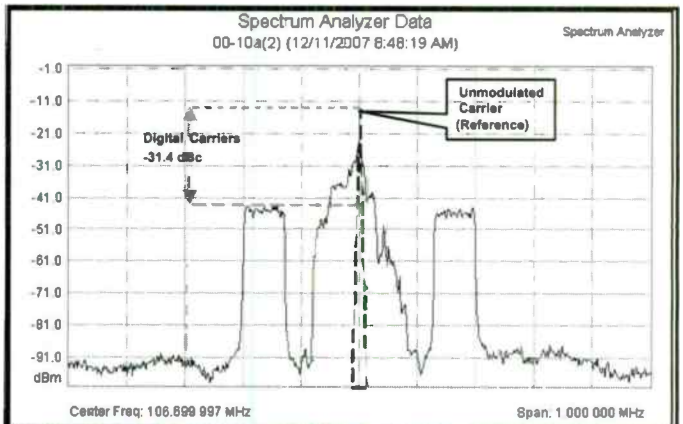


Channel Power Width: 400.000 kHz — Channel Power: -13.97 dBm

Span: 1.000 MHz — Channel Power Density: -69.99 dBm/Hz

Measurement Parameters

Trace Mode	Normal	Frequency Span	1.000 MHz
Reference Level Offset	0.0 dB	Reference Level	0.000 dBm
Input Attenuation	20.0 dB	Scale	10.0 dB/div
RBW	1.0 kHz	Serial Number	601128
VBW	1.0 kHz	Base Ver.	V1.60
Detection	RMS	App Ver.	V1.73
Center Frequency	106.700 MHz	Date	12/11/2007 9:00:33 AM
Start Frequency	106.200 MHz	Device Name	CBS-Los_Angeles
Stop Frequency	107.200 MHz		



Channel Power Width: 400.000 kHz — Channel Power: -13.78 dBm

Span: 1.000 MHz — Channel Power Density: -69.80 dBm/Hz

Measurement Parameters

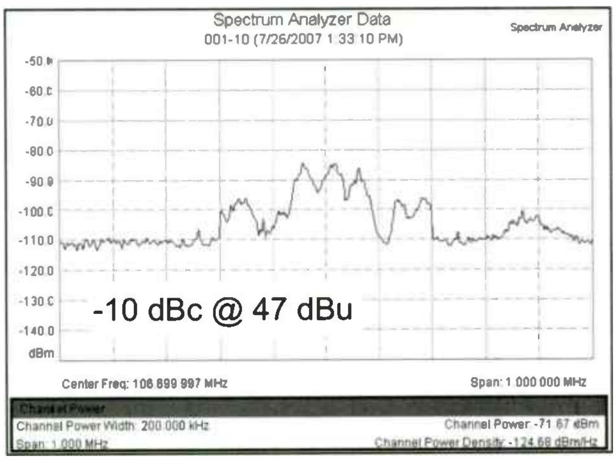
Trace Mode	Normal	Frequency Span	1.000 MHz
Reference Level Offset	0.0 dB	Reference Level	0.000 dBm
Input Attenuation	20.0 dB	Scale	10.0 dB/div
RBW	1.0 kHz	Serial Number	601128
VBW	1.0 kHz	Base Ver.	V1.60
Detection	RMS	App Ver.	V1.73
Center Frequency	106.700 MHz	Date	12/11/2007 8:48:19 AM
Start Frequency	106.200 MHz	Device Name	CBS-Los_Angeles
Stop Frequency	107.200 MHz		

Spectral Plots Harris Z-16HD Plus Output @ -20 dBc & -10 dBc

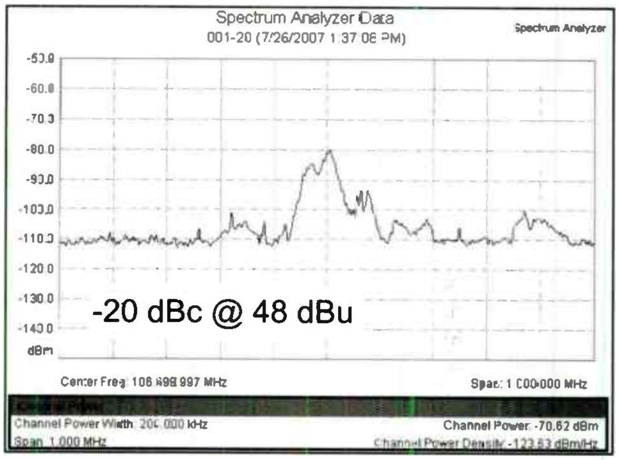
SITE 01 – KROQ STUDIOS / 5901 VENICE BLVD, LOS ANGELES, CA

(N34° 02.382' / W118° 22.116') 20.19 km to TX
Predicted FI (LR9090): 63 dBµ   Actual FI Outside: 65.7 dBµ   Actual FI Inside: 47.33 dBµ
Atten: 18.37 dB

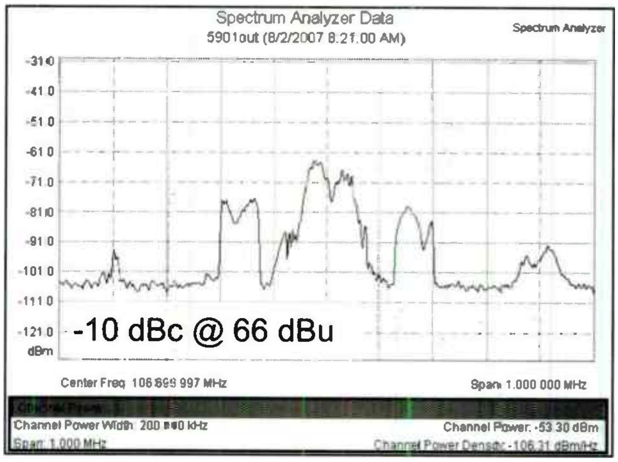
The KROQ studios are in a low-rise stucco building in West Los Angeles. The West Side is partially shadowed from the Verdugo Mountains, due to the Hollywood Hills which further west become the Santa Monica Mountains. About a mile north of this location, Wilshire Blvd. runs East-West between this location and the KROQ main transmitter site. Along Wilshire Blvd. are numerous high-rise buildings which cause additional shadowing as well as severe multipath conditions. Reception quality of KROQ on typical analog consumer receivers with their manufacturer-provided antennas ranges from fair to poor inside the building. The location chosen for our testing was in a conference room on the North side of the building. Analog reception here, on our Boston Acoustics Recepter radio was noisy but listenable. At -20 dBc digital reception was impossible. At -10 dBc the digital reception was solid.



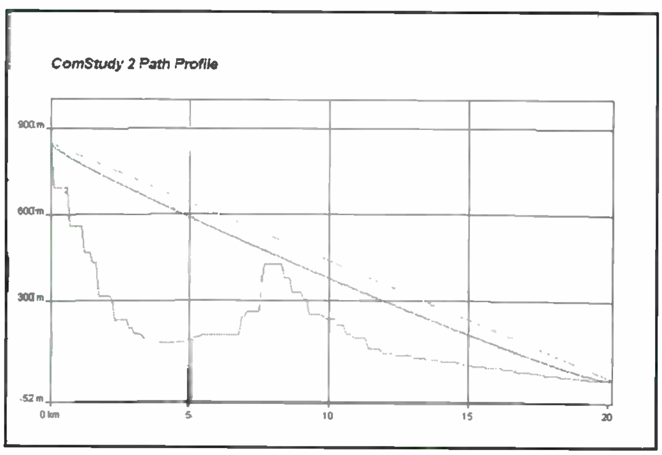
#1 KROQ Studios Conf. Rm -10 dBc



#1 KROQ Studios Conf. Rm -20 dBc



#1 KROQ Studios Outside -10 dBc

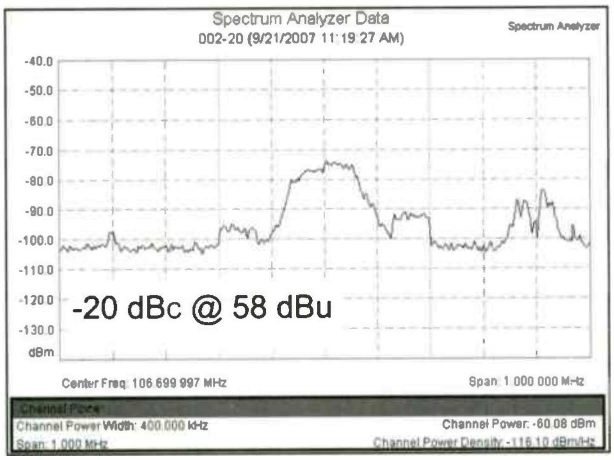


#1 Terrain (KROQ to KROQ Studios)

SITE 02 – THE 801 TOWER / 801 SOUTH FIGUEROA ST., LOS ANGELES, CA

(N34° 03.910' / W118° 14.774') 14.66 km. to TX
Predicted FI (LR9090): 67 dBμ Actual FI Outside: 70.07 dBμ Actual FI Inside: 62.06 dBμ
Atten: 8.6 dB

This was the only downtown LA high-rise building where we were able to arrange access; and then only to a perimeter office on the South side of the building (the side away from the KROQ main transmitter site). The building is of standard steel-and-glass construction. Analog reception here, on our Boston Acoustics Recepter radio was good although with some audible artifacts of multipath propagation. At -20 dBc the digital reception was intermittent. At -10 dBc the digital reception was solid.



#2 801 Tower (Inside @ Window -20 dBc)

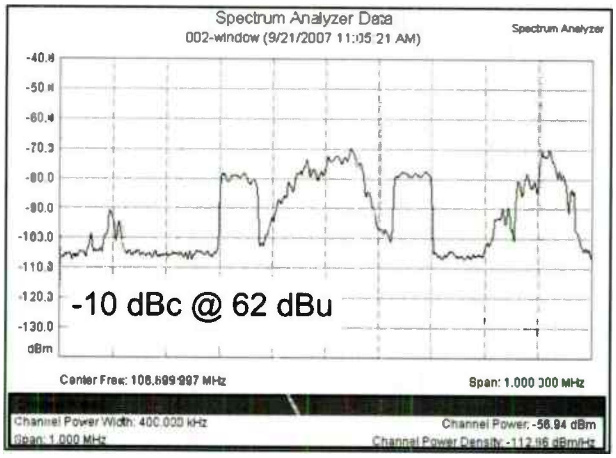


Figure 2 - #2 801 Tower (Inside @ Window -10 dBc)

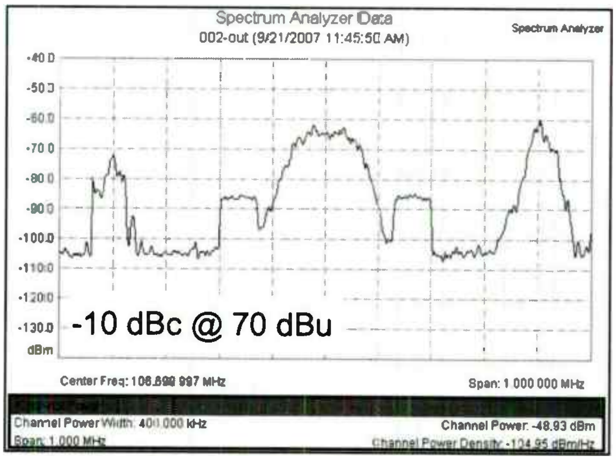


Figure 3 - #2 801 Tower (Outside -10 dBc)

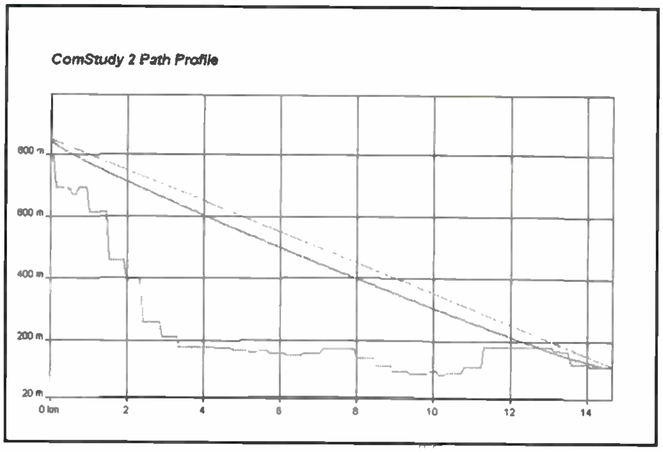
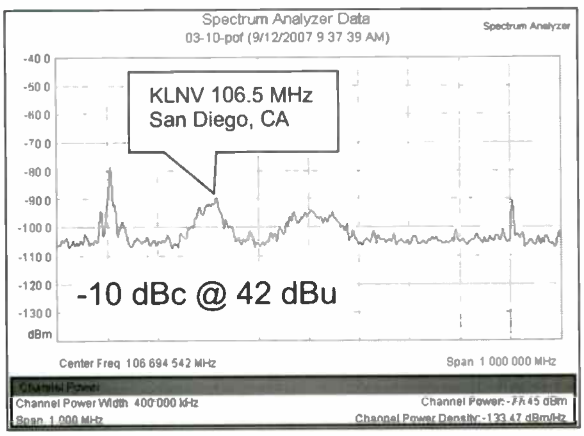


Figure 4 #2 Terrain (KROQ to 801 Tower)

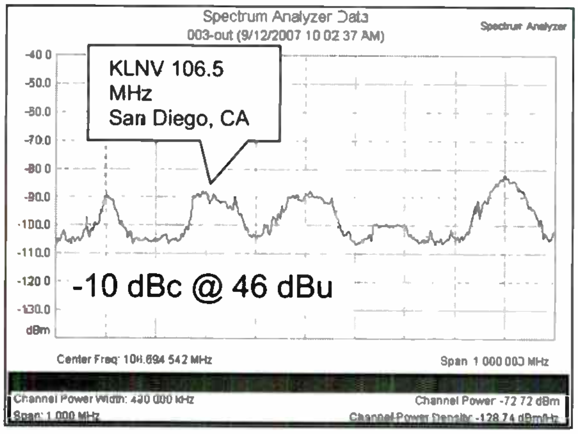
SITE 03 – 8 UNIT APT. BUILDING 4127 EAST 1ST ST., LONG BEACH, CA

(N33° 45.630' / W118° 08.973') 49.59 km to TX
Predicted FI (LR9090): dBμ Actual FI Outside: 46.28 dBμ Actual FI Inside: 43.08 dBμ
Atten: 3.2 dB

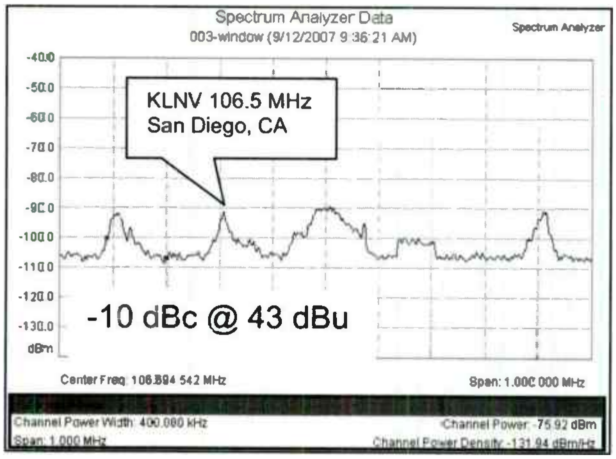
This is a typical small wood frame apartment building in a beachfront section of Long Beach CA locally known as "Belmont Shore", in the Southeast sector of the city only three or four miles West of the Orange County line. Reception of most FM stations is problematic here due to terrain features to the North including one in particular known as Signal Hill. Throughout the apartment the analog reception of KROQ is poor. At -20 dBc digital reception was impossible. At -10 dBc the digital reception was good although a few dropouts were noted as the antenna was moved or as people and household pets moved about, and once as a small airplane flew overhead.



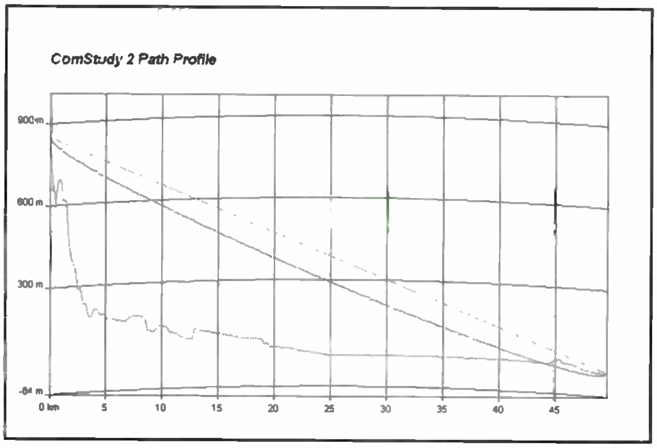
#3 Apt. Bldg. (Inside POF @ -10 dBc)



#3 Apt. Bldg. (Outside @ -10 dBc)



#3 Apt. Bldg. (Inside Window @ -10 dBc)

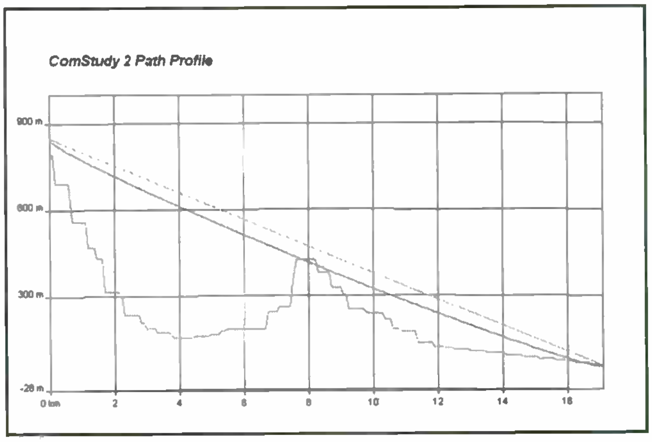


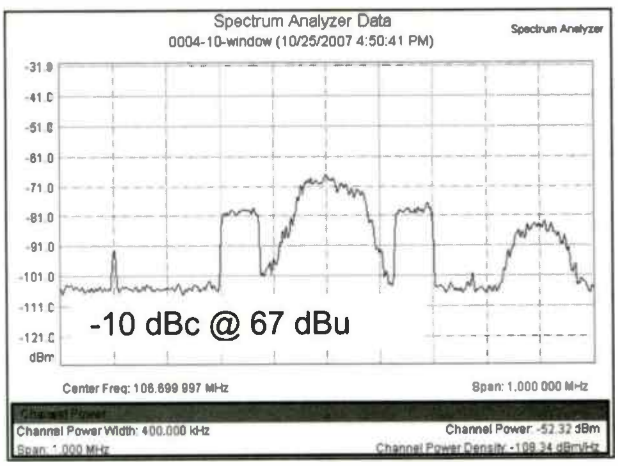
#3 Terrain (KROQ to TX)

SITE 04 – CBS RADIO / 5670 WILSHIRE BLVD, LOS ANGELES, CA

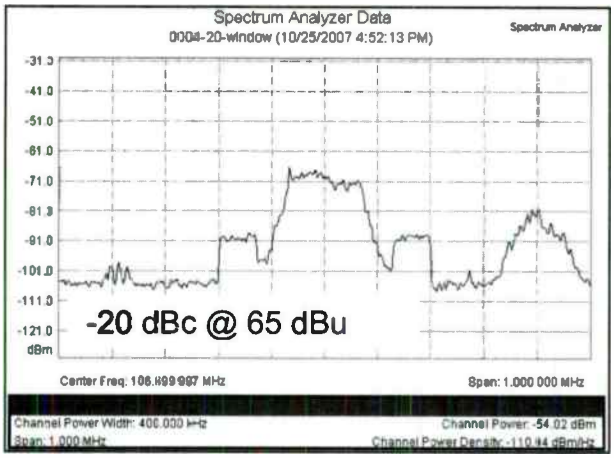
(N34° 03.754' / W118° 21.132') 17.24 km. to TX
Predicted FI (LR9090): 44 dBµ Actual FI Outside: 61.98 dBµ (Street Level)
Actual FI Inside (3th Floor): 66.69 dBµ Atten: NA

This is another high-rise office building, but located about five miles West of downtown Los Angeles in an area commonly referred to as the Miracle Mile District. The building is constructed of concrete, steel, and glass. It happens that CBS Radio leases space in this building for various studios and offices not related to KROQ, so we enjoyed free access to the second and third floors. Our test receiver, spectrum analyzer, and audio recording computer were first set up in a second floor conference room with a glass window facing North (toward the KROQ main transmitter site). As seen in spectrum analyzer trace 04-20-window, there was good digital carrier penetration even at -20 dBc. When antenna was moved approximately 30 feet into the core of the building core, with one intervening steel-framed wall, we abruptly lost both the analog and the -20 dBc digital signals. Upon increasing the digital carriers to -10 dBc the receiver immediately locked to the digital signal, and remained locked for the remainder of our observation time there

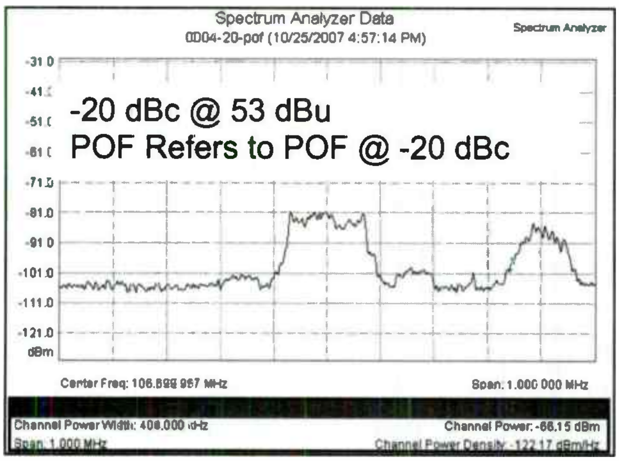




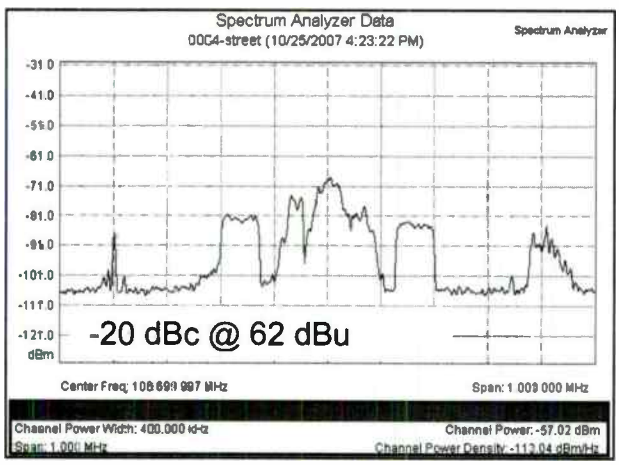
#4 CBS HQ (Inside @ Window -10 dBc)



#4 CBS HQ (Inside @ Window -20 dBc)



#4 CBS HQ (Inside @ POF -20 dBc)

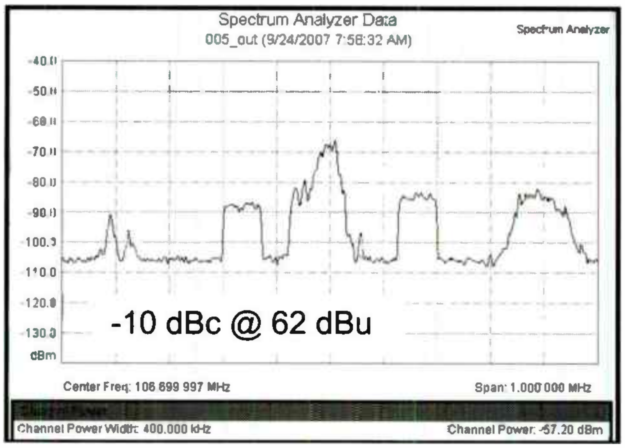


#4 CBS HQ (Outside @ -20 dBc)

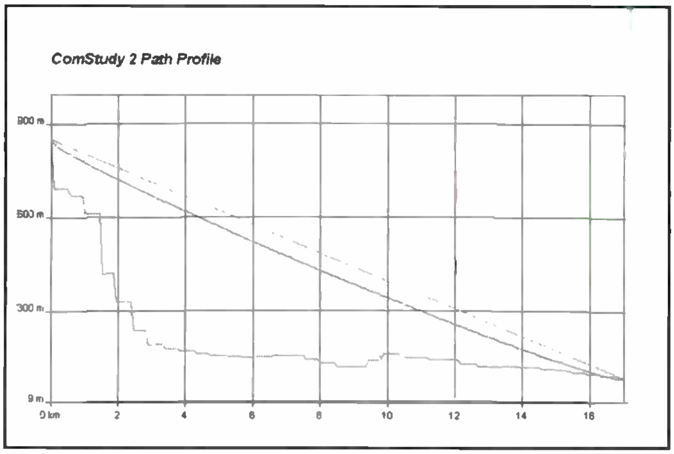
SITE 05 – SBS OFFICE BUILDING /10281 W. PICO BLVD, LOS ANGELES, CA

(N34° 02.578' / W118° 16.374') 17.13 km. to TX
Predicted FI (LR9090): 64 dBμ   Actual FI Outside: 61.8 dBμ (Measured Street Level)
Actual FI 33.08 dBμ   Atten: 28.7 dB

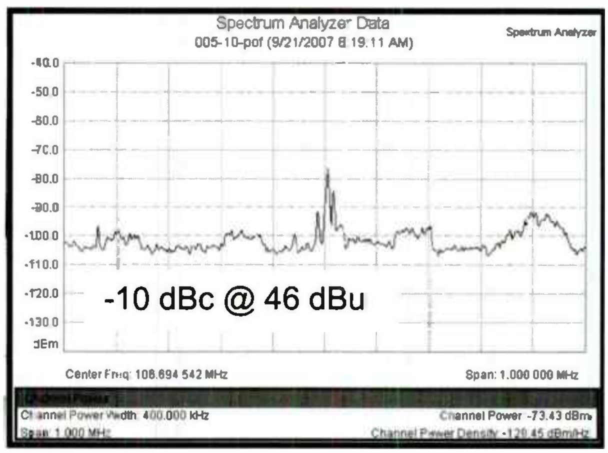
This is a medium-rise building but of construction similar to the two high-rises visited earlier. Like Site 01, this site is largely shadowed from the KROQ main transmitter site, by both natural terrain features and man-made obstacles. This was the most difficult site evaluated thus far for FM reception in general. We were only allowed into the building as far as the ground-floor reception desk and ground-floor indoor parking area. Digital reception at -20 dBc was not possible in any of these areas, and even analog reception was very poor. We finally located one point where with the receiver and test equipment in a window along the South side of the building, we were able to receive the KROQ analog carrier albeit very noisily. Again, when the digital subcarrier levels were increased, we were able to enjoy digital reception without dropouts. Since our earlier measurements, SBS has retaken possession of the second floor of their building at 10281 W. Pico Blvd, which had been leased to a tenant and was thus off limits. This afforded us a far superior location for the measurements, at the North side of the building, toward the KROQ transmitter. The North wall of the building is a shear wall, of reinforced masonry and concrete construction with only a few windows, those being glazed with glass containing a fine wire mesh. At -20 dBc, digital reception was only obtained five feet into the building. At -10 dBc coverage extended all the way to the building core.



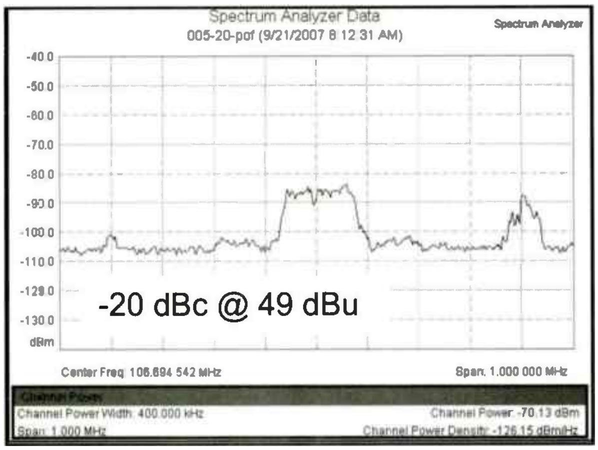
#5 SBS Bldg. (Outside @ -10 dBc)



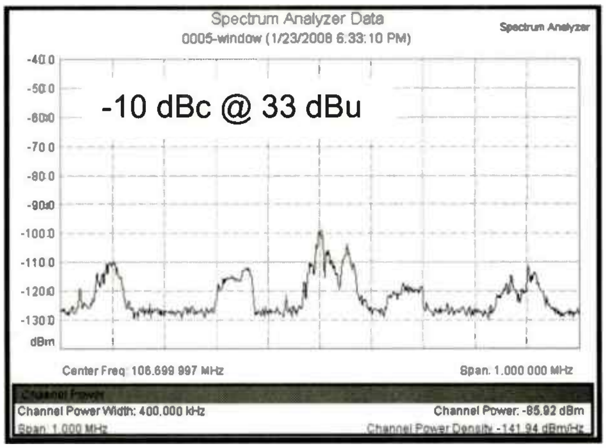
#5 (Terrain to TX)



#5 SBS Bldg. (Inside @ POF -10 dBc)



#5 SBS Bldg. (Inside @ POF -20 dBc)



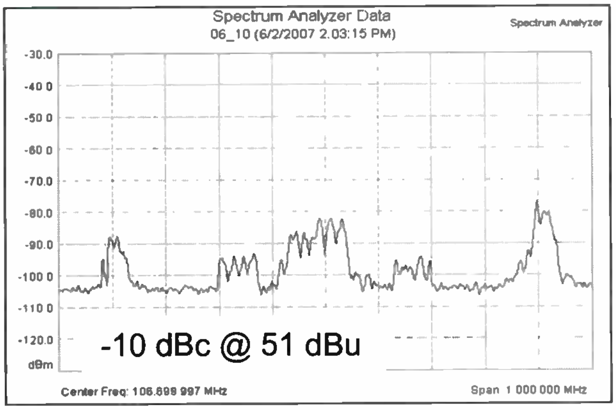
#5 SBS Bldg. (Inside @ Window -10 dBc)

SITE 06 – HOLLYWOOD HIGHLAND ENTERTAINMENT COMPLEX 6801 HOLLYWOOD BLVD, LOS ANGELES, CA

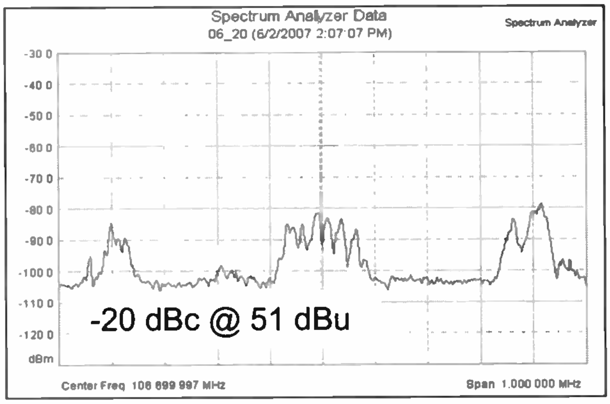
(N34° 06.693' / W118° 20.336') 12.93 km. to TX
Predicted FI (LR9090): 53 dBμ   Actual FI Outside: 63.77 dBμ   Approximately 51 dBμ Inside (see note)
Atten: 12.8 dB

(NOTE: In its first round of testing CBS failed to get indoor channel power measurements, unfortunately CBS radio's lease has expired and measurements are not possible. The signal levels in the first two figures are estimated based on the reference level of the digital carriers in the third figure.)

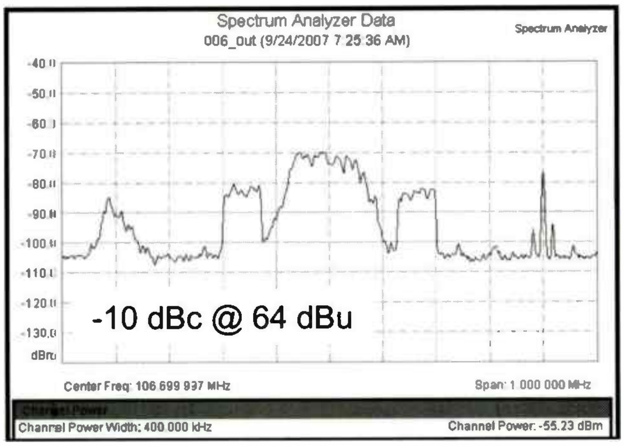
This site is in a commercial district in Hollywood, in a concrete-and-steel building with four floors of motion picture theaters, restaurants, retail stores, and even a bowling alley. The building covers an entire city block. The location is at the base of the previously-mentioned Hollywood Hills. It also is situated at a relatively high elevation and vestiges of interfering FM carriers from Orange County and even San Diego were sometimes visible on the spectrum analyzer. Two exterior photos have been provided. The photos show the front of the complex as viewed from across the street on Hollywood Blvd and the actual measurement location, a small penthouse on the top level of the complex, which houses a remote studio formerly utilized by KROQ for special broadcasts. On the outdoor terrace visible in the photo, analog and digital reception were both fair to good However once inside the steel-frame penthouse, we lost the -20 dBc digital signal altogether, and the analog FM reception was very poor. On the spectrum analyzer we observed a strong interfering signal below the assigned KROQ frequency; we identified this as KALI-FM in Santa Ana at 106.3 MHz. However, after the KROQ digital subcarrier ratio was changed to -10 dBc the digital reception was good.



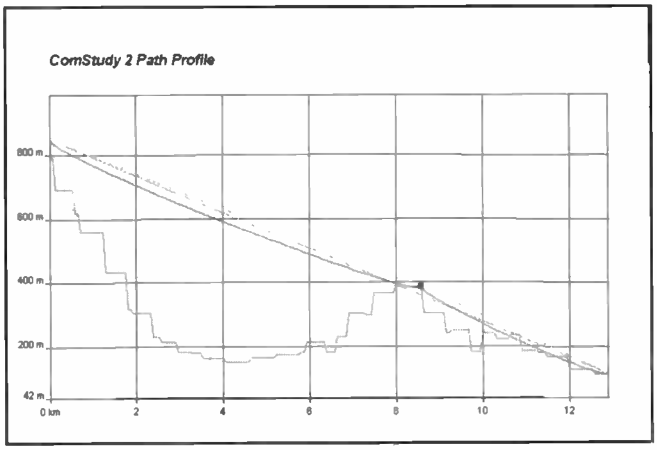
#6 Highland (Inside @ -10 dBc)



#6 Highland (Inside @ -20 dBc)



#6 Highland (Outside @ -10 dBc)

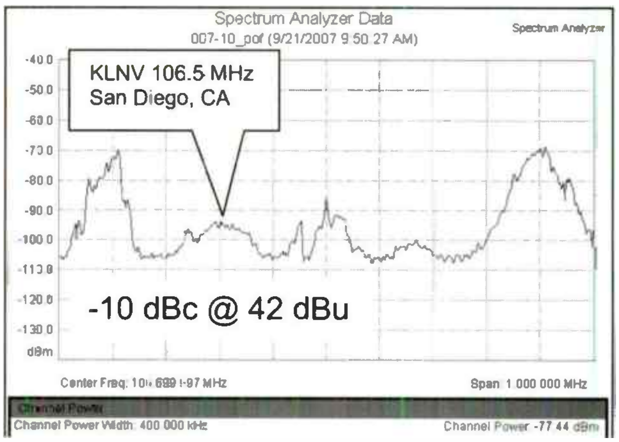


#6 (Terrain to TX)

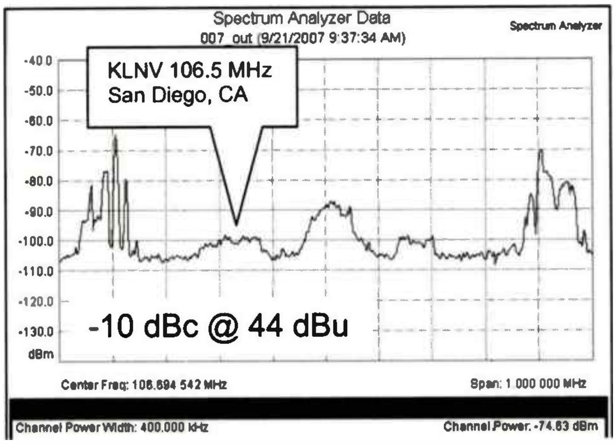
SITE 07 – SINGLE RESIDENCE 1313 35TH ST, SAN PEDRO, CA

(N33° 42.973' / W 118° 18.351') – 53.50 km. to TX
Predicted FI (LR9090): 34 dBμ   Actual FI Outside: 44.37 dBμ Actual FI Inside: 41.56 dBμ
Atten: 2.81 dB

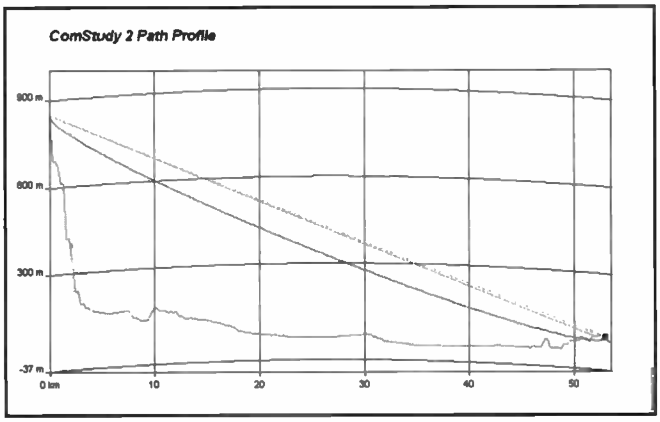
This is the most distant site from the KROQ main transmitter site. It is a single-family/single story residence of wood and stucco construction. It is located in a very hilly neighborhood overlooking the Pacific Ocean, and due to the over-water propagation path many San Diego FM and TV stations are routinely received here. The 106.3 MHz signal from Santa Ana was again very noticeable, and at none of various locations checked throughout the house was clean analog reception of KROQ possible; nor was digital reception possible anywhere in the house, with the digital carrier ratio at -20 dBc. Even when the carrier ratio was changed to -10 dBc, digital reception was marginal; we did locate one area in the living room where we did achieve good digital reception and that is the location where our measurements were made. It is interesting to note that because of the construction of the house the signal levels were mostly uniform throughout the structure.



#7 1313 (Inside POF @ -10 dBc)



#7 1313 (Outside @ -10 dBc)

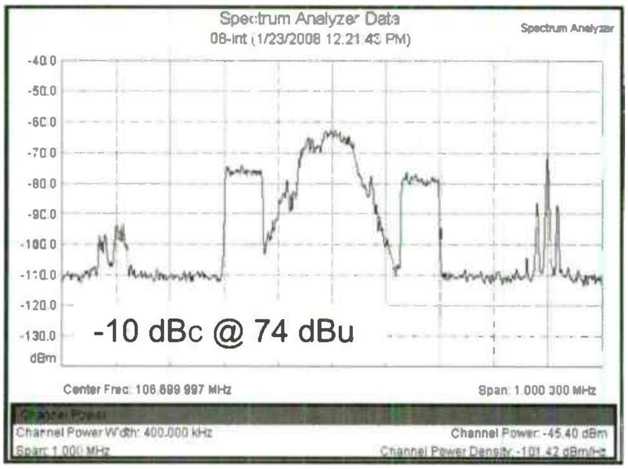


#7 1313 (Terrain to TX)

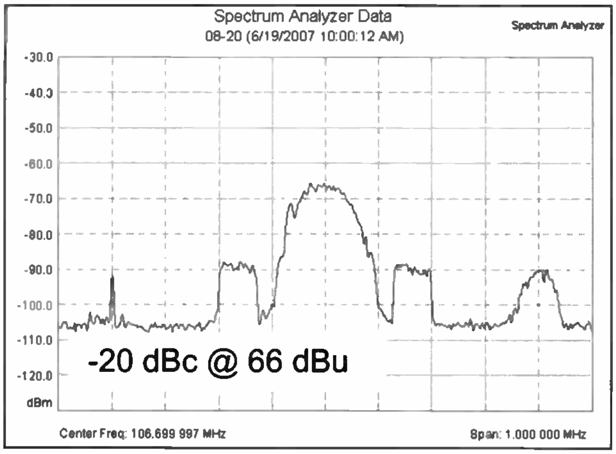
SITE 08 – 36 UNIT APT. COMPLEX 4646 CAHUENGA BLVD, TOLUCA LAKE, CA

(N34° 09.274' / W118° 21.667') 10.59 km. to TX
Predicted FI (LR9090): 69 dBμ   Actual FI Outside: 68.1 dBμ   Actual FI Inside: 73.6 dBμ
Atten: N/A

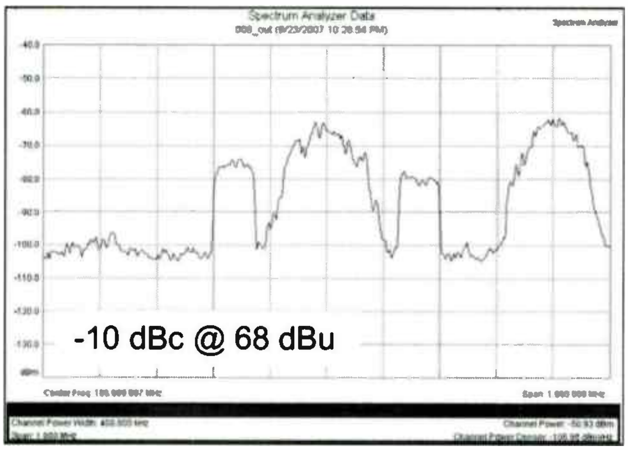
This is an apartment on the top (2nd) floor of a two-story apartment building in the San Fernando Valley, where KROQ enjoys excellent coverage. There was no loss of digital signal at either -10 or -20 dBc, anywhere in the apartment. As with Site 07 the construction of the building yielded nearly uniform attenuation throughout the structure.



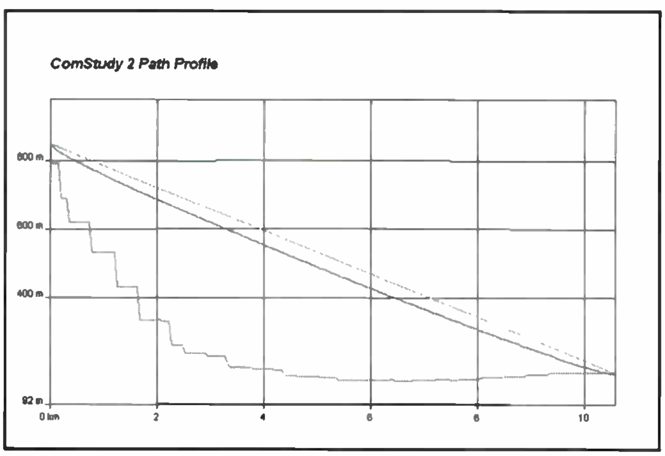
#8 4646 (Window @ -10 dBc)



#8 4646 (Inside @ -20 dBc)



#8 4646 (Outside @ -10 dBc)

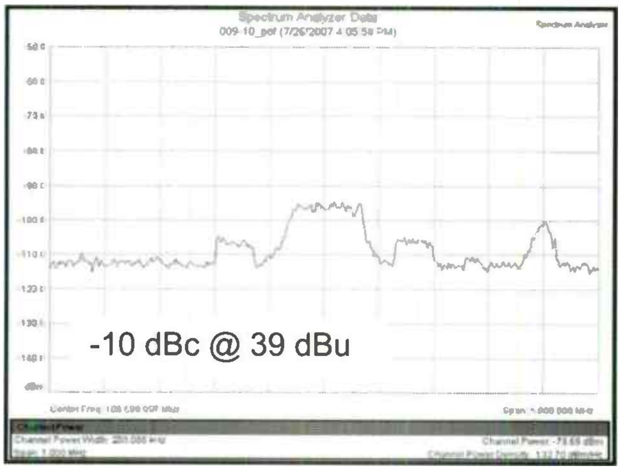


#8 4646 (Terrain to TX)

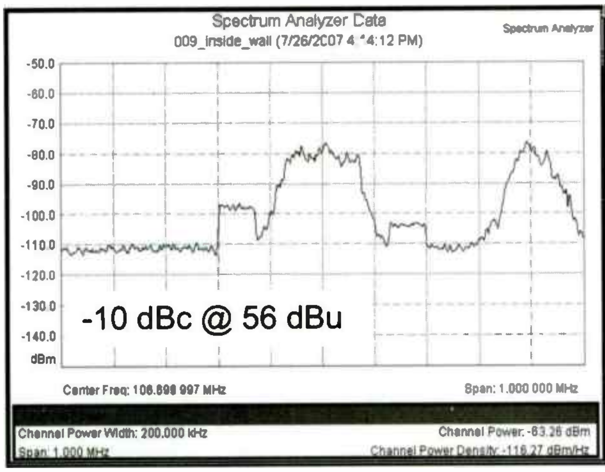
SITE 09 – PARKING GARAGE 11239 VENTURA BLVD., STUDIO CITY, CA

(N34° 08.507' / W118° 22.493') 12.37 km. to TX
Predicted FI (LR9090): 68 dBμ Actual FI Outside: 82.61 dBμ Actual FI Inside: 55.74 dBμ
Atten: 26.87 dB

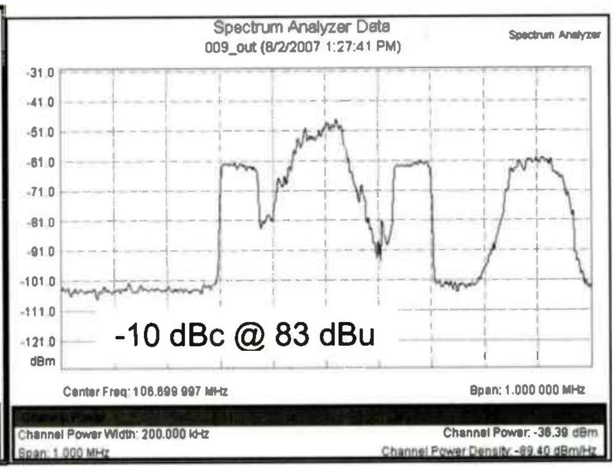
This is an underground parking structure, also in the San Fernando Valley in an area at the base of the Santa Monica Mountains known locally as Studio City. We found parking structures to be well suited to the measurement program due to ease of access. ability to keep our measuring equipment inside a vehicle, and ability to power the equipment from the vehicular electrical system using an inverter. The vehicle can be slowly driven in the direction away from the transmitter site, and a continuous decrease in received signal level is observed as the vehicle moves further into the structure. At the specified point of failure, the vehicle is stopped while spectrum analyzer traces are captured and photographs taken.



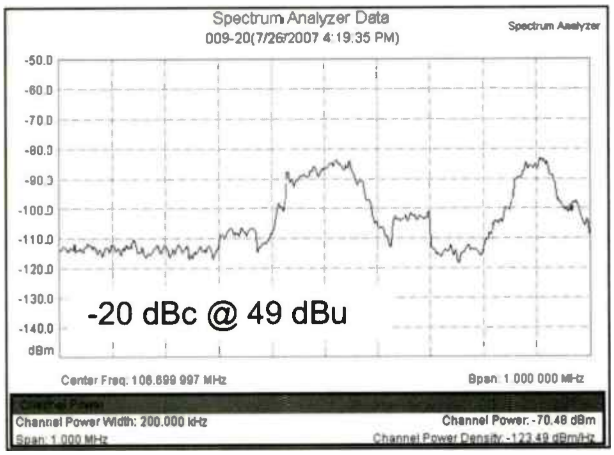
#9 Parking Garage (POF @ -10 dBc)



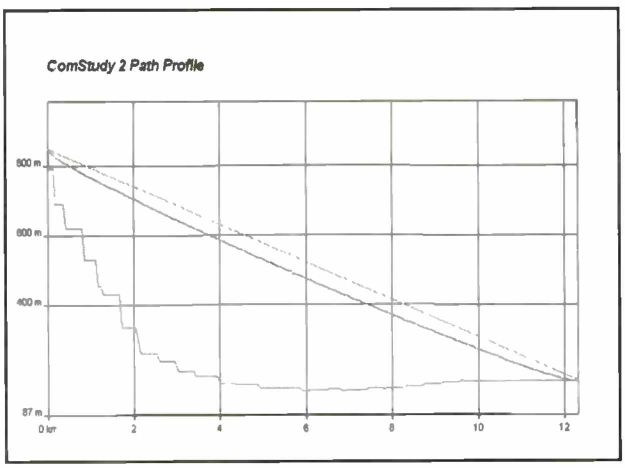
#9 Parking Garage (Inside @ -10 dBc)



#9 Parking Garage (Outside @ -10 dBc)



#9 Parking Garage (Outside @ -20 dBc)

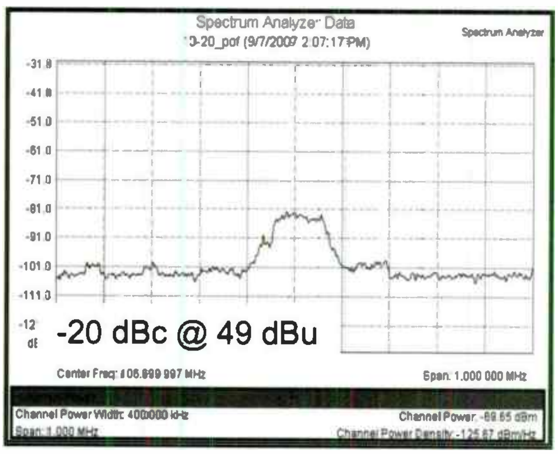


#9 Parking Garage (Terrain to TX)

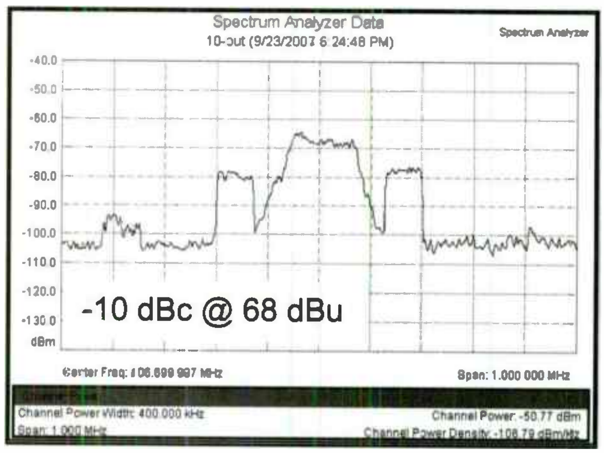
SITE 10 – PARKING GARAGE 211 N. GLENDALE AVE., GLENDALE, CA

(N34° 08.897' / W118° 14.769') 5.55 km. to TX
Predicted FI (LR9090): 75 dBµ Actual FI Outside: 68.23 dBµ Actual FI Inside: 49.35 dBµ
Atten: 18.88 dB

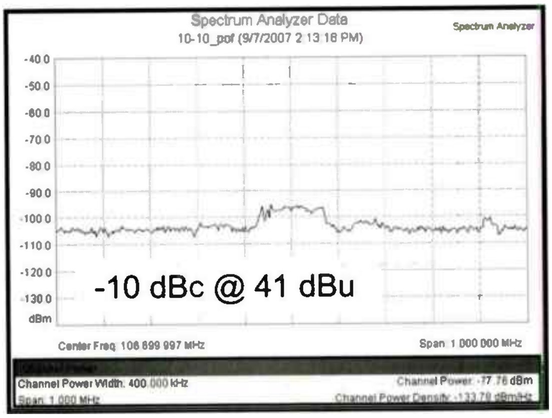
This site is similar to site 09; a concrete and steel commercial building with an underground parking lot within the commercial district of Glendale. This site is 5.5 km from the KROQ transmitter site. Outdoor reception of KROQ's analog and -20 dBc digital signal is good, however, when the test vehicle was driven more than 25–30 feet into the garage, the -20 dBc digital signal became unusable. With the digital subcarrier power increased to -10 dBc, the digital reception was solid up to 165 feet inside the facility. At that point, the analog reception had deteriorated to a degree that most listeners would not tolerate. The analog and digital signals both became unusable at distances greater than 165 feet.



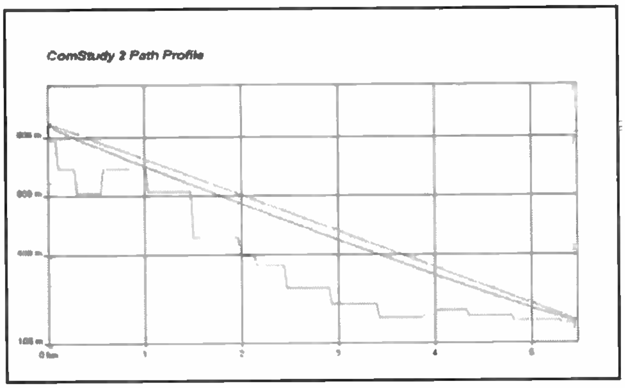
#10 Parking Garage (-20 dBc POF)



#10 Parking Garage (-10 dBc Outside)



#10 Parking Garage (-10 dBc POF)



#10 Parking Garage (Terrain to TX)

TABLE OF AUDIO CUTS

Follows a table of the analog recordings of made in conjunction with the test program. The cuts document the quality of the analog reception at or near digital failure in each of the test locations.

Site	Analog Cuts	Cut Label
00	DNA	[no audio recorded at transmitter site]
01	-20 POF	01-20-dBc_POF_Analog
02	-20 POF	02-20-dBc_POF_Analog
03	-10 Analog Poor	03-10-dBc_Analog
03	-10 POF	03-10-dBc_POF_Analog
04	-20 POF	04-20-dBc_POF_Analog
05	-20 POF	05-20-dBc_POF_Analog
06	-20 POF	06-20-dBc_POF_Analog
07	-10 outside	07-10-dBc_Analog
08	-10 POF	08-10-dBc_POF-Analog
09	-10 POF	09-10-dBc_POF_Analog
10	-20 POF	10-20-dBc_POF_Analog
10	-10 POF	10-10-dBc_POF_Analog

FREE SOFTWARE TOOLS FOR DESIGN OF AM ANTENNAS

VAN RICHARDS-SMITH
RadioTAB Network
AUSTRALIA

ABSTRACT

Many engineers employed in AM stations never get to design an antenna system. This is usually done by a consultant but they are expected to maintain the system.
This paper presents a set of software tools that may help station engineers understand the effect that any changes to components in the system may make.
This software is available to all NAB attendees Free of charge from the RadioTab FTP site. Details on how to do this are at the end of this presentation.

History of the Package

It is difficult for the engineer to convey to management what we are talking about when referring to the performance of the antenna system of our station. Working on the old adage that *picture is worth a thousand words* we developed the AntPlot Program .
It had its geneses in a very BASIC program for DOS that just did an on screen plot. Our head station was making a submission to the regulator for a power increase and had to provide the figures to justify our case.

The AntPlot Program

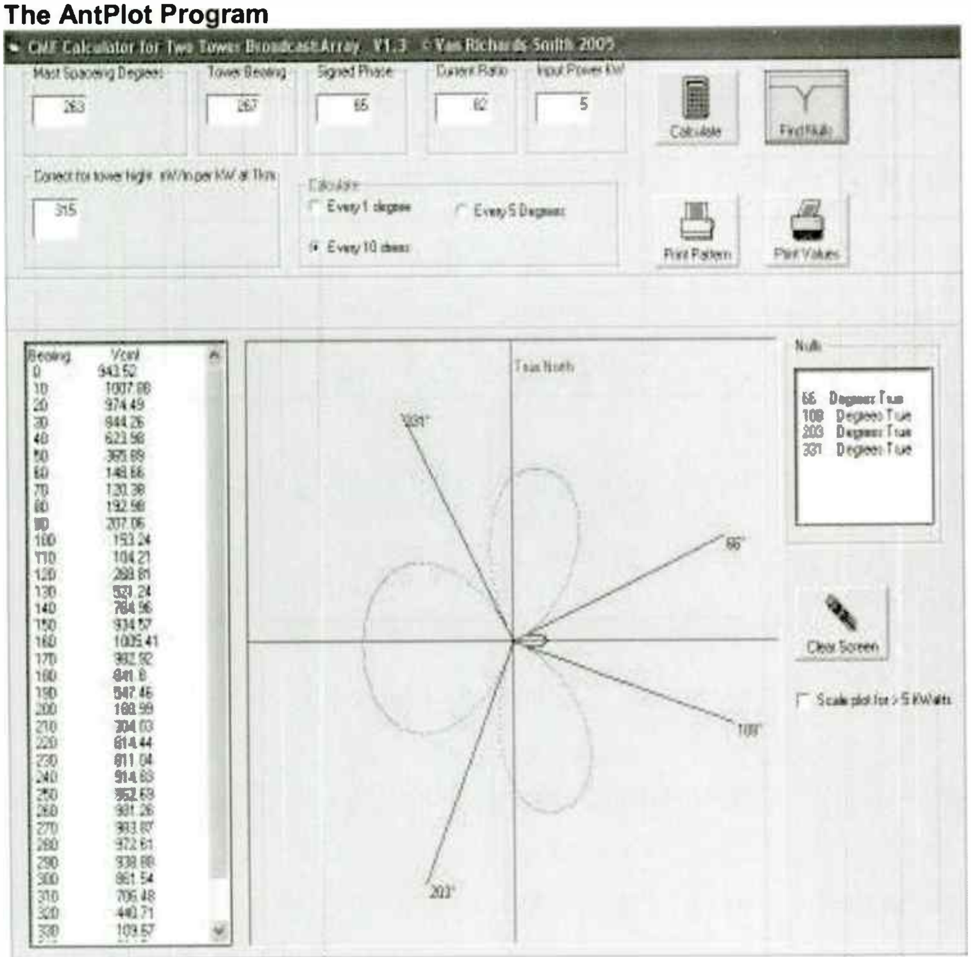


Fig 1 Sample Plot Print

This program allows you to enter the parameters of your stations two tower array and generate on screen the pattern relative to true north

When considering field strengths one of the often forgotten factors is the actual height of the towers

As you will see you can calculate the CMF results for 10°, 5° and 1° intervals.

The displayed plot can be scaled for high powers that over run the edge of the displayed area.

The nulls can be displayed on the plot by clicking *Find Nulls*

Printing the display uses your default printer.

There is provision in this program for inputting this information.

Where do I get this information from ?

Why not try our trusty NAB Handbook

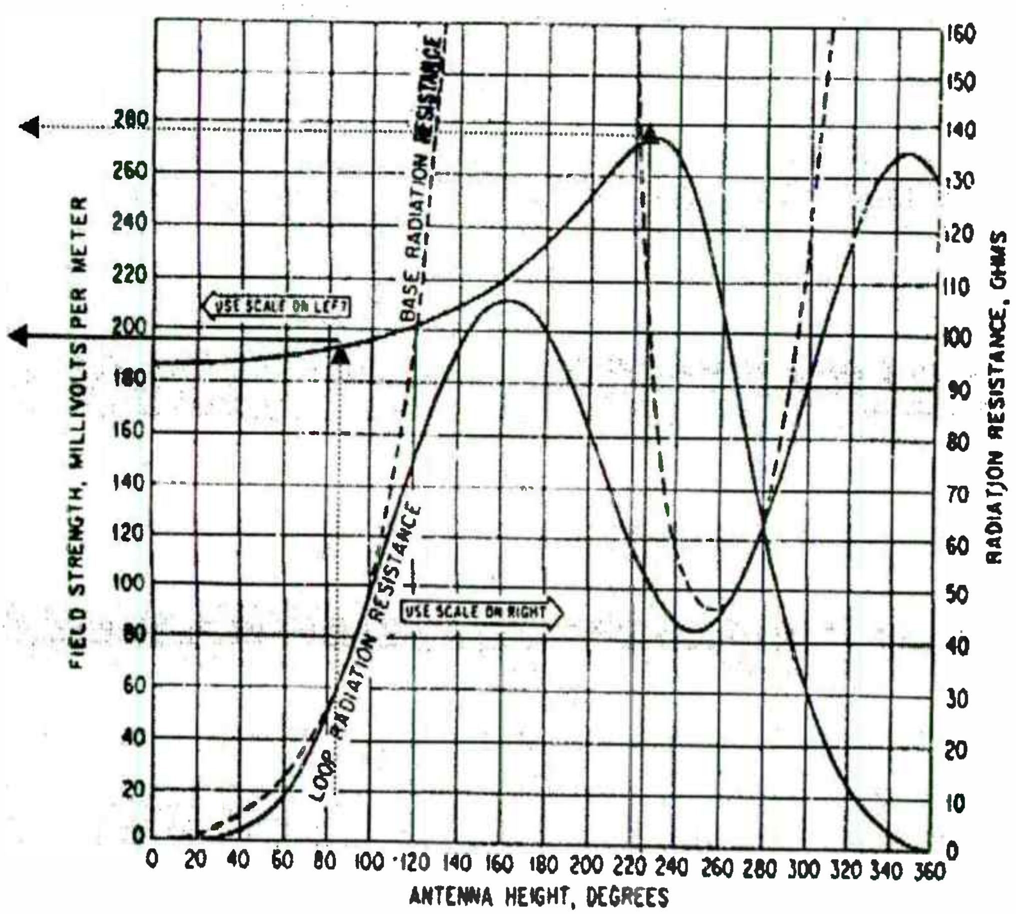


Fig. A-8. Inverse field strength at 1 mile for 1 kw, loop and base radiation resistance as a function of tower height over a perfectly conducting earth.

Extract from 7th edition NAB Handbook

This is a great tool for convincing management that the tall tower will give much better results than the cheaper alternative.
There is no substitute for metal in the air.

As the program was developed for the metric system , you have to convert the readings from the above graph which is in millivolts per meter at one mile to millivolts per meter at one kilometer by multiplying by 1.609 .

For a ¼ wave antenna the theoretical field strength for 1 kW would be 198 mV/meter at one mile or 315mV/meter at one Kilometer . This is the default value used by the program – but remember , If your towers are not 1/4wave you have to enter a different value.
On the graph you will see I have drawn two arrows The bottom one is for the 1/4wave case and the second is for a 5/8 wave tower.
The 5/8 tower gives a figure from the graph of 275 mV/m at one mile which converts to 442 mV/m at one kilometer.

When plotted on screen it is much easier to see the difference between the coverage of the two towers
.

Fault finding.

One of the practical problems Engineers face is when you suspect that the nulls of your directional array have shifted . With the aide of GPS, it is easy to locate where your null should be and to fix a position that is an exact distance from the center of the array which makes the mathematics easier.

Firstly you have to determine whether the antenna monitor is accurate or has the antenna system drifted ?
With this program you can plot what the phase meter is telling you and then check out the nulls to see if it is real or if you have a problem with the measuring system.

If the nulls are found to be in the correct place some of the problems in the metering system can be-

1. The Antenna Monitor is at fault

2. The monitor cables are at fault – they may have changed velocity factor due to ageing of the dielectric .

3. Bad connections in the sample cables or loops .

What effect has ground conductivity had on the signal

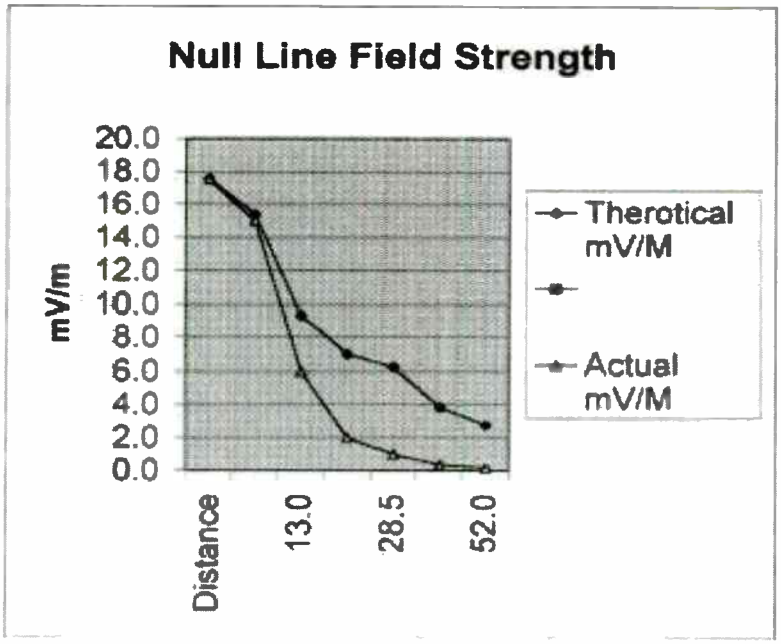
.
How do I check out how significant this factor is?

1. Take a field strength reading at a point on a bearing at a known distance D (in kilometers)
2. Divide the CMF figure generated by AntPlot for that bearing by the distance D and this will give you a figure X mV/m to compare directly with the field strength obtained in 1.
3. CMF / D = X mV/m

Field Strengths on 4IP Null Line of 203 degrees

200 V CMF

Distance	Km	11.4	13.0	21.5	28.5	32.0	52.0	73.0
Therotical	mV/	17.5	15.4	9.3	7.0	6.3	3.8	2.7
Actual	mV/M	17.5	15.0	6.0	2.0	1.0	0.4	0.2



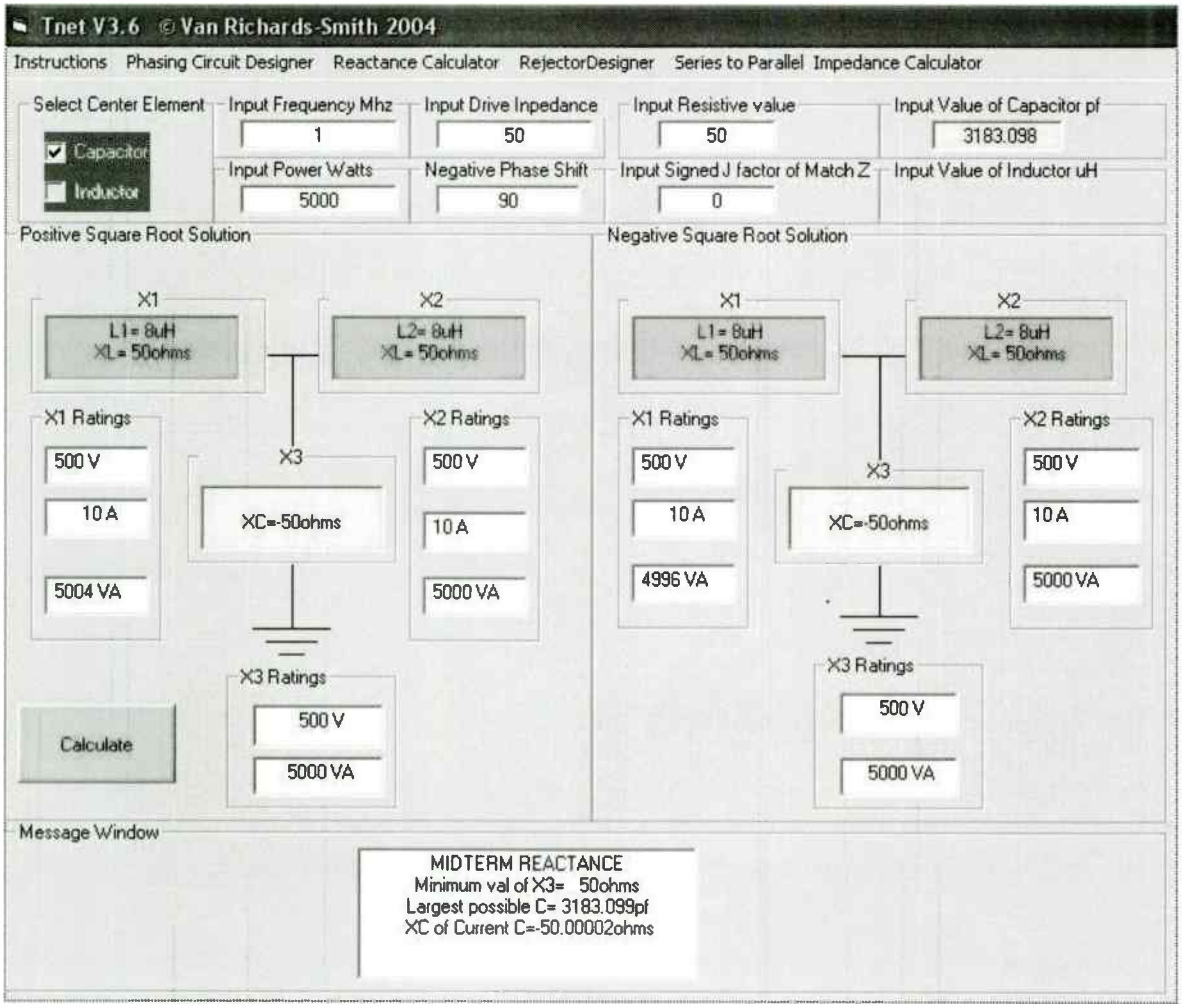
The above shows how the signal strength deviates from the theoretical over distance.

By presenting these figures to the regulator we were granted a 5kw FM translator near the town at the 50km mark to cover the deficiencies in coverage in one of the largest population growth areas in our coverage area.

Down to the Nuts and Bolts

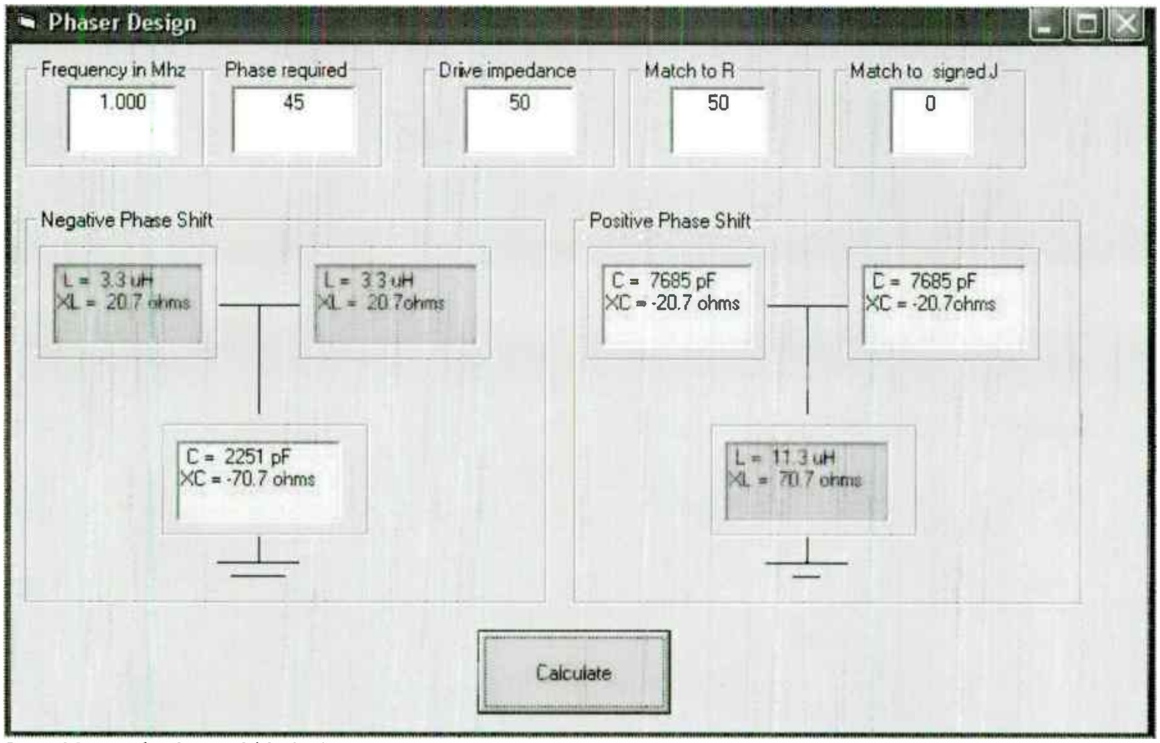
In recent years the world has revised its standards for Electro Magnetic Radiation and how it is treated in the work place. No longer is it acceptable for riggers to climb live towers or to have exposed transmission components where workers can come in contact with them. These changes required that our transmission facilities had to be re-designed to handle single tower operation as well a removing all exposed components such as knife switches and open line feeders.

In doing this we developed the T-Net program that calculates the values for T-Network components and calculates the all important ratings for each .

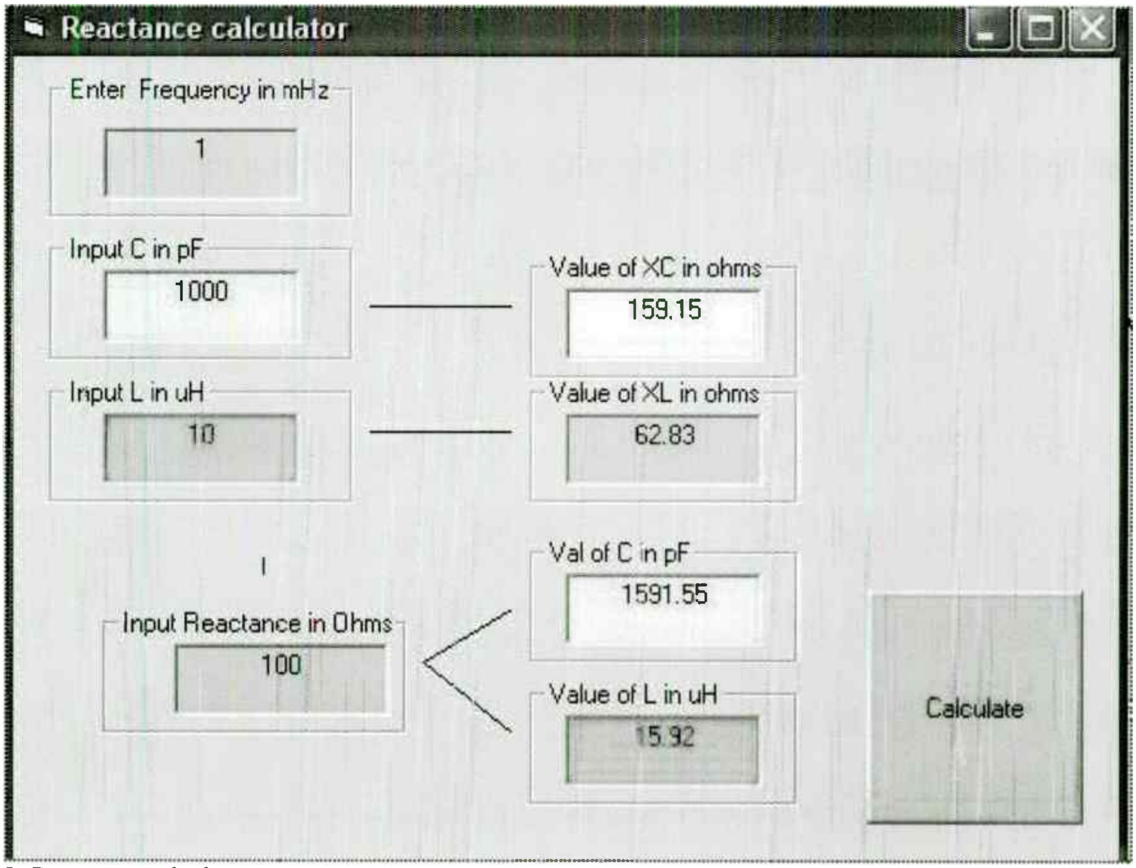


1. Main T-Net screen

Select The following from the Main screen menu

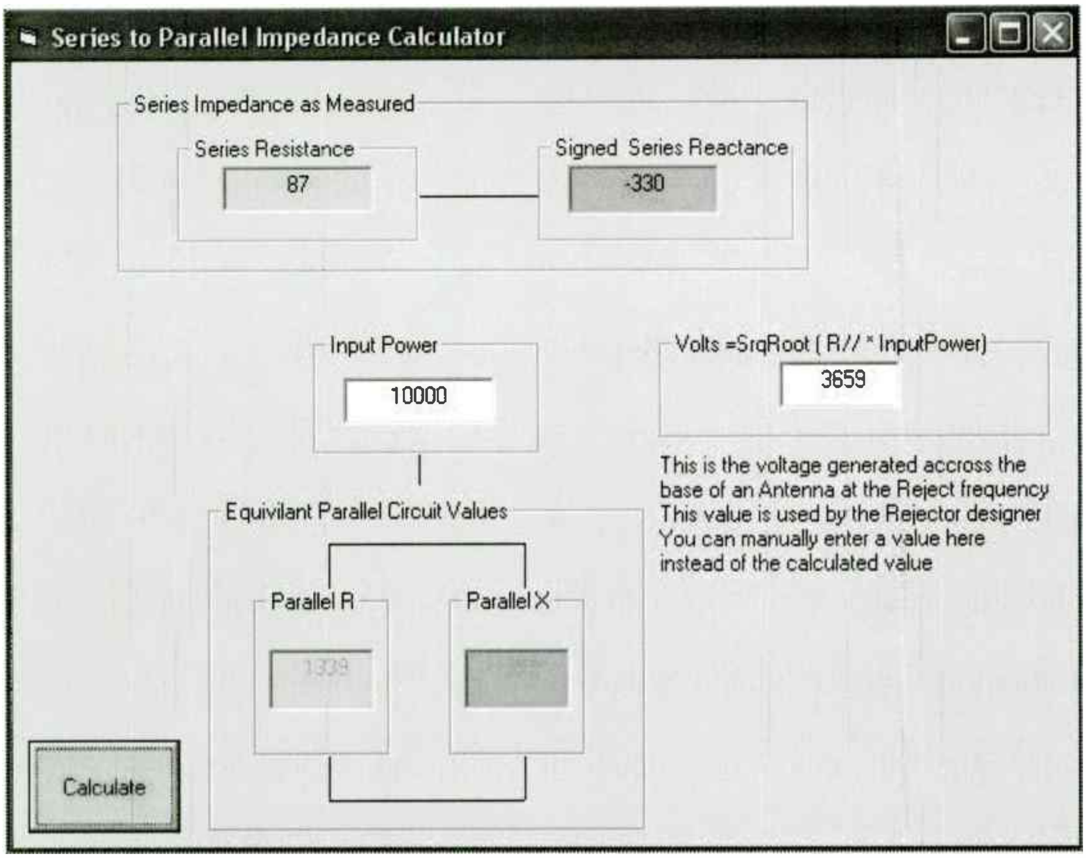


2. T-Network phase shift designer

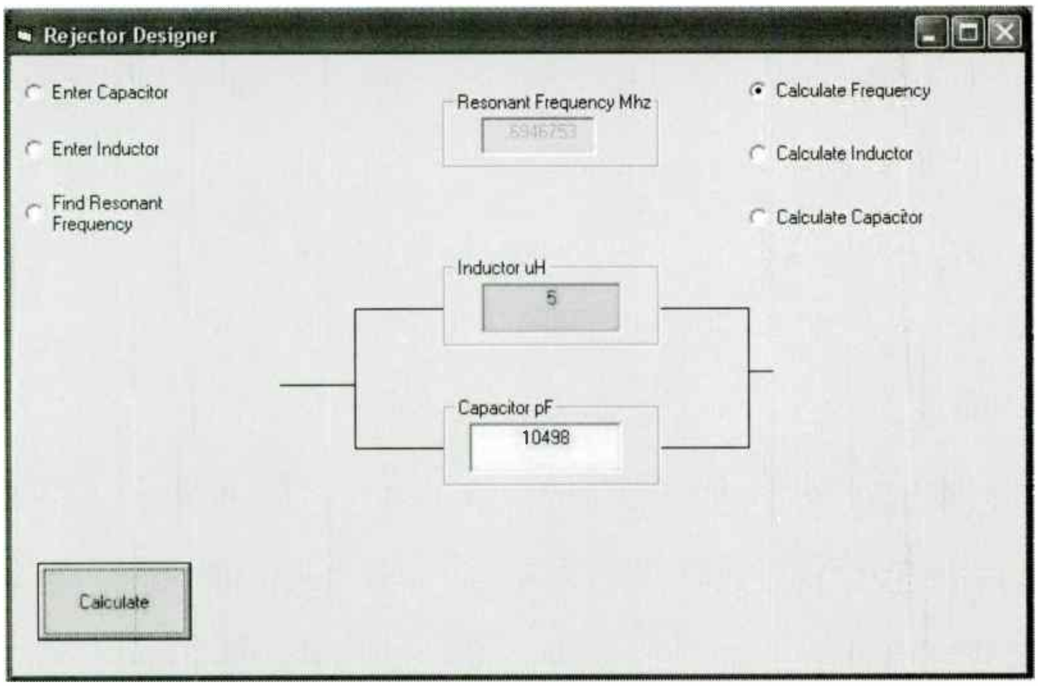


3. Reactance calculator

4 Rejector designer



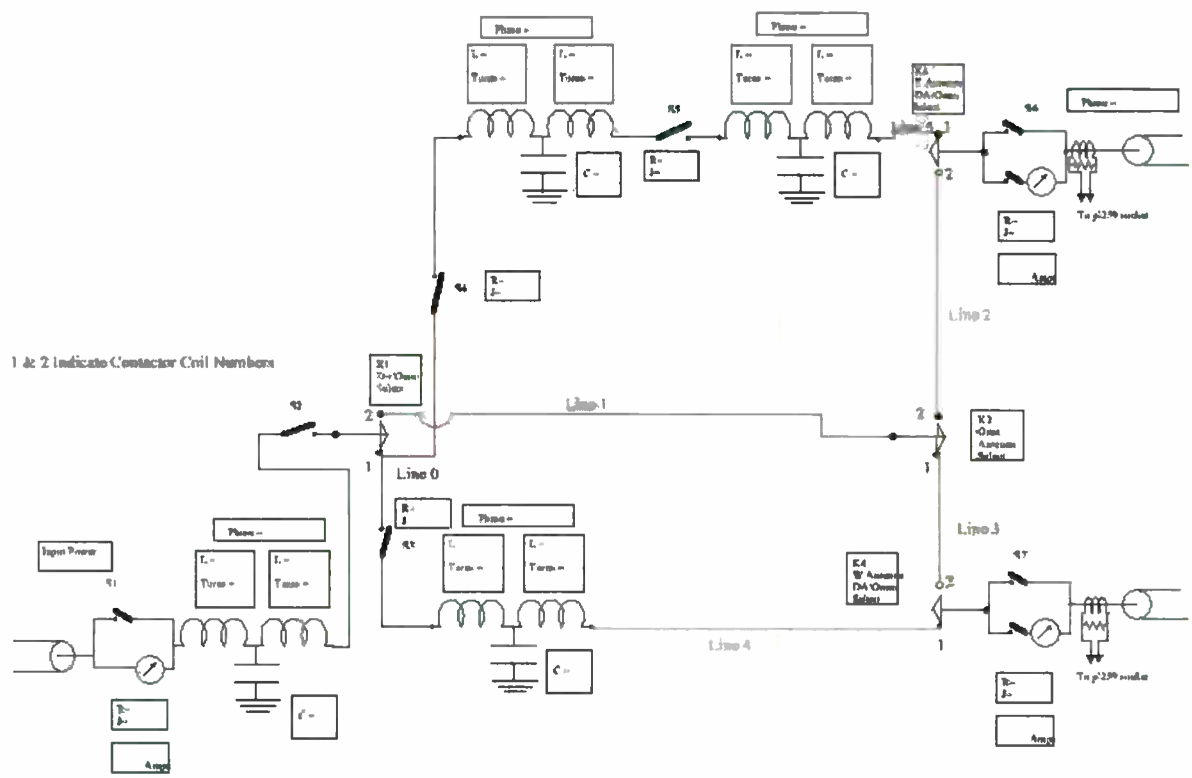
5. series to parallel impedance converter



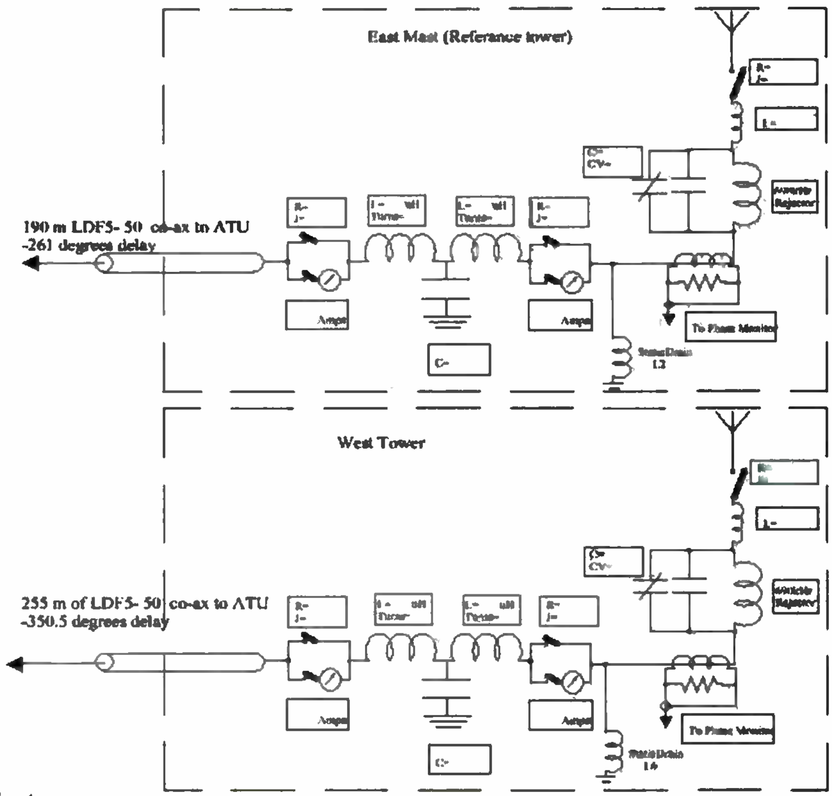
6 Parallel Resonant frequency designer

To illustrate the use of T-Net let us now work through a typical Antenna Design using the data from the chart

				Data Chart St Helana					
	4TAB	4KQ			4TAB			4KQ	
	1.008	0.693			R	J		R	J
Wave length meters 360 degrees	297.62	432.90		Input z east tower 28/9/01	75	-395		92	504
1/2 wave length meters	148.81	216.45		input z west tower 28/9/01	85	-375		155	353
1/4 Wave length meters	74.40	108.23							
				Line I	10.00			10.60	
Cable Wave velocity factor	0.88	0.88		East In				4.4	
length of cable for 360 degrees	261.90	380.95		East Ant	6.2			5.2	
Degrees per meter of cable	1.37	0.95		Power East	2883			2487.68	
Cable to East Tower length Meters	190.00	190.00							
Cable to West Tower length Meters	255.00	255.00							
East tower delay	261.16	179.55		West in	6.9			7.9	
West tower delay	350.51	240.98		West Ant	5			4.4	
west tower lags by degrees	-89.35	-61.43		Power West	2125			3000.8	
Required phase	65.00	-50.00		Power Ratio	0.73708			0.829006	
current ratio	0.82	0.90							
Required phase delay in east tower	-24.00								
Input z to Power devider	50.00	50.00							
voltage at common point	500.00	500.00							
Z1 of PDU East	110.98	105.56							
Z2 of PDU West	91.00	95.00							
input power Kw	5.00	5.00							
Power to east tower	2252.75	2368.42							
Power to west tower	2747.25	2631.58							



Radio 4IP St Helena Antenna System   Current Ratio = .8   Phase = 65



Radio 41P St Helena Antenna System

In Maintaining a Broadcast Antenna System

We need the correct tools –

1. An Impedance Bridge

2. An accurate Antenna Monitor

3. A GPS

4. A field strength meter

If your station does not have or access to these Items you can not be expected to keep the wheels turning efficiently.

The only thing that gets your signal to the listener is the transmission system

Sometimes Management forget where the Radio station really is.

I hope you find these programs useful in maintaining an efficient broadcast system.

Download the program packages

In you web browser type

Ftp:// 203.3.76.75

A logon window will open

The User Name is RADIOTAB

Password Is R@diot@b

Select AntPlotinstalle.zip or TNetinstaller.zip

Save to a temporary directory and extract files to two different directories and then run setup.

Presenters Email:
Van.richards-smith@RadioTab.com

In conclusion I would like to acknowledge the following excellent Published works of

Pythagoras

Robert A Jones .PE " Directional Antenna Handbook"

Jack Layton " Directional Antennas Made Simple "

The many contributors and the editors of the NAB Engineering Handbook

Technology Innovations

Thursday, April 17, 2008
9:00 AM – 11:30 AM

Chairperson: Charles Jablonski
Redwood City, CA

Super Hi-Vision Transmission Experiment in the 21GHz Band
Hisashi Sujikai, NHK Science & Technical Research Labs, Tokyo, Japan

HDMI as Television Application Platform for Interactive and More
Rainer Zwing, Thomson, Villingen-Schwenningen, Germany

11.88 GB/S SDI Continuing the Evolution of SDI
Gareth Heywood and Ryan Latchman, Connectivity Technologies, Gennum Corporation, Burlington, Ontario, Canada

Audio Mixing Requirements in Next Generation Broadcast Receivers for Audio Description and Other Enhanced Features
Roland Vlaicu, Dolby Laboratories, San Francisco, CA

HDTV System Onboard the Lunar Explorer Kaguya (SELENE)
Seiji Mitsuhashi, NHK (Japan Broadcasting Corporation), Tokyo, Japan

SUPER HI-VISION TRANSMISSION EXPERIMENT IN THE 21GHz BAND

Hisashi Sujikai, Yoichi Suzuki, Shoji Tanaka and Kazuyoshi Shogen
NHK Science & Technical Research Laboratories
Tokyo, Japan

ABSTRACT

NHK is studying ultrahigh-definition video system "Super Hi-Vision" as next generation broadcasting, aiming at the ultimate broadcasting system with a heightened sensation of reality. It has 16 times the amount of information of the current HDTV system and provides an overwhelmingly realistic viewing sensation with more than 4000 scanning lines. NHK is investigating a step toward the practical use of Super Hi-Vision as broadcasting. The 21-GHz band satellite broadcasting system is under development as a promising transmission channel to deliver Super Hi-Vision to individual homes because it has a wide RF channel bandwidth of 600MHz. This time we have developed prototypes of a 300-MHz wide-band modulator and a demodulator and carried out an indoor Super Hi-Vision transmission experiment through the 21GHz band experimental transponder with a single carrier to verify performance of the hardware and evaluate the wide-band transmission characteristic. In the experiment, a compressed Super Hi-Vision signal was modulated, amplified by a TWT, and went on the air in the 21GHz band. It was received by a parabolic broadcasting antenna whose diameter was equal to a usual home satellite antenna, demodulated and decoded, and displayed on a monitor. Through the experiment, we confirmed the possibility of the Super Hi-Vision broadcasting via a 21GHz band satellite.

INTRODUCTION

The ultrahigh-definition, large-scale video system is expected in a variety of fields such as broadcasting, digital cinema, medical care, archives, education and so on. To realize an extremely high-resolution imagery system whose audio and visual signals convey a strong sensation of reality to viewers, NHK is studying ultrahigh-definition video system "Super Hi-Vision" as next generation broadcasting[1]. The video format of the Super Hi-Vision comprises 7,680 × 4,320 pixels, which is 16 times the total number of pixels of an HDTV (High Definition Television) and the frame rate is 60 Hz with progressive scanning, and it has a 22.2 multi-channel sound system.

By using a Super Hi-Vision codec system which NHK have developed based on MPEG-2 and AVC/H.264 video coding standards, the Super Hi-Vision signal can be compressed into about 150 to 600 Mbps. Even though the signal is compressed, it is still difficult to send it in an existing broadcasting transmission path because it requires wide bandwidth. NHK is considering that the satellite broadcasting using the 21GHz band (21.4GHz to 22 GHz) is one of the promising ways to send the Super Hi-Vision to individual home. We have developed prototypes of a 300-MHz wide-band modulator and a demodulator, and carried out an indoor transmission experiment through the 21GHz band experimental transponder with a single carrier to verify performance of the hardware and evaluate the wide-band transmission characteristic. And we successfully transmitted the Super Hi-Vision signal through the air in the 21GHz band. This will be described in this paper.

SUPER HI-VISION SYSTEM

Super Hi-vision is the technology developed by NHK Science & Technical Research Laboratories that delivers the images so real that viewers feel as if they were actually at the site of the broadcast and find themselves attempting to touch what's on the screen. The image of the Super Hi-Vision is shown in Figure 1.

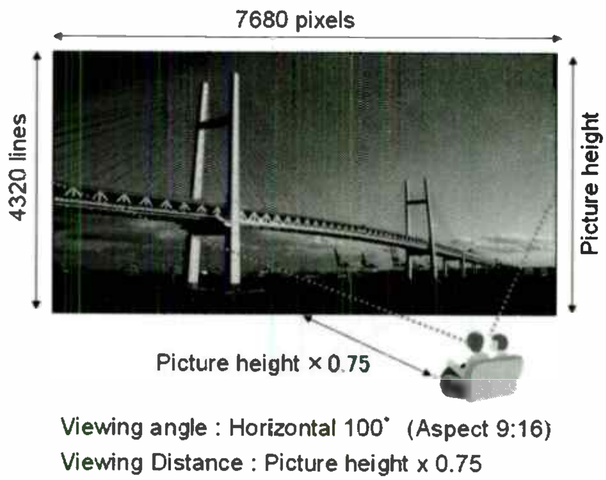


Figure 1: Super Hi-Vision

The Super Hi-vision system uses a video format with 7680 x 4320 pixels. It is designed for a viewing distance of 0.75 of a picture height, affording a horizontal viewing angle of more than 100 degrees in contrast to 30 degrees of the HDTV system whose ideal viewing distance is three picture heights from the screen. The individual scanning lines are not visually noticeable even when relatively close to the screen, reflecting the high resolution of the system. A wider viewing angle conveys a stronger sense of a reality.

This world's first video system with 4000 scanning lines that delivers ultra-clear, realistic three-dimensional images that can be achieved only by ultrahigh-definition technology. The large, wide-screen video images with the resolution equivalent to that of gravure printing strike viewers as a fresh surprise. On top of that, the new 3-D audio system with 24 loudspeakers dramatically enhances presence[2].

NHK has already developed cameras, projectors, disk recorders, and audio equipment for the Super Hi-Vision system. Several Super Hi-Vision programs have been produced with these pieces of equipment and were demonstrated in some exhibitions. An optical transmission system has been also developed. It can transmit uncompressed super hi-vision signal through an optical fiber hundreds kilo meters. NHK conducted a live relay involving optical transmission of an uncompressed Super Hi-Vision video signal and 22.2 multi-channel audio signals in 2005[3].

The data rate of an uncompressed Super Hi-Vision signal in the current system is approximately 24Gbps. In order to make the system suitable for practical use such as broadcasting services, NHK has developed several Super Hi-Vision codec systems based on MPEG-2 and AVC/H.264 video coding standards[4]. They can compress the Super Hi-Vision signals into about 150Mbps to 600Mbps. The new AVC/H.264 codec systems at 128 Mbps were demonstrated at NHK Science & Technical Research Laboratories open house in 2007.

21GHz SATELLITE BROADCASTING

The 21.4 to 22.0 GHz band is allocated for broadcasting satellite service (BSS) to Regions 1 and 3 (i.e. the European and Asian regions) by the International Telecommunication Union (ITU). Because of its wide bandwidth, the 21GHz band satellite broadcasting is expected to accommodate new broadcasting services like Super Hi-Vision broadcasting, stereoscopic television, and wideband download service. But there are some technical issues to be solved, for example, 21-GHz band suffers from about three times heavier rain attenuation in terms of decibels than the 12-GHz band. This means broadcasting services are interrupted much more frequently by rain attenuation in the 21-GHz band. NHK has proposed effective techniques of compensating for rain attenuation to provide reliable satellite broadcasting services in the 21-GHz band. The key technologies are a phased-array antenna to concentrate the RF power in a narrow area and a miniature TWT to realize a phased-array antenna system[5][6]. The concept of the 21GHz band satellite broadcasting is shown in Figure 2.

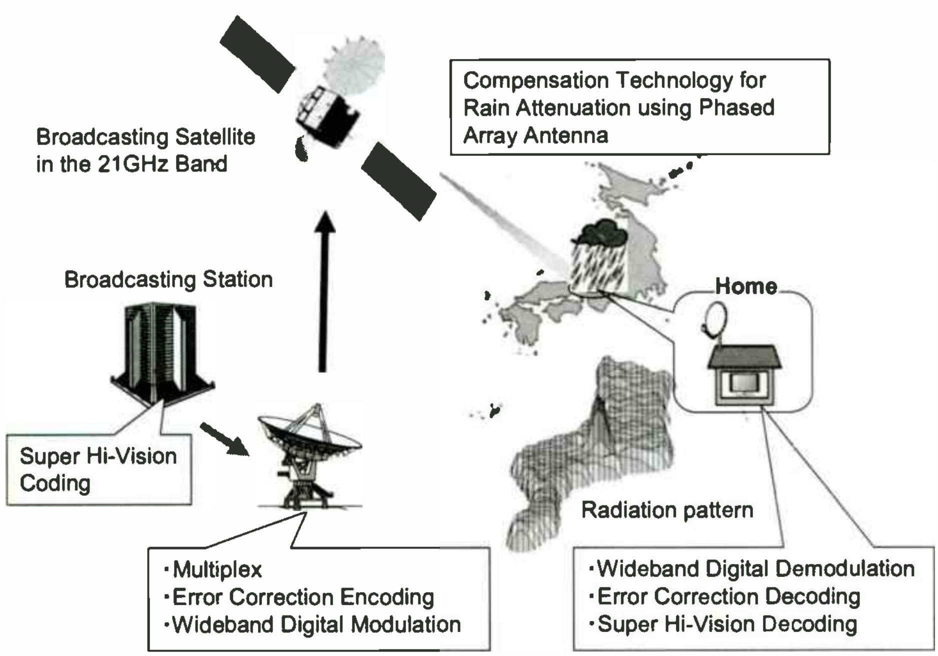


Figure 2: 21GHz Band Satellite Broadcasting

As the transmission capacity of a bandwidth totaling 600MHz in the 21GHz band is about 900Mbps in the QPSK modulation scheme with low density parity check code (LDPC code) at a coding rate of 7/8 for example[7], it will be able to accommodate multiple Super Hi-Vision programs by compressing the signals with the codec system. NHK is considering that 21GHz band satellite broadcasting is a promising medium for the Super Hi-Vision broadcasting service.

TRANSMISSION EXPERIMENT

Test system

To confirm the possibility of the Super Hi-Vision broadcasting via a 21GHz band satellite, we carried out a transmission experiment with a dummy satellite path indoors. The schematic diagram of the experiment is shown in Figure 3.

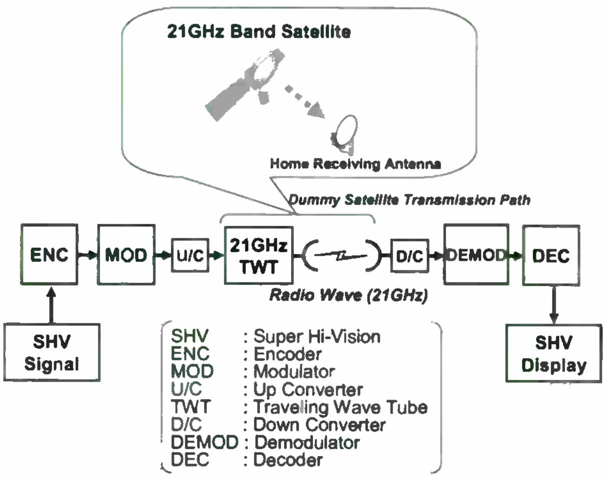


Figure 3: Schematic Diagram of the experiment

The Super Hi-Vision signal was compressed by an encoder and sent to a wideband modulator. The modulated signal was transmitted through a dummy satellite transmission path in the 21GHz band that simulated a broadcasting satellite, and received by a demodulator and then decoded. In the dummy satellite transmission path, a TWT (Traveling Wave Tube) was used as a function of a satellite transponder.

The decoded signal was displayed on a monitor. As the monitor that can display full 4000 scanning lines is still under development, the decoded Super Hi-Vision video image was down converted and displayed on a monitor with 2000 scanning lines.

Modulator and Demodulator

We have developed prototypes of a wide band (300-MHz bandwidth) modulator and a demodulator that have the capability of transmitting a compressed Super Hi-Vision signal. They have multiple settings of filter roll-off rates for transmission and reception to select from, to determine the optimum modulation parameters for a 21GHz-band satellite transmission. They can currently transmit maximum 500 Mbps data and make it feasible to transmit coded Super Hi-Vision signals for multiple programs simultaneously over a 21-GHz-band satellite transmission path. A picture of the prototypes of the modulator and the demodulator is shown in Figure 4. Both of them have the same outer frame and can be rack mounted. Functions are divided into some modules on boards. Those boards are inserted in the frame to realize the function.

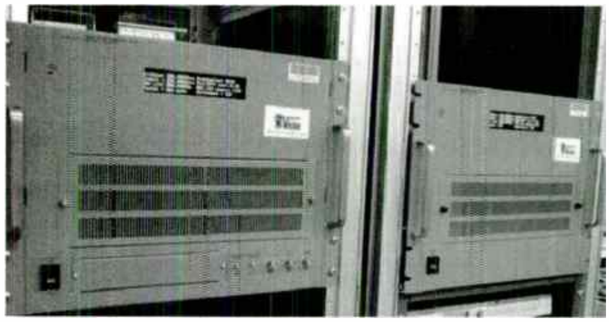


Figure 4: Modulator (Right) and Demodulator (Left)

The architecture of the prototype of the modulator is shown in Figure 5. Maximum 500Mbps digital serial data inputted to the modulator is mapped into QPSK scheme symbols, filtered by root raised cosine filter, converted to the analog signal, and directly modulated by a wideband quadrature modulator. The carrier signal (3GHz) is provided internally and can be provided externally. The roll off factor can be selected from 0.2, 0.35 and 0.5.

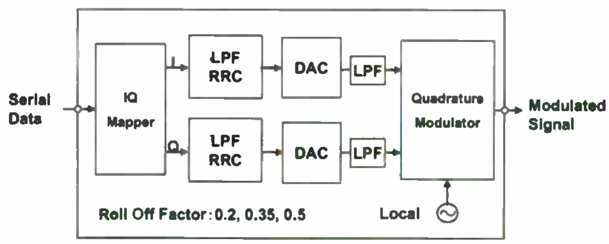


Figure 5: Wideband Modulator

The architecture of the prototype of the demodulator is shown in Figure 6. The modulated signal inputted to the demodulator is demodulated by a wideband quadrature demodulator module, and then derived I and Q symbols are converted to digital signals and digitally processed to compensate their phase errors. The demodulator can demodulate maximum 250Mbaud QPSK signal currently and 8PSK in the future.

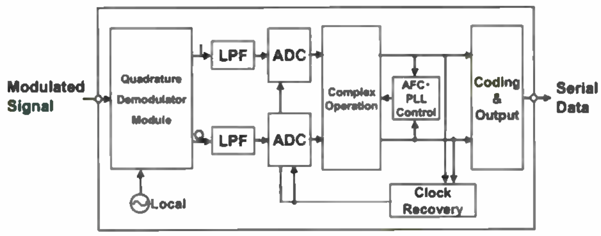


Figure 6: Wideband Demodulator

The prototypes of the modulator and the demodulator are designed to transmit and receive data to/from the WINDS (Wideband InterNetworking engineering test and Demonstration Satellite)[8] that is supposed to be launched in fiscal year 2007, and fitted to the interface of the earth station for WINDS.

Dummy Satellite Transmission Path

We built the dummy satellite transmission path in the 21GHz band indoors using a 21GHz TWT which NHK has developed [9]. It was a miniature TWT (mini-TWT) because it had been designed as the amplifier for an active array element in the 21GHz band. The TWT was used to evaluate the transmission performance via satellite, because the non linearity of the TWT is the main factor in affecting the transmission performance. Table 1 indicates the specification of the 21GHz band TWT. The output power and power efficiency as functions of input power for the TWT is indicated in Figure 7.

Table 1: 21GHz band TWT Specification

Center Frequency	21.7 GHz
Output Power	10.8 W
Gain	38.3 dB
Efficiency	48.1 %
Dimensions	15.3mm x 20mm x 300mm
Weight	270 g

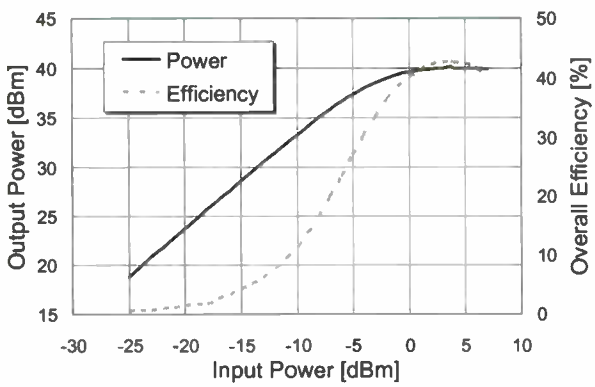


Figure 7: Output Power and Efficiency of TWT

The block diagram of the dummy satellite transmission path is shown in Figure 8. The modulated signal is up-converted to 21 GHz band and amplified by a miniature TWT which simulates a transponder in the broadcasting satellite. OBO (Output Back-Off) of the TWT is varied by a variable attenuator at the input of the TWT. The output of the TWT is connected to a 21GHz band pass filter. The filter is a 5-stage chebyshev filter and its 3dB bandwidth is 350MHz. The amplified and filtered signal is radiated from a horn antenna and after going through the air, it is received by a parabolic broadcasting antenna whose diameter is equal to a usual home satellite antenna, i.e. 45cm in diameter. The received signal is amplified, down-converted to IF frequency, and demodulated by the wideband demodulator. Finally the signal is decoded and displayed on a monitor. The picture of the TWT and the 21GHz BPF is shown in Figure 9.

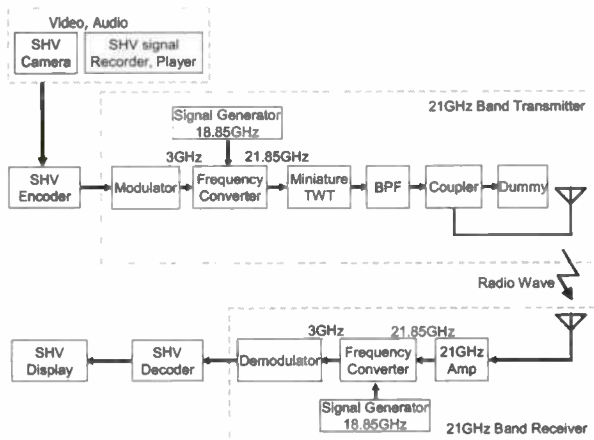


Figure 8: Block Diagram of dummy satellite transmission path

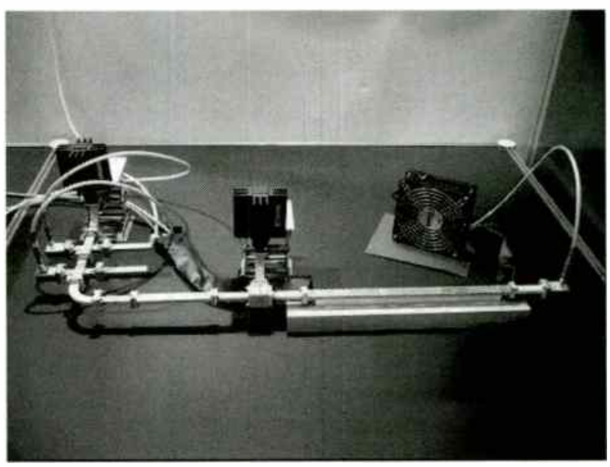


Figure 9: Miniature Traveling Wave Tube and Filter

A transmitting horn antenna, whose aperture size is 5.41cm x 3.96cm and antenna gain is 20dB, is attached on the top of a wall to shower the radio wave on the receiving parabolic antenna. The modulated signal is sent on the air and received by the parabolic antenna at a distance of about 3 meters. The pictures of the transmitting antenna and the receiving antenna are shown in Figure 10 and Figure 11 respectively.

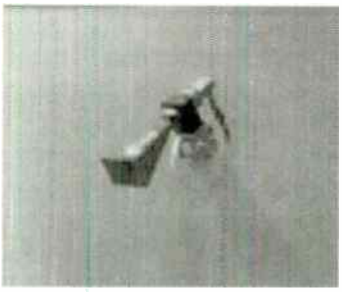


Figure 10: Transmitting Horn Antenna

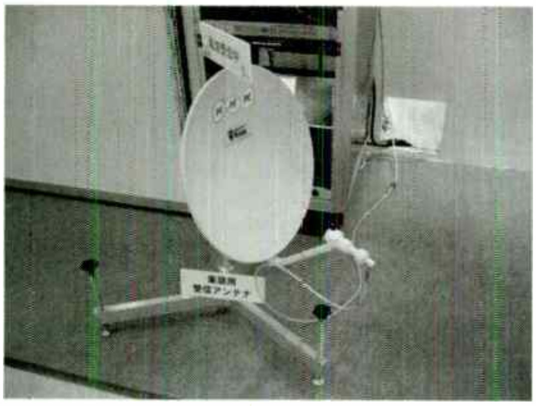


Figure 11: Receiving Parabolic Antenna

TEST RESULT

BER and Spectrum

To evaluate the modulator and the demodulator performance, BER (Bit Error Ratio) vs. C/N value was measured by adding AWGN (Additive White Gaussian Noise) from a noise generator on the signal in the 3GHz IF frequency band. The modulation scheme was QPSK and the symbol rate was 250Mbaud. As the modulator and the demodulator had not been equipped with FEC (Forward Error Correction) yet, the BER was measured without FEC between the modulator and the demodulator with the roll off factors varied. The BER curves were measured when the IF path was loopbacked and when the TWT was at saturation. The result is shown in Figure 12. As a reference, the theoretical curve of QPSK is also indicated on the graph. The implementation loss of the modem system was about 1 dB when IF path was loopbacked. When the signal was transmitted through the dummy satellite path, larger C/N value was required due to the distortion by the TWT. There was little difference in the BER curves with roll off factors when IF path was loopbacked, but there was deterioration especially with the roll off factor was 0.2 when the TWT was at saturation. In the low C/N value below C/N=6dB, the demodulator could not synchronize the carrier. We will improve the carrier recovery and symbol synchronization performance by modifying the modulation scheme like adding BPSK signal in front of the transmitting frame.

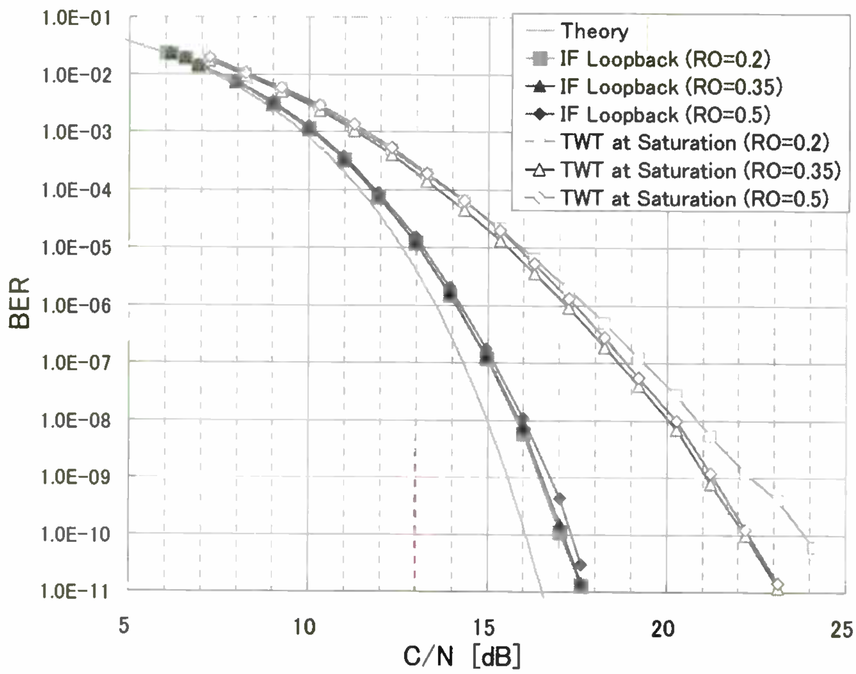


Figure12: BER result at 500Mbps data rate (QPSK)

HDMI AS TELEVISION APPLICATION PLATFORM FOR INTERACTIVE AND MORE

Rainer Zwing, Gerold Lübcke
Thomson, Hermann-Schwer-Strasse 3, 78048 VS-Villingen, Germany;

Scott Francis, Greg Herlein, Dan Kyte
Thomson/PRN, 600 Harrison Street, 4th floor, San Francisco, CA 94107, US

INTRODUCTION

HDMI has become the ubiquitous connection for CE device interconnection and in recent times also for PC products for video interconnection, replacing DVI Interface. To date, the CEC bus has been only lightly leveraged for system control. But new extensions can enable a wide range of television applications for professional system needs as well as interactive television. This system can be leveraged for unprecedented levels of control and interaction - giving ad hoc installations a degree of integration previously only available in professional installations. A case study of building, deploying and programming an interactive application will be presented along with possible additional future applications.

HDMI (High-Definition Multimedia Interface)

HDMI is a digital high speed multimedia interface which allows for video, audio and control information flow between media instruments.

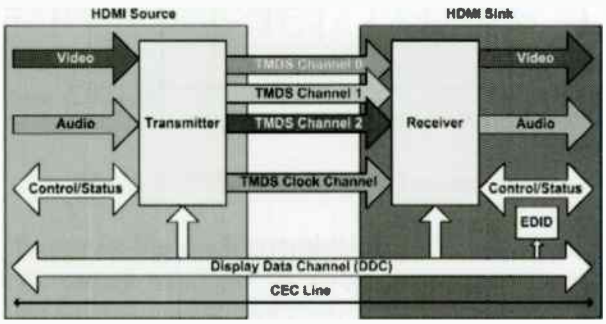


Figure 1

Figure 1 shows the basic elements of a single HDMI source and HDMI sink scenario. There are 3 differential pairs for video, audio and downstream information flow (red, green and blue lines). The forth differential line transmits the clock reference. The DDC (Display Data Channel) connection is used for upstream EDID (Extended Display Identification Data) data read in order to inform about sink capabilities and is based on I2C bus technology. In addition, in a content protection scenario for HDCP (High Definition Content Protection), the DDC line is used for keying and revocation interactions. Finally there is the one line serial bus CEC (Consumer Electronic Control) which permits bidirectional control data exchange between various devices in a CEC network.

CEC (Consumer Electronics Control)

CEC is a protocol that provides high-level control functions between all of the various audiovisual products in a user's environment. Although the CEC line specification was included in HDMI version 1.0 (Dec. 2002), the CEC feature set and Test Specification were not published until version 1.2a (Dec. 2005). Since then this control bus has slowly found its way into consumer products.

CEC commands can be segmented in three different groups:

Mandatory commands which must be implemented with CEC

Optional Public Commands to control devices

Optional Vendor specific commands, which allow a set of vendor-defined commands to be used between devices of that vendor or different vendors

The system described here is largely based on CEC functionality utilized for applications in the field of digital signage and more specifically for retail store infotainment. At a base level, this system uses a series of newly defined CEC commands which, when optionally implemented by manufacturers, allow for the control of televisions, audio receivers and other HDMI equipment when displayed in retail stores for the purposes of merchandising the equipment.

Digital Signage / Out of Home Advertising

Digital Signage and Out of Home Advertisement are largely growing markets with a rate of about 30% yearly between 2007 and 2011 over the entire industry. This includes mainly Advertisement, Systems, Displays and Software, where some of these will enjoy even higher growth rates.

According to Insight Reports in the US alone there are 2.7 million identified locations suitable for digital signage. This includes about 45,000 shopping malls about 800 airports, 1.12 million retail sites plus hospitals, and entertainment locations. Solutions will require integrated, low maintenance and software upgradeable technologies, which make use of general purpose hardware as much as possible. In addition to above mentioned locations Trade Shows and Corporate Communications installations will require intelligent systems that are easy to install and run.

Despite the wide spread usage of HDMI with its features for CE device interconnect that are ideal to configure systems for digital signage applications the application of HDMI in this area has been slow. In the following case study, a system is described which has been developed by PRN/Thomson for, but not limited to retail store communications including advertising.

Modern out of home advertisement starts with content aggregation, editing and scheduling. Distribution of the content over wide areas is frequently accomplished with satellite networks –Figure 2.

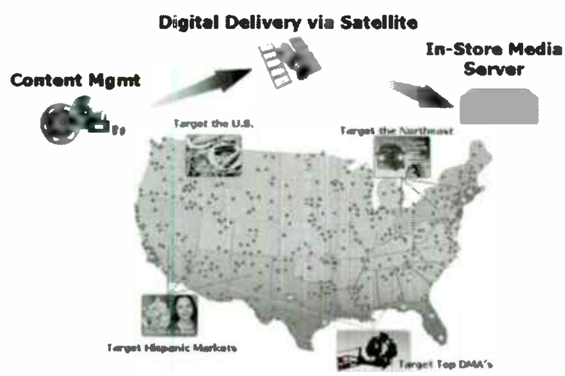


Figure 2

Within each store a media sever, running specific data management software, receives the media streams. From this central point content is dispatched within the store environment for various applications - Figure 3:

Figure 3

The topology of such in-store networks is often built on LAN networks combined with RF or analog base band (e.g. CVBS) for screen interconnections. This topology requires human monitoring to ensure the compliance of the system which also results in high maintenance costs as well as high mean time to repair values. Examples of some common problems are shown below - see Figure 4

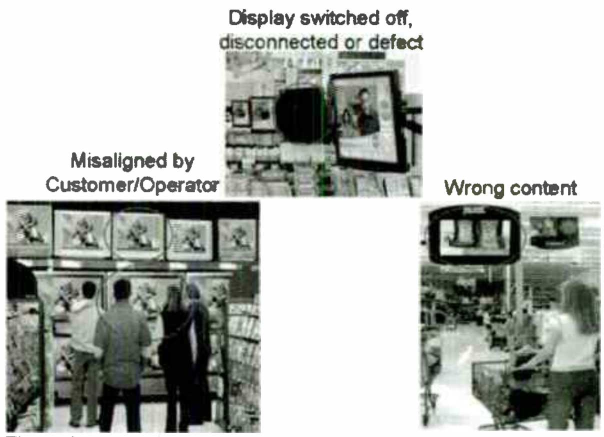


Figure 4

'The only compliant TV is a dead TV'

When televisions (often used as monitors) are installed for communications purposes such as digital signage, mood setting or other professional AV needs, as well as when on display in retail locations, it is not only desirable for the units to maintain certain settings and characteristics, it is critical. The degree to which a given unit or collection maintains this configuration is called compliance. Since these installations often leave the displays exposed to casual public action, this compliance is frequently compromised. For example, at the recent Consumer Electronics show a prankster turned off televisions during product demonstrations [http://gizmodo.com/343348/confessions-the-meanest-thing-gizmodo-did-at-ces]. When televisions are presented at retail for sale, the public is actually encouraged to "play" with the settings while making purchase decisions. This can lead to a very fragmented and unattractive merchandising display.

The only TV whose settings can be relied upon to be consistent is one that is broken; to paraphrase Clint Eastwood – the only compliant TV is a dead TV.

Additionally equipment on display for merchandising is subject to customer return. Common reasons for return of television and home theater equipment include

- Incompatibility with other equipment
- Failure to subscribe to HD service or use of otherwise inappropriate content
- Differences in experience between the in-store media presentation used to make purchasing decisions and the content feed provided in the home.

This can be overcome with customer education interactive screening of compatible devices, and simulation of various types of content feeds - e.g. DVD, DSB, and CATV. The product identification features provided by HDMI (EDID) provide the basic data needed to allow a system to present relevant information to a shopper. The unique media targeting abilities of the system described here provide useful feedback and comparison media to the shopper.

Of course several of those issues described above can be circumvented by using dedicated solutions with specific integrated TV screens or monitors, for digital signage applications. The usual solution is to use commercial displays and other specialized equipment - all of which command premium price points. However, this is not possible for applications like video walls where the retail product itself is part of the network.

Universal Control for all Vendors

It is possible to leverage the HDMI/CEC interface with a series of universal vendor commands to provide

- Power control
- Audio management
- Image management
- Lockout/Remap of buttons/features

With support from manufacturers for these commands, it is possible to construct installations where compliance can be controlled. This is particularly useful for retail installations.

By combining a LAN network for IPTV content management with HDMI for screen interconnections, professional quality controlled installations can be made using standard TV displays or monitors which incorporate the HDMI interface. The only required modifications at this point are incorporation of the CEC commands which can be utilized in the home environment as well. The required commands are described in Reference [4]. Implementation of which requires only software modifications for those vendors already using CEC.

Figure 5

Figure 5 shows a basic "singleton" topology: the content received by the server on a first level is transmitted over LAN using IPTV streaming protocols including control and topology monitoring commands. An IPTV set-top box is connected to each display via an HDMI cable.

By utilizing the universal vendor commands over CEC the network server can monitor, control and target each individual display where the IPTV set top box translates the LAN protocol into CEC commands and vice versa.

An alternative system set-up is shown in Figure 6 using a clustered approach grouping several displays together which are fed by a specific HDMI distributor and CEC switch-box.

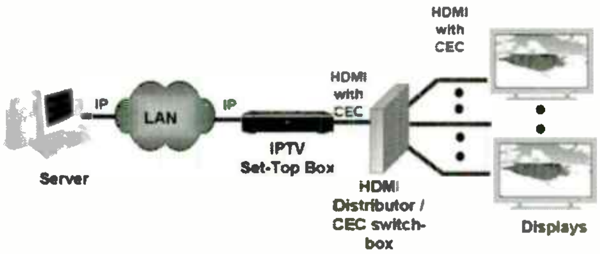


Figure 6

Compared with a singleton approach this solution can drive several displays with a single IPTV set-top box but requires compatible distributors. Units that combine both IPTV client and distribution in one box are also possible. The final and optimized topology decision depends on the local requirements and the concepts allows for various kinds of customization.

Television ID (ID)

HDMI EDID information provides the ability to uniquely identify each television in the system. Using this information, it is possible to provide reports on exactly what products are currently configured in a dynamic installation where the units may change. In addition, if the desired televisions are not connected, it is possible to provide alarms to operations personnel to rapidly correct the problem.

Television ID uses the standard HDMI DDC line to carry Extended Display Identification Data (EDID). This technology is already present on all televisions. To read this data from each TV requires an HDMI switch that supports the DDC line or one compatible video player/set top box per screen.

Television Control

If all televisions in an installation implement this universal command interface, individual screens or groups of screens can be powered on or off to accomplish significant power savings. Screens can be restored to default settings for contrast, brightness, sharpness, etc. This allows consumers to play with the controls yet ensures that all screens are returned to their optimum settings automatically.

Television Control uses the HDMI CEC line to carry control commands to the television. This technology is already present on many televisions. To affect control over each TV requires an HDMI switch that supports the CEC line or one compatible video player/set top box per screen.

Televisions that support the proposed universal vendor extension to the CEC specification can also obtain additional benefits. Such televisions can be placed into a mode of operation where their panel buttons are dynamically re-mapped to other functions such as triggering and operating a VoD menu. Other extensions include the ability for the television manufacturer to predefine operational profiles that are best for certain display scenarios (large retailer, brightly lit, etc) that can be easily triggered. A great example of this would be a display mode that shows on screen – perhaps split screen – the special image processing features that make that individual television model unique and more valuable than others.

Unique Media

With each television in an installation identified and uniquely controllable and addressable, it is possible to provide specific media to unique displays. Every television would be identified and controlled as previously described. However, unique and individual media could go to each screen. This allows entirely new programming to be deployed to the TV wall – such as the ability to support a dynamic shopper assistant that asks questions about customer preferences and switches televisions off that do not meet the customers' needs. Other applications would be the ability to play television manufacturer specific media only on those manufacturers screens, and the ability to provide detailed per television information about features and photos of the back panel connections on each television.

Implementation Requirements

To achieve above functionality following system requirements are envisioned.

System Overview

Video Network Manager (VNM) software suites will implement a Device Group Control Protocol (DGCP) controller and this will be the only controller in the system.

A block diagram of the system is shown in Figure 7.

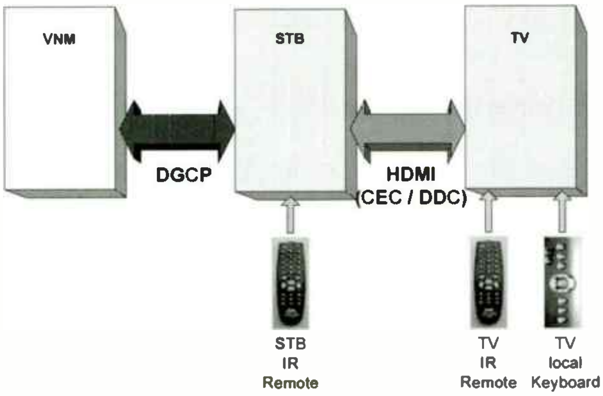


Figure 7

The DGCP controller will broker all group membership, GET and SET operations, and NOTIFICATION messages from DGCP clients. In the above application, the DGCP client will be implemented in the Set Top Box (STB). The STB provides the hardware interface to the HDMI CEC and DDC lines and the DGCP client in the set top box provides the network interface to these hardware lines.

The CEC is the line across which universal vendor commands are sent. Other CEC commands are also available across this channel as defined in the HDMI specifications. Some of these are used for Infra-Red (IR) code carriage (see IR receiver support) and other functions unique to particular hardware – vendor specific.

The DDC is the line across which EDID data is read. This line is standard technology on all current televisions and screens.

Infra-Red (IR) remote key codes for either the Television or the STB can be made available on the network via DGCP as well. In the case of the TV the codes are moved across the CEC line as vendor specific codes and translated into appropriate DGCP messages. STB IR codes are intercepted and translated directly to DGCP messages.

Set Top Box Requirements

The Set Top Box (STB) must meet the requirements as defined in the [5] IPTV Set Top Box Technical Requirements and also meet the additional requirements as defined in this section.

The STB must provide the DDC line and CEC line of the HDMI connector per the [1] HDMI Specification ver. 1.3a. DDC uses standard I2C technology and pose no technical barrier to implementation. The hardware must provide a means to implement the CEC protocol including the bit stream and timing.

The STB may optionally support the "IR Bypass" feature. This feature provides for a mode of operation where the IR codes arriving from a remote control are detected but not processed locally. These codes are packaged and sent as DGCP messages to the VNM. The VNM processes these messages and chooses actions to take for control of the STB (and possibly other components) and sends DGCP messages to the STB. The simplest case would be that IR codes be detected at the STB, packaged over DGCP to the VNM, sent back to the STB via DGCP and processed normally - Figure 8.

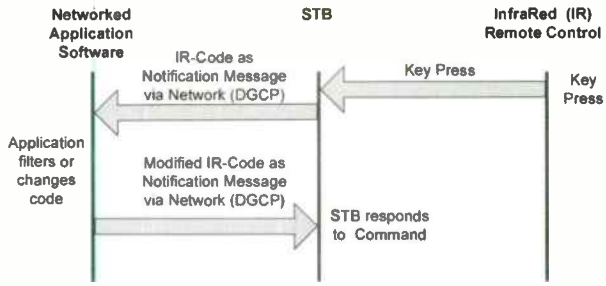


Figure 8

The Set Top Box (STB) will implement a DGCP client per the [2] Device Group Control Protocol together with the vendor specific extensions described in Reference [3] and Reference [4].

Television Requirements

The Television must provide the DDC line and CEC line of the HDMI connector per the [1] HDMI Specification ver. 1.3a. DDC uses standard I2C technology and pose no technical barrier to implementation. The hardware must provide a means to implement the CEC protocol including the bit stream and timing.

The Television may optionally support the "IR Bypass" feature. This feature provides for a mode of operation where the IR codes arriving from a remote control are detected but not processed locally. These codes are packaged and sent over HDMI messages using CEC. The STB would package these messages in DGCP and send them over the network to the VNM. The VNM processes these messages and chooses actions to take for control of the TV (and possibly other components) and sends DGCP messages to the STB. The STB would forward those messages over the CEC to the TV. The simplest case would be that IR codes be detected at the TV, forwarded to the STB, packaged over

DGCP to the VNM, sent back to the STB via DGCP, forwarded on to the TV and processed normally - Figure 9.

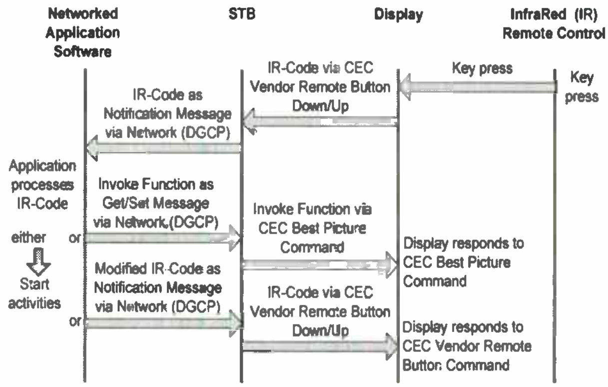


Figure 9

The Television should support the "Local Keyboard" feature. This feature provides for a mode of operation where the local keyboard activities are detected but not processed locally (Keyboard locked mode). Key presses are packaged and sent as CEC messages over HDMI. The STB would package these messages in DGCP and send them over the network to the VNM. The VNM processes these messages and chooses actions to take for control of the TV (and possibly other components) and sends DGCP messages to the STB. The STB would respond or forward those messages over CEC to the TV - Figure 10.

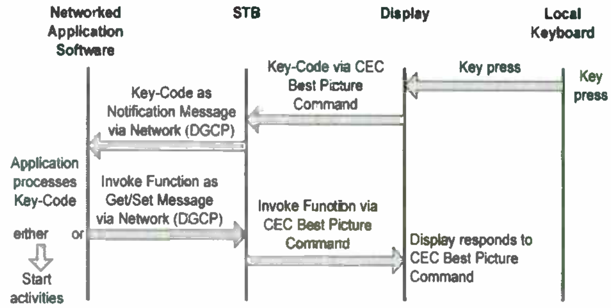


Figure 10

The TV must implement a CEC client in accordance with Reference [4].

In order to validate such functions and requirements of the different components and to allow for development at TV/STB manufacturer level, PRN provides development kits which emulate sink or source behavior.

Summary

The described set of CEC Commands, group control methods and IR Remote bypass functionality enable a new level of system control and interactivity. The changes required for televisions that currently implement HDMI and some form of CEC are minimal. The use of these controls in ad hoc installations such as retail television merchandising "TV Walls," can provide consistent imaging with a high degree of manageability.

References

[1] HDMI Specification ver. 1.3a, www.hdmi.org

[2] Device Group Control Protocol - PRN, 600 Harrison Street, 4th floor, San Francisco, CA 94107, US

[3] DGCP Payload Proposal - PRN, 600 Harrison Street, 4th floor, San Francisco, CA 94107, US

[4] HDMI based concept proposal - PRN, 600 Harrison Street, 4th floor, San Francisco, CA 94107, US

[5] IPTV Set Top Box Technical Requirements - PRN, 600 Harrison Street, 4th floor, San Francisco, CA 94107, US

11.88 GB/S SDI
CONTINUING THE EVOLUTION OF SDI

GARETH M. HEYWOOD & RYAN LATCHMAN
Connectivity Technologies, Gennum Corporation
Burlington, Ontario, Canada

ABSTRACT

With the advent of large-scale deployments of high definition television, and future migration to higher bandwidth 1080p50/60 transmission, the video industry will need to reduce the number of physical links, both electrical and optical, between facilities, equipment racks, and outside broadcast vehicles. This requirement will extend within large pieces of equipment, such as serial video routers, to reduce the size and complexity of high-speed interconnect.

By combining multiple HD signals, up to 1080p50/60, into a single optical fibre or coaxial cable link, the cost of installations can be significantly reduced, reducing the number of fibre runs, and providing a more efficient use of cabling resources.

In addition, the definition of higher bandwidth video formats, such as Ultra High Definition Television (UHDTV), requires interface capacities for the carriage of such formats to scale accordingly. UHDTV image formats require interface capacities ranging from 7.5 Gb/s, up to 72 Gb/s. Using current HD-SDI interfaces, operating at 1.485 Gb/s, UHDTV requires multiple links, from 8 to more than 48, depending on the image format and sampling structure. This is a costly and technically challenging solution. The real estate required for all the HD-SDI connectors is considerable, leading to increased system costs, and the cost and complexity of cabling is significant.

By utilizing a serial digital interface operating at 11.88 Gb/s, one can define the video data mapping protocol, the carriage of ancillary data, coding and, physical interface characteristics.

THE VIRTUAL INTERFACE

The SMPTE 425M[1] standard introduced the concept of the virtual interface. In the parallel data domain, the virtual interface provides a 20-bit data multiplex operating at 148.5 or 148.5/1.001 MHz clock frequency. The virtual interface shown in Figure 1, is multiplexed and serialized as per SMPTE 424M[2], to create the Serial Digital Interface (SDI) operating at 3 Gb/s (nominal).

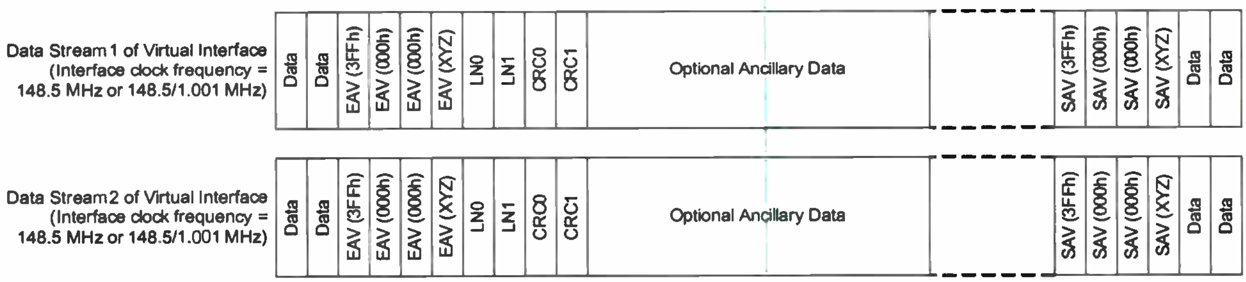


Figure 1: SMPTE 425M Virtual Interface

The virtual interface can be used to carry a number of video formats with various sampling structures and frame rate options, fully defined in SMPTE 425M. Once an image format exceeds the bandwidth of this 20-bit interface, the format must be carried on a serial interface capable of exceeding the current SDI maximum of 3 Gb/s.

The proposed 11.88 Gb/s serial digital interface is simply a 4x multiple of 3 Gb/s SDI (or an 8x multiple of HD-SDI, which operates at 1.485 Gb/s). In order to define a mapping structure for an 11.88 Gb/s interface, the SMPTE 425M virtual interface can be expanded by four, as shown in Figure 2.

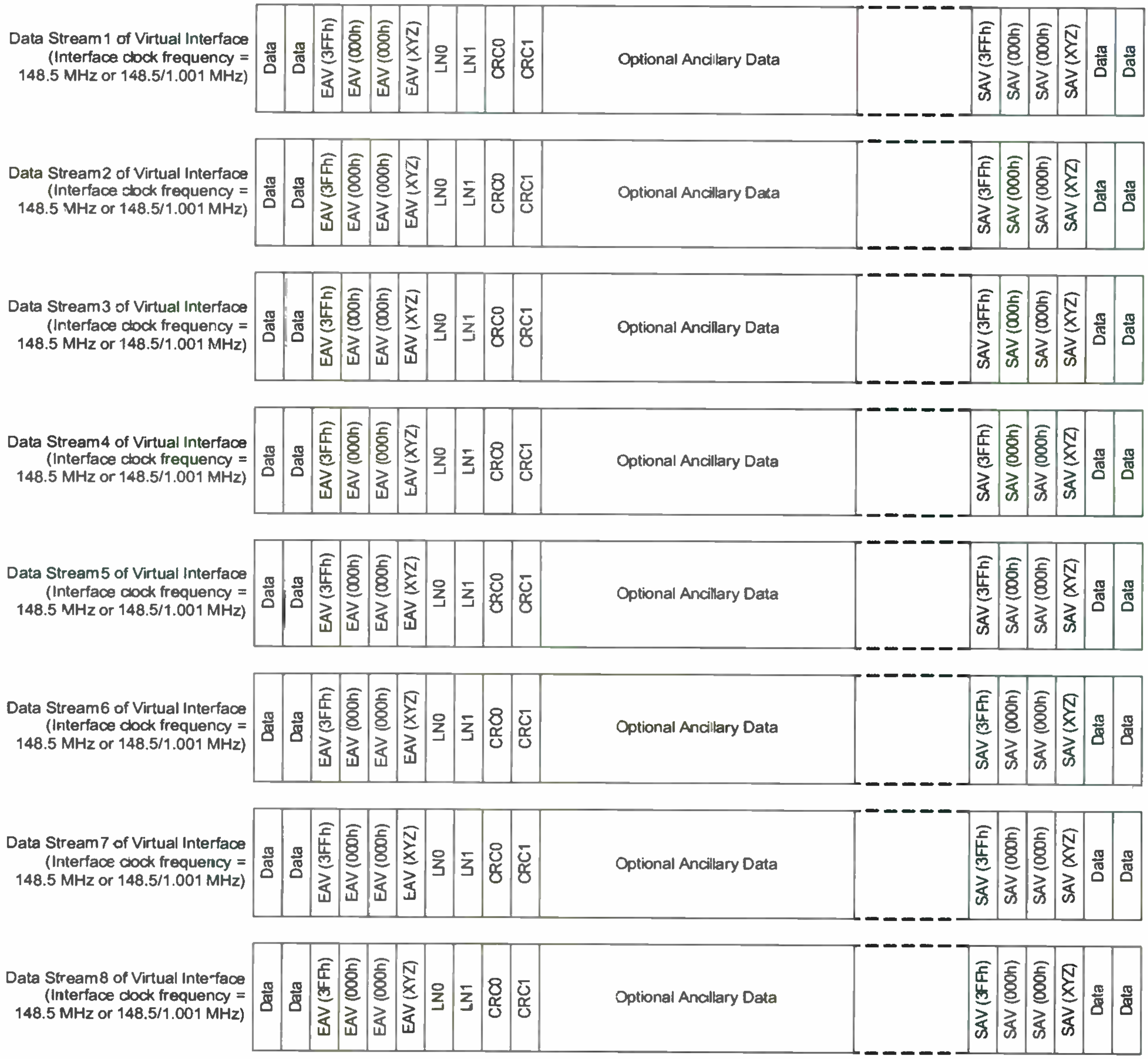


Figure 2: 11.88 Gb/s Virtual Interface

The most obvious application of the above data mapping is for the carriage of four 1080p50/60 10-bit 4:2:2 video streams. Each 1080p stream requires two data streams of the virtual interface.

The multiple data streams, or virtual interface, in Figure 2 above can also be considered as a container, capable of carrying a plethora of data. Later sections shall describe multiple applications for the 11.88 Gb/s SDI container.

By making the 11.88 Gb/s interface an integer multiple of existing SDI standards, advantage can be taken of existing clock and timing generation circuits and systems. It also reduces the need for additional buffering required for clock interchange between non-related data rates.

Another benefit is the existing knowledge-base within the broadcast design community, which is very familiar with established standards and industry practices. By leveraging the 11.88 Gb/s standard on existing standards, re-use of proven design resources and IP is possible, reducing development time and providing a time-to-market advantage.

DATA MULTIPLEXES

Each pair of data streams in the 11.88 Gb/s SDI virtual interface is further byte interleaved to create the same 10-bit multiplex that is serialized to create a 3 Gb/s SDI stream, according to SMPTE 424M. This 10-bit multiplex has an interface clock rate of 297 MHz or 297/1.001 MHz, shown in Figure 3.

Figure 3: SMPTE 424M 10-bit Multiplex

In order to create a 10-bit multiplex, which when serialized, will produce an 11.88 Gb/s serial data stream, each 10-bit multiplex shown in Figure 3 must be further byte interleaved, as shown in Figure 4.

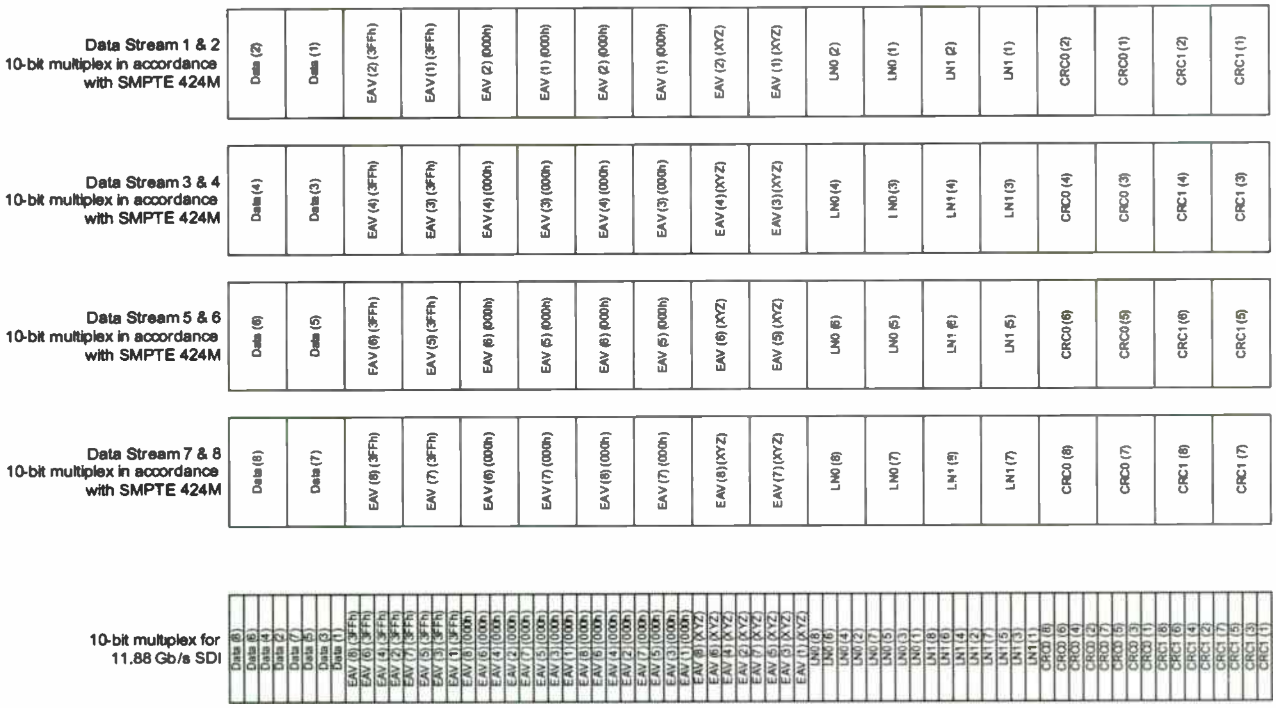


Figure 4: 10-bit Multiplex for 11.88 Gb/s SDI

SERIALIZATION TO 11.88 GB/S

In all SDI systems, operating from 270 Mb/s up to 3 Gb/s, the 10-bit parallel data words are serialized, least significant bit (LSB) first. The data is then scrambled using the linear-feedback shift-register polynomial in Equation 1:

Equation 1: $G_1(X) = X^9 + X^4 + 1$

The data is then converted from NRZ to NRZI form using the following polynomial:

Equation 2: $G_2(X) = X + 1$

The 10-bit multiplex shown in Figure 4 is serialized and scrambled according to the above equations, resulting in an 11.88 Gb/s serial data stream. The benefit of using the same scrambler, as used in previous generations of SDI, is that it does not carry any data overhead. In very high data rate systems, overhead caused by inefficient encoding schemes, such as 8B/10B encoding, can be significant. Encoding schemes requiring overhead can result in either an increase in the serial data rate by up to 20%, or part of the input data being discarded to reduce the input bandwidth by up to 20%.

Control of Pathological Conditions

As is well known, there are certain 10-bit video signals that, when scrambled using the polynomial in Equation 1, can cause repetitive pathological patterns containing long run lengths of 1's or 0's in the serial data stream.

These patterns are a consequence of repeating input patterns interacting with a specific initial state in the scrambler. The 1992 SMPTE Journal paper by Takeo Eguchi[3] provides an excellent description of the phenomenon, while SMPTE EG34[4] provides further background and system considerations.

Two commonly used checkfields are used to produce "equalizer (EQ) pathological" and "phase-locked loop (PLL) pathological". The EQ checkfield is produced by a 4:2:2 sampled input with the alternating chroma and luma values, 300h / 198h / 300h / 198h / etc. The PLL checkfield can be produced by the following chroma and luma values: 200h / 110h / 200h / 110h / etc.

Taking four inputs of full screen EQ checkfield or PLL checkfield at 1080p50/60, it can be seen in Figure 5 and Figure 6 that the 11.88 Gb/s 10-bit multiplex results in a data pattern, which when scrambled according to the polynomial in Equation 1, will not produce pathological conditions.

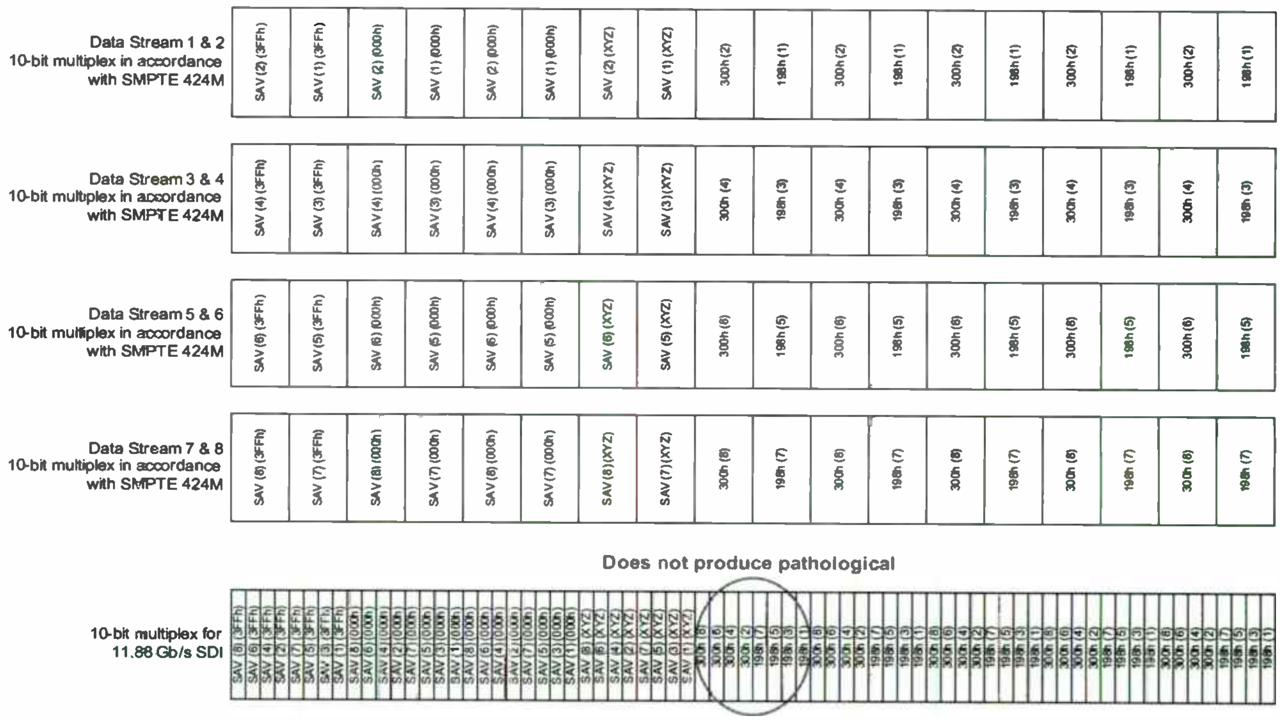


Figure 5: 11.88 Gb/s 10-bit Multiplex with EQ Pathological Checkfield

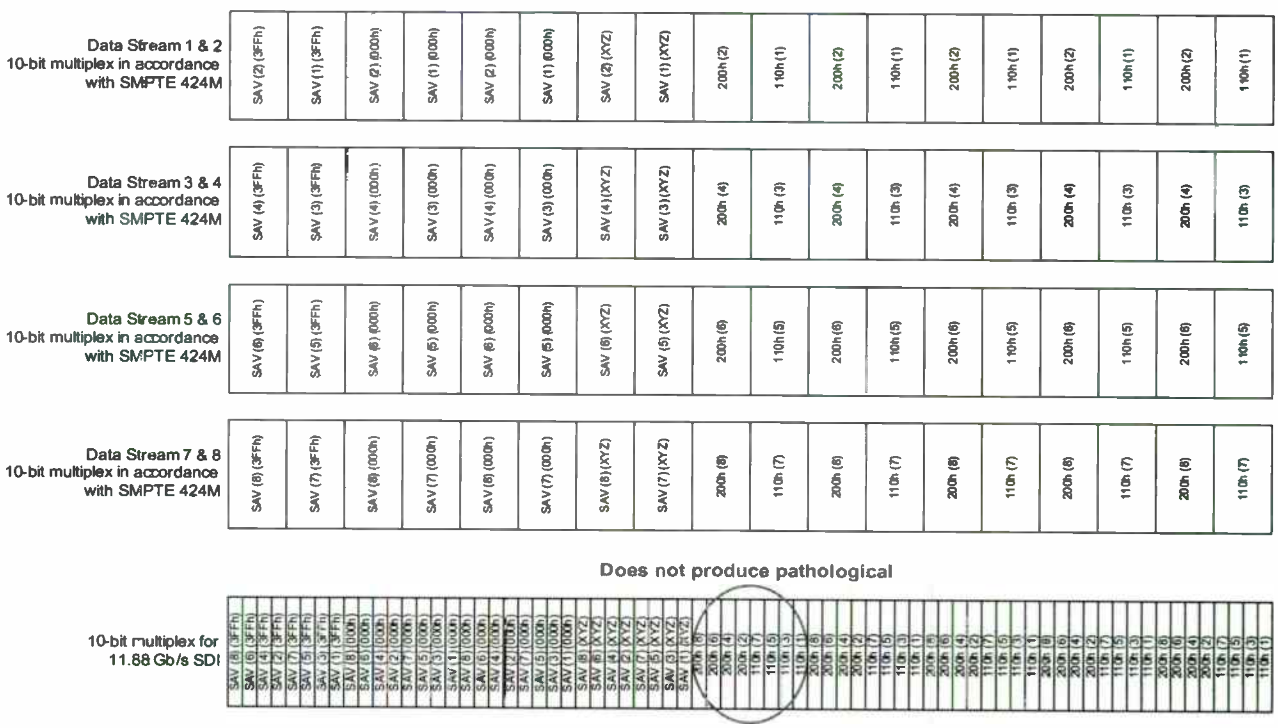


Figure 6: 11.88 Gb/s 10-bit Multiplex with PLL Pathological Checkfield

To further reduce the possibility of input video signals, particularly pathological-causing checkfields, from producing pathological conditions at 11.88 Gb/s, the 10-bit multiplex is modified such that some data words are "flipped", MSB to LSB, prior to scrambling. This results in a repeated pattern of four data words feeding the scrambler LSB first, followed by four data words MSB first. The 11.88 Gb/s SDI 10-bit multiplex with word flipping applied is shown in Figure 7.

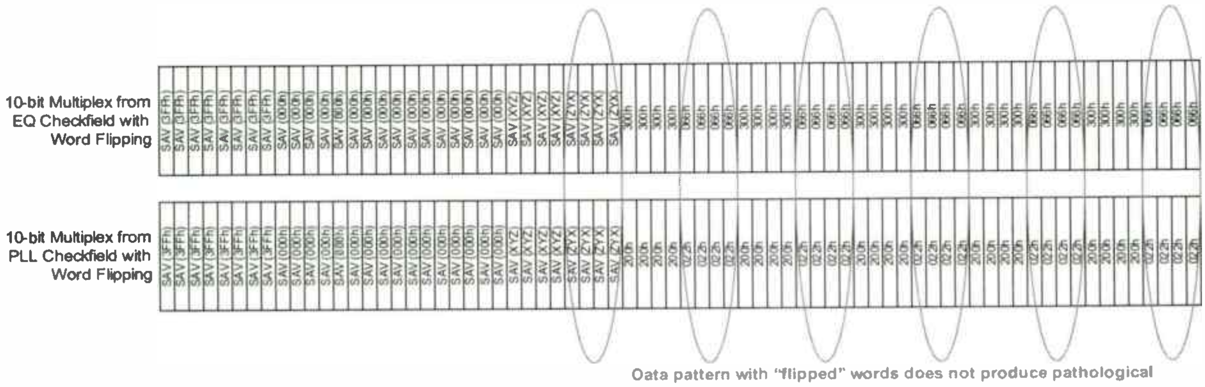


Figure 7: 11.88 Gb/s 10-bit Multiplex with Word Flipping

The word flipping process only reduces the possibility of repetitive pathological conditions caused by active picture checkfields from occurring. The preamble of the timing reference signals (TRS), with data words 3FFh and 000h, in the 10-bit multiplex for 11.88 Gb/s SDI are not affected by the word flipping process. The concatenated 000h data word values, when serialized, result in a long run of zeros feeding the scrambling polynomial.

Another phenomenon described in detail in Eguchi's 1992 paper, notes that with the scrambler seeded appropriately, there is always an input pattern that can clear its registers within ten serial clock cycles. If all scrambler registers are loaded with zero, then a zero is needed at the input to load another zero into the scrambler. Therefore, to determine the longest possible run of zeros out of the scrambler, it is only necessary to determine the longest sequence of zeros in the serialized video data at the scrambler input.

As can be seen in Figure 8, the concatenated TRS preambles result in 16 consecutive words of 000h. When serialized, this will result in 160 zeros at the input of the scrambler. If the scrambler registers are all at zero, this results in an additional 10 zero bits. Also, after the TRS preamble, a valid XYZ word can be 200h, which when serialized LSB first, results in another 9 zeros. Therefore, the maximum run of zeros output from the scrambling polynomial in Equation 1 is 179 (160 + 10 + 9). After NRZI is applied, this results in a long run of 179 zeros or ones in the 11.88 Gb/s serialized data stream.

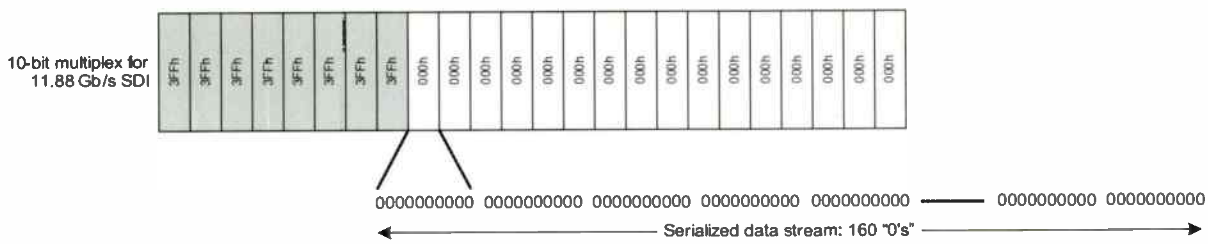


Figure 8: Concatenated TRS Preambles

These long runs, while infrequent, can cause unwanted DC shifts on the serialized link, or may look like a loss of signal to the receiving equipment. To prevent long runs of zeros and ones, a *"sync-bit"* is inserted into the two LSB's of the serialized data stream, prior to the scrambler, as shown in Figure 9.

The two LSB's are changed from 11b or 00b to 01b or 10b, the *"sync-bit"* symbols. This results in two possible values each for the 3FFh and 000h words of the TRS preambles. Since 3FFh and 000h can also occur in the ancillary data flag (ADF) preamble, in the combination of 000h / 3FFh / 3FFh, these data words will also be subjected to sync-bit insertion.

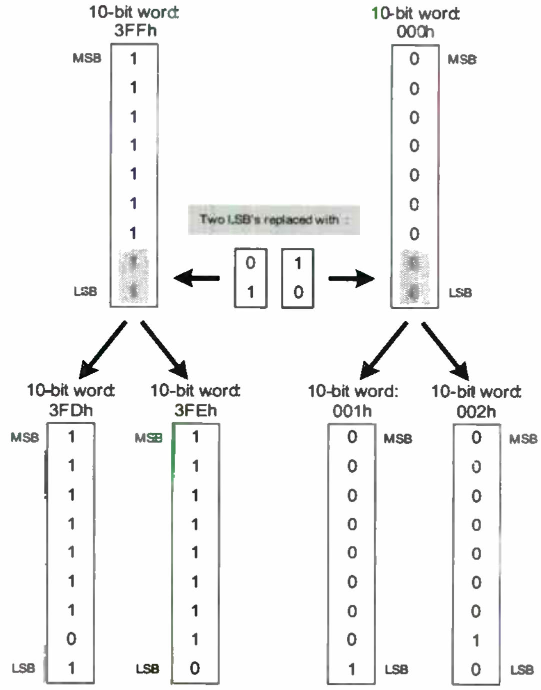


Figure 9: Sync-bit Insertion

The parallel data stream is modified such that sync-bit values are inserted alternatively, in the order of 01h followed by 10h for each data word, as shown in Figure 10. Once the sync-bit insertion has taken place, the serialized data stream feeding the scrambler, LSB first, will only contain a maximum run of 10 ones or zeros during the TRS preamble (or ADF preamble). After scrambling, the maximum run of ones or zeros possible is 29.

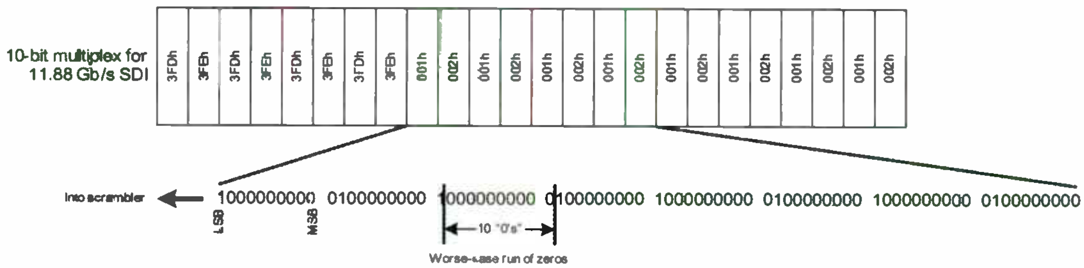


Figure 10: TRS Preamble with Sync-bit Insertion

Sync-bit insertion is only applied to the 3FFh and 000h data words, which uniquely occur in the TRS and ADF preambles. The modified preamble values, 3FDh, 3FEh, 001h and 002h, are still illegal video code words, therefore, they cannot appear within the active video data stream. These data values are still unique enough such that data stream synchronization using the TRS is still possible. Alternatively, TRS and ADF detect blocks need only look at the upper 8 bits of the 10-bit data words, which remain unchanged, in order to synchronize to the data streams.

By employing both word flipping and sync-bit insertion, the 11.88 Gb/s SDI can be realized with zero data overhead, while also avoiding known pathological conditions.

CARRIAGE OF 8 HD-SDI STREAMS

Each of the eight data streams in the 11.88 Gb/s SDI virtual interface may be used to carry a HD-SDI stream. To preserve the original timing reference signals of each HD-SDI stream, each data stream of the virtual interface is modified to be compatible with the 10-bit multiplex required for SMPTE 292[5], as shown in Figure 11. This will increase the number of embedded TRS in the 11.88 Gb/s SDI 10-bit multiplex by two, as shown in Figure 12.

Figure 11: SMPTE 292 10-bit Multiplex

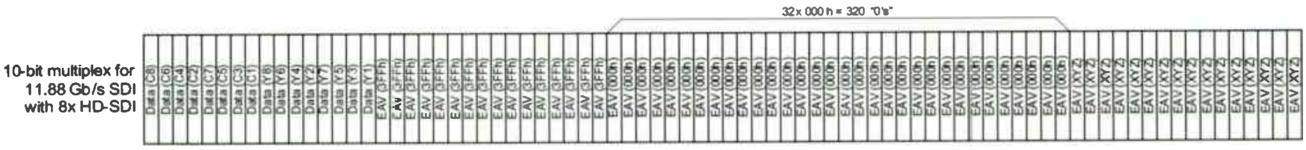


Figure 12: 11.88 Gb/s 10-bit Multiplex with 8x HD-SDI

The multiplex shown in Figure 12 will be subject to sync-bit insertion, as described in the previous section. This will reduce the maximum run-length of zeros feeding the scrambler from 339 to just 29.

It should be noted that all 8 HD-SDI streams must be synchronized at the TRS level before being multiplexed and serialized at 11.88 Gb/s. A typical application for this is trunking of HD signals between equipment or facilities. In the majority of broadcast installations, all HD video sources are synchronized to a common timing reference, usually studio sync. Therefore, within a routing system, 8 synchronized HD streams can be easily concatenated into a single 11.88 Gb/s SDI signal.

CARRIAGE OF 1080P 50/60 AT 12-BIT 4:4:4 SAMPLING

The 3 Gb/s SDI standard was primarily defined for the carriage of 1080p50/60 4:2:2 video sampled at 10-bit resolution. In applications such as digital cinema, telecine and high-end video production, there is a growing need to produce content at higher bit depths and at 4:4:4 sampling in both R'G'B' and Y'CbCr colour spaces.

In Figure 13, the source sampled data stream for 1080p50/60 at 12-bit R'G'B' 4:4:4 sampling is shown mapped to four data streams of the virtual interface.

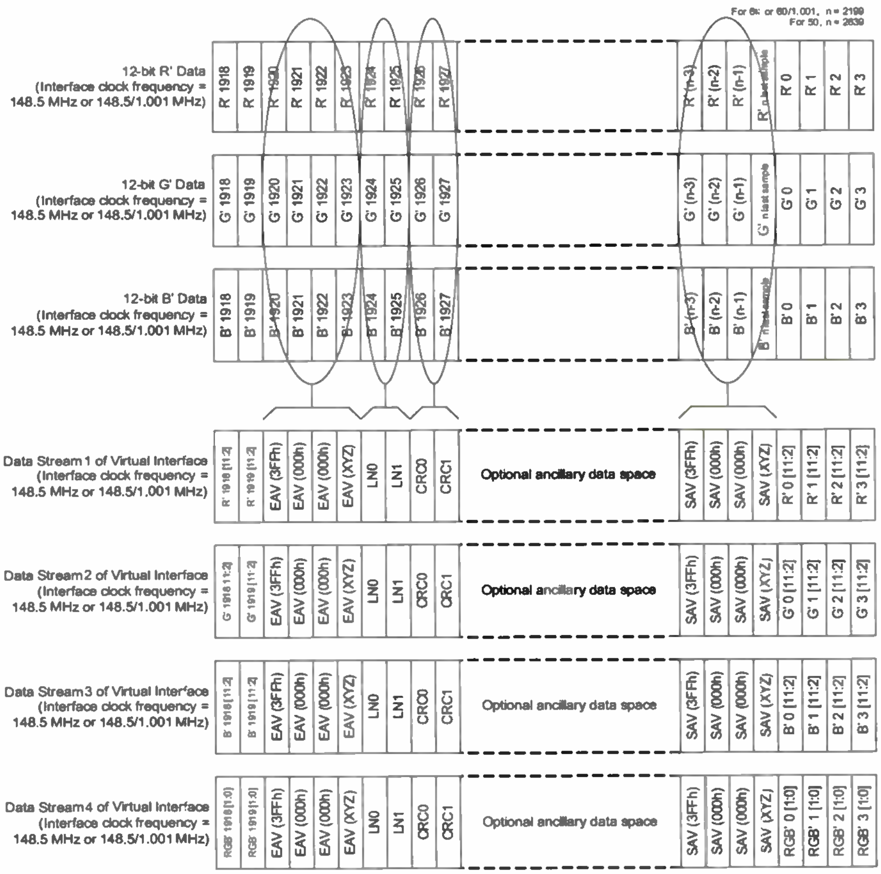


Figure 13: Virtual Interface Mapping for 1080p50/60 12-bit 4:4:4

Since only four of the 8 total data streams of the 11.88 Gb/s SDI virtual interface are used, a second source with the same sampling structure can also be transmitted over the 11.88 Gb/s link. If only one source is required, data streams 5 through 8 can be simply *"padded"* with a black signal, and ignored at the receive-end.

NEXT GENERATION HIGH DEFINITION TELEVISION

With the advent of larger resolution displays and the increasing efficiency of video compression schemes, the ability to *"go beyond"* the current HD resolutions will likely be the next step for the television broadcast community. Already, the Japanese national broadcaster, NHK, has announced plans for Ultra High Definition Television (UHDTV), based on two possible resolutions: 7680 x 4320 and 3840 x 2160. Both these image formats are direct integer multiples of the 1920 x 1080 HD format that is so prevalent today. UHDTV also defines progressive scan only at both 50 and 60 Hz frame rates.

This *"natural extension"* or evolution of HDTV aligns very well with the concept of 11.88 Gb/s SDI. By segmenting the UHDTV image format into 1080p sub-pictures, as shown in Figure 14, each sub-picture can be very easily mapped to the 8 data streams of the 11.88 Gb/s SDI virtual interface.

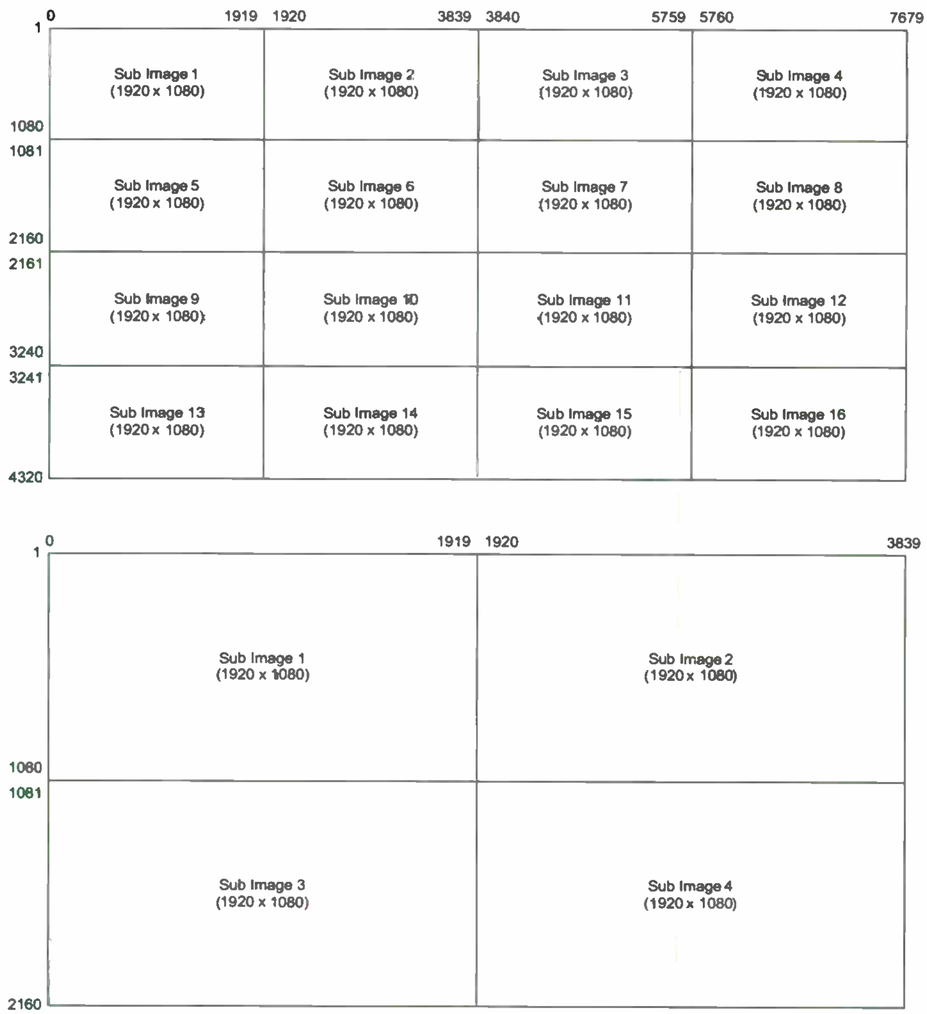


Figure 14: UHDTV Image Segmentation

UHDTV image formats require interface capacities ranging from 7.5 Gb/s up to 72 Gb/s. Using current HD-SDI interfaces, operating at 1.485 Gb/s, the transport of UHDTV requires multiple links, from 8 to more than 48, depending on the image format and sampling structure. This is not only a costly solution, but also technically challenging. All links need to be kept in synchronization, requiring additional hardware. The real estate required for all the HD-SDI spigots is considerable, leading to increased system costs. Finally, the cost and complexity of cabling becomes significant.

By mapping each UHDTV sub-picture into two or more data streams of the virtual interface, the transport of UHDTV over 11.88 Gb/s SDI provides a much simpler and more cost effective solution. By way of an example, if we take a 3840 x 2160 image at 60 Hz (progressive), sampled at 10-bit resolution, 4:2:2 sampling structure: after image segmentation, four 1080p60 data streams are mapped into the 11.88 Gb/s SDI virtual interface. A single sub-picture is mapped into two data streams of the virtual interface, as shown in Figure 15.

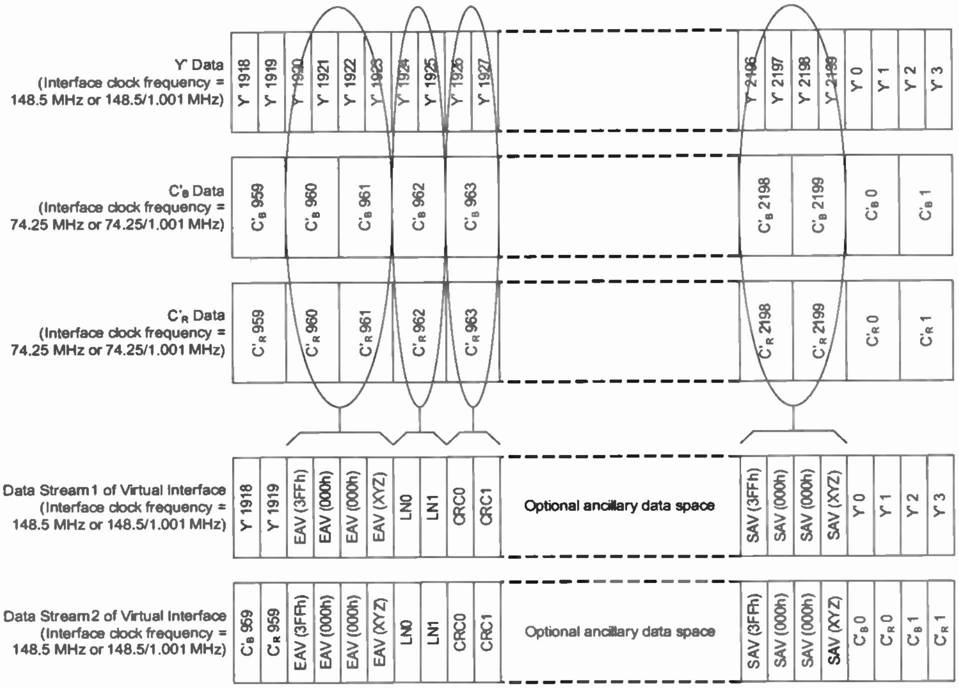


Figure 15: Format Mapping for UHDTV Sub-picture (1920 x 1080 @ 60p 10-bit 4:2:2)

Since there are four sub-pictures, all 8 data streams of the virtual interface are required. Thus, the entire 3840 x 2160 image can be carried over a single 11.88 Gb/s interface. By extension, a 7680 x 4320 image at the same 10-bit 4:2:2 sampling would require four 11.88 Gb/s SDI links.

All UHDTV image formats, which include bit-depth options of 10- and 12-bit, sampled at 4:2:0, 4:2:2 or 4:4:4, can be carried over 1 to 8 lanes of 11.88 Gb/s SDI. Also, all 24 channels of audio may be carried in the optional ancillary data space of each data stream in the virtual interface.

TIME DIVISION MULTIPLEXING

The 11.88 Gb/s SDI virtual interface was defined earlier in this paper (see Figure 2). As a container, this interface provides a payload capacity of 10.368 Gb/s in the *"active"* data periods. An additional 1.4472 Gb/s is available when the horizontal ancillary data space (HANC) is utilized.

The process of time division multiplexing (TDM) allows multiple streams of asynchronous video at lower data rates, to be carried over a higher data rate transport interface. For example, the carriage of multiple standard definition (SD) video payloads over a HD-SDI link. Using 11.88 Gb/s SDI, it is possible to carry a number of SD, HD and 3 Gb/s SDI streams, or combination thereof. Also, by discarding the blanking and TRS of each payload, more space is made available for additional payloads. Figure 16 illustrates the concept of TDM using the 11.88 Gb/s SDI virtual interface.

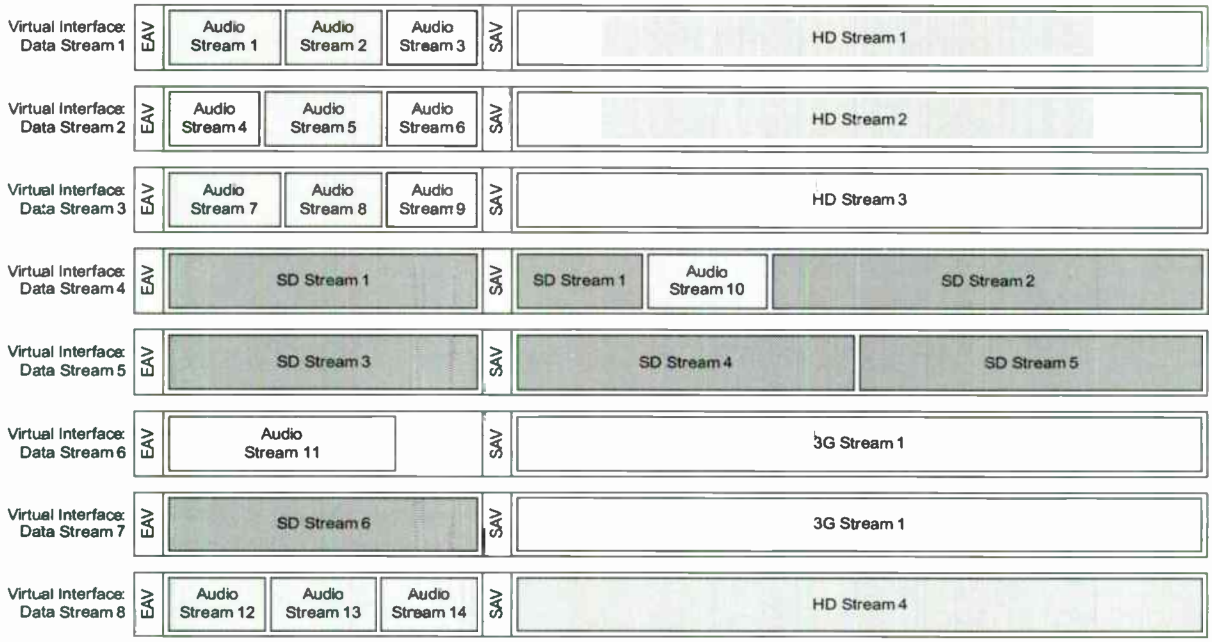


Figure 16: Time Division Multiplexing in 11.88 Gb/s SDI

The 11.88 Gb/s SDI container could also be used to carry compressed video and audio payloads, or transport streams. With this much bandwidth availability, the entire channel line-up of a DTV system could be carried in a single 11.88 Gb/s link!

ANCILLARY DATA

Each data stream in the 11.88 Gb/s SDI virtual interface provides sufficient HANC space to carry all currently defined ancillary data. The increased capacity, beyond current SDI solutions, allows larger HANC payloads, capable of carrying many channels of digital audio at very high sample rates.

The ancillary data space of a single data stream in the virtual interface is identical in structure to the space in an HD-SDI stream, allowing existing SMPTE standards and protocols for ancillary data to be applied directly to 11.88 Gb/s SDI. The following sections describe some applications of various HANC standards to 11.88 Gb/s SDI.

Embedded Audio

Digital audio can be embedded directly in the HANC space of each data stream in the 11.88 Gb/s SDI virtual interface, in the same way as is defined in SMPTE 425M and SMPTE 299M[6] (the HD audio embedding standard). Each data stream can carry up to 32 channels of AES/EBU format audio per line, at 48 kHz sampling. By directly referencing the SMPTE 299M standard, existing embedded audio IP can be re-purposed for 11.88 Gb/s SDI.

Figure 17 illustrates the carriage of up to 32 channels of audio in each data stream of the 11.88 Gb/s SDI virtual interface. Since each data stream represents a full line at 1080p60, the total HANC space available is 268 words (2200 total samples - 1920 active samples - 12 TRS/Line Number/CRC words).

At 60 frames per second, there is an average of 1 audio sample per channel of 48 kHz audio per HANC. SMPTE 299M defines an audio data packet size of 31 words to carry a single 24-bit sample for 4 channels of audio. Therefore, 32 channels of audio require 8 audio data packets at 31 words each, totalling 248 words. With an HANC capacity of 268 words, 32 channels of audio can be embedded in each data stream of the virtual interface, with 20 bytes to spare for other HANC data.

Most applications for embedded audio typically require a maximum of 16 channels. Therefore, only select data streams need to be used in these applications. As audio moves to higher sample rates, such as 96 kHz, there is still enough room for 16 channels of embedded audio per data stream in the 11.88 Gb/s SDI virtual interface.

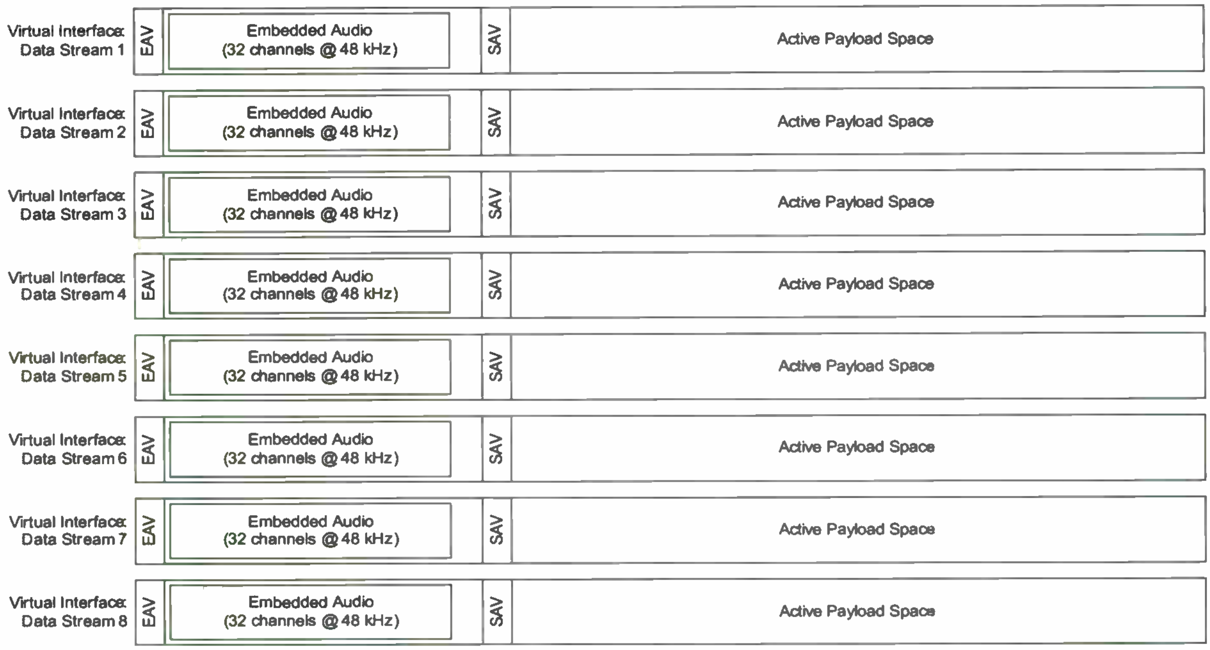


Figure 17: Embedding of Audio in the 11.88 Gb/s Virtual Interface

UHDTV calls for 24 channels of audio, at both 48 kHz and 96 kHz sampling. This can be easily accommodated in just two data streams of the virtual interface, leaving another 6 data streams of HANC space available for other applications.

Video Payload Identification

The 11.88 Gb/s interface simply adopts the SMPTE 352M[7] standard for video payload identification. Where SD, HD or 3 Gb/s SDI streams are carried over 11.88 Gb/s SDI, a 352M-compliant payload identification packet can be added to each stream, where byte 1 uniquely identifies the format as *"over 11.88 Gb/s"*. The other 3 bytes of the 352M packet can remain as currently defined in the standard.

New formats, such as UHDTV, can be defined in later revisions of SMPTE 352M with byte 1 payload ID's for carriage over 11.88 Gb/s SDI. The 352M packet could also be modified such that the content of each data stream of the virtual interface can be uniquely described. For example, *"data stream 1 is carrying 10 bits of Luma from sub picture 1"*.

With 11.88 Gb/s SDI, the use of SMPTE 352M packets is mandatory, in order to reconstruct the payload (or multiple payloads) at the receiving equipment. This is particularly important in UHDTV applications, where an image may need to be transported over multiple 11.88 Gb/s SDI links.

Other Ancillary Data

The SMPTE 291M[8] standard defines a generic ancillary data packet format, which many other SMPTE standards reference. By adopting SMPTE 291M as the ancillary data packet format for 11.88 Gb/s, both current and future applications and ancillary data types can be accommodated.

PHYSICAL INTERFACES

The increase in data rate to 11.88 Gb/s requires careful consideration at the physical layer to ensure links are robust. One of the most important functions of the physical layer in communication systems is to ensure digital information is interpreted correctly. A bit error rate (BER) of $< 10^{-12}$ over all operating conditions should be expected for optical systems (a BER $< 10^{-14}$ is recommended). A physical layer and link budget should therefore be defined, taking into account this overriding expectation. Although this paper focuses on an optical physical layer, electrical implementations leveraging coax are also possible.

High-speed considerations for the physical layer are of particular interest when communicating at a serial 11.88 Gb/s data rate. This is because non-ideal channel characteristics such as trace loss, connector discontinuities, jitter introduced by electrical-to-optical conversion, fibre dispersion, and jitter introduced by optical-to-electrical conversion become much more pronounced at higher rates. Figure 18 depicts critical points to ensure high-speed physical layer compliance.

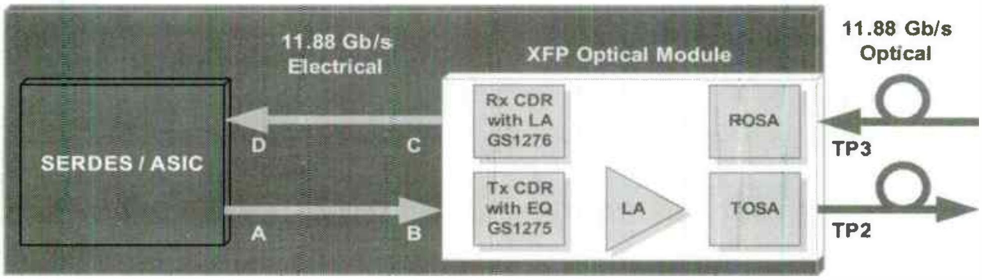


Figure 18: Physical Layer Compliance Points

Transmit Direction

At point A, the SERDES / ASIC is expected to provide a clean output for communication over the electrical channel. A reasonable expectation for the total jitter at the output of a SERDES / ASIC in this application is 0.3 Unit Intervals peak-to-peak (UIpp), which equates to 25ps. This high-speed signal then traverses a length of trace and encounters impedance discontinuities before entering the optical module. The total jitter can reach up to 0.61 UIpp at the connector of the XFP[9] module (compliance point B). Clearly this signal needs to be cleaned up before the electrical-to-optical conversion. To perform this clean-up function, an 11.88 Gb/s Transmit Clock and Data Recovery (CDR) circuit with trace equalization (such as the GS1275) is required. This device ensures a clean signal is used in the electrical-to-optical conversion process.

Since retiming is involved, it is critical to consider the input jitter tolerance and jitter transfer characteristics of the CDR. The required jitter tolerance is largely defined by the amount of jitter which can be presented to the input of the XFP module (in this case, 0.61 UIpp). Jitter transfer is an important consideration since jitter peaking in the transmit CDR will add jitter over certain frequency bands, and can break the overall jitter budget. Leveraging the recommended practices in SMPTE RP 184[10], Table 1 describes the jitter transfer requirements for the transmit CDR.

Table 1: Jitter Specifications

Parameter	Value	Description
Data Rate	11.88 Gb/s	
f1	50 kHz	Low frequency specification limit
fc	8 MHz	Upper band edge of jitter transfer
P	0.03 dB	Maximum peaking gain, f1 to fc (dB)
Test Signal	PRBS 2^{31}-1 or colour bars signal	

After the signal has been retimed by the transmit CDR, an electrical-to-optical conversion is performed using a laser driver (LD) and a Transmit Optical Sub-assembly (TOSA). The output at TP2 must then be measured for suitability for communication over optical fibre, and recovery by remote systems. To ensure proper link operation, intrinsic and output jitter measurements should be made. Leveraging recommendations in RP 184, Table 2 can be used to describe equipment output jitter.

Table 2: Intrinsic and Output Jitter

Parameter	Value	Description
Data Rate	11.88 Gb/s	
f1	10 Hz	Timing jitter, lower band edge
f3	4 MHz	Alignment jitter, lower band edge
f4	80 MHz	Upper band edge
A1	10 UI	Timing jitter
A2	0.10 UI	Alignment jitter
tm	1 minute	Measurement time
Test Signal	PRBS 2^{31}-1 or colour bars signal	

An optical eye diagram is also a useful tool in verifying physical layer compliance, particularly with respect to high frequency jitter. Figure 19 displays a normalized eye diagram used in high speed communication (see SMPTE 435-3[11]).

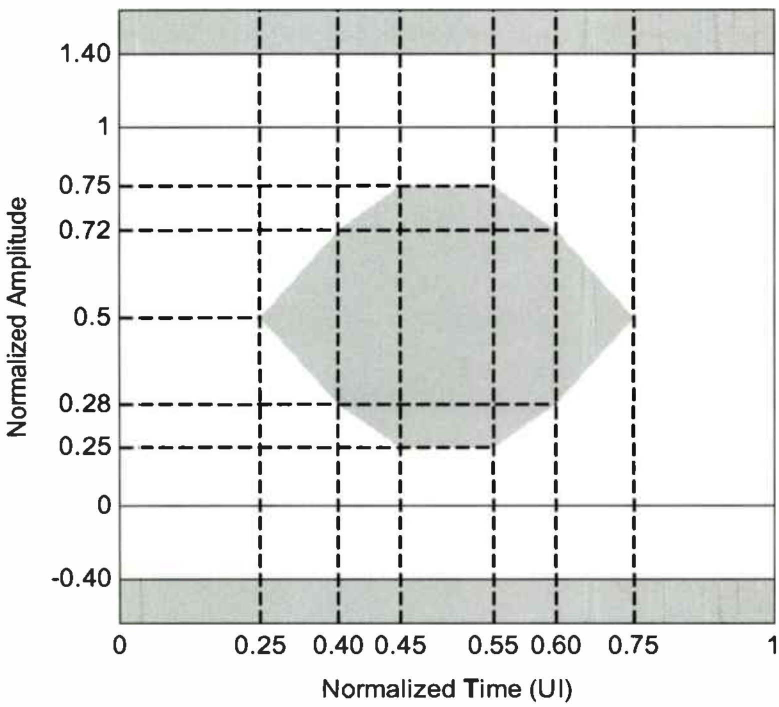


Figure 19: Transmitter Eye Mask Definition

In addition to jitter budgets and eye diagrams, optical source characteristics such as wavelength, average launch power ([P1+P0]/2) and extinction ratio (P1/P0) must also be defined to ensure proper communication. Leveraging SMPTE 435-3, Table 3 highlights optical source characteristics for 1310nm optics.

Table 3: Optical Source Characteristics

Parameter	Min	Typical	Max
Optical Wavelength	1260nm	1310nm	1355nm
Maximum RMS Spectral Width			1nm
Average Launch Power	-5.5 dBm		0.5 dBm
Extinction Ratio	6 dB		
Maximum Reflected Power			-12 dB
Electrical / Optical Transfer Function	Logic "1" = Higher optical power Logic "0" = Lower optical power		

Optical Channel

The optical channel is characterized by fibre type and connector. Single Mode Fibre (SMF) is used to communicate over longer links. In the 1310nm wavelength example given above, SMF can be used for distances from 2m to 2km. Multi-Mode Fibre (MMF) can be used with 850nm optics to achieve 2m to 300m links. When analyzing the supported distance, it is important to consider dispersive effects of the fibre, along with channel loss (attenuation).

Receive Direction

After traversing a length of fibre, the optical receive signal must be interpreted correctly by the module's receive path. To ensure this, a standard set of requirements need to be verified at TP3. Optical receiver characteristics, as defined in SMPTE 435-3, can again be used as guidance for the optical module (see Table 4).

Table 4: Optical Receiver Characteristics

Parameter	Min	Typical	Max
Average Receiver Power Maximum			0.5 dBm
Average Receive Power Minimum (BER 10^{-12})			-13.5 dBm
Detector Damage Threshold			1 dBm
Optical / Electrical Transfer Function	Logic "1" = Higher optical power Logic "0" = Lower optical power		

Jitter specifications are also required in the receive direction to ensure jitter budget closure. Table 5 defines potential Receiver Jitter Specifications and leverages recommendations outlined in SMPTE RP 184.

Table 5: Receiver Input Jitter Tolerance

Parameter	Value	Description
Data Rate	11.88 Gb/s	
f1	10 Hz	Low frequency specification limit
f2	20 kHz	Upper band edge for A1
f3	4 MHz	Lower band edge for A2
f4	80 MHz	High frequency specification limit
A1	10 UI	Low frequency jitter tolerance
A2	0.15 UI	High frequency jitter tolerance
Test Signal	PRBS 2^{31}-1 or colour bars signal	

This optical receive signal is then converted to an electrical signal by a Receive Optical Sub-assembly (ROSA), and recovered by a receive CDR with limiting amplifier (such as the GS1276). The limiting amplifier is used to ensure a small signal from the ROSA will be interpreted correctly (particularly at the low optical power levels), and the CDR removes the jitter accumulation from the channel, the electrical-to-optical, and optical-to-electrical conversions.

Since retiming is involved, it is important to again consider the input jitter tolerance, jitter transfer, and output jitter of the CDR. Input jitter tolerance requirements are largely determined by the optical jitter tolerance defined in Table 5 (plus any jitter added by the ROSA). A jitter transfer requirement similar to the transmit direction should be considered since jitter peaking in the receive direction is as important as the transmit direction. Table 6 defines receive jitter transfer requirements.

Table 6: Receive Jitter Transfer Requirements

Parameter	Value	Description
Data Rate	11.88 Gb/s	
f1	50 kHz	Low frequency specification limit
fc	8 MHz	Upper band edge of jitter transfer band-pass
P	0.03 dB	Maximum peaking gain, f1 to fc (dB)
Test Signal	PRBS 2^{31}-1 or colour bars signal	

Once the receive signal has been retimed, it is ready for communication over the XFP connector and host channel. Host providers can expect a relatively clean output from the XFP module, 0.34 UIpp. Once again, this signal traverses a channel with trace losses and discontinuities which must then be recovered by the host ASIC / SERDES. Depending on design, various levels of trace equalization can be employed to ensure robust signal recovery at compliance point D.

LEVERAGING 3 GB/S TO COMMUNICATE 11.88 GB/S OVER A SINGLE FIBRE

With the proliferation of 3 Gb/s SDI, numerous electrical and optical components have been developed to support 3 Gb/s video communication. It would therefore be beneficial to the industry if one could leverage this technology base to communicate 11.88 Gb/s over a single fibre. Fortunately, this is possible using optical wavelength division multiplexing (WDM). Using this technique, the multiplexing / demultiplexing of 4 x 3 Gb/s SDI occurs in the optical domain instead of the electrical domain. This is achieved by assigning each 3 Gb/s signal a specific wavelength which can be combined at the transmit end, and separated at the receive end (think of a prism and how it can separate the components of light). A diagram of this architecture is shown in Figure 20. Also, an example of the wavelengths which can be used around the 1310nm spectrum is outlined in Table 7.

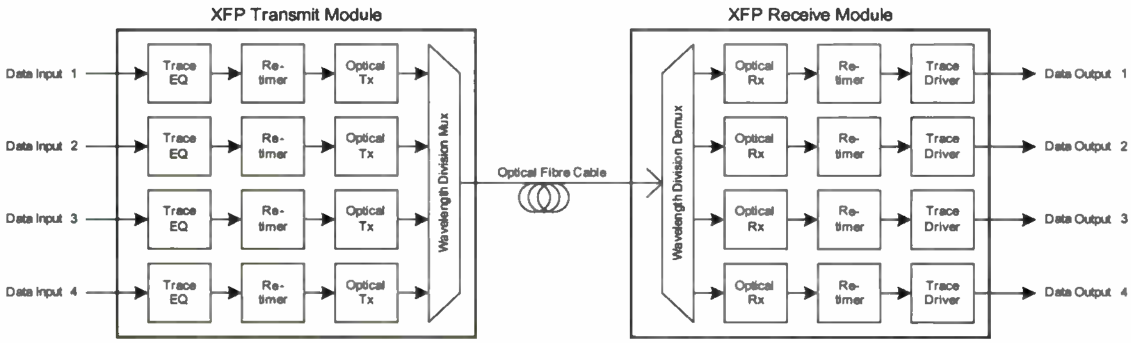


Figure 20: WDM Architecture for 11.88 Gb/s

Table 7: WDM Lane Wavelengths

3 Gb/s Data Input Number	Wavelength Ranges
Data Input 1	1269.0 - 1282.4 nm
Data Input 2	1293.5 - 1306.9 nm
Data Input 3	1318.0 - 1331.4 nm
Data Input 4	1342.5 - 1355.9 nm

This architecture is also supported using an XFP optical module by re-purposing some of the pins which are not necessary in this application. For example, one can take advantage of the following:

1. many of the communication links in Video and Broadcast applications are uni-directional;
2. the module can run off a single 3.3V supply;
3. CDR's such as the GS1275 and GS1276 do not require reference clocks.

Taking advantage of these points, an XFP module connector can be redefined as depicted in Figure 21.

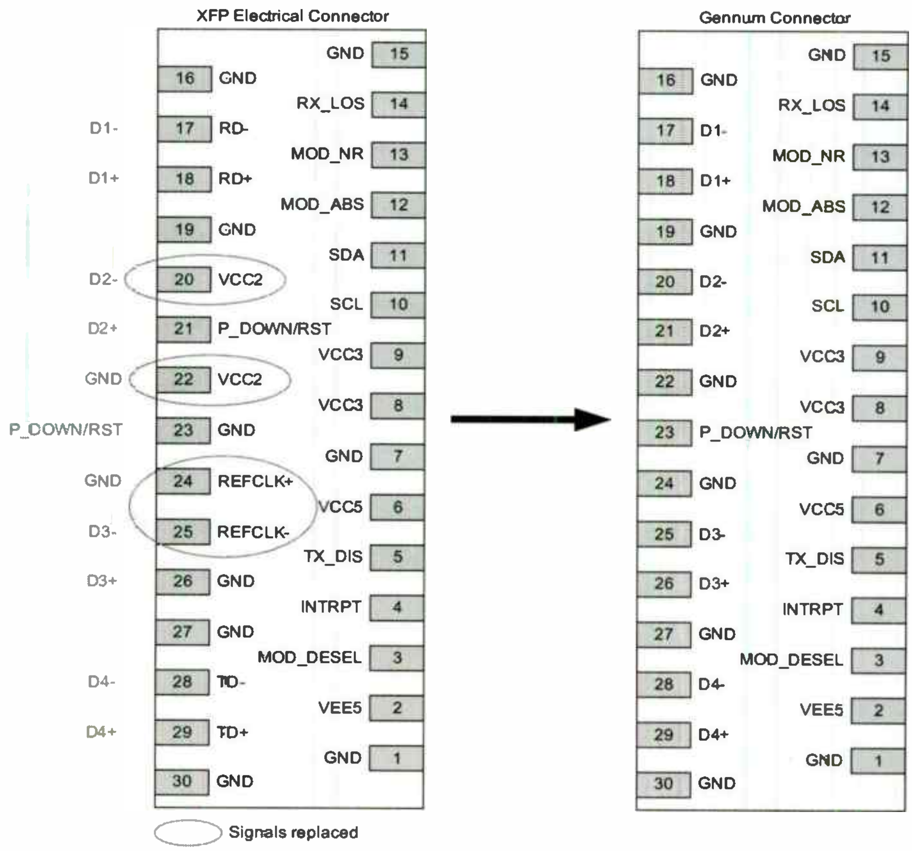


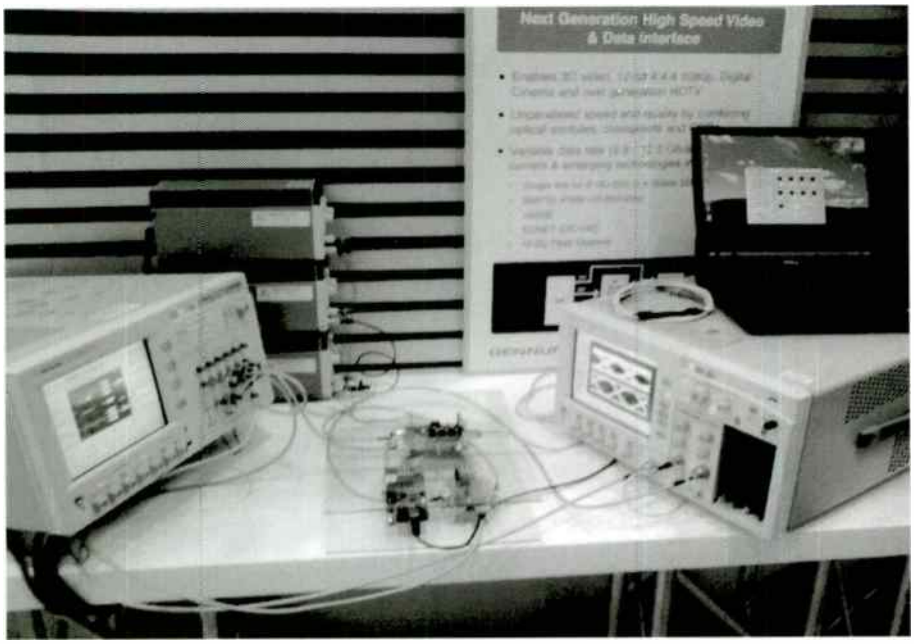
Figure 21: XFP Pin Connector to Support WDM

DEMONSTRATION OF FEASIBILITY

The high-bandwidth serial data interface described in this paper is specified to operate at 11.88 Gb/s and 11.88/1.001 Gb/s up to 2km distance using single-mode fibre.

A prototype physical interface operating at 11.88 Gb/s was demonstrated at both the SMPTE Technical Conference in October 2007, and the Inter BEE 2007 trade show in Japan. The Inter BEE demonstration is shown in the photograph below.

The demonstration features a 13.5 Gb/s BERT (Bit Error Rate Tester) which is used as a variable-bit-rate source and receiver. This signal is electrically transmitted to the input of the optical module where it is retimed by a CDR (the GS1275) to reset the jitter budget prior to optical transmission. The signal is then converted from electrical to optical and transmitted over fibre. The optical signal is then split into two optical signals. One signal goes directly to a high-speed oscilloscope for observation of the raw optical transmit eye. The other signal is sent to the optical receive portion of the module. The signal is received by the optical module and converted back to an electrical signal, where it passes through another CDR (the GS1276) which again resets the jitter budget. It then passes through the crosspoint and is sent to the oscilloscope. The oscilloscope shows the significant differences in performance between a signal that has passed through the CDR and one that has not, with the signal having passed through the CDR showing robust performance.



REFERENCES

1. SMPTE 425M-2006 3 Gb/s Signal/Data Serial Interface - Source Image Format Mapping

2. SMPTE 424M-2006 3 Gb/s Signal/Data Serial Interface

3. T. Eguchi, "Pathological Check Codes for Serial Digital Interface Systems", SMPTE J., 101:553-558, Aug. 1992

4. SMPTE EG34-2004 Pathological Conditions in Serial Digital Video Systems

5. SMPTE 292-2006 1.5 Gb/s Signal/Data Serial Interface

6. SMPTE 299M-2004 24-Bit Digital Audio Format for SMPTE 292M Bit-Serial Interface

7. SMPTE 352M-2002 Video Payload Identification for Digital Interfaces

8. SMPTE 291M-2006 Ancillary Data Packet and Space Formatting

9. XFP MSA, XFP (10 Gigabit Small Form Factor Pluggable Module) specification revision 4.5: 2005

10. SMPTE RP 184-2004 Specification of Jitter in Bit Serial Digital Systems

11. SMPTE 435-3-2007 10Gb/s Serial Signal / Data Interface

FURTHER READING

1. Pathological Revisited — David Brown, Connectivity Technologies, Gennum

2. Optical Fibre Transport for Digital Video — Nigel Seth-Smith, Connectivity Technologies, Gennum

ACKNOWLEDGEMENTS

The authors would like to thank their colleagues John Hudson, Nigel Seth-Smith and David Brown of Gennum Corporation for contributions to this work.

Audio Mixing Requirements in Next Generation Broadcast Receivers for Audio description and Other Enhanced Features

Roland Vlaicu

Dolby Laboratories
San Francisco, CA

ABSTRACT

The deployment of next generation broadcast platforms for HDTV and IPTV offers the opportunity to utilize advanced video and audio codecs, as permitted by recent revisions to ETSI specifications TS 101154 and TS 102005. In addition to bandwidth savings, new audio codecs for these applications will likely need to support additional features within their platform lifetimes, including 7.1-channel surround sound and improved provision for the visually impaired. Support for the latter feature is increasingly being mandated by broadcast regulators in Europe and elsewhere.

This paper specifically considers the requirements for audio mixing features within next generation broadcast audio codecs. Audio mixing enables enhanced services, such as audio description for the visually impaired or director commentaries, to be offered at efficient total data rates. This can be achieved by transmitting an additional bit rate-efficient channel of commentary, which is mixed dynamically with the main program audio within the audio decoder of a standard home set-top box.

It is concluded that mixing audio streams in the coded domain within a single standard audio decoder offers significant advantages over other approaches. These include simplicity of implementation, improved audio quality and connectivity to home audio equipment. The importance of implementing metadata control of the mixing feature is highlighted, and requirements for appropriate production and encoding tools are discussed.

1. INTRODUCTION

This paper discusses a new approach to offer support for audio description (AD) and other advanced services in broadcast receiver devices such as set-top boxes or TV sets through the introduction of a mixing component within the audio decoder.

Mixing solutions in broadcast receivers that support two-channel mixing have been available for quite some time now. The main and associated audio tracks are presented using one of the common broadcast audio codecs for stereo such as MPEG-2 Layer II and then mixed at the user's option upon decoding. Some broadcast standards such as DVB [2] use defined mixing metadata for this purpose in order to support codec-agnostic mixing in the receiver device. This approach is often referred to as receiver-mix. However, in these cases functionality is limited, usually requiring the user has to choose between the audio description services and multichannel surround sound. Although the syntax of the DVB specification describes how to pan an audio description source within a 5.1 surround signal, it has not yet been implemented.

Next-generation broadcast specifications around the world focus on two codecs to distribute multichannel surround sound: Enhanced AC-3 and aacPlus. Both codecs support surround sound configurations of 5.1 and more at very efficient bit rates. While aacPlus can make use of the mixing metadata defined in [6], Enhanced AC-3 comes with its own set of mixing metadata that is contained in the associated audio stream. This brings many advantages, including the ability to include the associated audio stream as an independent substream with the main audio service, all signaled in the transport stream under the same Audio-PID.

2. AUDIO DESCRIPTION AS A STANDARD REQUIREMENT FOR NEXT-GENERATION BROADCAST RECEIVERS

In many countries, broadcast services for visually impaired audiences are becoming mandatory, and are provided mostly through an associated audio stream containing both the main audio mix and an audio description soundtrack. Preparation of the AD track is time-consuming in post production, and requires approximately the same bandwidth on the transmission path as the main audio stream. Enhanced AC-3 offers a more efficient alternative.

By integrating a mixing module with the audio decoder in a broadcast receiver device such as a set-top box or television set, Enhanced AC-3 enables bit-stream mixing in the discrete cosine transform (DCT) domain. This allows mixing an associated audio stream of up to 5.1 channels, as well system sounds and interactive audio of up to 5.1 channels, with the main audio service implemented on the audio decoder level. User level adjustment parameters enable the user to control the balance between main and associated audio in the overall output.

Such a mixing module could enable a variety of new features such as audio description via text-to-speech rendering, which many European public broadcasters employ for director's commentary with video-on-demand services. It could also enable the playback or generation of interactive sounds on the set-top box, with the resulting audio signal provided via both two-channel outputs and digital IEC 61937 interfaces (S/PDIF or HDMI). It would also assist broadcasters such as SVT in Sweden who are looking into offering an associated audio track derived from a text-to-speech engine, which could allow a dramatic increase in programming with AD services.

3. RECEIVER MIX VS. BROADCAST MIX

Audio description has long been provided by broadcast services. The PAL system, for example, makes it possible for the second FM sound sub-carrier, otherwise used for the right channel in stereo transmissions, to carry an alternative version of the main soundtrack carried on the first FM sub-carrier. While this "broadcast-mix" approach limits the transmission to mono main and associated audio soundtracks, it is an efficient way to provide a basic service for the visually impaired. Figures 1 shows the symbol used by European broadcasters to mark transmissions offering Audio description.

Figure 1: Symbol used by European broadcasters for transmissions offering audio description (AD)

This broadcast-mix approach, whereby the main and commentary tracks are mixed in postproduction prior to transmission, was carried over into digital television DTV systems because early receivers offered only basic functionality, and the sources for digital transmissions were often still stored on analog media. It is also similar to that taken by the NTSC system where a secondary audio program (SAP) carries a pre-mixed version of main and audio description tracks as a mono signal. However, due to the affiliate broadcast structure in the US, at times technical impediments interfere with passing through a SAP signal on cable and satellite systems, making the service less accessible to viewers.

In Germany, public broadcasters offering a large portion of their programming with audio description services are facing the challenge of continuing to distribute their AD service "analog-style", i.e. leaving main audio on the left channel and AD audio on the right channel of what is basically the main broadcast audio track carried as a stereo signal but signalled as "dual-mono". This requires broadcast receivers to pass only the left channel or right channel to their audio outputs. However, since this is not a requirement within DVB, which mandates that different audio services be transmitted and signaled using different audio-PID, receiving and recording these services becomes a major challenge.

To this date, transmitting audio description services in the form of a broadcast mix is standard practice in many parts of the world. Efforts to offer a more efficient way to provide services for the visually impaired in the form of a "receiver-mix" approach have cumulated in the creation of a specification that is now part of DVB.

Annex E of [2] describes a basic approach to transmit an isolated mono signal containing the description track alongside the main soundtrack as a separate bitstream. A suitable broadcast receiver can then perform the mixing of the main soundtrack with the commentary track in the device itself, hence the term receiver mix.

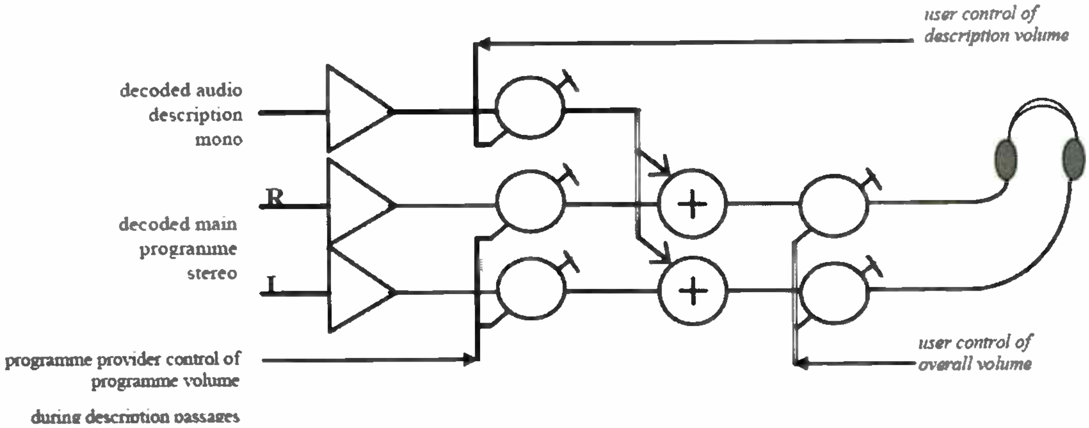


Figure 2: 2-channel Audio description as defined for DVB receivers in [2]

Figure 2 shows a diagram of the DVB approach. The DVB specification limits the associate audio stream to mono. While this approach is very valid to offer basic receiver-mix functionality to offer audio description services, it does not take into account other features that the mixer component in a broadcast receiver could support, such commentaries by several voices, or multichannel system sounds that could be overlaid onto a multichannel main soundtrack. The design of an enhanced mixing engine that supports these extended features is discussed in the next chapter.

4. PROPOSAL OF A NEW MIXING ARCHITECTURE

This chapter describes in detail a decoder-mixer-converter (DMC) for use in broadcast applications such as set-top boxes. The mixing of substreams with the main broadcast audio enables services such as audio description for the visually impaired or director's commentary, while at the same time maintaining the highest possible bandwidth efficiency.

Substream mixing negates the need for broadcasters to transmit several final mixes to support multiple audience preferences. It is necessary only to add low-bandwidth substreams containing the additional information along with mixing metadata that determines how the substream is combined with the main broadcast audio track in the receiving device. There is strong demand for this feature from public broadcasters in Europe, and it can add another differentiating feature to broadcast devices that will help to drive the success of next-generation digital television services around the world.

Audio Service		Codec	Substream usage	Channel configuration
Main		AC-3	N/A	Up to 5.1 channels
		Enhanced AC-3	Independent substream 0	Up to 5.1 channels
		aacPlus	N/A	Up to 5.1 channels
Associated				
	Delivered in same bitstream as Main audio service	Enhanced AC-3 (main audio must be AC-3 or Enhanced-AC-3)	Independent substream 1 Independent substream 2 Independent substream 3	Up to 5.1 channels
	Delivered in separate MPEG-PES	Enhanced AC-3 (main audio must be AC-3 or Enhanced AC-3)	Independent substream 0	Up to 5.1 channels
		aacPlus (main audio must be aacPlus)	N/A	

Table 1: Input formats for a proposed multichannel mixing device

Table 1 shows input formats to such a mixing engine that represent the predominant multichannel audio formats in current and next-generation broadcast systems.

This DMC design supports mixing one main audio service with one associated audio service. The associated service can be delivered in the same Enhanced AC-3 bitstream as the main service through the use of an additional independent substream, or as an Enhanced AC-3 bitstream carried in a separate MPEG PES stream.

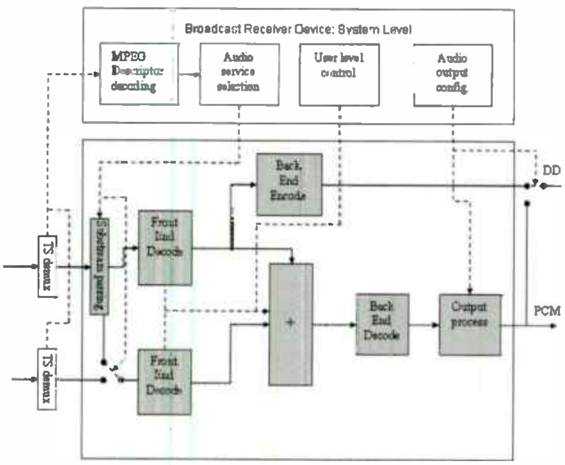


Figure 3: The structure of the decoder-mixer-converter (DMC)

In the case of aacPlus, the associated audio service can be delivered as an aacPlus bitstream carried in a separate MPEG PES stream. Information on the type of associated audio service is carried in the Enhanced AC-3 or aacPlus Descriptor that is found in the program map table of the incoming MPEG-2 Transport Stream.

Each Enhanced AC-3 or aacPlus bitstream used for the broadcast service has an identifying descriptor which the broadcast receiver device uses to inform the user which associated audio services are available. When the user selects a desired service, the DMC in the broadcast receiver device demultiplexes it from the MPEG transport stream.

Example 1

The user selects an associated audio service being carried in a separate MPEG PES associated with the main audio service through its Enhanced AC-3 descriptor. The broadcast receiver demultiplexes both Enhanced AC-3 streams from the broadcast transport stream, and passes them to both inputs of the DMC. The DMC receives information on the user's stream selection from the system layer of the broadcast receiver device and activates both inputs, then decodes and mixes the two streams. The mixing metadata is carried in the Enhanced AC-3 bitstreams containing the associated audio track.

Example 2

The user selects an associated audio service that is being carried in an additional independent substream in the Enhanced AC-3 main audio bitstream. The broadcast receiver demultiplexes only the main audio bitstream, while the DMC receives information on the user's associated service selection from the system layer of the broadcast receiver device and activates its main audio input only. The DMC then uses the substream parser in the main audio chain to separate independent substream 0 (main audio) from the independent substream carrying the associated audio. The associated audio service is then routed inside the DMC to the associated audio input of the mixer, and both independent substreams are decoded and mixed. The mixing metadata is carried in the Enhanced AC-3 bitstreams containing the associated audio soundtrack.

Example 3

The user selects an associated audio service that is being carried in a separate MPEG PES associated with the main audio service through its aacPlus descriptor. The broadcast receiver demultiplexes both aacPlus streams from the broadcast transport stream, and passes these streams to both inputs of the DMC. The DMC receives information on the user's associated service selection from the system layer of the broadcast receiver device, activates both inputs, then decodes and mixes the two aacPlus bitstreams. The mixing metadata is carried in the PES header information of the aacPlus bitstreams and has to be passed to the DMC from the system level parser of the broadcast receiver device.

4.1. Multichannel mixing and the inclusion of system sound effects in surround

As mentioned earlier in this paper, broadcast receivers can go beyond the minimum requirements set forth by DVB to support services for the visually impaired. Implementing a design as shown in Figure X, opens up even more applications for the broadcaster, including the bandwidth-efficient transmission of multiple 5.1 mixes of a program, or enabling a sports event mixed in surround to be enjoyed in different languages or with different commentators for different geographical regions.

A frustration for developers of next-generation DVR devices is that system sounds stored or created in these devices so far can make it out of the box only through their two-channel analog outputs. Typically, the multichannel bitstreams transmitted as part of a program are being passed to the digital outputs (HDMI or S/PDIF) unaltered. As a result, it has been impossible to include system sounds for those outputs. An even more extended design option of the DMC offers inputs for pre-produced system sounds in 5.1 multichannel surround sound. Figure 4 shows the system diagram of the DMC supporting this feature.

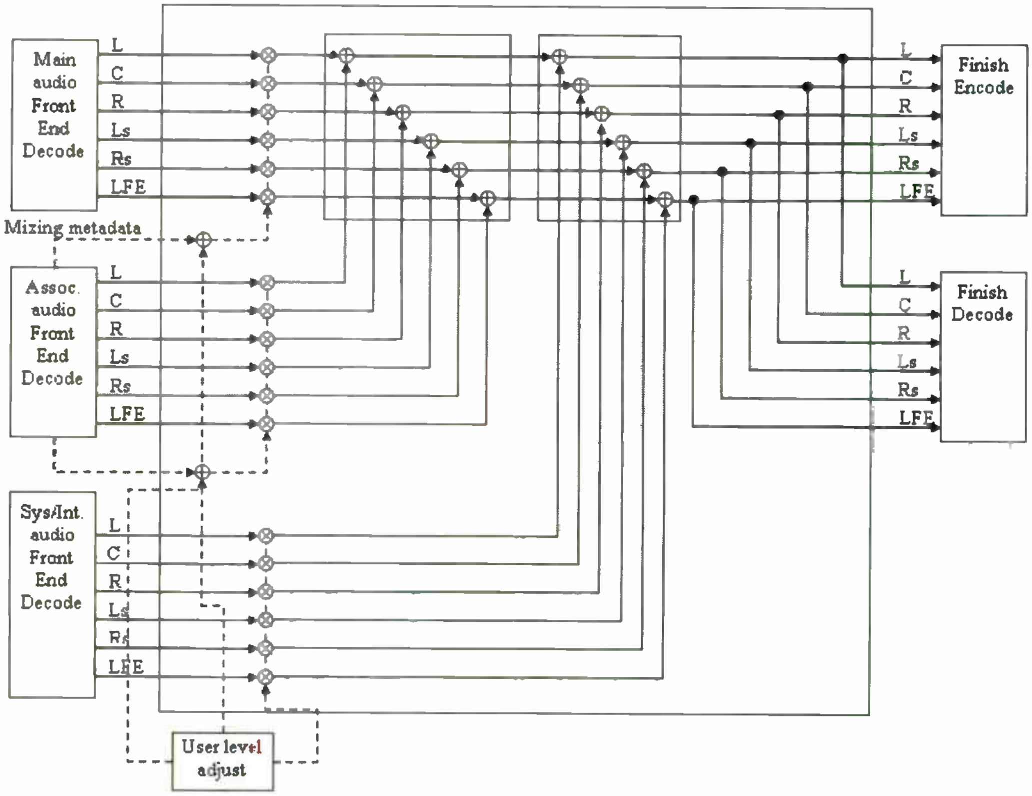


Figure 4: DMC with inputs for multichannel system sounds

This configuration supports mixing of one main audio service with one associated service, plus an additional input for system/interactive sounds which are be stored locally in the set-top box device.

The associated audio service can be delivered in the same Enhanced AC-3 bitstream as the main audio through the use of an additional independent substream, or can be delivered as an Enhanced AC-3 bitstream carried in a separate MPEG PES stream.

Information on the type of associated audio service being delivered is carried in the Enhanced AC-3 Descriptor that is found in the program map table of the incoming MPEG-2 Transport Stream. Each Enhanced AC-3 bitstream present in the broadcast service has an Enhanced AC-3 descriptor associated with it.

The broadcast receiver device uses the contents of each descriptor to inform the user of which associated audio services are available. The user selects the Enhanced AC-3 bitstream or substream carrying the desired associated service. In the case of aacPlus, both Main and Associated Audio services must be carried separate MPEG PES bitstreams.

System/interactive sound samples can either shipped with the broadcast devices, or in the case of a DVR, pushed to the device as data files contained in the MPEG transmission transport stream. This provides a very efficient way to supply high-end customers, who have multichannel home cinema systems and don't use two-channel analog outputs, with the same or even more exciting system sounds that can be mixed-in with the 5.1 broadcast audio.

5. PREREQUISITES FOR THIS NEW MIXING ARCHITECTURE

Support for multichannel mixing and enhanced features such as system sounds in surround call for significant changes in the audio decoder structure of broadcast receiver devices. Many next-generation audio codecs have been designed with these new requirements in mind. One of the two codecs discussed in this paper is the Enhanced AC-3 codec as defined in [2] and [5].

As a successor to the well established AC-3 codec used in many entertainment applications worldwide, Enhanced AC-3 provides a framework for new features like audio description. The following chapter illustrates how Enhanced AC-3 opens opportunities for broadcast operators to new and exciting services while maintaining maximum compatibility with the existing installed base of home cinema systems.

This chapter discusses the requirements for an audio coding scheme to address both short- and long-term technology challenges of next-generation broadcasting such as high-definition television (HDTV). A number of factors have been considered in determining the characteristics of a suitable audio coding scheme:

• Requirements for new broadcast services such as HDTV and services using advanced video codecs such as H.264 or VC-1.

• Opportunities for improving the audio performance of current broadcast services.

• Impact of a new audio coding scheme on broadcast production practices.

• Impact of a new audio coding scheme on consumer hardware and listening environments.

• Differentiation between high-end next-generation and competitive TV services.

6. REQUIREMENTS FOR EXISTING SERVICES, SHORT- AND LONG-TERM

The audio codecs previously specified in [2] offer solutions for many of the requirements facing DTV and IPTV operators—high-quality audio at low bit rates, multichannel audio services, guaranteed connectivity with consumer products through existing IEC 61937 interfaces (S/PDIF or HDMI), and vast improvements to the consumer listening experience through the use of audio metadata. Each codec previously specified in [2] meets some of these requirements, but until recently, there was no single codec that meets all of them.

For example, the AC-3 (Dolby® Digital) codec delivers up to 5.1 channels of audio, offers broadcasters and operators full control over the listening experience for all consumer environments through comprehensive metadata control, and offers standardized connectivity via IEC 61937 to over 40 million existing consumer A/V systems. However, the AC-3 codec is not optimized for low bit-rate performance.

When considering the feature set used by current broadcast services, a new audio codec should offer at least the following:

• Support for mono to 5.1- and 7.1-channel capability.

• Comprehensive metadata support, mandated in both encoder and decoder, with all parameters under encoder control:

- Dialogue normalization to ensure consistent listening levels between programs.
- Downmix capability to ensure backward compatibility with matrix surround, stereo, and mono playback systems.
- Control of dynamic range to ensure optimal reproduction for all consumer listening environments.

• Delivery of discrete 5.1-channel audio to current installed base of A/V receivers via IEC 61937 interfaces, and support for other emerging digital interface standards.

• Improved bit-rate efficiency compared with audio codecs currently in use in DVB and IPTV services, complementing efficiency gains of new video codecs.

• Licensing costs and terms in line with existing audio codecs.

• Encoder and decoder products subject to inter-operability testing to ensure consistent performance.

In addition to these requirements for core broadcast services, an opportunity also exists to improve the current provisions for deployment of audio description services for visually and hearing impaired. While relative levels between audio description (AD) and main program services can be controlled both by the broadcaster and the listener, variations in loudness and dynamic range between programs leads to a need for regular adjustments to listening levels by the consumer.

A new audio codec should meet the following requirements to deliver improved AD services:

• Metadata control of dialogue levels to ensure a consistent relative level between main and AD programs.

• Metadata control of the dynamic range of the main program to ensure that AD services are clearly audible at all times.

• Metadata to control mixing of main program and AD services in a broadcast receiver should be supported, to remove the need for frequent manual adjustment of levels in the broadcast receiver. Support for mixing AD services with multichannel as well as stereo program content should also be available.

• Ability to deliver both main program and AD services as a single stream that can be decoded and mixed with a single decoder in the broadcast receiver to simplify implementation in a broadcast receiver.

7. REQUIREMENTS FOR NEW SERVICES

Standards and technologies are being developed for IPTV services and the next generation of DTV broadcasts. New audio codecs must adapt to the requirements of these new technologies. Applications such as high-definition television and interactive services present new opportunities for audio services. A new audio codec should satisfy at least the following requirements to meet the demands of future broadcast services:

• Ability to deliver audio quality improvements to match video quality improvements of high-definition IPTV broadcasting.

• Flexibility to deliver more than 5.1 channels of audio to match future motion picture mixing formats.

• Support for mixing interactive audio content with main program audio, including multichannel audio content.

• Deployment of multiple programs in a single stream, enabling multiple languages, director's commentaries or other advanced services, all controlled by mixing metadata, to be decoded using a single decoder in the broadcast receiver.

7.1. Impact on Broadcast Production Environment

The adoption of a new audio codec for final broadcast should have minimal affect on a broadcaster's working methods, and should not adversely affect or complicate them. In the case of stereo audio services, the process of creating audio content does not differ greatly from codec to codec—the selection of encoding settings need be done only once, based upon the target quality of a broadcaster's service and the capability of the codec selected. If the selected codec supports control of program loudness and dynamics through metadata, this should always be factored into the production process.

When considering multichannel audio services, the workflow of a new emission codec through the broadcast production environment must be carefully considered. Multichannel audio presents a number of challenges to a broadcaster. These challenges include distribution of 5.1- (or higher) channel audio content within a broadcast facility equipped only for stereo content, and the task of creating audio metadata to ensure that 5.1 programming can deliver optimal backward compatibility with all listening environments.

When considering the integration of a new audio codec with a broadcast production environment, it is important that the level of metadata functionality should at least match and preferably exceed that of current codecs, to maintain a broadcaster's ability to deliver consistent audio quality. A simple interface between multichannel production equipment and transmission encoders for both audio and metadata is desirable.

Today, the Dolby E format is predominantly used in multichannel audio production. This format can handle up to eight channels of audio in a single AES/EBU carrier and integrates the complete metadata functionality as described in [1]. As a consequence, most contribution of 5.1 or 7.1 surround sound audio in broadcast production is done in Dolby E. IPTV operators will be able to benefit from its flexibility for both linear turn-around of existing broadcast channels as well during the preparation of assets for non-linear offerings such as video-on-demand.

In order to support audio description services, Dolby E can be used to store both the main audio soundtrack (up to 5.1 channels) along with an audio description track (up to 2 channels) within the same Dolby E bitstream. This bitstream can be stored on any carrier offering transparent AES/EBU audio capabilities at 20bit resolution.

7.2. Impact on Consumer Products and Listening Environments

A new audio codec should offer performance improvements for the consumer, while ensuring simple integration into the current consumer listening environment, and offer flexibility for future developments in consumer product enhancements. To ensure a consistent listening experience for all consumers, a new audio codec should meet the following requirements:

• Decoders must maintain compatibility with existing consumer A/V receivers and IEC 61937 interfaces when delivering discrete 5.1-channel content, without introducing excessive complexity to the decoder design.

• All decoders must be able to receive and decode multichannel audio services to deliver either a matrix surround, stereo, or mono downmix as required, removing the need for audio simulcasting (the new audio codec should not remove the possibility of audio simulcasting, only the need).

• Decoder complexity should be in line with current designs for an equivalent feature set.

• Metadata created during the production process must be supported by all decoders. If it is not, the benefits of using metadata are usually lost.

• The new codec should be compatible with emerging and future digital interface standards without introducing excessive complexity to the decoder design.

• Licensing costs and terms in line with existing audio codecs.

8. MEETING THESE REQUIREMENTS

In consideration of these requirements, Dolby developed Enhanced AC-3 (Dolby Digital Plus) for use in next-generation applications. This coding scheme has already been included in [5] and has been selected as mandatory technology for HD DVD players and as optional technology for Blu-ray Disc players. The scheme has also been standardized by DVB for next- generation broadcast and IPTV services in [2] and [4][4].

Enhanced AC-3 is well suited in applications that include lower data-rate carriage of audio and its conversion to the AC-3 coding standard for playback on today's installed base of audio/video entertainment equipment. It also supports combining streamed content with a main audio program in interactive multimedia; the reproduction of greater than 5.1 channels for playback of both existing and future cinema content; and the efficient transcoding of AC-3 program content to lower data-rate Enhanced AC-3 bitstreams, and conversion back to AC-3 for playback on the very large installed base of consumer Dolby Digital decoders.

8.1. Compatible Lower Data-Rate Carriage

Growing number of applications not only require lower data rates, but also compatibility with the existing broadcast-reception and audio/video decoding infrastructure.

The Enhanced AC-3 system is an excellent solution for these applications because of its inherent lower tandem coding losses compared to AC-3 and its greater coding efficiency provided by new coding tools. This results in minimizing quality losses when large content libraries are transcoded to other formats for IPTV or other advanced service environments. To ensure compatibility with the large installed base of Dolby Digital decoders, Enhanced AC-3 enables low-loss conversion to standard AC-3 over a digital audio interconnect such as S/PDIF and decoding by a standard Dolby Digital decoder.

The conversion stage is a special form of transcoder that minimizes quality degradations resulting from tandem coding losses, which is possible by using the same filterbank, transform block alignment, bit-allocation process, and basic framing structure as conventional AC-3.

8.2. A Next-Generation Broadcast Receiver Device

The next-generation broadcast receiver application is very similar to the conventional AC-3 reception paradigm, except the need for greater video channel capacity for HDTV services requires the transmission of audio programming at lower data rates than AC-3 applications. Traditionally, AC-3 has been deployed at 192–256 kbps for stereo and 384–448 kbps for 5.1-audio applications. The use of the new coding tools in Enhanced AC-3 allows for lower data rates while permitting efficient conversion to a conventional AC-

3 bitstream at 640 kbps for tested compatibility with existing home theaters. Figures 1 through 4 show the different configurations and use cases of this converter/decoder.

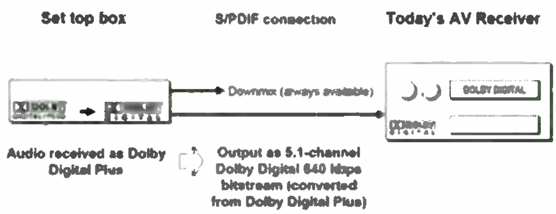


Figure 5: Scenario 1: Broadcast receiver device decoder-converter: Enhanced AC-3 input signal, S/PDIF connection (mixed or unmixed)

In scenario 1, the Enhanced AC-3 decoder-converter converts an incoming Enhanced AC-3 signal into a standard AC-3 signal at 640 kbps for output over IEC61937 interfaces. This conversion maintains all metadata information within the Enhanced AC-3 bitstreams and passes it on to the AC-3 bitstream. Channel configurations between mono and 5.1 remain unchanged. If the incoming audio had a 7.1-channel configuration, the decoder-converter would extract the 5.1-channel core from the signal and pass it on as AC-3 in a 5.1-channel configuration at 640 kbps.

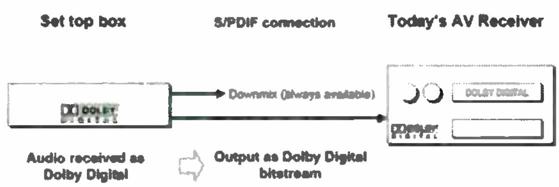


Figure 6: Scenario 2: Broadcast receiver device decoder-converter: AC-3 input signal, S/PDIF connection (mixed or unmixed)

In scenario 2, the Enhanced AC-3 decoder-converter acts as a standard AC-3 decoder and provides a two-channel downmix from the incoming audio while simultaneously passing the incoming stream to the IEC61937 interfaces unchanged as the AC-3 multichannel decoding capability can always be assumed in any digital A/V receiver.

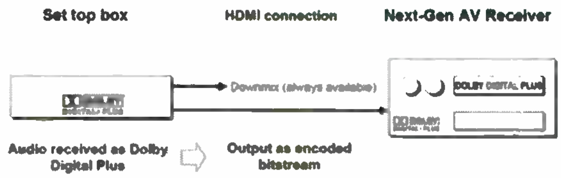


Figure 7: Scenario 3: Broadcast receiver device decoder-converter: Enhanced AC-3 input signal, HDMI connection (mixed or unmixed)

Scenario 3 is only possible in a situation where the multichannel home theater device supports internal Enhanced AC-3 decoding. In this case, no conversion is necessary and the set-top box passes the Enhanced AC-3 stream unchanged while simultaneously creating a two-channel downmix for its stereo interfaces. This scenario enables the use and playback of 7.1-channel configurations. A/V receivers that feature HDMI input without internal Enhanced AC-3 decoding capability will automatically communicate this to the set-top box device via the negotiation protocol on the HDMI link. Hence, the Enhanced AC-3 decoder-converter will automatically provide a standard AC-3 bitstream on the HDMI interface.

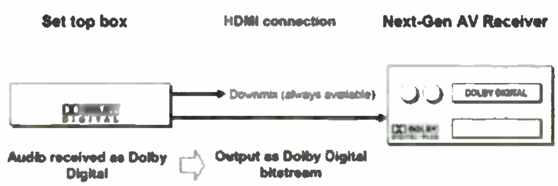


Figure 8: Scenario 4: Broadcast receiver device decoder-converter: AC-3 input signal, HDMI connection (mixed or unmixed)

Scenario 4 replicates scenario 2 with the exception of an HDMI interface connection between the set-top box device and the A/V receiver. There is a standard AC-3 input signal in both cases.

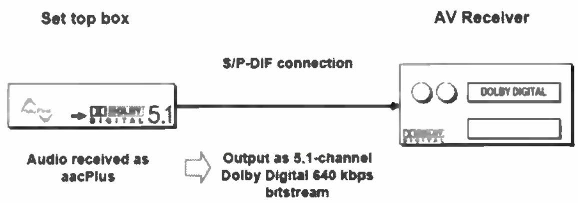


Figure 9: Scenario 5: Broadcast receiver device decoder-converter: aacPlus input signal, S/PDIF connection (mixed or unmixed)

The devices shown in Figure 5 through Figure 9 accept both the Enhanced AC-3 and AC-3 bitstreams and always output the appropriate version of AC-3 for the respective A/V receivers as the consumer switches programs within the TV service. The greater efficiency of the Enhanced AC-3 system allows for a greater number of programs within the broadcast system, while preserving full functionality for legacy receiver hardware and enabling new and extended functionality for new receiver hardware.

9. CONCLUSION

Audio description services are becoming a standard feature on next-generation broadcast platforms. Some European countries even mandate AD services on an increasing number of programs offered by broadcasters, operators and programmers. This paper shows how the support for these services through an enhanced mixing module integrated into a multichannel decoder for next-generation broadcast codecs used in broadcast receivers offers features that go beyond the mandate. These new features offer full home cinema compatibility and therefore enhance the entertainment experience for all audiences.

At the same time, 5.1 surround sound is an important component of the HDTV experience. Existing MPEG2-based HDTV services in the U.S., Europe, and Australia already offer 5.1 surround sound using Dolby Digital technology as documented by DVB and ATSC. With the move to next-generation video coding systems, broadcasters and operators also have many new requirements for audio delivery. This paper shows how new features of the Enhanced AC-3 aacPlus systems meet these requirements while also maintaining compatibility with the more than 40 million existing consumer home cinema systems.

10. REFERENCES

[1] Standards document ETSI TS 102 366, V1.1.1, "Digital Audio Compression (AC-3, Enhanced AC-3) Standard," February 2005.

[2] Standards document ETSI TS 101 154, V1.8.1, "DVB: Implementation Guidelines for the Use of Video and Audio Coding in Broadcasting Applications," July 2007.

[3] Louis D. Fielder et al., "Introduction to Dolby Digital Plus, an Enhancement to the Dolby Digital Coding System," 117th AES Convention, October 2004.

[4] ETSI TS 102 005 V1.3.1 "DVB: Specification for the Use of Video and Audio Coding in DVB Services Delivered Directly over IP Protocols".

[5] Advanced Television Systems Committee: Digital Audio Compression Standard (AC-3, E-AC-3) Revision B; Document A/52B, 14, June 2005.

[6] ITU-T Recommendation H.222.0 / ISO/IEC 13818-1: "Information Technology - Generic Coding of moving pictures and associated audio information: Systems"- and ITU-T Recommendation H.222.0 / Amendment 3 (2004)/ISO/IEC 13818-1: 2000/Amendment 3:(2004): "Transport of AVC video data over ITU-T Recommendation H.222.0 /ISO/IEC 13818-1 streams".

HDTV System onboard the Lunar Explorer Kaguya (SELENE)

Seiji Mitsuhashi[1], Junichi Yamazaki[1], Junichi Tachino[1], Masahito Yamauchi[1], Kazuo Tanimoto[2], Rie Honda[3], Yuichi Iijima[4], Hironori Maejima[4], Hisashi Otake[4], Shinichi Sobue[4]

[1] NHK (Japan Broadcasting Corporation) [2] Meisei Electric Co.,Ltd.
[3] Kochi University [4] JAXA (Japan Aerospace Exploration Agency)

Introduction

The first Earth-rise image was taken by the Apollo 8 in 1968. This still image made an impression on the people all over the world, and it's so beautiful that it has been used as the major image of the lunarscape. And the Apollo 8 shot moving image of the lunar surface by a monochrome tube camera as well. So NHK (Japan Broadcasting Corporation) had been trying to develop the camera system to get a high definition moving image of the Earth-rise.

On September 14, 2007, the lunar explorer Kaguya (SELENE) of the Japan Aerospace Exploration Agency (JAXA) was launched by H-IIA Launch Vehicle No. 13 (H-IIA F13) from Tanegashima Island, taking a lead the "Second Era of Lunar Exploration". The "HDTV (High-definition Moon Camera)" developed by NHK was onboard Kaguya with 13 other scientific observation instruments and started test operations before the lunar explorer was injected into a lunar orbit. The world's first high-definition images taken by the camera onboard spacecraft, including a receding Earth, crater terrains captured from a very short distance, and "Earth-rise" (Fig.1), where the Earth is rising from the Moon Pole, are highly reputed not only in Japan but also across the globe.

Fig.1 Earth-rise from the Moon pole taken by HDTV camera.

Overview of Kaguya

"Kaguya" is a large-size satellite that measures 2.1 meters in width and depth and 4.8 meters in height and weighs approximately three tons, including fuel. Consisting of two sub-satellites - relay satellite and VRAD satellite (50 kg each), 13 observation units, and a High Definition Moon Camera, Kaguya is intended to do scientific research focusing on the "origin and mysteries of the Moon" (Fig.2).

Fig.2 Lunar explorer KAGUYA

The attitude control system onboard is triaxially fixed, and the maximum power generated by solar cells is 3.5 kW. This lunar explorer is in a polar orbit, which is drawing wide attention for the possibility of the presence of ice and in which a satellite passes above both poles (orbital inclination: 90 degrees). It orbits the Moon at an altitude of approximately 100 km, and takes about two hours to complete its orbit. While it is making an orbit, it shifts one degree to the west, meaning that it returns to the original orbital longitude at the 360th orbit (30th day) and the sunshine condition of the Moon surface changes 30 degrees. So, the sunshine condition becomes the same as topography about six moths and 12 months later. Additionally, since the inclination of the Moon's axis of rotation affects shades in the polar areas where the altitude of the Sun is low, it is necessary to observe the Moon for a year.

For this reason, the onboard observation instruments operate almost 365 days and explore the entire Moon in depth. The frequency bands used for the communication system are S-band and X-band. S-band is applied to satellite TT&C (Tracking, Telemetry and Command) operation, whereas X-band communication via a parabola antenna called a high-gain antenna is used principally for the transmission of observation data.

Design policy

The HDTV system is suspended by spacers, screwed to the outside wall of the lunar explorer, and exposed directly to space. It is required to adapt itself to the cosmic radiation environment and the thermal vacuum environment and, at the same time, withstand strong vibration and shock at the time of the launch of the lunar explorer on which the system is installed. On the other hand, large-scale integrated circuits (LSIs) and other small parts must be used due to limited mass and volume, but these have never been used in space. We consequently selected the design policy of repeating various experiments on parts that have been used on the ground and correcting the defects or problems discovered through these experiments. We also decided not to use any electric motor, which is a source of magnetic fields, taking into account the Electro-Magnetic Interference to the onboard instruments for the measurement of faint magnetic fields of the Moon, and to give up the use of a pan-tilt camera platform and a zoom mechanism.

Overview of the HDTV system

The high-definition system (Fig.3) is mounted on the lower module of the lunar explorer. It comprises a camera, a data compression unit, a recorder, a power supply unit, and a control unit, all of which are contained in a single enclosure and thus isolated thermally from the lunar explorer. Inside the enclosure are two heaters equipped with a thermostat, and the outside is covered with multi-layer insulation (MLI) (Fig.4). The power supply unit receives up to 52 V DC from the solar cells of the lunar explorer, and generates and distributes a stabilized voltage of 12 V±5 V to each circuit. A temperature or voltage sensor is installed in each area.

Table1 Specification of HDTV system

Term	Specification
CCD	2/3inch IT-CCD 1920x1080
Compression Mode	HDCAM 144Mbps
Recording Mode	EEPROM 1GB
Transmission I/F	CCSDS packet 10Mbps
Dimensions	460(W) x 280(H) x 420(D)
Weight	16.5kg
Power Consumption	50W (max)

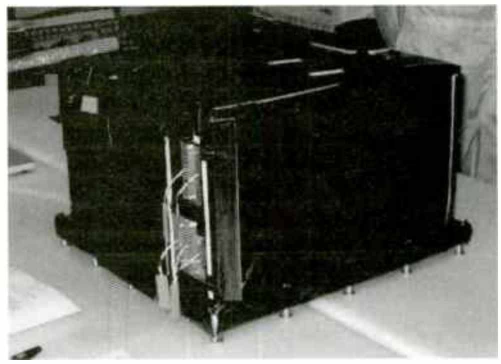


Fig.4 HDTV system

The system has a wide-angle camera (HDTV-WIDE) and a telephotographic camera (HDTV-TELE) aiming at the backward horizon and the forward horizon, respectively. 2/3-inch and 2-million-pixel three CCDs for RGB are mounted in these cameras. All images taken by the cameras are digitally compressed (DCT-compressed within the frame) and recorded into the flash memory (EEPROM). They are available in four recording modes: the standard mode (x1) and three interval recording modes (x2, x4, x8). The memory has a capacity of 1 GB, capable of storing one minute of High-definition moving image.
Each of the compression unit and the recorder is provided with two systems for redundancy. Recorded data are transmitted to the lunar explorer via a high-speed data interface circuit. Command signals from the lunar explorer are received and converted into control circuits by an RTU (Remote Terminals Units) interface circuit and then sent to each unit (Fig.4). The weight and power consumption of the entire system is 16.5 kg and up to 50 W.

Transmission and ground systems

Data are transmitted from the lunar explorer to the Earth in high- and low-speed modes. The high-speed transmission mode is capable of transmitting one-minute moving image data in about 20 minutes at a transmission rate of approximately 10Mbps. The 100-kpbs low-speed transmission mode is a transmission standby mode for still images. The ground equipment installed on the Earth is made up of a telemetry display (Quick Look) responsible for monitoring the working state (temperature, voltage, commands, and camera parameters) of the onboard instruments and a demodulator that decompresses image data transmitted from the lunar explorer (Fig.5). Observation data from "Kaguya" are mainly received by the Usuda Deep Space Center operating in the city of Saku, Nagano Prefecture, and transmitted to the SELENE Operation and Analysis Center (SOAC) located in the city of Sagamihara, Kanagawa Prefecture, through a ground line.

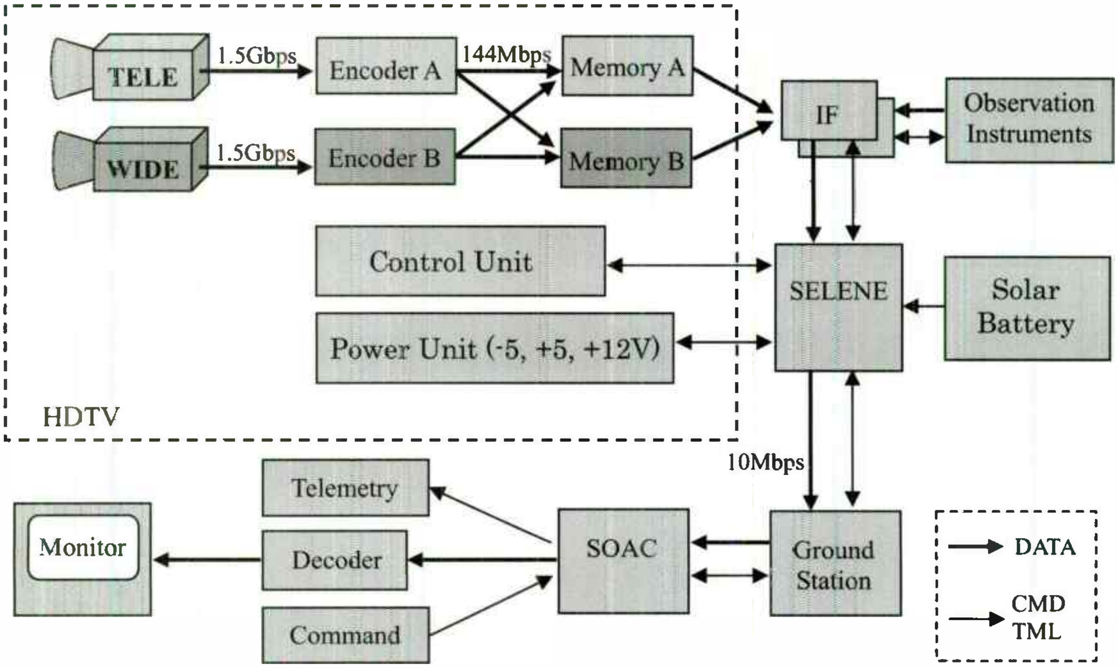


Fig.3 Schematic diagram of HDTV system

Fig.5 HDTV Ground System

Environment and measures

In order to verify its performance, the high-definition system first underwent experiments using individual parts actually adopted on the ground and simulating the space environment and then tests with a single unit incorporating corrective measures. In these experiments. we found that the lens became yellow due to the effect of gamma rays, and added 1-cm-thick quartz glass to its front face. We set an infinite distance without the focus ring, and replaced the aperture regulating ring with an F5.6 fixed round aperture to make the system resistant to vibration and shock. Lens-integral type color conversion (CC) and neutral density (ND) filters were also adopted (Fig.6, Table 2). Exposure is adjusted by the electronic shutter of the camera. A cooling fan is incorporated in cameras used on the ground to cool the inside, but we selected a structure designed to radiate heat to the case with a heat transfer plate because the system was to be used in a vacuum (Fig.7). We made the structure more resistant and robust to vibration and shock by increasing the number of clamping points, and introduced a flexible cable to the areas where different vibrations collide because it can relieve force. The camera is capable of issuing test signals and adjusting images, including electric gain, gamma, pedestal, and white balance, with the aid of command control. Particularly important items can also be serial-controlled and parallel-controlled (Table 3). As measures against cosmic radiation, software was stored in three locations in the memory, and triple modular judgment processing was also incorporated. The high-definition system contains highly integrated LSIs. These LSIs include a digital signal processor (DSP) processing images, an encoder chip (ENC) that compresses digital signals, and a central processing unit (CPU). If high-energy particles traveling around in space collide with these LSIs, an abnormal phenomenon (image destruction or runaway) may occur or an over current may continue flowing during operation. To deal with it, we added a current limiting circuit or timer circuit. We discovered a pixel defect (flaw) with the CCD imaging device, and developed a white flaw correction unit that restores demodulated images on the ground.

Table2 Specification of lenses

Term	Telephoto lens	Wide-angle lens
Focal length	10mm	35mm
View angle	50.1(H)x29.5(V), 56.3(D)	15.5(H)x8.7(V), 17.8(D)
F number	F5.6 (fixed)	F5.6 (fixed)
ND filter	1	1/8
Dimensions	112.5(L)x60(D)	81.5(L)x60(D)
Weight	330g	260g
Mounting Angle	22.5 deg.	18.5 deg.

Table3 Function of picture control

Term	Specification
Electric Gain	-6, -3, 0, +3, +6, +12, +18dB
Exposure Method	Electric shutter
Auto Mode	Peak / Average control (1/63.4 ~ 1/1,983sec)
Manual Mode	1/63.4 ~ 1/16,000sec
Adjustment item	Knee point, Knee slope Gamma, Black Gamma Pedestal, Auto White Balance Auto Black Balance
Test signal	Color Bar, CAL signal
Super Impose	Central coordinate Shutter speed
Control signal	Serial / Parallel signal

Table4 Experiments for HDTV system

Terms	
Radiation experiment	Total dose : Gamma ray
	Single event (LET Max37.4MeV) Nuclide H, N, Ne, Si, Ar, Kr
Temperature experiment	-30 ~ +60C +70C
Magnetic field test	
Vibration experiment	X,Y, X axis Sine wave : 10G, 22G Random wave : 9G, 16G, 22G
Shock test	X, Y, Z axis 1,000G
Thermal vacuum test Thermal equilibrium test	Temperature cycle test etc.
EMC test	30Hz ~ 40GHz

Fig.6 The lens modified for HDTV system (left) and the normal lens (right)

Fig.7 The compact HDTV camera modified for HDTV system (left) and the normal camera (right)

Comprehensive test

Using onboard experimental equipment (flight model) completed after modifications, we conducted a random vibration experiment in which it was subjected to strong vibration of the equivalent level to the vibration at the time of the launch of a rocket (overall 22 G) for 80 seconds, and an shock test of 1,000 G, which is the shock caused when a satellite is separated, and verified that it withstood them. We also carried out thermal vacuum and magnetic field tests on the equipment alone, as well as a mutual interference test with each observation instrument on the equipment combined with the lunar explorer besides these tests, and discovered no problem (Table 4).

Imaging programs and imaging tools

The high-definition camera is used for three programs - "Base Line," "Option," and "Checkout." "Base Line" is the setting of goals to achieve and the program for taking Earth-rise from the horizon of the Moon when the Earth is "full" like a full moon (this opportunity comes twice a year). The "Option" program is carried out with the effect on the other scientific observation instruments minimized and the data transmission time adjusted. The "Checkout" program is used for periodical operation checking and white flaw measurement. Since opportunities for high-definition imaging are limited and the period of continuous imaging is short, simulations that calculate the good timing of imaging with a high degree of accuracy were indispensable for satisfactory results. In this project, we developed software (operation support tool) capable of narrowing down targets to be imaged by inputting "Kaguya's" orbital elements and physical relationship with the Sun and the conditions relating to the high-definition system, such as the view angle. On the PC screen, a lunar map (data from the Clementine satellite) was displayed, and the orbit of the lunar explorer was expressed in a line and the altitude of the Sun in a colored line. The field of view at the start and end of high-definition imaging was shown as a trapezoid (Fig.8). In addition to this screen, we created screens for displaying the Moon's horizon and physical relationship with the Earth and other planets (Fig.9). They are all intended chiefly to take images of "Earth-rise" and "Earth-set." With these imaging tools, we succeeded in actually taking and transmitting images of about 30 selected objects between late October and mid-December. During this period, the other observation instruments were verified as initial function check one by one, and we could use this time for the transmission of high-definition images. Since mid-December, however, all science mission instruments have been in regular operation and transmitted observation data, and the frequency of the transmission of high-definition images has been reduced.

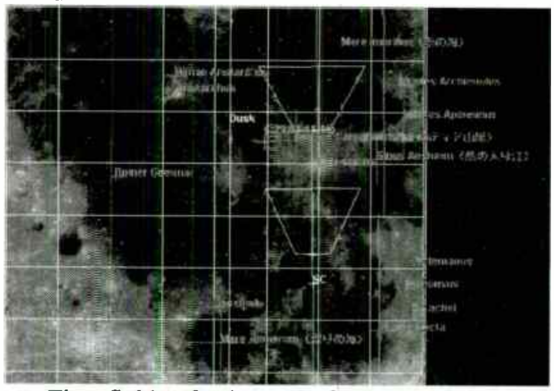


Fig.8 The field of view at the start and end of HDTV-WIDE imaging.

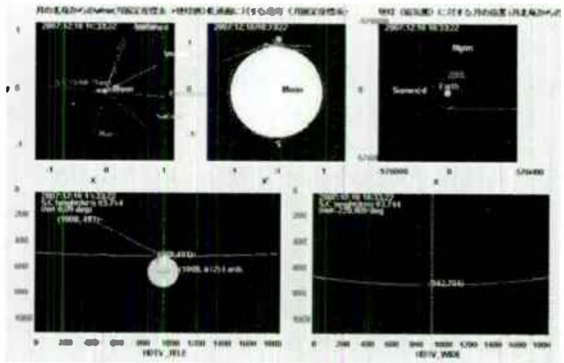


Fig.9 Moon's horizon and physical relationship with the Earth and other planets.

Imaging procedure

The high-definition system is operated at SOAC in Sagamihara. Although an imaging plan is formulated using the operation support tool, with the length of shadows taken into account, there is an error in predicted values of orbit information (the position and posture of the lunar explorer). So, the time is corrected based on the latest orbit information, and imaging commands are created accordingly. These commands are executed by "real-time command operation" in which commands are carried out while communication is linked, or by "stored command operation" in which commands are automatically fulfilled at a specified time. At present, "stored command operation" is mainly used to register imaging commands with the lunar explorer a day before imaging. "Downlink," which means transmitting imaged or recorded data from the lunar explorer to the Earth," is executed by "real-time operation" in a time zone in which the lunar explorer is visible from Japan (visible zone). At the Usuda Deep Space Center and Uchinoura Deep Space Center, the receiving stations, 64-meter-diameter and 34-meter-diameter receiving parabola antennas are installed respectively. This antenna are, however, shared with other satellites, including "Hayabusa,". Data received and accumulated at Usuda and Uchinoura are sent to SOAC in Sagami through a ground IP line. It takes about an hour to demodulate these data to complete a one-minute high-definition image from the start of downlink.

Obtained images

On September 29, 2007, we succeeded in taking an image of a "receding Earth" with the telephotographic camera at a distance of 110,000 km from the Earth when "Kaguya" was on its way to the Moon (Fig.10). In this process, the lunar explorer changed its attitude by gas injection (thruster) and maintained it with a flywheel (reaction wheel). After October 29, images of the Moon were taken from a lunar orbit altitude of approximately 100 km. As the lunar explorer approaches the Moon, the craters on its surface look like they are getting closer to us. "Earth-rise" and "Earth-set" (Fig.11) in which the Earth looks as if it comes up from and goes down below the horizon of the Moon are also observed for the same reason. The crater indicated by an yellow allow of Fig.11 is Shackelton (19km in diameter, located at 89.9 deg S, 0.0 deg E), which is considered to have the permanent shadow inside. Our data will be utilized for studies of crater walls at the polar region, to search for relativity bright icy patches inside the permanent shadow.

The producer requested us to introduce many topographic features in a short period of time and show the reality of how fast the lunar explorer flies.

As an answer to this request, we took images of the Moon in x8 interval recording mode using the wide-angle camera, and captured images showing that the lunar explorer dynamically traveled a wide area, or about 24 degrees in latitude (Fig.12). The prominent crater of Fig.12 is Jackson (71km in diameter) which lies on the far side of the Moon at 22.4 deg N, -163.1 deg E. An image of Fig. 13 is the close-up view of crater Jackson. It enables the detailed observation of complex from the central peak and the crater wall.

These images were uploaded to JAXA's website. NHK introduced them in a special program titled "Lunar orbit explorer 'Kaguya' fathoming the mysteries of the Moon", and they were also broadcasted on Canada's Discovery Channel.

Fig.10 An image of the "receding Earth" at a distance of 110,000 km from the Earth.

Fig.11 An image of Earth-set taken by HDTV-TELE. The crater indicated by an yellow arrow is Shackelton.

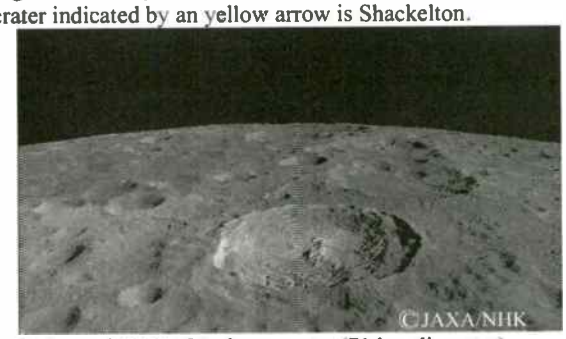


Fig.12 An image of Jackson crater (71 km diameter).

Fig.13 The close-up view of the wall and central peak of crater Jackson.

Future activities

This is the first time in the world that the high-definition system was carried into outer space 380,000 km away from the Earth, and the images captured created a sensation not only in Japan but also all over the world. The imaging of the Moon using the high-definition system will continue for a year.

Acknowledgement

The high-definition television system onboard Kaguya proved that even high-definition equipment for ground use can withstand the harsh space environment if proper modifications and tests are carried out. This result shows the technological possibility of loading the high-definition system into future lunar excursion modules. The high-definition system carried in "Kaguya" is a fruit of Japan's high technology and of the cooperation of many manufacturers, including Fujinon, Ikegami Tsushinki, Sony, Fujitsu, and NEC.

Reference

Initial Results of Imaging of Lunar Features by High-Definition Television (HDTV) on board SELENE(KAGUYA), R. Honda et al. , LPSC 39, 2008 (submitted)

Codecs, Compression Systems & Scaling for Video

Thursday, April 17, 2008
9:00 AM – 12:00 PM

Chairperson: Graham Jones
NAB, Washington, DC

10 Bit High Quality MPEG-4 AVC Video Compression

Matthew Compton, TANDBERG Television, Part of the Ericsson Group, Southampton, UK

Practical Applications of Compression Standards

Todd Roth, NEXIO Server Systems, Broadcast Communications Division, Harris Corporation, Mason, OH

Providing Spatial Scalability Using Scalable Video Coding to Mobile Broadcasting

Sangjin Hahm, Korean Broadcasting System, Seoul, Korea

Scalable Video Coding (SVC) and Broadcast Delivery of 1080P

Elie Sader, Harmonic Inc., Sunnyvale, CA

Understanding and Implementing JPEG 2000 Compression for Long-form EFP Acquisition

John Naylor, Thomson, Beaverton, OR

Bridging the Gap with HD Transcoding

Tim Simerly, Texas Instruments, Dallas, TX

10 bit high quality MPEG-4 AVC video compression

Matthew Compton
TANDBERG Television - Part of the Ericsson Group.
Southampton, UK

Abstract

HD MPEG-4 AVC is a standard that is well established in the broadcast industry for over 3 years. It offers premium compression performance at significantly lower bitrates then MPEG-2. However, there are still aspects of the H264 AVC toolset which are not widely used, in particular High 10 Profile (Hi10P) - which supports full resolution 10 bit video encoding. From contribution and distribution (C&D) to satellite direct-to-home there are applications for 10 bit HD operation that can enhance performance.

Direct-to-home broadcasters are competing with the high definition DVD market and therefore require optimum compression performance. In particular, areas for improvement are plain backgrounds, which can suffer from colour contouring etc. This is also becoming increasing relevant as consumer displays migrate to 10 bit technologies, therefore there will be a demand for ultimate dynamic range. Likewise, C&D markets desire 10 bit HD video through the entire broadcast production chain, this can be achieved in addition to 4:2:2 colour processing with the H264 AVC High 4:2:2 Profile (Hi422P).

This paper explores the advantages and disadvantages of HD MPEG-4 10 bit encoding at a variety of bitrates on different types of content. Comparisons are made at the different operating points, demonstrating where gains maybe achieved.

Introduction

The high compression efficiency of MPEG-4 AVC makes it an attractive solution compared to MPEG-2. Broadcasters have the choice of increasing the channel density through a given transmission medium (compared to MPEG-2) or increasing video channel quality. In addition, as technology advances and MPEG-4 AVC compression is better understood, incremental quality improvements are achieved over time.

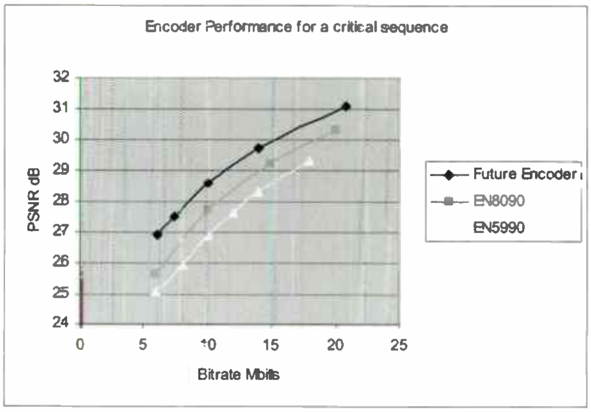


Fig. 1 showing PSNR comparison of MPEG-4 AVC equipment.

Fig. 1 shows the compression performance improvements of existing equipment over the last three years. An indication is also given to what might be expected in the future.

The coding gain of MPEG-4 AVC can be attributed to a number of new coding tools which are not available in MPEG-2. Detailed descriptions of the new coding tools have been published [1, 2]. A list of the most important coding tools is shown below:

- 9 4x4 intra prediction modes
- 4 16x16 intra prediction modes
- 7 inter prediction modes from 16x16 down to 4x4 block sizes
- 1/4 pel motion compensation
- advanced bi-directional (B) pictures
- motion compensation from outside of the picture
- multiple reference pictures
- integer transformation
- in-loop de-blocking filter
- Context-Adaptive Binary Arithmetic Coding (CABAC)

Measuring Image Quality

Perceived video quality can be by its very nature subjective, especially if there is any degradation from the source due to compression or video processing.

In order to make comparisons in this paper, a mechanism needs to be selected to compare sequences compressed at different bitrates and/or with different coding tools.

Common mechanisms are:

- **PSNR** (Peak Signal to Noise Ratio) is most commonly used, and is built into the JVT's reference software (JM 12). In addition it is widely used through the broadcasting industry. PSNR effectively measures pixel by pixel differences that are not necessarily visible to the human eye. This is considered to be an objective measurement

- **JND** (Just Noticeable Difference) is a general psychophysical measurement unit defined as: 'The difference between two stimuli that (under properly controlled experimental conditions) is detected as often as it is undetected' [3] to model human perceived video quality. This is considered to be a subjective measurement.

- **DMOS** (Differential Mean Opinion Score) likewise this is a measurement based on the human vision system and is considered to be subjective [3].

Primarily, this paper is considering the best quality and dynamic range that can be achieved during the compression process. It is not expected that there will be any obviously compromising artefacts. Therefore, for the moment PSNR and visual inspection are the sole measurements.

In addition, the PSNR metric may consider chroma if calculated using the following weightings:

$$PSNR = (0.8xY_{PSNR}) + (0.1xU_{PSNR}) + (0.1xV_{PSNR})$$

This is used for all curve comparisons in this paper.

Levels, Profiles and Operating Points

ISO/IEC 14496-10 otherwise known as MPEG-4 AVC defines a number of profiles applicable to video conferencing, broadcast and streaming applications:

- The **Baseline Profile** is mainly intended for video conferencing and streaming to mobile devices. Its simplistic applications mean that it does not support bi-directional frames, interlace or CABAC entropy encoding.
- The **Main Profile** allows bi-directionally predicted (B) frames with two direct modes: spatial and temporal and weighted predictions. Furthermore, it supports all interlace coding tools including Picture Adaptive Field/Frame coding (PAFF) and Macro-Block Adaptive Field/Frame coding (MBAFF) as well as CABAC. Main profile is most widely used for HD and SD broadcasting applications.
- The coding tools of MPEG-4 profiles which go beyond Main Profile are summarised as Fidelity Range Extensions [4]. In particular the **High Profile** allows adaptive 8x8 integer transforms, intra 8x8 predictions modes and scaling lists.
- The **High 10 Profile** allows coding of 4:2:0 video signals with 10 bit accuracy and
- the **High 4:2:2 Profile** allows coding of 4:2:2 video signals with 10 or 8 bit accuracy.

This paper considers two applications for **HDTV** MPEG-4 AVC. Here is a description of how they generally operate today:

Direct-To-Home

Typically direct-to-home satellite broadcasters (DTH) use **Main Profile** (i.e. 4:2:0 8bit video) at bitrates ranging from 6-20Mbits. Delay is not a critical issue for these applications, therefore a 15Mbit encoding buffer and 3x B pictures are considered quite acceptable.

Contribution and Distribution

Contribution and Distribution applications are varied and cannot be generalised. However, there is a common characteristic; in that end-to-end delay is critical. Therefore the number of B pictures and encoding buffer size is kept to a minimum. Bitrates can vary in the range of 8-90Mbits depending on the quality point that is required. In addition, **Hi 422 Profile** is used to achieve 4:2:2 8 bit video compression.

In both cases, 10bit processing is explored to demonstrate the gains that maybe achieved.

The importance of 4:2:2 video compression

Traditionally C&D applications have demanded 4:2:2 MPEG-2 video compression. This prevents the need to repeatedly down-sample and up-sample the chroma channels. As mis-matches in spatial positioning of the chroma filters and/or soft filter roll-offs can significantly degrade the chroma quality.

In addition, repeated concatenation of video encoding processes can impede compression performance through the later stages in the chain. Therefore, preserving optimum video quality is desirable from the very first stage.

Furthermore, 10 bit 4:2:2 is the most commonly used studio and production format. Therefore, from a compression efficiency prospective, it is important to know at what bit rate the picture quality of 10 bit 4:2:2 would be noticeably better than 8 bit 4:2:2.

Comparison of 8 and 10 bit encoding for DTH applications

Main Profile is compared with High 10 Profile for a set of sequences. Both use the following common tools:

- CABAC
- Inter prediction modes 16x16, 16x8, 8x16, 8x8
- Intra prediction modes 16x16, 4x4 (8x8 in Hi10P)
- 3x B pictures
- Field coding
- De-blocking filter on
- RDO (Rate Distortion Optimisation)

A software model is used to compress the sequences, which represents a viable realtime encoder.

The sequences selected are reasonably clean from noise and have large plain areas as well as detailed areas. They contain fade-to-black or subtle movement which can cause contouring or posterisation in the plain regions.

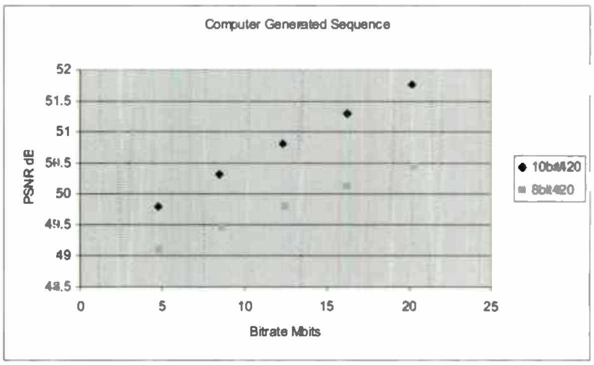


Fig. 2 a computer generated red landscape that subtly changes.

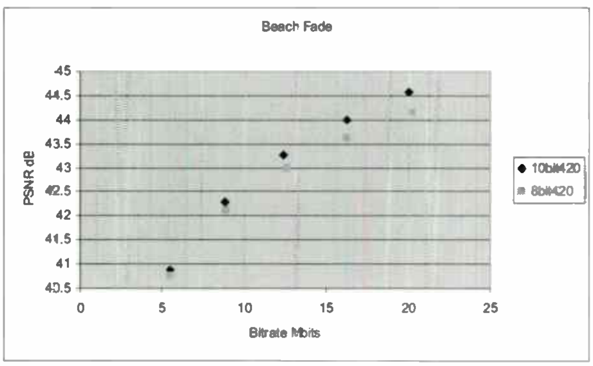


Fig. 3 a beach scene with blue sky, running horses and fade to black.

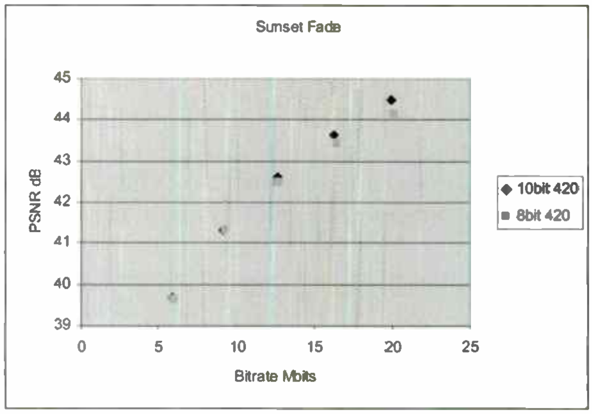


Fig. 4 a sunset scene with palm trees and fade to black.

Clearly all three fig. 2,3,4 show an improvement in terms of PSNR. But also there are subjective improvements which are visible on LCD and CRT screens. The differences are within the plain areas and contain less blocking or stripe artefacts in the 10bit version compared to the 8bit.

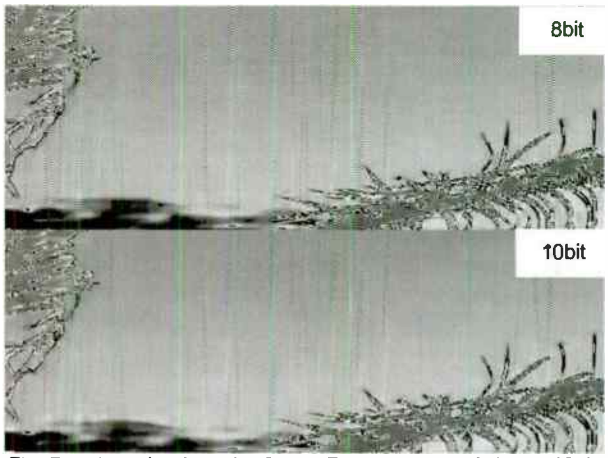


Fig. 5 a x4 section from the SunsetFade sequence 8bit and 10bit.

In order to appreciate the visual differences, fig. 5 shows an amplified (x4) section of the sky from the 9Mbit point of the SunsetFade sequence. The macroblock boundaries are much more noticeable in the 8bit version.

Further inspection reveals that there is only a luma level difference of 1 or 2 (in 8 bit terms) between neighbouring blocks, but because the entire macroblock is so plain the eye is very sensitive to this change. When this data is presented to the in- loop de-blocking filter, there is nothing the filter can do to hide this boundary. However, in the 10 bit case, the in loop de-blocking filter has two extra bits of precision, therefore intermediate levels can be created and hence soften the macroblock boundary. This can be verified by switching the de-blocking filter off in the 10 bit experiment.

It is important to note that the difference between 8 bit and 10 bit pixel precision is not noticeable when comparing the source. The justification for this is that 8 bit source images contain enough noise to hide any contouring which might otherwise be visible in such areas (i.e. a dithering effect). However, the inherent noise reducing properties of video compression can expose the luma and colour graduations and lead to a posterisation effect, even at relatively high bit rates.

Interestingly, there are two mechanisms in MPEG-2 which unintentionally help to hide such artefacts: DCT inaccuracy and mismatch control. Mismatch control is required as a result of DCT inaccuracy. The combined effect of these two mechanisms is to inject a small amount of noise at the output of MPEG-2 decoders. This noise acts like a dither signal. Therefore, although MPEG-2, like most compression algorithms, eliminates some of the original source noise due to the low-pass characteristics of its weighting matrices, it reintroduces enough DCT noise at the decoder output to reduce posterisation.

The disadvantage of MPEG-2 DCT inaccuracies is, of course, decoder drift. This limits the number of predictions that can be made from previous predictions before a refresh is required, i.e. it limits how many P frames can be coded between I frames. MPEG-4 AVC does not suffer from DCT inaccuracies because its integer transformations have no rounding differences between encoder and decoder. Therefore, it has no need for mismatch control and there is no limit to the number of predictions that can be made from previous predictions.

Finally, common to all three sequences, the PSNR difference (8bit vs. 10bit) approaches >0.25 dB when the average encoding QP for the sequence is in the order of <20.

Comparison of 8 and 10 bit encoding for C&D applications

High 4:2:2 profile 8 bit is compared with High 4:2:2 profile 10bit for a set of sequences. Both use the following common tools:

- CABAC
- Inter prediction modes 16x16, 16x8, 8x16, 8x8
- Intra prediction modes 16x16, 8x8, 4x4
- 1x B pictures
- Field coding
- De-blocking filter on
- RDO (Rate Distortion Optimisation)

The first sequence is a 10bit CrowdRun scene provided by SVT [5]. It contains an area of high motion, and a plain blue sky. In addition, the sequence is relatively noisy. The second sequence is 10bit cricket scene with grass and a plain sky, this is relatively free from noise.

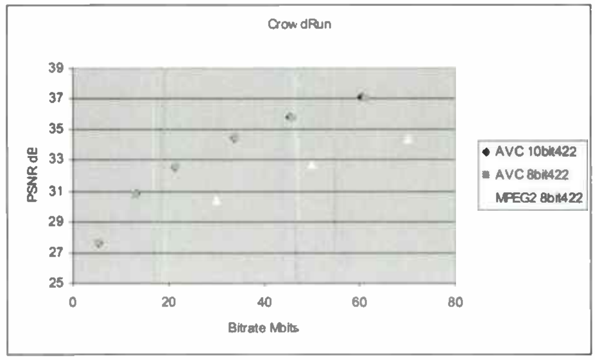


Fig. 6 a crowd of people running in a marathon, with a blue sky.

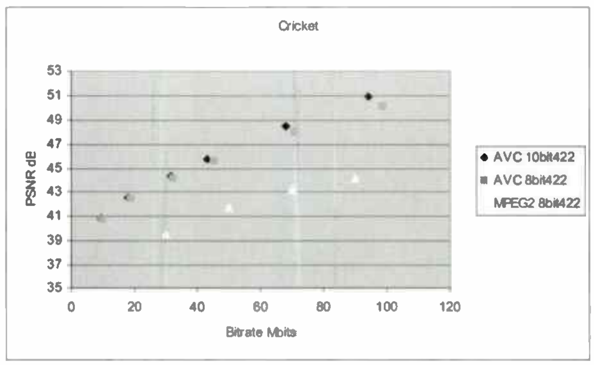


Fig. 7 a cricket scene with a plain sky.

The CrowdRun sequence in Fig. 6 shows little gain using 10bit precision. Of course, the PSNR measurement used is including the likeness of the noise between source and compressed sequences.

As discussed previously, the significance of noise in the source material is important. A) it provides a dither in the plain areas that may hide step changes in luma. B) the transform and quantisation process used in the MPEG-4 AVC algorithm removes noise at higher QPs.

Based on these points, there is no need to have 10bit precision in the de-blocking filter, and there is enough dither in the picture. Secondly, the QP would have to go very low (i.e. high bitrate) in order to preserve the exact noise pattern in the source and hence not showing the benefit of 10 precision in terms of PSNR (for the given bitrate operating range). It is quite noticeable visually that across the bitrate range, the

MPEG-4 AVC sequences have less noise than the source (at 60Mbits the average QP is 23).

The MPEG-2 curve is added for comparison and is significantly under the MPEG-4 AVC curves. Visually, there are some MPEG-2 compression artefacts in the high motion area. Fig. 8 shows an example of processing noise at 50Mbits.

Fig. 8 artefact comparison between MPEG-2 and MPEG-4 AVC at 50Mbits.

Conversely the cricket sequence in Fig. 7 is similar to the sequences used for the DTH experiments. The difference in performance can be measured at bitrates within the operating range. In PSNR terms, MPEG-4 AVC 10bit yields better performance than MPEG-2 and MPEG-4 AVC 8bit.

In summary, the advantage of using 10bit compression in C&D applications, is that: when noise is minimal, there is opportunity for 10bit precision to be preserved during the compression process. This will be useful for downstream processing (e.g. gamma or colour correction) where luma/chroma levels are changed in a non-linear fashion Hence, the extra precision could potentially mask any perceivable artefacts.

LCD Screen technology

Currently, most LCD screens available are 8bit resolution. However, several leading consumer electronics manufactures are producing 10bit capable technology.

The DTH sequences discussed in this paper were also viewed on a Sony 8bit LCD screen, and there is still a perceived difference between 8bit and 10bit MPEG-4 AVC compression. It can only be assumed that whist the final display is 8bits, intermediate processing (i.e. non linear gamma and colour correction) are on the full 10bit resolution. Hence, avoiding large step changes in 8bit terms.

In addition, 10bit LCD screen owners may further benefit from 10bit broadcast Television transmission.

Conclusions

In this paper, advantages of 10bit MPEG-4 AVC for C&D and DTH applications have been illustrated.

In the case of C&D, it has been shown that the High 4:2:2 Profile of MPEG-4 AVC can deliver a level of picture quality (within a practical bitrate range) that could not have been achieved in MPEG-2 (since MPEG-2 does not support 10 bit coding) and MPEG-4 AVC 8bit. Also, at the points where 10bit coding does not offer any advantage, there is not a degradation in compression performance. This makes High 4:2:2 10bit profile an attractive and safe solution.

In the case of DTH, it is unlikely that the bitrate ranges used will show a PSNR difference between 8bit and 10bit precision. However, the extra precision that 10bit offers gives more scope for the de-blocking filter to hide macroblock and transform boundary processing.

Further work maybe to investigate if it is possible to manipulate the MPEG-4 AVC encoding process to choose the appropriate coding mode, QP and/or transform - that maximises the amount of noise (i.e. dither) that gets through the compression process for a given macroblock.

References

[1] T. Wiegand, et al, 'Overview of the H.264 AVC Video Coding Standard', IEEE Transactions on Circuits and Systems for Video Technology, July 2003.

[2] M. Goldman, A.M. Bock, 'Advanced Compression Technologies for High Definition', IEE Conference 'IT to HD', London, Nov. 2005.

[3] K. Ferguson, 'An Adaptable Human Vision Model for Subjective Video Quality Rating Prediction Among CIF, SD, HD and E-Cinema', International Broadcasting Convention, Amsterdam, Sept. 2007.

[4] D. Marpe, et al, 'H.264/MPEG-4 AVC Fidelity Range Extensions: Tools, Profiles, Performance, and Application Areas', Proc. ICIP, Genova, Sept. 2005.

[5] L. Haglund. (2006, The SVT high definition multi format test set). Swedish Television, Stockholm.

Acknowledgements

Many thanks to Alois Bock and Greg Baggott for assisting with the preparation and analysis of data for this paper.

Practical Applications of Compression Standards

Todd Roth, VP Technology, NEXIO Server Systems
Broadcast Communications Division
Harris Corporation, United States of America

1. Introduction

The growing need to store and transmit more and more media in a cost-effective manner is the primary driver behind video compression. Although at a basic level, compression is a tradeoff between image size and quality, the mechanisms of degradation vary with differing compression standards and implementations.

Compression efficiency measures the reduction in bit rate at a given image quality. Unmeasured in compression efficiency metrics is image facility — the workflow-oriented characteristics that may limit the application of a given codec.

As shown in Table 1, even the traditional baseband interfaces such as SDI, HDI, DVI and HDMI introduce a level of image compression by chroma sub-sampling; for instance, moving from the 4:4:4 color space to 4:2:2 compresses the video signal by one third (3:2).

Table 1: Data Rates of Common Video Formats (NTSC Rates)

					Color Space			Data Rate	Via	Color Space			Data Rate
Format	W	H	fps	bit	Y	Cb	Cr	Mbps	SMPTE	Y	Cb	Cr	Mbps
SD	720	480	29.97	10	4	4	4	**310.73**	259 (270 Mb)	4	2	2	**207.15**
480p	720	480	59.94	10	4	4	4	**621.46**	292 (1.5 Gb)	4	2	2	**414.31**
720p	1280	720	59.94	10	4	4	4	**1657.22**	292 (1.5 Gb)	4	2	2	**1104.81**
1080i	1920	1080	29.97	10	4	4	4	**1864.37**	292 (1.5 Gb)	4	2	2	**1242.92**
1080p	1920	1080	59.94	10	4	4	4	**3728.75**	424 (3 Gb)	4	2	2	**2485.83**

Digital acquisition formats typically use compression ranges between 4:1 (HDCAM SR) and 50:1 (HDV). Along with various codec choices, data-reduction techniques such as horizontal sub-sampling, intraframe vs. interframe compression (motion estimation) and 10- vs. 8-bit resolution all affect image quality, media utility and impact workflows. Preservation of compressed media quality is best achieved by avoiding unnecessary transcoding operations.

Coupled with modern IT network and storage infrastructure, media accessibility through a workflow is becoming more and more file based. Achieving practical data rates for efficient storage and file-based workflows requires substantially more compression. This added dimension to the media file format often complicates the choice of codec. Other confusion can occur by confusing the video acquisition format with the actual codec.

Compression techniques also balance real-time performance, the complexity of encoding and decoding, as well as the "availability" of the codec algorithm. Newer codecs typically achieve higher compression efficiency, but also have increased processing requirements, which increase the cost of encode and decode devices. Proprietary codecs may achieve better results, but might make the compressed media less accessible to other applications and devices. An often overlooked consideration is long-term availability of a codec to decode archived material.

Compression and Workflow

Specific codec choices are best understood by examining application requirements at different stages of the workflow (see Appendix 1: Media Formats). Some principles — such as, quality that is lost cannot be later recovered, and minimizing generation loss from re-encoding — are simple to grasp. Other practices are better tailored to the stage of the workflow. Since compression, specifically color-space reduction and motion estimation can degrade image quality in a non-linear manner when material is decompressed and recompressed, it is important to minimize these operations and associated generational loss.

1.1. *Digital Intermediate (DI)- First Generation*

Video for telecine transfer application aspires to preserve the most image quality and facility. While usually uncompressed, very light compression (mathematically or visually lossless) can often be tolerated.

- Must be intraframe coded
- Full color space (RGB 4:4:4)
- 10-bit (or better)
- 1920x1080p (or better)
- Target data rate: > 300 Mbps – high-performance storage subsystems

1.2. *Contribution & Acquisition - First Generation*

High-quality video over the SMPTE 292/434 interface that can be further compressed to aid in workflow and storage requirements. The strict requirement for intraframe coding can be relaxed if the post-production workflow is carefully managed.

- Should be intraframe coded
- Broadcast color space (YUV 4:2:2)
- 8- or 10-bit
- Full resolution: 1280x720p, 1920x1080i or 1920x1080p
- Target data rate: 100 > 300 Mbps – acquisition media, Gigabit LANs

1.3. *ENG - First Generation*

Quality Video for electronic news gathering. Since ENG originated material usually goes through a simplified post-production process generational loss is less of an issue. Higher compression ratios supported by interfame coding and color-space reduction can be tolerated to reduce cost.

- Can be interframe coded
- Reduced broadcast color space (YUV 4:2:0)
- 8-bit
- Full or reduced resolution: 1280x720p, 1440x1080i or 1920x1080i
- Target data rate: SD - 25 > 50 Mbps, HD - 35 > 100 Mbps – acquisition media, Gigabit LANs, WANs

1.4. *Mezzanine & Post Production - First Generation IN, Third Generation OUT*

While a standard mezzanine format across post may sound attractive, it is only a source on generational loss. With modern file-based editors, and by keeping timelines in native formats (always referencing the source material for layered operations), it is possible to keep media within a couple of generations of original throughout the post process. Even in the case of interframe compression, when a simple splice can result in recoding of new reference frames, generational loss can be limited by maintaining reference to the source material.

Following this model, the final output from post will likely require a single transcode operation to the desired distribution compression format. The goal of this should be no worse than third generation.

Preserving as much source material quality as possible is achieved by compressing the video only enough to efficiently share media across multiple platforms (real-time access).

- Seeks to minimize generation loss; can use editing tools that support all native compression formats
- Keeps material as native as possible
- Can be intra- and interframe coded
- Broadcast and reduced color space (4:2:2 and 4:2:0)
- 8- and 10-bit
- Full and reduced resolution: 1280x720p, 1440x1080i or 1920x1080i
- Target data rate: 25 > 150 Mbps – Gigabit LANs, WANs, SANs

1.5. *Distribution & Archive – Second to Third Generation*

Distribution of material from networks to affiliates or content aggregators over fixed bandwidth links and wide area networks requires maintaining image quality at mid-range bit rates. This essentially means interframe compression is required.

Preserving enough image quality to allow limited further manipulation (e.g., splicing, logo insertion) can be performed without degradation visible to the consumer. These requirements also generally fulfill archiving needs.

- Material should be in common formats
- Should be interframe coded
- Broadcast and reduced color space (4:2:2 and 4:2:0)
- 8-bit
- Full and reduced resolution: 1280x720p, 1440x1080i or 1920x1080i
- Target data rate: 25 > 50 Mbps – Satellite links, WANs

1.6. *Broadcast Transmission – Fourth Generation*

Full-resolution transmission to home consumers requires significant compression. Here, the lowest bit rate to reproduce full-resolution, quality images demands compression efficiency.

As consumer high-definition receivers get better —supporting full 1980x1080p resolution, improving contrast ratios, and 120 Hz refresh rates — the 1080i and 720p transmission resolution standards are becoming the weak link. On better-quality receivers, consumers notice more image quality problems and are becoming dissatisfied; poor broadcast practices and lack of uniform HD service only makes matters worse.

- Needs to be ATSC/DVB format (currently MPEG-2, future AVC)
- Must be interframe coded
- Reduced color space (4:2:0)
- 8-bit
- Full-resolution: 1280x720p and 1920x1080i
- Target data rate: > 20 Mbps – ATSC/DVB

1.7. *High Definition DVD – Fourth Generation*

For the same reasons as transmission, high-quality, full-resolution images are required.

- Needs to be MPEG-2, AVC or VC-1
- Must be interframe coded
- Reduced color space (and 4:2:0)
- 8-bit
- Full and resolution: 1280x720p and 1920x1080i
- Target data rate: 8 > 25 Mbps – HD-DVD, Blu-ray

1.8. *Internet Downloadable Content & IPTV – Fifth Generation*

Getting an acceptable full-resolution image streamed over the Internet is an achievable but challenging goal. Media delivery products for connecting the computer to the television, as well as stand-alone products such as Apple's I-TV are commonly available. Television sets with Ethernet interfaces for direct IP media playback (i.e. YouTube) even have appeared at this year's Consumer Electronics Show. Highest possible compression efficiency and a software supported codec is a requirement.

- Needs to be software supported codec, AVC or VC-1
- Codecs with rights management/content protection layers (MPEG-4/VC-1)
- Must be interframe coded
- Reduced color space (and 4:2:0)
- 8-bit
- Full resolutions
- Target data rate: 1 > 14 Mbps – real time streaming (IPTV) as well as non-real-time down load

1.9. *Mobile & Handheld – Fifth Generation*

Reduced resolution transmission to Internet and mobile consumers requires scaling and significant compression. Sub-megabit data rates and requirements to support simpler playback codecs are made possible by reducing resolution.

- May require multiple resolutions for multiple target devices (QCIF, CIF, QVGA, etc.)
- Codecs with rights management/content protection layers (MPEG-4/VC-1)
- Must be interframe coded
- Reduced color space (and 4:2:0)
- 8-bit
- Reduced resolutions
- Target data rate: >1 Mbps – mobile broadcasting, ADSL

1.10. *Transcoding vs. Transrating*

Since the transcode process involves decompressing and recompressing an image, additional image fidelity is lost in the process. This is especially true when moving from one codec format to another. However, when staying within a given codec format it is possible to preserve and reuse certain internal information to maintain compression efficiency while reducing bit rate.

This process is known as partial decode transrating.

2. Compression Mechanisms

The basis for video compression is derived from the characteristic that most images contain redundant information in the form of correlation between neighboring pixels. Pixels that are either the same as, or similar to, adjacent or nearby pixels can be represented with less data. Differences that are imperceptible to the eye can also be removed. There are three general mechanisms for redundancy elimination:

- Spatial – similarity between nearby pixels
- Spectral – similarity in different color space
- Temporal – similarity between nearby frames

2.1. *Quantization and Colorspace*

Digital video formats represent or quantize pixels in either 8 or 10 bits per color component. This represents the dynamic range, an 8-bit representation has 256 possible levels (and a maximum signal to noise of 50 dB) while a 10-bit representation has 1,024 possible levels (and a maximum signal to noise of 62 dB). By simply throwing away the two least-significant bits of the 10-bit video data (5:4 compression) the bit rate to the codec can be reduced by 20%. More importantly, since most processors work most efficiently on byte (8-bit) aligned data, codec processing is greatly simplified. Most MPEG- and DV codec implementations are 8-bit.

Instead of the familiar red/green/blue color components, video is transformed into the Y-Cr-Cb color space, where luminance is represented by Y and the color differences by Cr and Cb. This is done to better model the way color is perceived by the eye (remember rods and cones from biology), and to take advantage of compression by color sub-sampling. Y-Cr-Cb sub-sampling is represented in the following ways:

- 4:4:4 – Full color resolution, Y, Cr and Cb are sampled equally – true uncompressed video
- 4:2:2 – Half color resolution (3:2 compression*), Y is sampled every pixel, Cr and Cb are sampled every other pixel – common video format for most interfaces and low compression codecs.
- 4:1:1 – Quarter color resolution (2:1 compression*), Y is sampled every pixel, Cr and Cb are sampled every fourth pixel – used in DV25 and high-compression codecs.
- 4:2:0 – Quarter color resolution (2:1 compression*), Y is sampled every pixel, Cr and Cb are sampled alternately every other pixel – used in most high-compression codecs.

  * compression ratio is calculated by dividing the sum of the components by 12

2.2. *Reduced Horizontal Resolution*

Another technique for reducing the amount of video information is to reduce horizontal resolution by changing the pixel aspect ratio. Unlike NTSC and PAL with CCIR-601 pixels, HD formats such as 720p (1280x720) and 1080i/p (1920x1080) use square pixels. Some 1080i formats adopt a 1.33:1 rectangular pixel to reduce video resolution to 1440x1080 (4:3, or 25% compression).

2.3. *Lossy vs. Lossless*

Compression is often described as either lossless or lossy; however, there are more accurate classifications and ways to describe image loss.

- Mathematically lossless
  - Totally reversible
  - No data loss between encode and decode
  - Usually limited to about 2:1
  - Unpredictable performance
  - No cumulative loss
- Mathematically approximately lossless
  - Some loss due to rounding/truncation (i.e., 10- to 8-bit conversion)
  - Usually limited to about 3:1
  - Unpredictable performance
  - No cumulative loss
- Visually lossless
  - Data that is lost is imperceptible to the human eye
  - Can sometimes reach 6:1
  - Unpredictable performance
  - Cumulative loss may become perceptible if done multiple times
- Visually lossy
  - Data that is lost is perceptible, either in the form of artifacts, loss of resolution or combination
  - Can reach 100:1
  - Can produce a constant or a constant average bit rate.

2.4. *The Basic Codec*

The block diagram of a basic, intraframe, lossy image compression system is shown in Fig. 1. It consists of three connected components: Source Encoder, Quantizer and Entropy Encoder. Compression is accomplished by applying a reversible transform to convert the image data out of the visual or time-domain, quantizing the resulting transform coefficients, and entropy coding the quantized values.

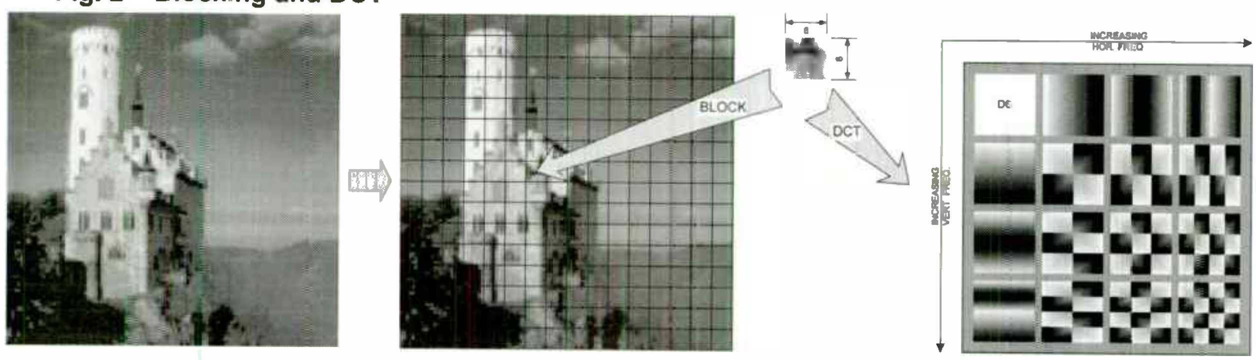
Fig. 1 – General Intraframe Codec Block Diagram

2.4.1. Source Encoder

The source encoder performs the initial linear transform on the incoming image. Although many reversible linear transforms have been developed over the years, the most common ones in use today are the Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT), each with its own advantages and disadvantages. The output of the linear transform is a set of binary coefficients.

The more common DCT transform converts blocks of pixel data to the component frequencies (see Fig. 2). This allows the less noticeable higher frequency data to be compressed by the quantizer.

Fig. 2 – Blocking and DCT



The type of transform selected corresponds to the nature of the artifacts introduced by the compression process. For example, since DCT based transforms operate on a rectangular "block" of pixels, DCT-based artifacts tend to appear as "blocking" within the image. However, by itself the transform is reversible and therefore, in most cases, can be lossless.

2.4.2. Quantizer

Image loss is introduced by the quantizer, which by reducing the precision of the transformed coefficients, reduces the number of bits needed to store of those values. Quantization can be performed on each individual coefficient, which is known as Scalar Quantization (SQ). Quantization can also be performed on a group of coefficients together, a process that is known as Vector Quantization (VQ). Both uniform and non-uniform quantizes can be used depending on the desired results.

Since the quantization process is analogous to selectively "rounding off" certain values, it is lossy and the main source of within-frame compression in an encoder. The amount of data thrown out by the quantizer determines the severity of the artifact type introduced by the initial transform.

2.4.3. Entropy Encoder

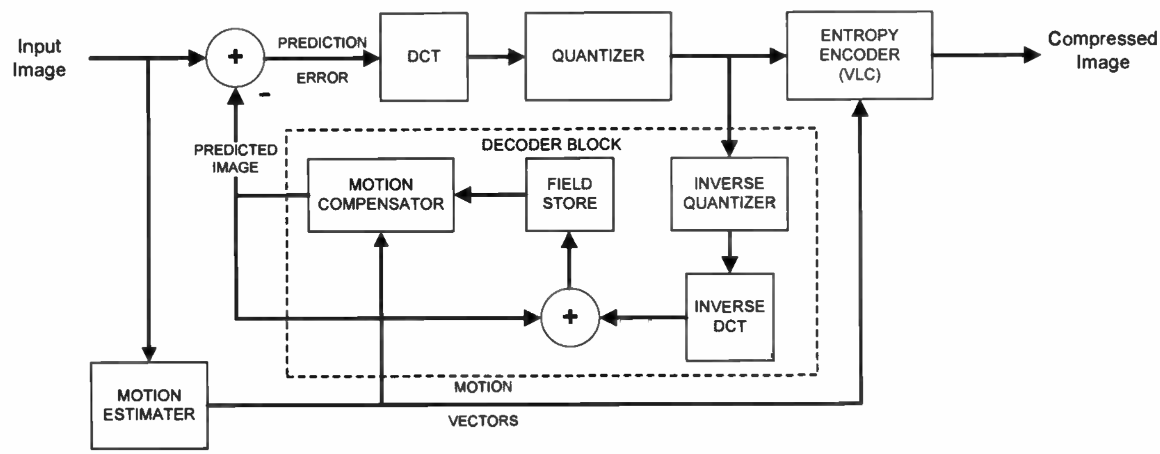
An entropy encoder further compresses the quantized values losslessly to give better overall compression. It uses a model to accurately determine the probabilities for each quantized value, and produces an appropriate code based on these probabilities so that the resultant output code stream will be smaller than the input stream. Simply put, more common values are encoded with fewer data bits then less common values.

The most commonly used entropy encoders are the Huffman Variable Length Coder (VLC) and the simple Run-Length Encoder (RLE). More advanced entropy encoders are adaptive such as the Context Adaptive Binary Arithmetic Coder (CABAC), the Context Adaptive Variable Length Coder (CAVLC) and Embedded Block Coding with Optimized Truncation (EBCOT).

2.4.4. Motion Estimation

Codecs that perform motion estimation are called "interframe." These codecs (shown in Fig 3) also use the block-based DCT; however, they add motion estimation and compression to enable interframe compression. Essentially, an interframe combines the basic intraframe encoder, an intraframe decoder, and block-based motion estimation.

Fig. 3 – General Interframe Codec Block Diagram



Interframe compression removes temporal redundancy and achieves significantly better image quality at lower bit rates than I-frame compression. Since motion estimation is performed on blocks of pixels, a block based transform (such as the DCT) is required; since the DWT in not block-based, it cannot be used in interframe codecs.

Motion compression produces a group of pictures (GOP) bordered by I-frames; frames within the GOP are predicted, and therefore not self-contained. GOP length is set as an encoding parameter and can vary from a few frames to a few seconds; longer GOP's produce higher compression efficiency, shorter GOPs recover faster from data loss and channel switching. Most long GOP encoding uses a 12- or 15-frame GOP length; the ATSC standard recommends an I-frame at least every half second.

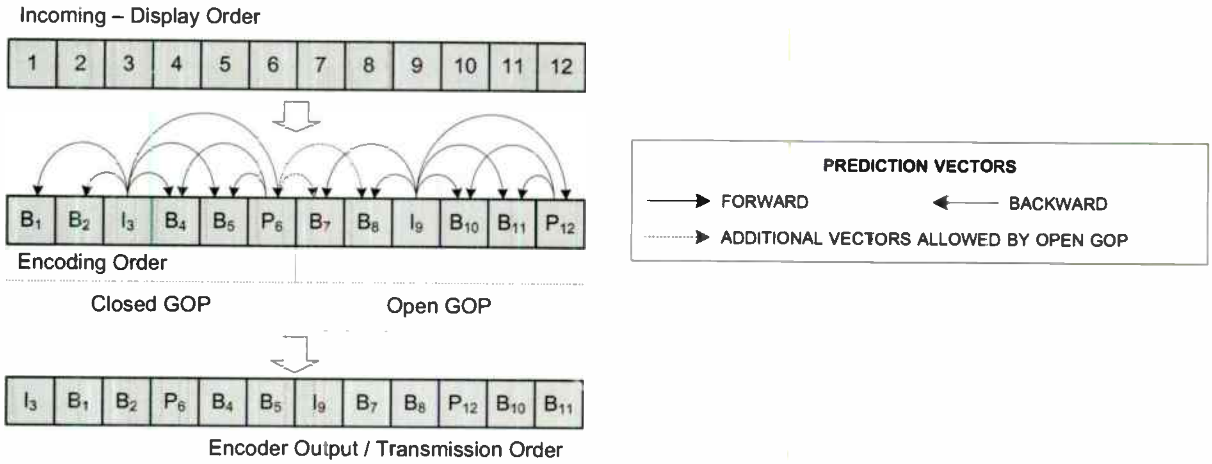
Within a GOP there are three kinds of frames: I, P and B.

- I-frames are standalone.
- P frames are predicted from the previous anchor frame (I or P frame), and are about half the size of I frames.
- B frames are bi-directionally predicted from both previous and future anchor frames, and are about half the size of P frames.

The compression efficiency improvement added by motion estimation can range from 3:1 to 5:1, depending on search range and GOP size.

A GOP can be either closed (not relying on external prediction vectors) or open (relying on vectors from previous anchor frames (see Fig. 4).

Fig. 4 – GOP Composition



2.4.5. Interframe vs. Intraframe

Intraframe (I-frame) compression results in compressed frames that are all fully self-contained, and can be decoded without data from other frames. Because temporal redundancy cannot be removed by intraframe compression, it is less efficient than intraframe. All DV formats and JPEG2000 are intraframe only, while other codecs can operate in an intraframe or interframe mode.

Key characteristics of intraframe compression are mostly related to *Image Facility.*

- Easier to edit – frames can be removed, replaced or concatenated without affecting neighbors
- Can create constant sized compressed frames, easier to index
- Not susceptible to motion induced artifacts, good for high-motion content (e.g., sports)
- Easy to recover from data loss - only frames containing lost data are disrupted
- Encoders and decoders are symmetrical and usually about equal in processing requirement

Key characteristics of interframe compression are mostly related to *Image Size.*

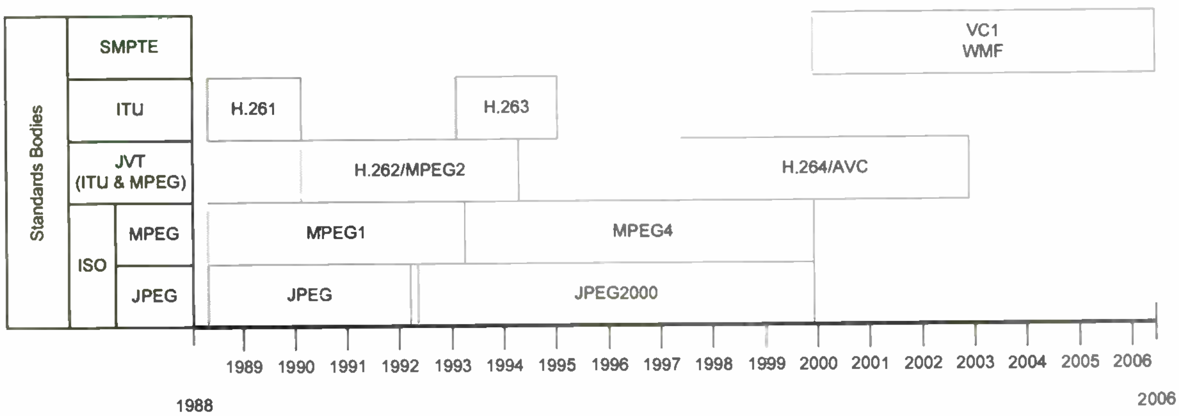
- Similar image quality to intraframe at a much lower bit rate (~ 3:1)
- Harder to edit – non GOP aligned sequences need to be recoded to create the appropriate anchor frames.
- Can create constant average bitrates but not constant sized frames, harder to index
- Susceptible to motion induced artifacts – especially quick camera pans
- Longer to recover from data loss – whole GOP's can be lost due to a disrupted anchor frame
- Encoders and decoders are asymmetrical, encoder requiring much more processing than decoders

Both intraframe and interframe codecs can operate in interlaced (field) and de-interlaced (frame) modes. However, only interframe codecs can gain efficiency by offering tools designed for interlaced video as well as 3:2 pull-down (cinema) handling.

2.5. *Codec Development and Standards*

Most codec development stems from three main standards bodies, The International Telecommunications Union (ITU), the International Standards Organization's (ISO) Joint Photography Experts Group (JPEG) and Motion Pictures Experts Group (MPEG). Some work was accomplished as part of a joint ITU and MPEG team, known as the Joint Video Team (JVT). Microsoft's Windows Media codec was standardized by SMPTE as well.

Fig. 4 – Codecs and Standards



2.6. *Popular Codecs and Block Diagrams*

General codec features and capabilities are shown in Table 2. A more detailed feature comparison can be found in Appendix 2.

Table 2 – Codec Features

GENERAL FEATURES	JPEG	DV	MPEG-2	MPEG-4 ASP	H.264/AVC	AVS	VCI	JPEG2K
Standard	ISO		ISO	ISO	JVT	China	SMPTE	ISO
Typical Minimum Compression Ratio	4:1	10:1	5:1	4:1	4:1	50:1	12:1	2:1
Typical Maximum Compression Ratio	10:1	10:1	60:1	60:1	100:1	100:1	100:1	14:1
Resolution (bits)	8	8	8	8	8/10/12/14	8	8	8/10/12
Supports Intraframe Only Coding	X	X	X	X	X			X
Supports Interframe Coding			X	X	X	X	X	
Supports Lossless Encoding								X
Supports Near Lossless Encoding	X					X	X	X
Reduced Color Space (4:2:0, 4:1:1)	X	X	X	X	X	X	X	X
Broadcast Color Space (4:2:2)	X	X	X	X	X			X
Full Color Space (4:4:4)	X				X			X

2.6.1. DV – DV25, DVCAM, DVCPRO50, DVCPRO100

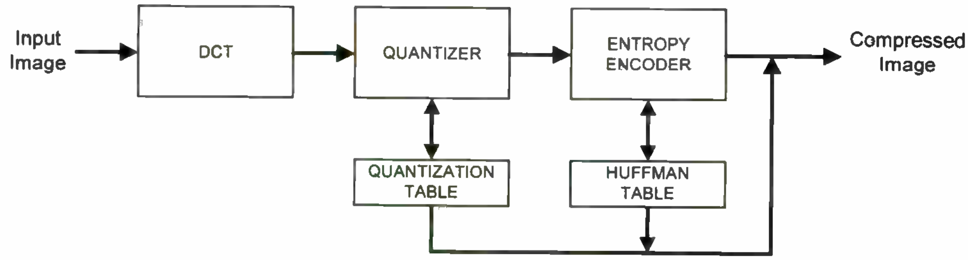
The DV family of codecs (shown in Fig. 5) are based on the block-based DCT and variable quantization.

All DV codecs are 8-bit, feature I- frame compression and are aimed at producing constant size frames suitable for tape recording.

DV codecs offer comparable compression efficiency to motion JPEG and I-Frame MPEG.

Manufacturer-specific DV varieties include Sony's DVCAM and Panasonic's DVCPRO, DVCPRO50 and DVCPRO100 HD. In the case of DV formats, the specified data rates are total (video plus audio and metadata).

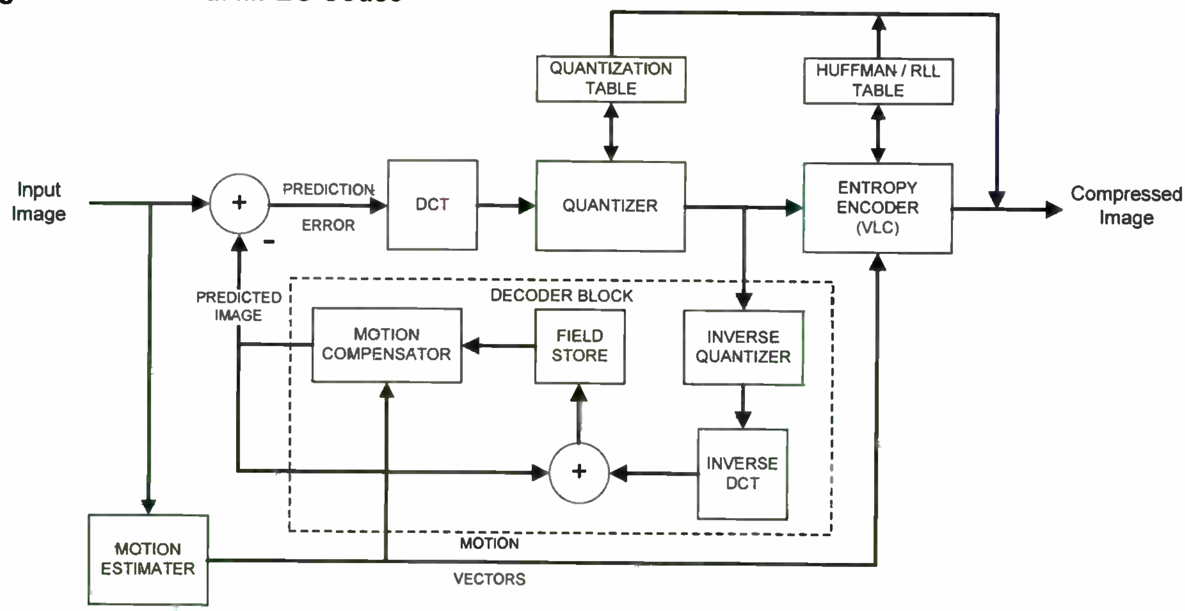
Fig. 5 – The DV Codec



2.6.2. MPEG-1/MPEG-2/MPEG-4 Part 2

The MPEG family of codecs (shown in Fig. 6) also use the block-based DCT; however, they add motion estimation and compression to enable interframe compression. The MPEG-1/2/4 Part 2 codecs are 8-bit based and feature both constant and variable bit rate modes. They do not however produce constant frame sizes.

Fig. 6 – The General MPEG Codec



In addition to the video codec, the MPEG-1 specification includes support for audio codec layers. MPEG- 1 has been used primarily in Internet and web-streaming applications; it is however being largely replaced by MPEG-4 Part 2.

MPEG-2 adds the concept of "levels" and "profiles" (see Table 3) to optimize compression for specific applications. MPEG-2 codecs are among the most widely used in full-resolution media applications (see Appendix 1); they are used in acquisition formats (IMX, XDCAM, XDCAM HD), delivery formats (DVD, HD-DVD, Blu-ray) and transmission formats (ATSC, DVB).

Table 3 – MPEG-2 Profiles and Levels

Profile @ Level	Resolution	FPS	Sampling	Max. Bitrate.	Typical Application
				Mbps	
SP @ LL	176 x 144	15	4:2:0	0.1	Mobile Sets
	352 x 288	15			
SP @ ML	320 x 240	24	4:2:0	0.384	PDA's
MP @ LL	352 x 288	30	4:2:0	4	Set top boxes, Cable
	720 x 480	30			
MP @ ML	720 x 576	25	4:2:0	15	SD-DVB, DVD (9.8Mbps)
	1440 x 1080	30			
MP @ H-14	1280 x 720	30	4:2:0	60	HDV (25 Mbps), XDCAM HD (35Mbps)
	1920 x 1080	30			
MP @ HL	1280 x 720	60	4:2:0	80	ATSC (19Mbps), HD-DVB
	720 x 480	30			Sony IMX (I-frm Only 30, 40, 50 Mbps)
422P @ ML	720 x 576	25	4:2:2	50	General "Broadcast"
	1440 x 1080	30			
422P @ H-14	1280 x 720	30	4:2:2	80	
	1920 x 1080	30			
422P @ HL	1280 x 720	60	4:2:2	300	XDCAM HD (50 Mbps)

MPEG-4 is a multi-part (23 sections) standard that includes not just the codec, but also parts that define such things as the delivery framework, file format, content protection and conformance testing procedures. The actual codec is in part 2.

MPEG-4 Part 2 actually has a more complex profile/level structure than MPEG-2; however, the most common ones are Simple Profile (SP) and Advanced Simple Profile (ASP). MPEG-4 Part 2 SP is comparable in features and application to MPEG-1, and used for low-resolution/low-bit-rate Internet and web streaming applications. ASP is similar to MPEG-2. ASP is used in HD-CAM SR compression, and is the basis for the proprietary DivX codec. MPEG-4 P2 ASP has encountered resistance to more widespread adoption due to several reasons:

- Substantially increased processing requirement (relative to MPEG-2)
- Slightly improved compression (relative to MPEG-2)
- Initial licensing issues/per content fees
- And, of course, the development of MPEG-4 Part 10, or AVC

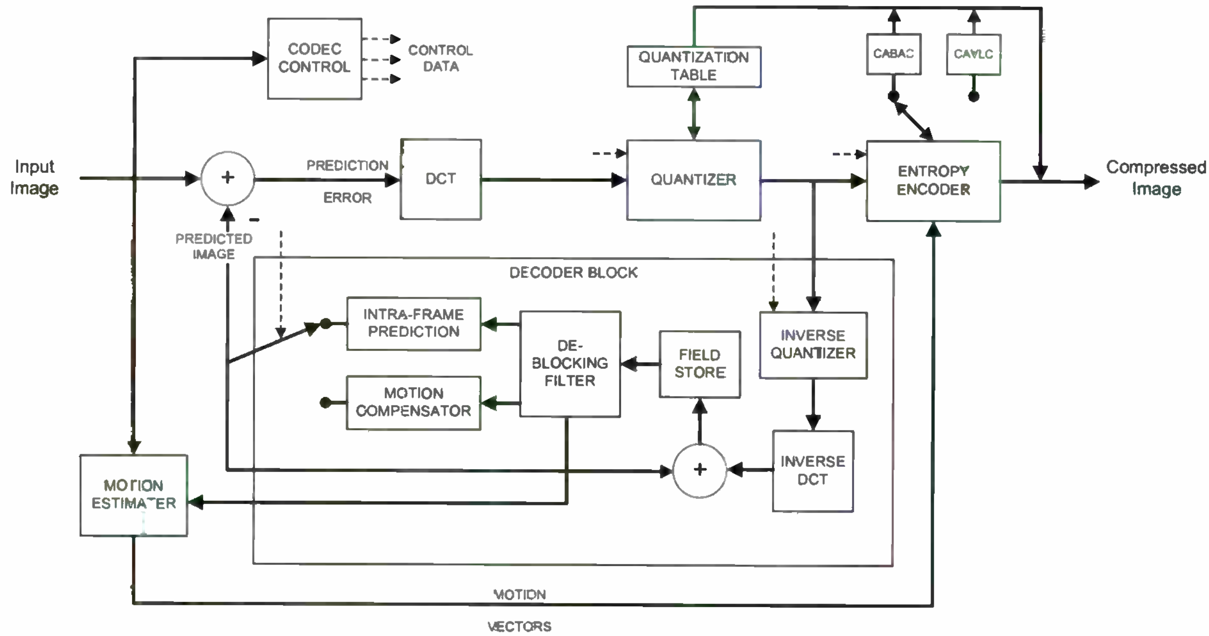
2.6.3. AVC/H.264

Also know as MPEG-4 Part 10 and H.264, the Advanced Video Codec or AVC (shown in Fig 7) offers an approximate two times encoding efficiency (the same video quality as half the bit rate) over MPEG-2 and MPEG-4 Part 2 ASP.

Like MPEG-2, AVC supports both intraframe and interframe compression; however, it adds support for 10-bit video processing via 16-bit DCT and internal processing. The major tradeoff of AVC is the increased real-time processing requirement, which can range to 16 times for encode and 4 times for decode over MPEG-2. Also, unlike other codecs, decision tree's during the encode process can lead to a wide range in performance based on codec implementation. An AVC encoder might have a less complex decision tree, or not support CABAC, and therefore offer less compression efficiency than one that does.

AVC targets applications ranging from VHS-quality SD over ADSL speed (600Kbps), to HD DVD's at 10 Mbps and all the way to HD acquisition at 100 Mbps. Panasonic's AVC-I codecs use 10-bit intraframe AVC compression; AVC-I 50 uses CABAC, while AVC-I 100 uses CAVLC. AVC is also one of the mandatory codecs for the HD-DVD and Blu-ray disk formats. Apple is also targeting AVC as a contribution level codec for Quicktime.

Fig. 7 – The AVC Codec



The AVC standard contains seven basic profiles, four intraframe-only profiles (see Table 4), and 16 resolution/data-rate levels (see Table 5). As can be seen from the following tables some of the operating points fall outside of practical ranges (i.e QCIF @ 240Mbps and HD @ 80Kbps).

Table 4 – AVC Profiles and Typical Applications

Profile	Data Rate Range	Sampling	Bits	Typical Application
Baseline	64-768Kbps	4:2:0	8	QCIF, CIF, QVGA, Video Conferencing, Mobile
Main	2-4Mbps	4:2:0	8	SD - Consumer products (SD)
Extended	768Kbps-4Mbps	4:2:0	8	QCIF-SD - Streaming, IPTV
High	5-25Mbps	4:2:0	8	HD - Consumer products
High 10	12-60Mbps	4:2:0	10	HD - Professional
High 10 Intra	12-60Mbps	4:2:0	10	HD - Professional (Panasonic products)
High 4:2:2	40-200Mbps	4:2:2	8-10	HD - Broadcast
High 4:2:2 Intra	40-200Mbps	4:2:2	8-10	HD - Professional and broadcast (Panasonic products)
High 4:4:4 Predictive	200-960Mbps	4:2:2	8-14	HD - 4Kx2K, Digital Cinema
High 4:4:4 Intra	200-960Mbps	4:2:2	8-14	HD - 4Kx2K, DI, Film replacemant

Table 5 – AVC Levels (Practical Ranges Highlighted)

Level Number	Typ. Picture Size	Typ. Frm. Rate	Maximum Bitrate @ Profile (Mbps)					
			Base	Main	Extended	High	High 10	High 4:2:2, 4:4:4
1	QCIF	15	0.064	0.064	0.064	0.08	0.192	0.256
1b	QCIF	15	0.128	0.128	0.128	0.16	0.383	0.512
1.1	CIF/QCIF	7.5/30	0.192	0.192	0.192	0.24	0.576	0.768
1.2	CIF	15	0.384	0.384	0.384	0.48	1.152	1.536
1.3	CIF	30	0.768	0.768	0.768	0.96	2.304	3.072
2	CIF	30	2	2	2	2.5	6	8
2.1	HHR	30/25	4	4	4	5	12	16
2.2	SD	15	4	4	4	5	12	16
3	SD	30/25	10	10	10	12.5	30	40
3.1	720p	30	14	14	14	17.5	42	56
3.2	720p	60	20	20	20	25	60	80
4	720p/1080i	60p/30i	20	20	20	25	60	80
4.1	720p/1080i	60p/30i	50	50	50	62.5	150	200
4.2	1080p	60p	50	50	50	62.5	150	200
5	2kx1k	72	135	135	135	168.75	405	540
5.1	2kx1k/4kx2k	120/30	240	240	240	300	720	960

2.6.4. AVS

The Audio Video Coding Standard (AVS) is a national standard being developed in China to compete against AVC and VC-1 as the successor to MPEG-2. Early efforts are underway for ISO standardization as well.

Although designed to avoid the MPEG royalties of AVC, AVS and AVC share a similar block diagram. AVS compression approaches AVC efficiency with a generally lower processing requirement.

AVS1 Part 2 (see Table 6), targeting SD and HD resolution, contains two profiles: the Jizhun Profile for television broadcast, transmission and delivery (including satellite and cable) broadcast; and the Zengqiang Profile for DVD, HD-DVD and Blu-ray disk storage. AVS1 Part 7, also known as the Jiben Profile, targets mobile broadcast with reduced resolutions.

Table 6 – AVS1 Part2 Profiles and Levels

Level	Resolution	FPS	Sampling	Max. Bitrate. Mbps	Typical Application
4.0	720 x 480	30	4:2:0	10	SD - Transmission and Delivery
	720 x 576	25			
4.2	720 x 480	30	4:2:2	10	SD - Transmission and Delivery
	720 x 576	25			
6.0	1920 x 1080	30	4:2:0	30	HD - Transmission and Delivery
	1280 x 720	60			
6.2	1920 x 1080	30	4:2:2	30	HD - Transmission and Delivery
	1280 x 720	60			

AVS generally seems targeted at high-volume consumer devices such as set-top boxes, disc players, mobile media devices, etc., as well as at media distribution. China produces over 30 million DVDs and 10 million set-top boxes a year, with MPEG royalties exceeding $100 million a year[i]. AVS development is largely a cost reduction strategy for Chinese manufacturing. Broadcast applications are not specifically targeted.

2.6.5. VC-1

The VC-1 codec is based on development by Microsoft (Windows Media 9 Video Codec) and the SMPTE 412M standard from 2006. Similar in structure and compression efficiency to AVC; VC-1, however, only supports the 4:2:0 color space and interframe compression.

Like MPEG- 2, VC-1 offers application targeted profiles and levels (see Table 7), however VC-1 profiles are more oriented to classifying the complexity (and corresponding processing requirement) of the targeted codec.

Aside from Windows Media applications, VC-1 is one of the mandatory codecs for the HD-DVD and Blu-ray formats.

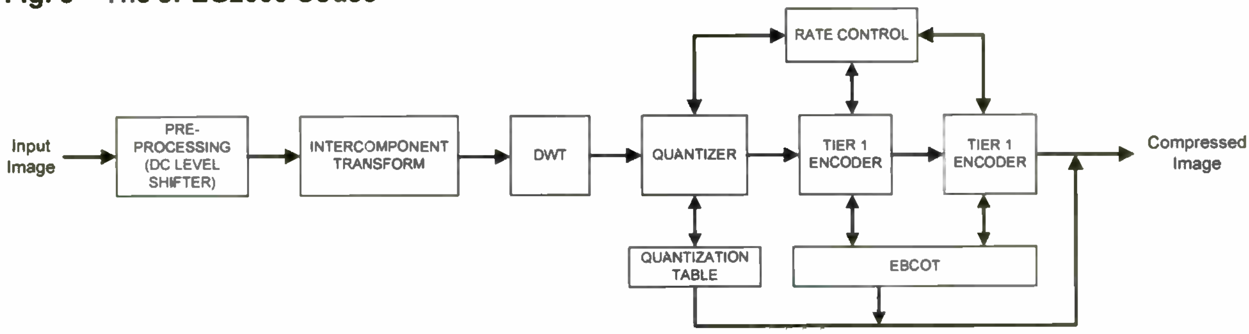
Table 7 – VC-1 Profiles and Levels

Profile @ Level	Resolution	FPS	Sampling	Max. Bitrate.	Typical Application
SP @ LL	176 x 144	15	4:2:0	0.096	QCIF
	240 x 176	30	4:2:0		
SP @ ML	352 x 288	15	4:2:0	0.384	CIF
MP @ LL	320 x 240	24	4:2:0	2	QVGA
	720 x 480	30	4:2:0		NTSC / 480P
MP @ ML	720 x 576	25	4:2:0	10	PAL / 576P
MP @ HL	1920 x 1080	30	4:2:0	20	HD
AP @ L0	352 x 288	30	4:2:0	2	CIF
	720 x 480	30	4:2:0		SD NTSC
AP @ L1	720 x 576	25	4:2:0	10	SD PAL
	720 x 480	60	4:2:0		480P
AP @ L2	1280 x 720	30	4:2:0	20	720P
	1920 x 1080	24	4:2:0		1080P
	1920 x 1080	30	4:2:0		1080I
AP @ L3	1280 x 720	60	4:2:0	45	720P
	1920 x 1080	60	4:2:0		1080P
AP @ L4	2048 x 1536	24	4:2:0	135	Film

2.6.6. JPEG2000

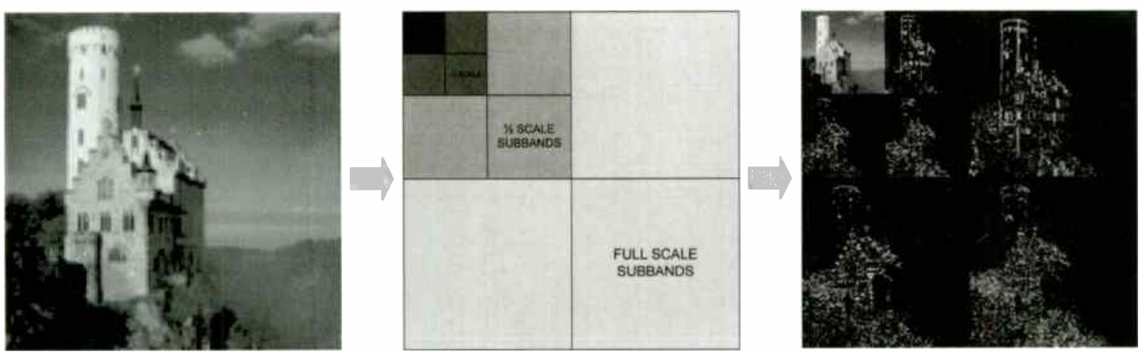
The advanced version of the JPEG codec breaks away from using the DCT as the basis for its linear transform. By using instead the Discrete Wavelet Transform (DWT) (see Fig. 8) the familiar blocking and ringing artifacts of DCT-based codecs are replaced by non-linear blurring. Generally, the compression efficiency of JPEG2000 is only about 20-40% better than JPEG.

Fig. 8 – The JPEG2000 Codec



The DWT decomposes the image into scaled sub-bands (see Fig. 9), which can be independently quantized and entropy encoded. Each successive sub-band represents the difference between the image at the previous scale and the next scale. This leads to the feature of scalable decoding; if a quarter-scale image is needed, only the sub-bands up to quarter scale need to be decoded.

Fig. 9 – Wavelet Image Decomposition



JPEG2000 supports both lossless and lossy compression; 8-, 10- and 12-bit resolution; and offers intrinsic scalability based on the multi-resolution decomposition performed by the DWT. JPEG2000 only offers intraframe compression and therefore cannot match the compression efficiency of interframe codecs. JPEG2000 also has much higher processing requirements for both the encoder and decoder (6 to 10 times) while offering only a 20-40% improvement over JPEG.

Applications for JPEG2000 target digital intermediate for film, digital cinema and high-end contribution.

2.7. *Avid's DNxHD*

Seeking standardization by SMPTE as VC-3, Avid's DNxHD is a proprietary codec targeted for use in their post production systems. DNxHD appears to be an intraframe codec that uses the 4:2:2 color space and operates at either 145 Mbps (8-bit) or 220 Mbps (8- or 10-bit).

3. Processing and Real-Time Consideration

3.1. *Symmetric vs. Asymmetric Codecs*

Generally, intraframe codecs are symmetric and have the same processing requirements for encode and decode. On the other hand, interframe codecs are asymmetric and require much more processing for encode. For example, MPEG-2 interframe encoding is roughly four times more processing intensive than decode.

Having an asymmetric codec allows for a simple, lightweight and low-cost decoder. This eases software implementation and allows systems of arbitrary performance to decode video. This requirement is obviously necessary in any transmission-oriented codec, but is also useful in any codecs that are part of a file-based or shared media workflow.

3.2. *Software Applications*

The days of a viable codec with a hardware only or proprietary implementation are passed. Any compression system considered for use in a modern file-based workflow should be standards-based and offer, at least, a software-only decoder implementation. This is important in preserving the future value of media. A software codec, whether it be a Windows Media Player filter, a plug-in for QuickTime or a standalone application, can always be saved as a file along with the media that it is associated with.

3.3. *Processing Direction*

Codecs can be implemented in specialized discreet chips (ASICs), field programmable gate array's (FPGS), software that runs on computer processors (CPUs), digital signal processors (DSPs) and graphics processors (GPUs), often in some combination. However, in all cases more complexity requires more cost.

Real-time codecs that can be implemented without the need for specialized hardware are one of the driving forces behind the consumer media explosion. Camcorders that create files instead of tapes or disks require PCs to play the files back, or share them over the Internet, or even to burn them to disk. Since CPUs and GPUs are the only hardware available to the off-the-shelf PC, CPU and GPU manufacturers are competing to add media and improve functionality to their chips. These improvements have the ripple effect of improving codec performance in the professional and broadcast markets.

What once took specialized and costly hardware can now be done on a PC or workstation. This trend will continue and likely accelerate.

4. Summary and Conclusion

While it is obvious no single codec addresses all requirements for all applications, broadcasters that limit types and levels of compression within their facility will find it easier to manage workflows, and will wind up with generally better image quality.

Newer codecs such as AVC promise to replace older codecs with a "Swiss Army Knife" of compression tools (see Table 8); however, even counting on advances in processing performance, integration with existing equipment and transmission systems still stand in the way.

Table 8 – Codec Application Review

APPLICATIONS	DV	MPEG-2	MPEG-4 ASP	H.264/AVC	AVS	VCI	JPEG2K
Digital Intermediate				X			X
Contribution & Acguisition		X	X	X			X
ENG	X	X		X			
POST	X	X		X			X
Distribution & Archive		X		X			
Broadcast Transmission		X		Future	China		
High Definition DVD		X		X	China	X	
Internet Download & IP TV			X	X	China	X	
Mobile and Handheld			X	X	China	X	

Occasionally we hear announcements of new proprietary codecs that promise tremendous improvement in compression efficiency; few of these work as well as promised or stick around long enough to be recognized by standardization organizations. This is partially due to the random nature of video signal content. What works well in a controlled lab often fails in real-life application.

In reality we may have reached to point of diminishing returns of improving codec compression efficiency with respect to processing overhead. Increasing the processing power does not translate into proportional increases in efficiency, nor do increasingly complex algorithms yield dramatically improved performance. Today's work is mostly focused on implementing and putting into practice the codecs defined over these past few years.

5. Appendix 1 – Media Formats

Media	Raw	Video	Capacity	Video					
	Data Rate	Data Rate			Color	Res.	Comp		H
Acquisition Media HD	**Mbps**	**Mbps**	**min.**	**Codec**	**Space**	**Bits**	**Ratio**	**GOP**	**Samp.**
SONY HDCAM SR	600	440	120	MPEG-4 p2	4:2:2	10	2.7:1	I-frm	1920
SONY HDCAM SR	600	440	120	MPEG-4 p2	RGB 4:4:4	10	4.2:1	I-frm	1920
SONY HDCAM SR	880	880	60	MPEG-4 p2	RGB 4:4:4	10	2:1	I-frm	1920
SONY HDCAM	144		124	DCT Based	3:1:1	8	6:1	I-frm	1440
Panasonic D5-HD	360	323	155	DCT Based	4:2:2	8/10	4-5:1	I-frm	1920
Panasonic D5-HD w/ external codec	360	323	155	External JPEG2K	4:4:4	8/10/12	5:1	I-frm	1920
Panasonic AVC-I 100 P2 (32GB)	100		32	AVC-I	4:2:2	10	12.5:1	IBP	1920
Panasonic AVC-I 50 P2(32GB)	50		64	AVC-I	4:2:0	10	25:1	IBP	1440
Panasonic DVCPro 100HD	100		30	DV100	4:2:2	8	12.5:1	I-frm	1440
Sony XD-CAM HD		50	40	MPEG-2 422P/HL	4:2:2	8		IBP	1920
Sony XD-CAM HD		35	60	MPEG-2 MP/HL	4:2:0	8		IBP	1440
HDV	36	25		MPEG2 MP/HL	4:2:0	8	50:1	IBP	1440
Acquisition Media SD									
Panasonic DVCPro 50	50			DV50	4:2:2	8	5:1	I-frm	
Panasonic DVCPro	25			DV25	4:1:1	8	10:1	I-frm	
Sony DV CAM / DV25	25			DV25	4:1:1*	8	10:1	I-frm	
Sony IMX		30,40,50		MPEG2 422P/ML	4:2:2	8	3-6:1	I frm	
Distribution Media									
Blu-Ray Disc (50GB dual sided)	36-54	12-40	550	MPEG-2 HL/ AVC / VC-1	4:2:2	8		IBP	
HD-DVD (30GB dual sided)	36	8-28	480	MPEG-2 HL/ AVC / VC-1	4:2:2	8		IBP	
DVD (8.5Gb dual sided)	11	5	210	MPEG-2 ML	4:2:0	8		IBP	
Devices and Interfaces	**Data Rate**	**Data Rate**	**GB**						
Data Storage									
Very High Performance Hard Disk (15k RPM)	1312		450						
High Performance Hard Disk (10k RPM)	707		300						
Desktop Hard Disk (7200 RPM)	600		1000						
T1000 Data Tape	960		500						
SAIT-2 Data Tape	360		800						
Media Interfaces									
HDMI 1.3	10200	10163							
HDMI 1.0	4900	4863							
SMPTE 424	3000	2482							
SMPTE 292	1500	1242							
SMPTE 259	270	207							
Data and Network Interfaces									
10Gb Ethernet	7000								
Gb Ethernet	700								
USB 2.0	480								
Firewire (1394)	400								
OC-192	10000								
OC-48	2488								
OC-12	622								
OC-3	155.52								
OC-1	51								
T-3	44.7								
T-1	1.5								
ATSC	19.2	18							

6. Appendix 2 – Codec Features

FEATURES	JPEG	DV	H.261	MPEG1	MPEG2	MPEG4 SP	H.264/AVC	VCI	AVS	JPEG2K
GENERAL										
Typical Maximum Compression Ratio	10:1	10:1			30:1	30:1	60:1	60:1	60:1	14:1
Video Bit Resolution	8	8	8	8	8	8	8/10	8	8	8/10/12
Loop De-blocking Filtering			X	X	X	X				
Adaptive Loop De-blocking Filter							X	X	X	
Interlaced Coding Tools					X	X	X	X	X	
Supports Loss-less Encoding										X
INTRA-FRAME										
Macro Block Based Processing	8x8	8x16	16x16	16x16	16x16	16x16	16x16	16x16	16x16	
Intra-Frame Coding	X	X	X	X	X	X	X	X	X	X
Intra Prediction and Coding							X		X	
DCT (8bit floating point)	8x8	8x8	8x8	8x8	8x8	8x8	8x8			
Integer DCT (8bit Integer)							4x4		8x8	
Integer DCT (16bit Integer)							4x4			
Switchable DCT/IDCT								8x8,8x4,4x8		
DWT										X
Intra DCT Prediction						X	X	X	X	
Perceptual Quantization	X	X								
Fixed Quantization			X							
Adaptive Perceptual Quantization				X	X	X	X	X	X	X
Quantization and Transfor Coeficient Scanning							X		X	
Run-Length Coding	X	X	X	X	X	X	X	X	X	
VLC	X	X	X	X	X	X	X		X	
CAVLC							X		X	
CABAC							X			
2D- VLC									X	
Multiple VLC Tables								X	X	
EBCOT										X
INTER-FRAME										
Inter-Frame Coding			X	X	X	X	X	X	X	
Inter Prediction and Coding										
P- Frames			X	X	X	X	X	X	X	
B- Frames				X	X	X	X	X	X	
Motion Estimation (pixel)			1	1/2	1/2	1/2	1/4	1/4	1/4	
Motion Compensation			X	X	X					
Widened Search Range					X	X	X	X	X	
Unrestricted Motion Vectors						X	X	X	X	
Variable Block Size Motion Compensation						16x16, 8x8	16,8,4x16,8,4	16x16, 8x8	16x16, 8x8	
Multiple Reference Frame Selection							X		X	

7. References

1. Subhasis, Saha, "Image Compression – from DCT to Wavelets: A Review," *Crossroads The ACM Student Magazine*, 2000, http://www.acm.org/crossroads/xrds6-3/sahaimgcoding.html

2. Waggoner, Ben, "Understanding HD Formats," *Windows Media*, 2004, http://www.microsoft.com/windows/windowsmedia/howto/articles/UnderstandingHDFormats.aspx

3. Golston, Jerimiah, "Comparing Media Codecs fro Video Content," *Embedded Systems Conference*, 2004

4. Rabbani, Majid, "JPEG2000: Background, Scope, and Technical Description," *Cornell University*, 1998

5. Loomis, Jay, "VC-1 Technical Overview," *Windows Media, 2007* http://www.microsoft.com/windows/windowsmedia/howto/articles/vc1techoverview.aspx

6. "H.264/AVC Intraframe Coding (AVC-Intra) Specification," Version 1.0, Matsushita Electric Industrial Co., LTD, 2007

7. Liang, Fan, "Overview of the AVS Standard," *2004 IEEE International Conference on Multimedia and Expo*, IEEE, 2004

8. Clendenin, Mike, "Chinese AVS codec, rivaling H.264, is rising," *EE Times*, Feb. 2006 http://www.digitaltvdesignline.com/news/180205429

9. Ngan, King Ngi, "Audio Video Coding Standard (AVS) of China)," Department of Electrical Engineering, The Chinese University of Hong Kong, November 2007

10. "XDCAM Video Operating Points; update for HD422," Sony Corporation, 2007

11. McGrath, Dylan, "Hi-Def Surprise: Widespread Consumer Unhappiness With HDTVs," *Information Week*, Dec 2007, http://www.informationweek.com/hardware/showArticle.jhtml?articleID=205201889

12. "Recommended Practice: Guide to the Use of the ATSC Digital Television Standard," Advanced Television Systems Committee, Dec 2006, www.atsc.org

---

[i] Ngan, King Ngi, "Audio Video Coding Standard (AVS) of China," Department of Electrical Engineering, The Chinese University of Hong Kong, November 2007

Providing Spatial Scalability using Scalable Video Coding to Mobile Broadcasting

Sangjin Hahm[1], Changseob Park[1], Byungsun Kim[1], Injoon Cho[1], Keunsoo Park[1], MunChurl Kim[2]
[1]Broadcast Technical Research Institute, Korean Broadcasting System, Seoul, Korea
[2]Multimedia Computing Communications and broadcasting Lab, ICU, Daejeon, Korea

ABSTRACT

Terrestrial mobile broadcasting has been on air since 2005 in Korea and is named Digital Multimedia Broadcasting (DMB). The video resolution of DMB is 320x240 pixels for handheld devices. In mobile broadcasting, higher resolution for better video quality is needed in some cases like in-car. Scalable Video Coding (SVC) can be applicable in mobile broadcasting environment due to the flexibility of spatial, temporal and quality scalability. Recently, SVC technology becomes mature rapidly but its reference SW encoder isn't optimized yet. Therefore, we have developed a real-time SW SVC encoder for broadcasting. In this paper, we show our SVC encoder that can provide two spatial layers: QVGA (320x240) and VGA (640x480). The base layer can be fully compatible with H.264/AVC. Our encoder is performing real-time operation on a normal PC by optimizing SVC algorithm
Keywords-component; SVC, H.264/AVC

1. INTRODUCTION

H.264/MPEG-4 Advanced video coding (AVC) is used in mobile broadcasting and next-generation optical video disc such as HD-DVD and Blu-ray for high compression rate and good video quality. In Korea, Digital Multimedia Broadcasting (DMB) also adopted AVC. The video resolution for mobile broadcasting is usually low because the mobile terminal such as cellular phone has small display for mobility and low power consumption. Recently, demands for higher video resolution in mobile broadcasting environment are increasing. This requirement leads to the use of scalable video coding which can provide both low and high resolution together.

Now, the standard reference software of SVC called Joint Scalable Video Model (JSVM) [1] isn't implemented efficiently just only to verify SVC tools in the perspective of standard conformance. It is far from real-time encoding. So, we designed a SVC encoder only with spatial scalability for the real-time application of mobile broadcasting. Our SVC encoder meets the requirements of real-time implementation and acceptable performance with tolerable PSNR value drop and bit-stream increment by optimizing SVC algorithm.

2. DMB and SVC

2.1 DMB

Digital Multimedia Broadcasting (DMB) is a digital transmission system for sending multimedia (radio, TV, and data) to mobile devices such as cellular phone in Korea. DMB is broadcasted in two type of transmission – satellite and terrestrial. Terrestrial DMB (T-DMB) transmitted by Korean Broadcasting System (KBS) was started on December 1 in 2005. T-DMB service consists of 7 TV channels, 13 digital radio channels and 8 data channels.

Because DMB is based on the Eureka 147 Digital Audio Broadcasting (DAB) standard, T-DMB is transmitted on radio frequency bands band III (VHF).

The video resolution of T-DMB is 320x240(QVGA) pixels. It is encoded by MPEG-4 Part 10 (H.264) with baseline profile and level 1.3. The audio of T-DMB is encoded by MPEG-4 part 3 BSAC. The quality of audio is higher than FM because of stereo 48 kHz sampling. The audio and video is encapsulated in MPEG-2 TS. Because of OFDM-DQPSK modulation, T-DMB has strength for the channel effects such as fading and shadowing. T-DMB works in vehicles traveling up to 100km/h.

T-DMB broadcasting stations research and consider many different kind of service: Interactive DMB, 3D video DMB, higher resolution DMB and so on. The service for DMB of higher resolution is required in in-vehicle such as bus and train where viewers should see video in the distance and wan to see lager size of display. The need for higher resolution DMB is increasing now. SVC can be the optimal video codec for DMB of high quality resolution because the video codec of DMB is H.264/AVC and SVC is compatible with H.264/AVC. It is not need to develop and adopt new video codec.

2.3 SVC

SVC is a scalable extension of H.264/AVC being developed by JVT(Joint Video Team) co-established by MPEG(Moving Picture Expert Group) of ISO/IEC and VCEG(Video Coding Expert Group) of ITU-T [2]. The SVC standard aims at providing the technologies for flexible representation of its compressed bit-stream to make it possible to cope with various display sizes and wide range of network bandwidth etc. SVC provides three scalabilities: spatial, temporal and quality [3].

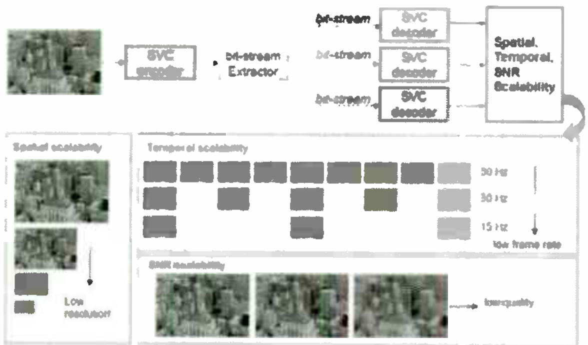


Figure 1 Scalability of SVC

- Spatial scalability: video is coded at multiple picture sizes.
- Temporal scalability: Coded video can drop a frame or frames from the bit-stream
- SNR/Quality/Fidelity scalability: video is coded at different qualities

Figure 1 shows the three scalabilities.

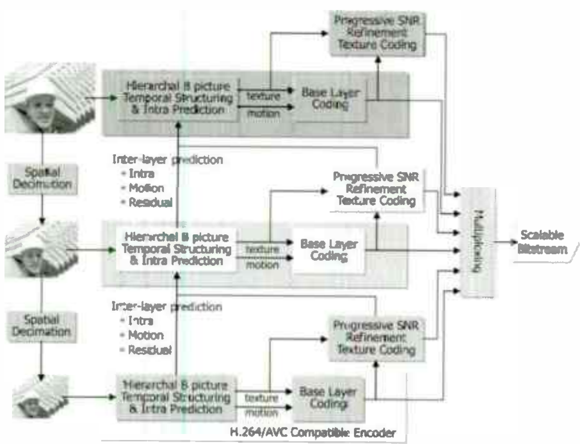


Figure 2 Structure of SVC Encoder

SVC can encode a bit-stream with three combined scalabilities so that the structure of SVC encoder would be like figure 2. In SVC, the method of encoding a video bit-stream with spatial scalability is up-sampling by one dimensional 4-tap FIR filters for luma component and a bilinear filter for chroma component. During encoding for spatial scalability, the inter-layer(between the layers) motion prediction is needed because the motion vector and texture should be changed according to picture size. Inter-layer prediction used in SVC is composed of three predictions: Inter-layer texture prediction, Inter-layer motion prediction, Inter-layer residual prediction. Each prediction can be used according to how base layer is encoded. In SVC, the method of encoding a video bit-stream with temporal scalability is hierarchical B- picture. The concept of hierarchical B-picture is that temporal scalability can be very efficiently provided with dyadic temporal enhancement layers. In SVC, the method of encoding a video bit-stream with quality scalability is a special case of spatial scalability with identical picture size for base and enhancement layer. To provide a variety of quality and bit rate of coded bit-stream, SVC supports two method of quality scalability: Coarse grain quality scalability (CGS) and Medium grain quality scalability (MGS).

SVC represents three scalabilities using layer structure. In each scalability layer, the first layer is called the base layer and all higher layers, called enhancement layers, are built on top of the base layer. Figure 3 show a coded SVC bit-stream combined three scalabilities. According to the bandwidth of network, the size of display or the computational power of device, A base layer and enhancement layers of three scalabilities can be extracted from SVC bit-stream.

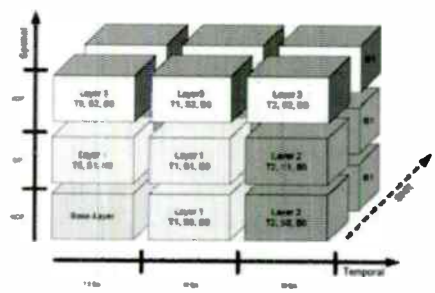


Figure 3 Structure of SVC bit-stream

3. SVC Encoder

3.1 Proposed SVC in mobile broadcasting

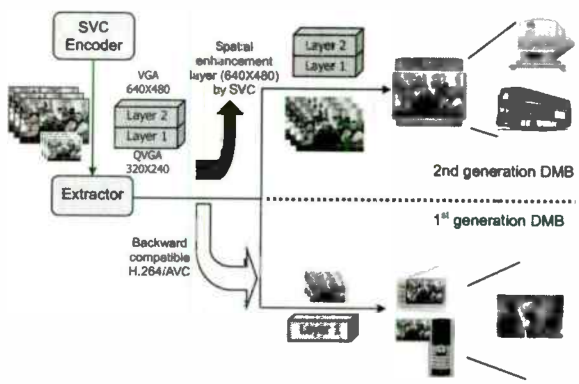


Figure 4 Proposed DMB service using SVC

In T-DMB service environment, there is no variation of transmission bandwidth like internet streaming and there must be no variation of the quality of video and the number of frame per second because T-DMB is a broadcasting service.
In this paper, we only considered the spatial scalability for mobile broadcasting application – higher resolution DMB. Now, the video resolution of DMB is 320x240 pixels (VGA) and would be the base layer of lager video resolution. We designed that one enhancement layer of lager video resolution would be 640x480 (VGA). The video resolution of VGA is similar to that of SD (720x480) in NTSC. Figure 4 shows the service model of SVC bit-stream for the lager video resolution of DMB. SVC bit-stream of DMB would be decoded only base layer in existing DMB receiver so that SVC bit-stream would be fully compatible with existing DMB receiver.

3.1 Structure of SVC encoder

For spatial scalability coding, SVC incorporates interlayer prediction and independent AVC between the base and enhancement layer. The base layer of SVC is compatible with AVC. For the enhancement layer, the inter-layer prediction coding includes inter-layer texture prediction, inter-layer motion prediction and inter-layer residual prediction between two layers. The encoding information such as texture, motion and residual data from base layer is also used in encoding enhancement layer.

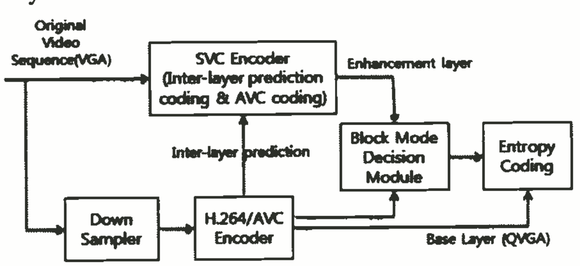


Figure 5 SVC encoder for one spatial enhancement layer

Figure 5 shows the architecture of developed SVC encoder for real-time application.
The spatial enhancement layer is encoded using the inter-layer prediction coding and independent AVC coding. The AVC coding in enhancement layer is the same as that of the base layer. The inter-layer prediction is performed based on Macro Block (MB) mode selected in the base layer. Between modes in the independent AVC coding and inter-layer prediction coding, the mode with the least rate-distortion cost is chosen as MB mode in the enhancement layer.

3.2 Algorithm for optimization

For real-time encoding, we optimized H.264/AVC (the base layer) encoding algorithms using our developed fast intra-prediction, fast zero motion block detection and fast inter-layer mode decision.
Intra prediction of H.264/AVC is also used in the base layer encoding in SVC. It is known that the intra prediction in AVC and SVC achieves a 3-4 dB improvement of PSNR [4]. However this improvement causes the large increment of computational complexity. For fast intra prediction, we developed an efficient fast intra prediction method of AVC and SVC with negligible amount of PSNR drop and bitrates increase.
Sim et. Al [4] proposed an algorithm for the fast intra mode decision suing the probabilistic characteristic of context adaptive information by an off-line training scheme.

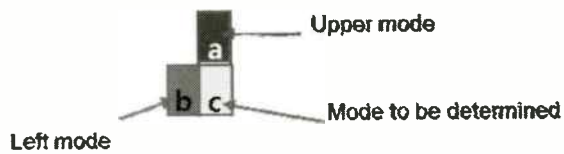


Figure 6 Current block C and its adjacent block a and b

As shown in Figure 6, the current 4x4 block(C) to be intra-coded has high correlation with selected candidate modes at block a and b. As this relation among a, b and c block, we can create the probability table stating which candidate modes are the most probable mode (MPM) by referring to its adjacent blocks. This algorithm can reduce the 9 directional modes to 3~6 candidate modes during the intra prediction process. However, in many cases, due to the data which is sampled form several sequence, the number of reduced candidate modes is limited.
From our intensive observation on the intra prediction statistics for the mode decision, it is founded that the MPM is determined by the equation - $P_c(c|a,b)$, but depends on sequence types and the QP (Quantization Parameter). Also, the probability of MPM varies a lot and each sequence has its own visual pattern. Therefore, proposed algorithm includes the content/context dependent probability and MPM that need to be renewed periodically. The proposed algorithm can obtain almost 80% reduction of the complexity for the 4x4 intra luma prediction.
For the inter-layer motion prediction, the motion vector field including the MB_{BL} partitioning is scaled up into the enhancement layer [5]. In this paper, the block modes predicted from the base layer into the enhancement layer is denoted as $MODE_{BL\_PRED}$. The 8x8, 8x16, 16x8, 16x16 and Intra modes in the base layer correspond to the 16x16 candidate mode of $MODE_{BL\_PRED}$ in the enhancement layer. Notice that the final mode decision on 16x16 candidate mode of $MODE_{BL\_PRED}$ is made by comparing its R-D costs with the ones for 16x16, 8x16, 16x8, 8x8, 8x4, 4x8 and 4x4 modes of the macro blocks in the enhancement layer(MB_{EL})by the independent H.264/AVC coding.
For fast zero motion block detection, we found the

statistics that shows the frequencies of the possible $MODE_{BL\_PRED}$ modes which are predicted from the base layer to the enhancement layer. The statistics were collected from 10000 frames of our test sequences and MPEG test sequence. Here, we define the $MODE_{BL\_PRED}$ into following types:

- Bi-predictive Zero Motion Block (BZMB) is defined as the MB_{EL} with its corresponding MB_{BL} having zero motion vector in bi-direction
- Uni-predictive Zero Motion Block (UZMB) is the MB_{EL} with its corresponding MB_{BL} having zero motion vector in uni-direction
- Zero motion block (ZMB) is defined as the block for either BZMB or UZMB type
- Non-Zero Motion Block (NZMB) is the MB_{EL} with its corresponding MB_{BL}having non-zero motion in bi-directions
- A zero coefficient block (ZCB) for the residuals between the current MB and a motion-compensated block is defined as the blocks with all zero coefficient values of Integer Transform (IT) in the enhancement layer

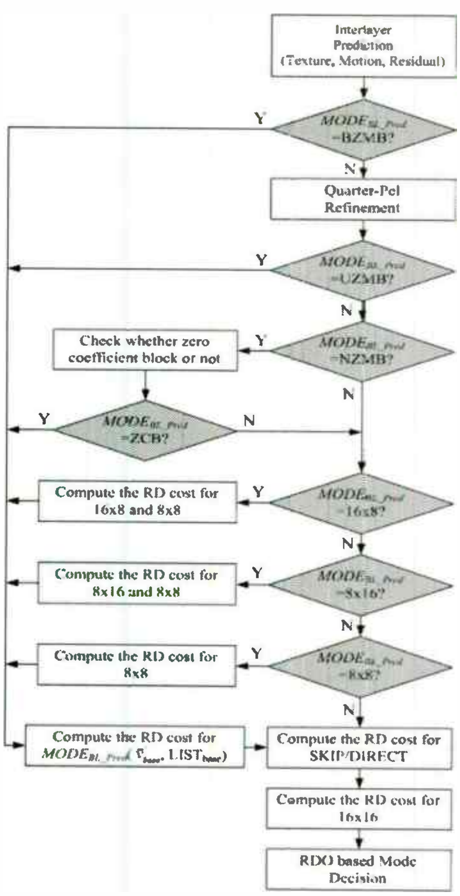


Figure 7 Fast inter-layer mode decision algorithm

After a lot of test, we found the statistics that the portions of the ZMB are very high, compared with those of non-16x16 modes for different QP values in stationary-very low motion video sequences. The ZMB occurs more frequently as QP value becomes larger since the detailed texture is removed for large QP values. For very high motion sequences like a sequence of sports, ZMB is less likely to happen due to high motion. Therefore, the block modes with zero motion vectors at the base layer correspond to the ZMB with large portions of zero motion vectors at the enhancement layer. It also can be noticed that ZCB occurs in large portions for NZMB. This is because ZCB occurs for the blocks with NZMB when its motion activity is small and large quantization parameter values are applied. Motivated from these observations, a fast inter-layer mode decision method can then be developed for spatial scalability coding of SVC by detecting ZMB and ZCB from the base layer.

Figure 7 illustrates the proposed scheme in which the inter-layer prediction for the four corresponding MB_{EL} partitions with $MODE_{BL\_PRED}$ from their corresponding MB_{BL} is performed for the residual or reconstructed texture of an MB_{BL} of nth base layer frame after bi-predictive motion estimation.

The proposed fast interlayer SVC encoding scheme shows that the total encoding time saving is achieved up to 72% in maximum and the degradation of R-D performance is negligible with PSNR drop of 0.25 dB in maximum and 1.73% bit rate increment.

3.3 Implementation of real-time SVC encoding

Figure 8 User interface of SVC encoder

Figure 8 shows user interface of developed SVC encoder. During encoding, the left two windows show input video for encoder and the right two windows show decoded video. Therefore you can compare the quality between original and SVC decoded video. You can also recognize real-time processing. Our encoder has several encoding options: quantization parameter, method of motion estimation, search range for motion estimation, number of frames for encoding and etc.

4. Experimental result

The experiment was performed with video sequences with various motions and textures. For the performance evaluation of our SVC encoder, the encoding time (FPS: Frame per Second) was measured for each test video in conjunction with PSNR drop and bit-stream increment.

The Conditions for our experiments:

- CPU : Intel Core 2 Duo E6600, 2.40Ghz
- Memory: DDR2 800Mhz 4GB
- Video Resolution : Base layer-QVGA(320x240)
  Enhancement layer-VGA(640x480)
- Encoding threads : 4
- Frame number : 3000
- Motion estimation method : Diamond search

Table 1 shows that the real-time processing of SVC encoding is successfully achieved with good performance in terms of PSNR value and encoding speed.

QP	20	24	26	28	30	40
FPS	30.2	32.6	35.8	36.4	39.3	39.54
PSNR	43.3	40.6	38.8	36.9	35.5	27.87

Table 1 Test results of SVC encoder

5. CONCLUSION AND FUTURE WORK

SVC can be applicable in mobile broadcasting environment due to the flexibility of spatial, temporal and quality scalability. In this paper, we propose a service model of mobile broadcasting that provide two spatial scalability using SVC and show a real-time SVC encoder with two layers of spatial scalability. In order to achieve the real-time encoding capability, we optimized algorithm of H.264/AVC for base layer encoding and SVC for enhancement layer encoding with our proposed fast algorithms. The experimental results show the encoding speeds from 50 to 70 frames per second with acceptable PSNR quality.

REFERENCES

[1] ISO/IEC JTC 1/SC29/WG 11, Joint Scalable Video Model JSVM-10, N8963, San Jose, CA, USA, April 2007.

[2] ISO/IEC JTC 1/SC29/WG11:'Text of ISO/IEC 14496-10:2006/PDA M3 Scalable Video Coding', MPEG76/N8015, Montreux, April 2006.

[3] Heiko Schwarz and Detlev Marpe, "Overview of the scalable video coding extension of the H.264/AVC standard," IEEE Tans. Circuit Syst. Video Technol., vol. 17, No 9, Sep. 2007.

[4] D.G. Sim and Y.M. Kim, "Context-Adaptive mode decision for intra block coding in H.264/MPEG-4 Part10," Real Time Imaging. vol.11, 2005, pp.1-6.

[5] He Li, Z. Li and Changyun Wen, "Fast Mode Decision Algorithm for Inter-Frame Coding in Fully Scalable Video Coding", IEEE Tans. Circuit Syst. Video Technol., Vol. 16, no. 7, July 2006

ACKNOWLEDGEMENTS

This work is a joint project between KBS and Information and Communication University (ICU), Daejeon, Korea

SCALABLE VIDEO CODING AND BROADCAST DELIVERY OF 1080P

ELIE SADER
Harmonic Inc.
Sunnyvale, California

ABSTRACT

The new 1080P high resolution format is gaining momentum in production environments and consumer applications. Broadcast deliveries of these formats pose challenges for service provider infrastructures. With 1080P infrastructure equipment still in its "infancy", how can an operator plan to broadcast this rich content? Millions of set-top boxes are deployed and designed for specific codecs; what are the strategies to maintain continuity with the deployed base? This paper will discuss the challenges faced by broadcasters and operators related to bandwidth constraints and the set-top installed base; compression strategies for SVC; and solutions available to deliver new video formats using the current service broadcast model.

INTRODUCTION

Not so long ago it used to be so simple to compress TV: one format (NTSC or PAL), one resolution, one codec (MPEG-2), and one way to carry it (ASI). Those were the good old days! However, in the past few years two things have been radically changing: the kind of video people watch, and how they watch it. Today, a compression system to send full resolution HD in MPEG-4 AVC (H.264) High Profile over satellite has very little in common with a QVGA system for DVB-H or CIF for streaming to PCs. The multiplicity of video formats, codecs, and transmission protocols is likely to grow even more as the number of receiving device types grows to encompass everything from big screen HDTVs to mobile personal media players and cell phones.

In this already complex context how can we improve the current workflow and deliver new formats efficiently? We will explore both sides of the issue: from understanding the technical challenges to exploring different solutions that can be put in place.

ONE STREAM DOES NOT FIT ALL

Exciting times

As of July 2006, most major film studios have stopped releasing new movie titles in VHS format after 30 years of good service. In one decade we have already had to face the arrival of 3 new major video format: Digital TV and DVD, HD and 1080P, respectively carried by 3 codecs: MPEG2 SD, MPEG2 HD and AVC. With the talks of 3D TV and Ultra HD, we can sense that there are only more to come.

The sheer quantity of content formats is slowly pushing the industry to streamline the production and primary distribution processes. In an ideal world, one would want to create and store one and only one source or asset for each piece of content. GXF or MXF format will wrap all the editing information and metadata around the audio and video raw content,from the full screen 16:9 HD to the center-cut low resolution version.

Anyone that has ever watched a Blu-ray movie on a 50" full HD screen knows that it looks amazing! We can't get enough of it. As a matter of fact, some of us don't even watch regular HD broadcasts like we used to a few years ago, much less SDTV… except perhaps for unique events like the Olympic Games. The user experience is changing in terms of accessibility as well. What if you've missed the game and need to leave on a business trip? Should you purchase it on VOD when you get back, or download it to a PSP or iPod for the flight? No more battery on the iPod? No big deal, it can be watched again online on a laptop computer from the comfort of the hotel room. For users, these are exciting times.

As a result of the proliferation of viewing platforms, it is paramount to engineer a smarter and more robust workflow. Keeping tabs on the content proliferation throughout the production, compression and delivery lifecycle is going to get increasingly difficult. Keeping the costs of operation and maintenance under control can make or break a business case.

New competitors

The Internet video revolution has begun and is quickly gaining momentum. As IP networks become faster and content storage space cheaper, the over-the-top (OTT) model presents a growing threat even to large cable MSOs. What is to prevent content owners from streaming content directly to PCs? A new breed of IP set-top boxes are already hitting the market; they allow viewers to stream video from their broadband connections to a TV set instead of a PC.

Both the brand new user-generated content, and legacy broadcast "long-tail" programming, offer a more unique and user-centric experience to be offered by the operator. They are particularly well accommodated by Internet portals such as Google's YouTube.

Internet content has one other advantage in that it is inherently portable and can only get better. Netflix and iTunes are both proposing online rental services including high definition content. They are slowly bridging the "quality gap" with broadcast TV as we know it. Consequently, traditional operators need to stay on top on the innovation curve and move faster towards new formats like 1080P and a further streamlined end-user experience.

More subscribers that want more

Consumers want to watch any content, anywhere, anytime and on any device:

- HDTV at home, including multiple programs in multiple rooms
- TV on the road on personal media players, in cars, airplanes, stadiums, etc.
- Anytime viewing, including pausing live TV that can then be resumed on a portable device

Consumers are more tiered than ever:

- Knowledgeable trend-setting consumers have high end "true HD" big screen displays, PVRs and DVRs, Video iPods or PSPs
- Some have only a "regular" standard television, and they do not know if they watch analog or digital TV
- Some have multiple broadband access points and high data rates, while others have limited broadband or a limited data rate in the case of xDSL eligibility.

Operators must adjust to all of these scenarios and re-invent themselves in order to effectively compete and gain new subscribers.

For example, a satellite operator's HD headend today is VBR in nature and therefore not designed to handle CBR or Capped VBR outputs for IPTV. Yet in order to grow their footprint and subscriber base, a satellite operator has no choice but to diversify and address wireline applications. As a matter of fact, more and more legacy satellite operators are transforming into multi-service aggregators. It's all about leveraging the immense coverage of satellite to offer pre-packaged IPTV services to Tier 2 or Tier 3 telephone companies who cannot afford to build dedicated headends.

1080P: is it that easy?

When HDTV first started, a fierce battle emerged between the proponents of 1080i or 720p with eventually both formats being aired. Progressive definitely offers the best viewing experience, especially for fast moving content like sports and does not suffer from interlacing artifacts. On the other hand 1920x1080 is a better fit for large displays.

Both sides are now coming together pushed, interestingly, by the advent of 1080P capable TV sets and high definition players. With more consumer having 1080P capable devices, some broadcasters are looking at differentiating their service offering with 1080P content. But this is not as easy at it seems.

Firstly, 1080P is a sound byte that doesn't accurately refer to the nature of the material. Are we talking about 1080P at 24 frames per seconds (exactly 23.976), also known as 1080PsF, or are we talking about 1080P at 60 frames per seconds?

1080PsF is typically a format compatible with motion picture content and can be routed over a regular HD SDI interface per SMPTE 292M (1080i, 720p) at 1.5 Gbps. Compatible with the existing headend infrastructure, many current AVC HD set top boxes are also able to decode this format, provided a firmware upgrade. We also find this format in Blu-Ray discs for instance.

But 1080PsF is not suited to broadcast a live sport event, as 24 frames per second is not enough to adequately render the high motion content; even on a 1080P 120Hz TV set that would replicate the 24 frames which could lead to some undesirable processing.

1080P at 50 or 60 frames per second is the true convergence format for broadcast distribution. The catch is that the infrastructures for this format are still in infancy. Equipment such as cameras that offer non-objectionable performances, servers and routers that can robustly process 3 Gbps streams per SMPTE 292M (1080i, 720p) is required. When preparing the world's first demonstration of 1080P 50Hz AVC encoding at IBC 2007, I found it very challenging to actually gather compelling content, reliable servers and capable monitors. Typical pitfalls included 1080P monitors with 1080i internal circuitry, or dual link 1.5Gps HD SDI servers instead of single link at 3Gps. The conundrum for operators is that existing, deployed, AVC set-top boxes are not 1080P50/60 capable. New silicon will be ready soon, but that entails replacement of the current "new" generation of boxes.

While current hardware like Harmonic's Electra™ 7000 HD AVC encoder can already process 1080PsF content for encoding of motion picture (24 fps) content, it is predicted that the first broadcast of 1080P50/60 will happen in the late 2009 or early 2010 timeframe.

One stream does not fit all

To stay ahead of the pack, an operator today must prepare multiple copies of the same content:

- 1080P version of a motion picture or a network show for ultra high-end users
- Regular HDTV for prime-time broadcast in 1080i or 720p
- Standard TV 480i version to be carried to the mainstream installed base
- QVGA version for PC streaming or ready-to-go PMP asset
- Low resolution for picture-in-picture (PIP) or mosaic applications
- On-demand assets (in various resolutions) that are server friendly in order to fit in least amount of disk space

Owing to this complexity, one can understand the increased scrutiny regarding the production and distribution systems' ability deliver this new rich content with a coherent workflow. Also it is clear that, without even broaching the subject of compression or bandwidth, an inherently scalable solution is required.

ONE NETWORK MUST FIT ALL

Networks can choke

The very nature of broadcast video is that it quickly requires huge amount of bits to send live programming. Regardless of the medium, all operators have a finite amount of bandwidth to serve all of their users:

- Satellite has only so many transponders available on expensive satellites
- Terrestrial TV has a given set of regulated airwaves
- Cable and IPTV have only some many home passed and bandwidth limitations for broadcast services
- Phone networks are arguably the most constrained network in terms of bandwidth and have end devices with limited capabilities and battery life
- Internet is limited by the shared and heterogeneous bandwidth on the operators' backbones. as well as the enforcement of "network neutrality"

Some are predicting the doom of data networks within a couple of years. This impending choking is not yet a reality but video is clearly pointed to as a leading concern.

Video delivery has a cost

At Harmonic we know this dilemma only too well. The scarcity of bandwidth resource drives a huge investment in research and development of bandwidth-efficient video delivery systems. For our customers, the cost of owning frequencies or airwaves rights, owning satellites or copper networks, leasing transponders or backbone links, amounts to a significant annual cost that dwarfs the compression equipment costs.

The other end of the delivery system has a cost too, perhaps the highest for pay TV operators. There are millions and millions of set-top boxes deployed in the field today. Their limited capabilities in term of demodulation and decoding mean that introducing a new format like 1080P has a huge opportunity cost . Typically, a traditional MPEG-2 SD STB is not able to decode or display high definition. Or, the STB may only handle MPEG-2 SD and HD, but not MPEG-4 AVC (H.264), as in most cable systems.

One network must fit all

Moving to a new video format like 1080P leads to three major consequences for the delivery network:

1. Additional network bandwidth required for the new format

2. Simulcast of new and old formats to support both the old and the new STBs

3. Progressive replacement of millions of STBs to eventually eliminate the need for simulcast and reclaim bandwidth for future applications

NEW COMPRESSION STRATEGIES

Scalability is key

Over the years many technical innovations have emerged in order to maximize the use of every bit available to video delivery:

- Digitization of the airwaves to send multiple videos per chunk of spectrum

- MPEG compression using spatial and temporal redundancy, as well as encoder-to-decoder variable buffers to allow variable picture size

- Variable bit-rate compression and statistical multiplexing

- Multi-pass encoding and filtering

- Remote statistical multiplexing by locating encoders at the source and using IP WAN transports

While all of these strategies are undeniably extremely effective, and while more are being researched, they all fit the same category: delivery of content in one resolution.

Scalability—being able to address multiple applications / resolutions out of one generated copy or stream—is a missing ingredient that could solve many problems.

Scalability also means backward compatibility as a core requirement, as one must support the past formats and what is yet to come.

Without scalability the current industry paradigm is to encode a given video multiple times at different bit-rates and resolutions (let us call this method simulcasting for a lack of better definition) and that creates a few significant problems:

1. Bandwidth crunch on the operator's delivery network since bandwidth is not infinite

2. An increased amount of headend equipment like encoders, as well as increased complexity in control and management

3. A need for a switching mechanism to provide the right copy of the video to the end user

Scalable video coding: "SVC"

A potential solution is scalable video coding (SVC). The prime driver for SVC technology is to provide ways for Pay TV operators to maximize the use of their delivery bandwidth. Thus SVC makes complete sense for someone looking at delivering new video formats like 1080P, and new services like PC or mobile streaming of the same video content. For instance, 1080P/60 will introduce a disruption to the current 720p or 1080i HD formats, requiring new STBs capable of decoding it.

Managing the millions of (already!) legacy H.264 STBs is much more efficiently handled by a SVC encoding approach.

A multicasting scenario would create a bandwidth crunch: the operator would fully encode a 720p stream (legacy STB) and a 1080P stream (new STBs required). Using SVC, however, a video signal can be coded as a base layer for a given format, and incremental enhancement layers for additional formats. For example, a base layer can be in 720p, which can be decoded by HD set-top boxes already in the field, and an enhancement layer to 1080P for future or newer generation set-tops. The main advantage over simulcast is bandwidth savings, since the enhancement layer contains only the "information difference" between 720p and 1080P. Another possible use case is a QVGA or CIF base layer for mobile and an enhancement layer for SDTV.

The typically touted bandwidth savings for SVC is around 40%: one can encode an H.264 AVC base layer stream in 720p, compatible with the current STBs, as well as encode an SVC enhancement layer permitting those new STBs to decode both layers and produce the 1080P video. The trick is that the enhancement layer is only encoding the missing data from the base layer.

The bottom line is that operators looking at offering those new applications must deploy new STBs. Therefore, proponents would argue that those new STBs should be SVC compatible to help their bandwidth efficiency.

SVC is a standard now

The Joint Video Team (JVT) of the ISO/MPEG group and the ITU-T ratified the SVC extension to MPEG-4 AVC (H.264) in November 2007. SVC is a new Annex to H.264. SVC uses AVC to code its "base layer" video.

The adopted technique did the best job in extending MPEG-4 Part 10 (AVC) – in a backwards-compatible way – to support scalable video compression.

The aim of the scalable encode is to allow the reconstruction of a high quality video bitstream from one or more elementary bitstreams that can themselves be decoded, analogous to Russian dolls.

An elementary bitstream can represent a lower spatial or temporal resolution or a lower quality video signal (each separately or in combination) compared to the whole transport from which it is derived.

Scalability is achieved in multiple ways with SVC:

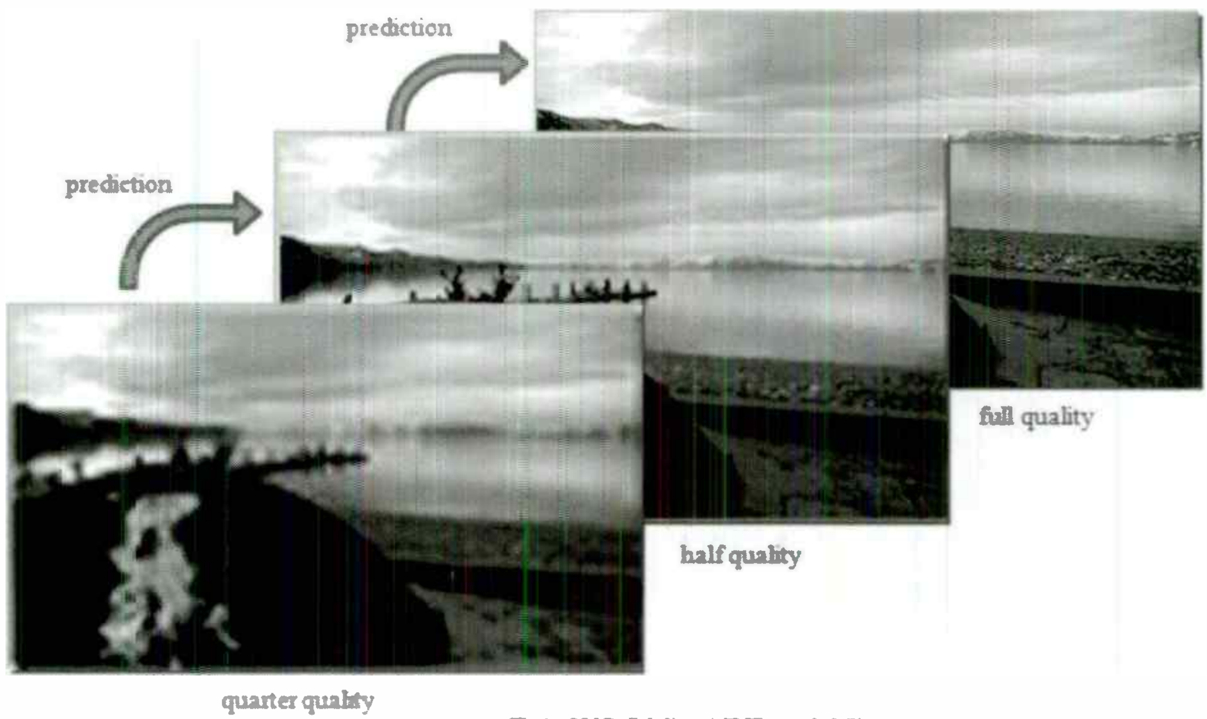


Fig1. SVC fidelity / SNR scalability

- Temporal Scalability: Complete picture frames can be dropped from the bitstream.
- Spatial Scalability: The video is encoded at multiple spatial resolutions. The elementary bitstreams of lower resolutions are used as base layer to predict the data for higher resolutions. This reduces the overall bit-rate to code the various resolutions compared to simulcast.
- Fidelity/SNR Scalability (fig. 1): The video is encoded in a single spatial resolution but at different qualities. The elementary bitstreams of lower qualities are used as base layer to predict the data for higher qualities. This reduces the overall bit-rate to code the various qualities compared to simulcast.

The SVC specification, which is found in Appendix G in the latest editions of the AVC standard, introduces the new tools/syntax elements used for more efficient encoding of a single video sequence with two different resolutions. The smaller resolution is typically termed the "base layer", and is coded with regular AVC modes and can be decoded with regular AVC decoders. The higher resolution is termed the "scaled layer" and is coded with the regular AVC modes in addition to three new MB modes that exploit the redundancy between base-layer and scaled-layer to add extra compression efficiency compared to multicasting. These three new inter-layer MB modes are:

1. Inter-layer Intra (base layer in Intra mode)
2. Inter-layer Inter (motion, refs and partition info reused)
3. Inter-layer Residual (the residual is scaled and coded, not the video)

Some studies have found that through careful use of all three of these modes, along with very sophisticated rate-distortion balanced mode decision and rate control, it may be possible to code a video signal scalably with only a 10% penalty relative to the cost to code the same signal with AVC. As an extension to H.264, SVC will work in conjunction with H.264 which will remain the codec of choice for typical broadcast video formats. SVC is attractive when one considers the added bandwidth cost to bring "exotic" video formats (1080P, QVGA) into a typical SD/HD service architecture, but not as a pure codec.

SVC: spatial scalability

Let's return to the previous example of a user viewing a live HD broadcast on a big screen TV at home, who decides to pause live TV and wants to resume the viewing at a later time on a portable device.

The user wants nearly instant portability of the content to the personal media player. He doesn't want to sit through minutes or hours of file transfers and file transcoding.

From the broadcaster's perspective, what the user needs is both the full resolution HD content and a lower resolution version that is a spatially scaled down version of the former. The broadcaster also needs a way to manage the data transfer for the user and keep the content secure.

SVC promises a simple, fast and elegant workflow in this scenario. The base layer is the lower resolution stream and the enhanced layer stream is the missing spatial information of the HD version. Both streams are sent in one transport and easily managed by the various consumer devices.

Another example of SVC's spatial scalability is IPTV's problem of eligibility over xDSL networks. Users that reside far from the DSLAM might not be eligible for HDTV service. What should the operator do? Only deliver an SDTV service? Multicast in SD and HD? Is that granular enough? Is that bandwidth efficient?

With SVC , one transport could be generated containing the SD base layer and one or more HD enhancement layer. Therefore with minimal bandwidth overhead the operator could provide:

- The SD base layer to those who are the farthest
- The SD base layer and a half resolution HD enhancement layer for those subscribers with medium reach
- The SD base layer and the full resolution HD layers for the closest subscribers

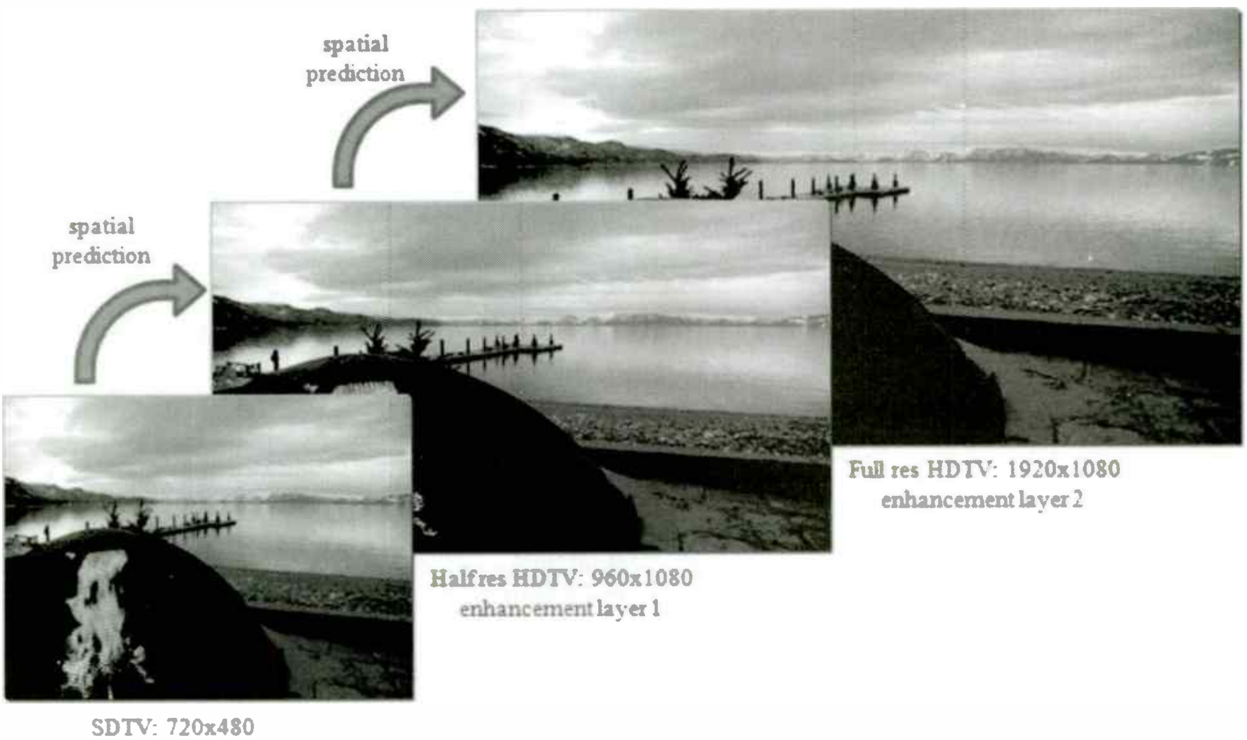


Fig2. SVC spatial scalability

This example (fig. 2) illustrates the flexibility brought by a scalable approach. More intelligence is however necessary in the delivery network to appropriately route the correct streams.

Once can compute the projected savings (fig. 3) and the more the operator is trying to enhance the service offering the more the savings add up.

SVC: better quality of service?

Some proponents of SVC will also advocate that SVC provides an additional benefit in terms of quality of service (QoS). Let's return to our last example in which an IPTV operator has a base layer in SD and multiple enhancement layers for higher resolutions. In the event of a network failure that would take down part of the delivery network's bandwidth, the service could be maintained by only routing the base layer SD streams through what is left of the available bandwidth.

While such implementations are years away, this is definitely the type of advanced scenarios that need to be adequately addressed.

Scalable AVC as an alternative to SVC?
Non-SVC proprietary scalable video coding techniques might be simpler to implement than SVC--both on the decode and encode sides--and might outperform SVC if only a subset of the SVC toolkit were implemented.

This scenario is rather unlikely, as it would require a joint push from encoder and set top manufacturers to have any chance of winning their acceptance at major operators.

The preferred way for both manufacturer and operators is to work through standard bodies' long term, open, robust, solutions.

Roadblocks to the scalable model
An HD-capable full SVC encoder is very complex to implement in terms of technology. For instance, there are no silicon chips available that can process this encoding or decoding today, and frame accurate synchronization is not a simple task. The industry is still very much focused on its "traditional" HD deployments, and it is believed that the availability of 1080P/60 content will actually drive the go-to-market timing of the SVC technology both on the professional and consumer sides. However, compression improvements in H.264 and new bandwidth efficiency schemes in the coming years could make the simulcast scenario somewhat a more straightforward approach for many operators and simply defeat the SVC scenario in the short term (say through 2013)

There is no doubt that in the long term all new video format and codecs will have to provision for scalability.

SUMMARY OF ADVANTAGES AND BENEFITS

In the future, can video from the same full HD source be intelligently scaled and efficiently distributed to devices of different kinds over different networks? We believe so. It will take a revolution as big as the advent of MPEG-2 compression 15 years ago. In the encoding world, advanced transcoding and SVC are steps in the right direction. But it will require much more than compression technology improvements to make it all happen. It will take new levels of QoS to make the network "format aware", and asset management systems tightly linked to edge devices with far greater intelligence. The opportunities lie ahead, now it's up to us to shape the future of our industry.

Harmonic is engaged with customers to explore the requirements – including timing – for scalable video products. Additionally, Harmonic is researching implementation architectures and encoding technologies to enable the best possible products.

ACKNOWLEDGEMENTS

All my estimated colleagues at Harmonic Inc, especially:

- Arnaud Perrier, Senior Product Marketing Manager Encoders
- Patrick Waddell, Manager, Standards & Regulatory
- Paul Haskell, Senior Director, R&D

Scenario	200 SD + 100 HD		200 SD + 100 HD with SD simulcast (enhanced eligibility)		200 SD Ch + 100 HD with SD and Half res HD simulcast (enhanced eligibility ++)	
Bandwidth requirements	Non SVC 200x2+100x 8=1.2 Gbps	SVC 200x2+100x 8=1.1 Gbps	Non SVC 200x2+ 100x8+100x 2=1.4 Gbps	SVC 200x2+ 100x7+= 1.1Gps	Non SVC 200x2+ 100x8+100x 4+100x2= 1.8 Gbps	SVC 200x2+ 100x3.5+100 x4+= 1.15Gps
SVC savings Mbps	0 Mbps		300 Mbps		650 Mbps	
SVC savings in %	0 %		21 %		36 %	

Fig3. SVC generated savings

Understanding and implementing JPEG2000 compression for long-form EFP acquisition

John R Naylor and David Bancroft
Thomson, Beaverton, OR

While the concept of digital compression is now universally recognised, there is still widespread concern and confusion about which codec is appropriate at which stage of production and delivery - and what degree of compression is acceptable.

Compression, incidentally, is not just an electronic process. The human eye has approximately 130 million photo-receptors but only one million nerve connections to the brain. Given that more than 50% of the human brain is involved in visual processing, we are pretty good at compression and decompression ourselves!

Returning to digital technology, in an ideal world the production workflow (at least) would not use compression at all. Unfortunately, "technically ideal" uncompressed HD requires a data rate of close to 1.5Gb/s[1]. While this can be handled in some special circumstances, it is simply too much data for most applications. In particular, recording and replaying such data rates in real time is very challenging.

Conversely, applying compression, by its very definition, implies affecting the quality of the image. In fact, you are throwing away some of the data. Those who develop compression schemes have to find the balance between the amount of data that is discarded and the impact on image quality.

That balance is just one decision to be made when developing or selecting a compression scheme. Another is its computational complexity. It goes without saying that a video codec has to operate in real time, certainly in playback - but what sort of processor power is required to achieve that?

MPEG-2 and now H.264 MPEG-4 are particularly suited to delivery because they are asymmetrical in their computational demands. Encoding calls for considerable processing power, especially if high compression ratios are required, whereas decoding is simple enough to be performed by a low-cost chip embedded in a consumer-priced set-top box or receiver.

But that asymmetry is not necessarily appropriate for professional applications where equal processing power can be included in each device. There the requirement is more dependent upon quality, or upon other factors such as ease of editing.

So there is no single answer to the compression question. In fact there are at least four, as can be seen in figure 1. The four stages of capture, edit, playout and delivery all have different requirements, and therefore logically should use different codecs.

This paper considers only the production workflow, so focuses only on the capture stage and how it interacts with the edit stage.

---

[1] 1920 x 1080 x 10 bits x (1 x Y + ½ x Cr + ½ x Cb) x 30 frames per second = 1,244,160,000 bits per second of payload data, plus audio, control and network overhead

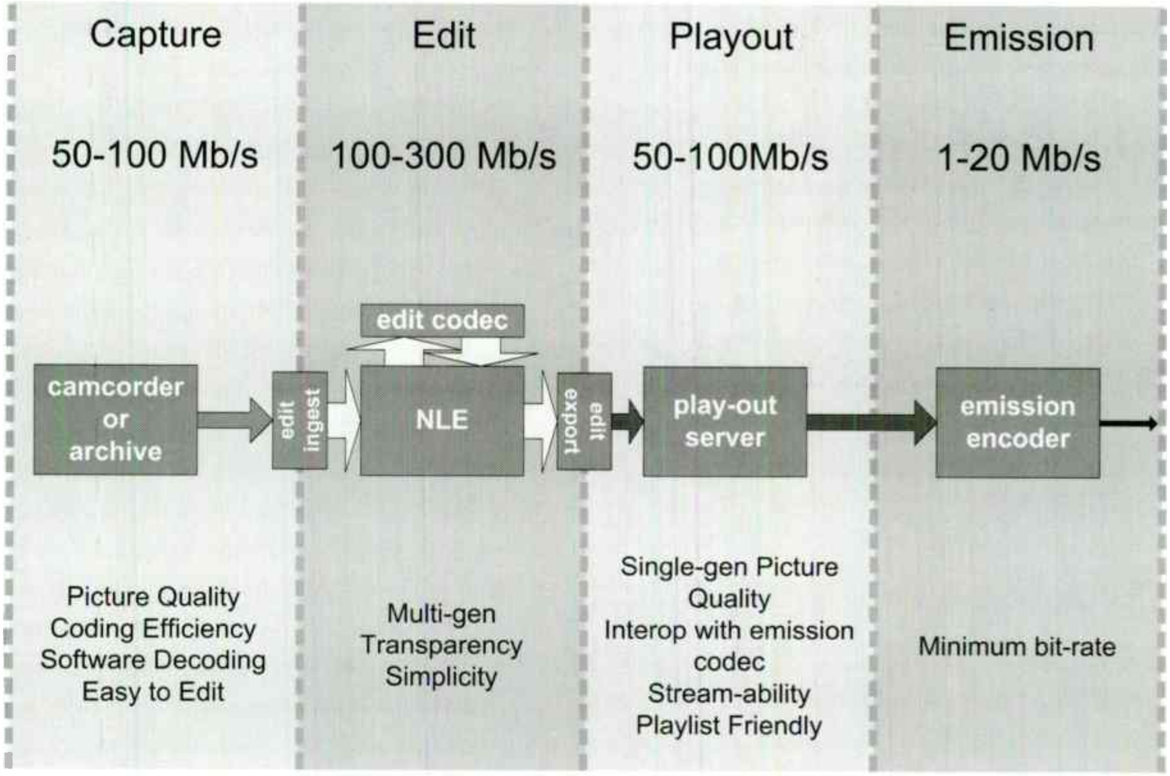


Figure 1 - The Four Stages of the Production Workflow

CAPTURE

"Garbage in, garbage out" is a long-established principle in computing, and it applies equally well to video production. If the image in the camera is poor then nothing you do downstream is going to improve it.

Even considering something to be "good enough" is a dangerous assumption. Today's recordings are tomorrow's archives. Outside the United States, many production companies are shooting now in high definition. This is not because the programme will be broadcast in HD immediately, but because it future-proofs the content.

The capture stage is not a place where picture quality should be traded off lightly. Yet that has been common practice, and compromised solutions continue to be championed.

The SMPTE standard 274M encodes HD video at a resolution of 1920 x 1080 pixels, usually using 10 bits per sample and 4:2:2 colour sampling. Note that, while being regarded as the gold standard, this encoding in itself includes a compromise in a form of compression: the colour is sub-sampled at half the rate of the luminance. This closely mimics the operation of human vision, and is generally regarded as visually lossless.

Some capture systems introduce further compromises, including:

- a sub-sampled raster, where only 1440 pixels horizontally are captured, rather than 1920 pixels, throwing away 25% of the available information
- more aggressively sub-sampled colours, processing 4:2:0 (another 25% loss) or worse; and/or,
- truncated samples, using only 8 bits rather than 10 bits - another 20% of the signal gone forever.

Camera systems using legacy codecs such as MPEG-2 and the DV family have typically made at least one - and sometimes all three - of these additional compromises, all of which can accumulate to a potential overall information loss of 55% before you even get to a codec.

These compromises were initially implemented because of the restricted recording bandwidth available, particularly on compact tape formats for use in camcorders. As developments in media technology progressed, so the spotlight shifts to the record time available: there is always pressure to capture more on a single tape or other unit of media.

The use of advanced codecs such as JPEG2000 and H.264/AVC changes the balance slightly. They provide higher compression efficiency, which can either improve quality at a given data rate, or in some cases provide lower bitrates for an equivalent quality. These new formats are, however, computationally complex.

There is an important issue here. You can tackle the issue of computational complexity in two ways. You can throw processor power at it – which becomes power-hungry with all the problems that entails – or you can develop a dedicated chipset to implement the coding in hardware.

For those of us in the broadcast industry, it is rare to have the opportunity to develop such an application-specific chip, because the volumes are simply not there to make it an economic proposition. JPEG2000, though, is not just a broadcast codec. Indeed, it was deliberately developed to have a wide range of applications and therefore a much greater volume. Those applications are likely to be in consumer devices every bit as much as professional environments, so the economies of scale in developing dedicated silicon are readily available.

Analog Devices, for example, is now on its third generation of JPEG2000 coding chips, the ADV212, and is finding applications in CCTV and surveillance systems, and in digital still cameras, as well as in broadcast quality camcorders.

In developing the Infinity digital media camcorder, the Thomson design team made the decision to offer the user a choice of codecs: MPEG-2, DV as well as JPEG2000 using ADV212 chips. The legacy codecs are provided to allow our customers to migrate in an orderly fashion from existing workflows, with the recognition that JPEG2000 delivers better images.

We see JPEG2000 as the codec for the future, particularly in HD, and I will devote the balance of this paper to outlining the practical benefits that JPEG2000 brings.

JPEG2000

The original JPEG format was developed more than 20 years ago. As its name suggests (JPEG stands for Joint Photographic Experts Group), the format's original application was as a digital photography standard. It is only thanks to determined video engineers that it also became a motion standard. When JPEG set out to create a new standard for the 21st century, moving images were an integral part of the design from the outset.

JPEG and MPEG each depend on a technique of dividing each image up into blocks, then applying discrete cosine transforms (DCT) to each block to achieve the compression required. This has the advantage of making relatively modest processing demands in the encoding stage, but it has an equally enormous disadvantage in that, when the compression is pressed hard, the coding produces discontinuities at the block borders. This is the blocking we have all seen in compression. Because it manifests itself as perfectly horizontal and vertical straight lines – which are virtually unknown in nature - it is immediately and objectionably visible.

Many implementations of MPEG for professional video also use temporal compression to get the bitrate down. A group of pictures will include one I-frame with full picture information, and the rest of the group will be references to that I-frame.

This approach makes editing a nightmare, for the inescapable fact is that, in a group of pictures, some will look better than others, and if you need to cut on one of the not so good pictures you have a real problem. Another issue is that, should there be a data corruption in an I-frame, it will potentially be visible on every frame in the GoP.

Taking advantage of advances in the science of image coding, the developers of JPEG2000 have been able to tackle both these issues.

First, rather than DCT, JPEG2000 uses wavelet transforms to process the image and reduce the amount of data. Do not worry: we will not delve deeply into the mathematics of transforms here. The point we need to underline here is that, because wavelet transforms code both frequency and location information at the same time, there is no need to divide the image up into blocks: images up to 4k resolution can be processed in one pass.

The simple way to think of wavelet transforms is as a digital filter which is passed over the image, separating high and low frequencies, horizontally and vertically. After the transform you end up with four elements, wavelet coefficients representing the high horizontal and high vertical frequencies (conventionally described as HH), high horizontal and low vertical frequencies (HL), low horizontal and high vertical (LH), and low horizontal and vertical (LL).

Figure 2 – original image[2]

As you can see from figure 3, the LL element is a half-size, half-resolution version of the original image. The wavelet transform can be applied multiple times. Figure 4 shows the result of the second pass on the LL element of the first pass, with the very top left now a quarter size (in each direction) version of the original.

[2] Figures 2 thru 4 are reproduced from *Digital Video Compression* by Peter Symes with the permission of the author.

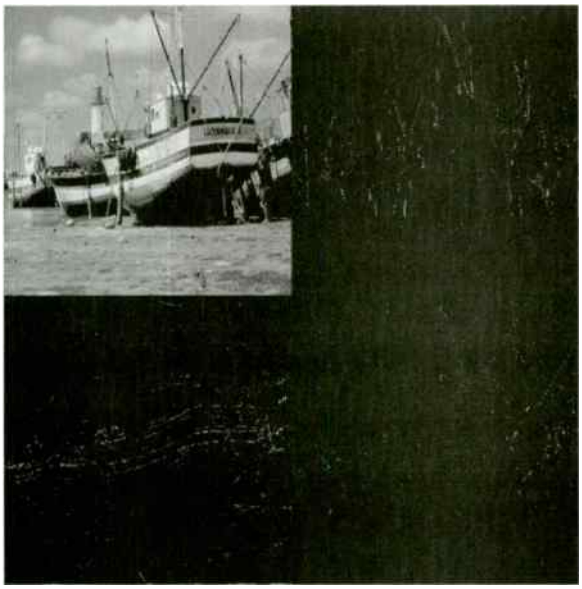


Figure 3 – first wavelet pass, with LL top left, HL top right, LH bottom left and HH bottom right

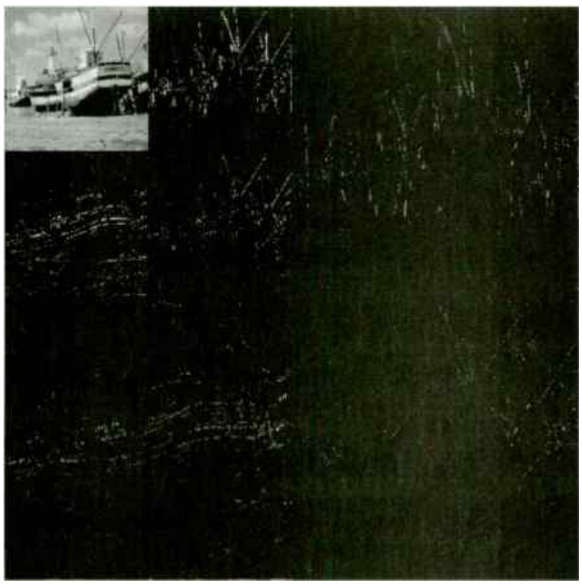


Figure 4 – second wavelet pass, showing the original is now at one quarter resolution

Part of the "intelligence" in video compression – and this applies to all coding schemes - is that coefficients representing only high frequencies may be quantised more coarsely without significant change to the perceived image quality. So it is clear that this iterative process produces a signal that is progressively easier to compress.

That is critical to the success of JPEG2000, and it also leads to a side benefit that is of enormous importance in practical professional video workflows. Each wavelet transform includes a copy of the original, half the size of the previous pass. So when compressing a 1920 x 1080 image, wavelet compression automatically generates a 960 x 540 version, then a 480 x 270 version, and so on.

In the decode process you can choose to pull out the LL version at any stage. So you have a thumbnail version of any image without any additional processing. In video terms, you have a browse version created as an integral part of the compression scheme itself.

So if you are shooting HD on location and want to start editing on a laptop, you simply work on a 960 x 540 version, or even a 480 x 270 version. It does not take any more processing to create this low-resolution version: in fact it takes less processing! And because the browse version is forever embedded into the full resolution signal the two can never be separated so you are always working on the right content at the right moment, potentially right up to the moment of delivery.

In fact, the way that the JPEG2000 encoding is implemented in the Infinity camcorder, and in Thomson Grass Valley Aurora and EDIUS nonlinear editors, is that the data is layered at the time of encoding, as can be seen in figure 5.

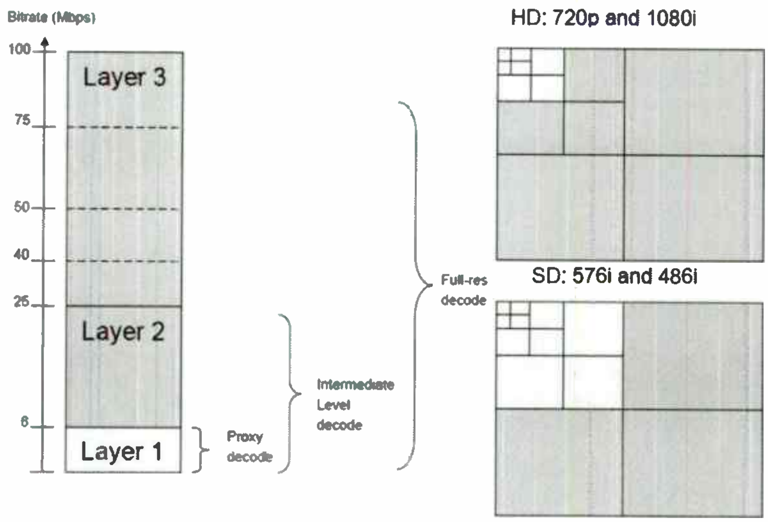


Figure 5 – JPEG2000 Layering Scheme for Infinity Content Format

Layer 1 is a 6.25Mb/s stream for browse quality: quarter resolution HD or half resolution SD. Layer 2, together with layer 1, totals 25Mb/s and provides half resolution HD or OK SD. Layer 3 brings the full signal (in Infinity it is selectable between 50, 75 and 100Mb/s) for full quality.

One important benefit of the ready access to these lower bitrate versions of the identical signal is that they can be used in editing on platforms with limited power. That means, for example, that journalists can edit stories in the field using a laptop editor.

That is not the only way to implement the progressive nature of wavelet compression within JPEG2000. In news, for instance, it might be important to get the full size picture home as soon as possible, so you could re-organise the bitstream so that a soft but full resolution image is delivered and decoded first, getting progressively sharper as more data is decoded.

IMAGE QUALITY

The clever things that can be done with the data structure of JPEG2000 are of course important, but are meaningless if the codec does not deliver good quality.

Subjective and objective testing, by Thomson and others tends to confirm that visually satisfying reproduction is achieved in JPEG2000 when the bitrate of the compressed signal corresponds to between one and two bits per pixel in the original image.

In developing the Infinity digital media camcorder, the design decision was taken to offer the user high definition JPEG2000 settings of 100Mb/s, 75Mb/s and 50Mb/s. In 29.97 frames per second video these represent 1.6 bits per pixel[3], 1.2 bits per pixel[4] and 0.8 bits per pixel[5] respectively. 100Mb/s delivers excellent quality and 75Mb/s good HD. While 50Mb/s should be regarded as marginal, it may be acceptable in some applications as a trade-off between quality and recording time.

For 25 frames per second markets, the yields are 1.9 bits per pixel at 100Mb/s[6], 1.4 bits per pixel at 75Mb/s[7] and 0.96 bits per pixel at 50Mb/s[8] – even closer to the ideal.

---

3 100 x 1,000,000/(1920 x 1080 x 29.97) = 1.609

4 75 x 1,000,000/(1920 x 1080 x 29.97) = 1.207

5 50 x 1,000,000/(1920 x 1080 x 29.97) = 0.805

6 100 x 1,000,000/(1920 x 1080 x 25) = 1.929

7 75 x 1,000,000/(1920 x 1080 x 25) = 1.447

8 50 x 1,000,000/(1920 x 1080 x 25) = 0.9645

As already noted, the layered structure adopted for the Infinity camcorder also includes a "quarter HD" stream at 25Mb/s and a "quarter quarter HD" stream at 6.25Mb/s, which are created as part of the inherent multipass wavelet compression of JPEG2000.

Image distortion introduced by excessive wavelet compression tends to take the form of a softening of detail or slight smearing, which is psycho-visually much less obtrusive than blocking. The complete absence of macroblocking ensures that JPEG2000 compression degrades under pressure much more gracefully.

This also has an impact downstream. Blocking artefacts from the acquisition compression can stress the transmission compression encoder, so even if MPEG errors in acquisition are not visually disturbing on the studio monitor, they may cause a significant reduction in the quality of experience once compressed for delivery.

In comparison, any softening from JPEG2000's wavelet compression process would not cause problems for the transmission encoder. Indeed, any slight softening of the image would actually reduce the load on the downstream MPEG compression.

That wavelet compression does not generate offensive artefacts when challenged has been used as a benefit by other manufacturers. The new entrant into the digital cinematography market, RED Digital Cinema Camera Company, has based its proprietary REDCODE™[9] on wavelet compression. As already noted, JPEG2000 is also suited to other video applications such as surveillance, and has been implemented by Ikegami among others. At Thomson, we employ wavelet compression, in a scheme based on JPEG2000, in our digital HD wireless camera systems.

---

9 REDCODE is a trademark of the RED Digital Cinema Camera Company

Perhaps the best assurance that JPEG2000 delivers excellent image quality comes from the fact that the Digital Cinema Initiative, the consortium of seven top Hollywood studios, selected this codec after thorough evaluations as its standard for delivery of movies. The DCI's stated aim was to select the standard which guaranteed the high presentation standards possible in the digital cinemas of the future.

JPEG2000 IN THE WORKFLOW

After capture, content is passed to edit workstations, and so it is important that any codec implemented in a camcorder can be used downstream.

It is important to make the point that JPEG2000 is an Intraframe compression scheme. There are no temporal issues to make life difficult for the editor.

As we saw earlier, the requirements for a compression scheme in an editor are very different to those in a camcorder. In a fully-featured craft editor, there is a need to balance very high quality to ensure transparency after multiple passes with the requirement for a relatively simple codec which allows very many layers to be combined with complex effects.

For this reason, most edit suppliers have developed their own internal codecs to which camcorder formats are transcoded on ingest. Examples are EDIUS HQ and Apple ProRes 422. Some manufacturers have already developed a JPEG2000 input, and the others are sure to follow, particularly as it is in the interest of the editor to encourage the use of a high-quality, non-compromised, intraframe only origination format.

The latest version of the EDIUS craft editor supports JPEG2000 directly on the timeline without the need to transcode.

As already noted, the bit-layering approach inherent in JPEG2000 makes it possible to chose lower quality layers when the CPU power or disk performance is limited. For example, with editing on a laptop computer, the edit process is just as quick and convenient as on a desktop, and when the edit is finished the system conforms the high resolution material for transfer as a file to a playout system.

The Thomson Grass Valley Aurora news editor – which provides basic editing functionality on a journalist's workstation – can also access this JPEG2000 embedded proxy and work directly on it, including jog, shuttle, and sub-clip selection, again with the decisions made conformed on the full resolution material before delivery.

Conclusion

As a codec for acquisition, JPEG2000 offers a number of important advantages:

- Full resolution, full colour capability;
- Superior compression performance, with freedom from artefacts and the complete absence of blocking;
- Lossless and lossy compression, with visually satisfying performance at practical bitrates of 100Mb/s or less;
- Graceful degradation under severe pressure;
- Intraframe coding for freedom from temporal constraints;
- Multiple image resolutions created as part of the compression process; and,
- Progressive transmission by pixel and resolution accuracy.

Implementing this codec in a camcorder is a challenge as it is computationally complex. However, because JPEG2000 is an open standard developed to be applicable to a broad range of markets. ASIC manufacturers have sufficient interest from other industries to recognise the high volumes needed to justify the development of dedicated encoding chips. This has allowed manufacturers of relatively specialised equipment, such as broadcast camcorders, to build systems which are within the accepted norms of power consumption, weight, operational practicality and affordability.

JPEG2000 is seen as the codec to meet many of the challenges and requirements we face in acquisition, production and archiving master content developed for both broadcast and digital cinema applications. Today's users will migrate towards it as they introduce new products like the Infinity digital media camcorder alongside legacy equipment.

The benefits are enormous as JPEG2000 provides a cost-effective, full HD resolution compression solution for camcorders, and also one that offers increased flexibility for editing. Furthermore, it is scalable to even higher resolutions and visually lossless performance should that be needed in other applications – making it a key component in future-proofed solutions.

Bridging the Gap with HD Transcoding

Tim Simerly
Texas Instruments
Dallas, Texas

Introduction

Since the early 90's, the MPEG-2 systems and audio/video (AV) codecs have been the predominant mode of transporting and storing AV content for cable, satellite and digital storage medium (DSM) applications. For cable and satellite, the MPEG-2 Transport Stream (TS) has been the preferred method of transport while for DSM applications like DVDs, it has been the MPEG-2 Program Stream (PS). While the TS and PS container formats have served the industry well, they are not the preferred format of the portable devices. With the portable devices, it is the MPEG-4 MP4 file format that is the most often used by such devices as cell phones, personal data assistants (PDAs), gaming devices (e.g., play station portables (PSPs)) and personal AV players (e.g., iPODs). In addition, the networking protocols that are used most often by these devices in streaming media applications are multicasting, HTTP with MP4 and RTSP/RTP.

And with the growth of these portable devices, there have been a huge number of new applications that have become commonplace in the home. Some examples include multiplayer gaming, true video-on-demand (VOD), iTunes, social networking (MySpace), video sharing (YouTube), and video home editing. Figure 1 provides an illustration of how some of these applications are already being used in the home.

Figure 1: AV Media Applications in the Home

And with the introduction of AV portable devices, a new era of AV media content is arriving. It's no longer just a dedicated device such as a STB or DVD player providing one type of service (i.e., digital MPEG-2 TV) in the home. It is turning into many devices with many different types of services and applications being made available to the consumer with a broad range of content (i.e., MPEG-2, MPEG-4 Simple Profile (SP), H.264, VC-1, On2, DivX). With the consumer, content is becoming king, and they want it anywhere, any time and on any device.

The amount of content is growing exponentially along with a corresponding increase in the number of ways of accessing it. The internet and cellular networks are also being used to access AV media and it is being accessed and viewed all over the home, not just in the living room. Figure 2 provides an example illustration where we see content being viewed and accessed in separate bedrooms, dens, and even from the car.

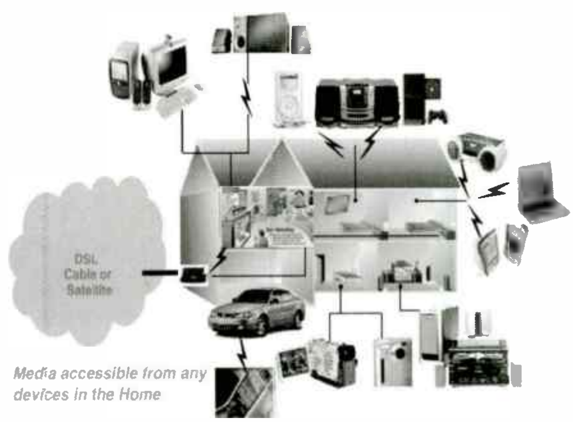


Figure 2: Consumer AV Media Devices in the Home

With AV media content as king, the consumer wants to easily move the content from device to device depending on the location where it will be used and be able to do this quickly at no additional cost, often in real-time or even faster than real-time. In this dynamic ever changing environment, the devices available to the consumer, which are many, must be able to exchange the content that is stored or viewed on them such that they can be viewed on alternative devices at different times by potentially different consumers, all within the confines of the home originally, but eventually taken outside the home.

Finally it's not just the different ways of accessing the content that is changing. Whereas the STB and DVD have similar processing and display capabilities, there is a wide disparity in capabilities between these and portable devices. This disparity widens immensely with the introduction of HD into the home. With over six times the number of pixels as SD, transporting HD content in the home becomes costly. And with the added complexity of H.264 over MPEG-2, the disparity between processing power of the STB/DVD and portable devices widens even further. To allow these devices to interoperate, a means of transforming the video and compressed formats between these different devices is required. Transcoding is the process that performs this much needed operation.

Traditionally, transcoding has referred to the coding and recoding of digital AV content from one compressed format to another in order to allow for the transmission and playback by various devices. This becomes increasingly important as HD video is more widely available and viewed by the consumer. Over time, HD will become more of a necessity instead of a luxury. This will create a dichotomy in the home where the older or less capable devices will not be able to access it, which will further increase the need for transcoding. Many devices, especially portables, will not have the processing power or display capability to process the HD content. Hence the content will have to be transcoded to a different compression format or scaled to a lower image size, bit rate, and possible frame rate in order for the portable devices to process it.

For transcoding, there are basically three main types.

- **Transcoding (traditional reference)** - involves the conversion from one compression format to another, such as when converting MPEG-2 to H.264. This method involves the most changes to the original content; codec tools, image size, frame rate, and bit rate.
- **Transrating of content** – using the same compression format, but lowering the bit rate of the original content to allow it to be transmitted, stored, or used by a less capable device.
- **Transcoding of image parameters** – using the same compression format (same profile, different level), but reducing the original image size and frame rate to allow for playback on less capable devices.

There are several reductions that result from transcoding, namely the following:

- **HDD storage:** for STB with PVR functionality, transcoding MPEG-2 content to H.264 can potentially reduce the storage capacity by 50%. For MPEG-2, good quality can be attained at compression ratios from 30:1 to 50:1 while for H.264 similar quality can be obtained between 60:1 and 100:1. Hence a 50% improvement in compression efficiency results in a 50% reduction in HDD storage space. Although the consumer will always reach capacity on any HDD, transcoding will double this capacity over MPEG-2 allowing the consumer to store and access more content on the PVR over time.
- **Bandwidth utilization:** for bandwidth limited pipes, this will allow either more channels to be transmitted or will enable transmission that would not be possible otherwise.
- **Need for storing multiple files:** AV media servers often store multiple files of the same content at different image resolutions and bit rates to allow viewers to access the content depending on the capabilities of the device and available bandwidth. An example of this is seen when playing movie trailers where the user is asked to select small, medium, or large for the desired viewing preference. Being able to transcode in real time, eliminates the need for multiple files being stored thereby reducing file storage on the server and unnecessary complexity on the part of the server by knowing apriori which file to use and transmit.
- **Need for supporting multiple formats:** over time, different compression formats continue to emerge, both standards based and proprietary. And it's a moving target. With PC applications tending to be more proprietary and broadcast applications tending to be more standards based, there is a wide range of codecs that are already in use. In time, this will continue to increase with newer standards continuing to evolve such as China's AVS and follow-on by the ITU for the next H.26x.

These reductions are further pronounced with HD where the savings from transcoding are very substantial and in some cases (i.e., bandwidth limited pipes) will enable capabilities that would otherwise be inoperable.

This paper will address how the transcoding process can be implemented into a single silicon device. We will first briefly look at the MPEG-2 and H.264 codecs and discuss similarities and differences and the associated advantages of moving to H.264. Next we will address the different transcoding methodologies and what can be done in the transcoding process to provide the best possible use of the compression tools in order to arrive at the best quality while minimizing the amount of processing cycles during this process. This will become increasingly important with HD, since the consumer wants all of this capability at the lowest possible cost and allow transparent use of the AV content whether it is being displayed on a cell phone or an HD monitor. Finally we will address how to best implement this into a consumer appliance. In particular, Texas Instrument's DM6467 HD DaVinci system-on-a-chip (SOC), can be used to perform all the necessary processing in a single device in order to enable a much needed capability for future consumer products to enjoy a long life in this dynamic environment of the home.

The MPEG-2 vs H.264 CODEC

For H.264, there are four profiles that are available. For broadcast and DSM applications, the Main and High Profile are used primarily because of their support for interlaced video, bidirectional (B) frames and more efficient VLC and transform operations. Figure 3 provides an illustration of the tools available among the four profiles of H.264. Many of these tools are common between MPEG-2 and H.264.

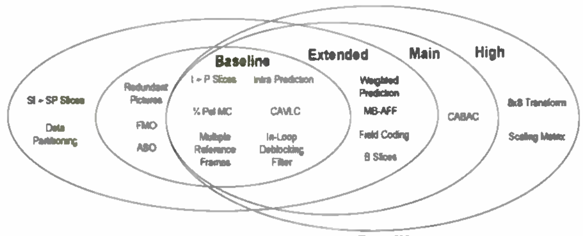


Figure 3: The H.264 Profiles

The improved compression efficiency of H.264 is a result of the introduction of newer and more advanced tools. Before addressing the transcoding operation, we will first look at where most of the compression efficiency is obtained by using these newer tools of H.264 in lieu of MPEG-2. To get the best results, it will be important that these newer tools are used wisely in order to get the best compression efficiency possible while maintaining and/or improving the resultant quality.

For comparison purposes, we will first look at some of the significant differences between an MPEG-2 and H.264 decoder. Figures 4 and 5 provide a block diagram of the MPEG-2 and H.264 decoder, respectively.

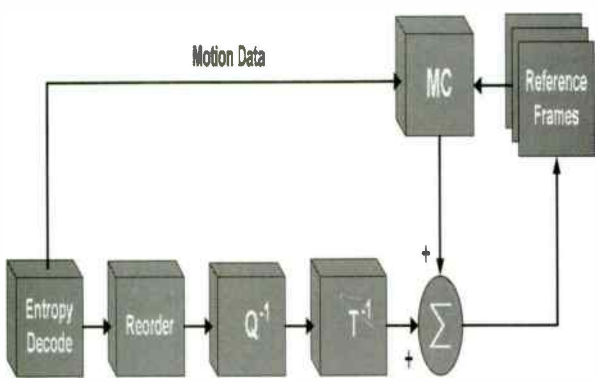


Figure 4: MPEG-2 Video Decoder Block Diagram

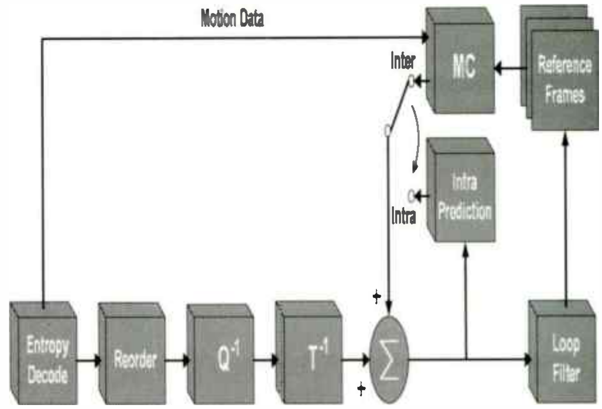


Figure 5: H.264 Video Decoder Block Diagram

By comparing these two diagrams, there are two noticeable differences. The H.264 decoder has an added Intra Prediction and Loop Filter module. In addition, the Entropy Decode module for MPEG-2 is a variable length Huffman decoder (VLD) while H.264 (Main/High Profile) uses a context-based adaptive binary arithmetic coder (CABAC). Also H.264 uses an integer based transform in lieu of the DCT used by MPEG-2, thereby preventing mismatches in the inverse transform between the encoder and decoder that occurs with MPEG-2. These differences account for noticeable improvement in bit rate reduction of H.264 over MPEG-2. As a general rule, Table 1 provides a rough estimate of the percentage improvement of using these H.264 modules in lieu of MPEG-2. This is very content dependent with the actual amount of bit rate reduction being dependent on the amount of frequency and motion content in the sequence being encoded.

Table 1: Bit Rate Reduction Percentage Improvement of H.264 over MPEG-2

Module	Percentage Improvement
Intra Prediction	10 – 20%
Inter Prediction	30 – 40%
Loop Filter	5 – 10%
CABAC	5 – 10%
8x8 Transform	5 – 10%

We see similar differences in the modules that make up an encoder. Figures 6 and 7 provide a block diagram of the MPEG-2 and H.264 encoder, respectively.

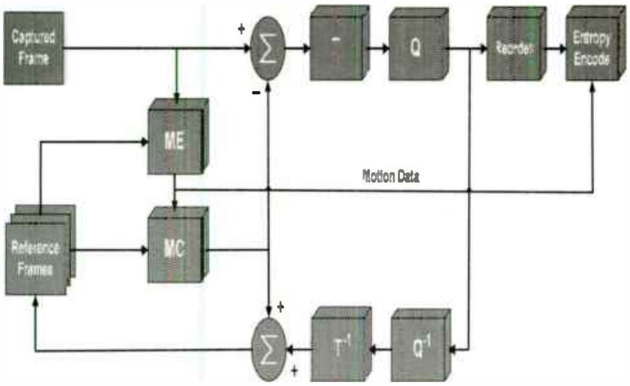


Figure 6: MPEG-2 Video Encoder Block Diagram

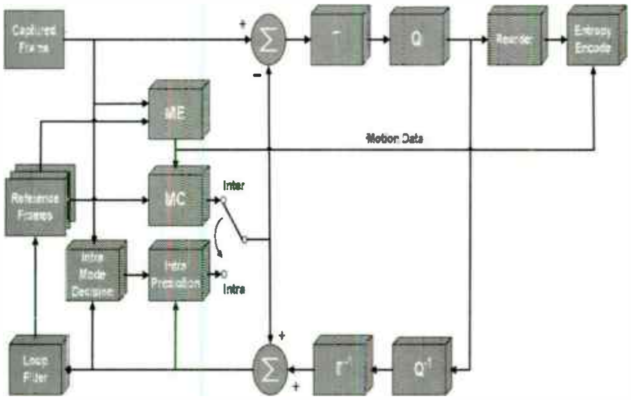


Figure 7: H.264 Video Encoder Block Diagram

Here again, we see the introduction of the Loop Filter and Intra Prediction modules with an Entropy Encode replacing the Entropy Decode module. With the motion compensation (MC) module of the H.264 encoder, there is additional bit rate reduction of about 10% to 15% by using weighted motion vectors (MV) and using a ¼ pel MC along and multiple reference frames. Overall, the major important enhancements in using H.264 over MPEG-2 are the following:

- Entropy encoding enhancements
- Small block transforms with exact matching in the forward and inverse transform
- Motion prediction enhancements
- Inserting the loop de-blocking filter in-loop

For H.264 Main and High Profiles, it is possible to attain more than a 50% improvement in compression efficiency at the expense of upwards of a 16X increase in computational processing cycles.

However putting an H.264 Main/High Profile encoder into a consumer product is impractical due to size and cost constraints. The challenge is to perform a transcode operation in a single piece of silicon that could easily go into a STB or DVD player in the home. Since the AV content to be transcoded has already been encoded, most likely with good quality by a broadcast MPEG-2 encoder, the intent is to make use of the ME and frame/macroblock (MB) mode decisions that have already been made by the broadcast encoder and pass this information off to the H.264 encoder during the transcode process. Within the transcoder, the MPEG-2 decoder needs to retain the MVs and frame/MB mode decisions and hand this data off to the H.264 encoder. By reducing the amount of decisions being made by the H.264 encoder (i.e., MVs, frame/MB mode decisions), this will significantly reduce the amount of redundant processing involved and allow this transcode operation to be implemented in a single piece of silicon whereby the MPEG-2 decoder and H.264 encoder are working together and not independently. In addition, there may be a need for video scaling and removal of MPEG-2 artifacts prior to handing the MPEG-2 decoded video frames to the H.264 encoder. With this as background, we are now ready to discuss the much needed transcode operation.

Transcoding

Among the different processes involved in enabling the interoperability amongst AV devices in the home, transcoding is by far the most complex and computational intensive. For HD content, this task is well beyond the reach of a typical host processor relegating it to a video coprocessor that is capable of handling the massive computational requirements associated with decoding and encoding an HD bit stream. But it's not just about massive computations. There is much algorithmic work involved such that decisions that are made during the transcoding process are well thought out and attention is applied to each processing step along the way. As an example, Figure 8 provides an illustration of a brute force technique where a full decode is followed by a full encode.

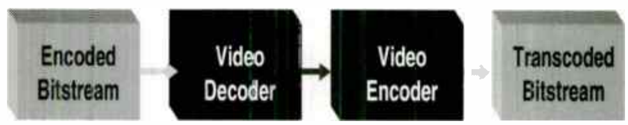


Figure 8: Brute Force Transcoder

This on the surface looks simple, yet it comes at the expense of additional computational complexity creating additional artifacts within a frame and propagating them to additional frames that follow, resulting in an overall loss of image quality. This is particularly significant with low bit rate content when the quality is degraded and significant MPEG-2 artifacts are fed forward from the MPEG-2 decoder to the H.264 encoder.

With a brute force implementation, the primary increase in computational complexity is a result of the encoder having to do a full motion search in its motion estimation process since it has no knowledge of the MVs that were used in the decoding process. One of the primary losses of image quality is a result of inaccurate frame types being used between the decoder and encoder. For example, if the encoder decides to perform an intra (I) frame encode on a frame from the decoder that was decoded as a bidirectional (B) frame, the resulting image quality will be degraded. In this case, a B frame will be used as a reference frame for past and future frames of the encoder propagating artifacts and image degradation for many frames to follow.

This is best illustrated by considering an example scenario. Consider the case of a cell phone desiring to playback content that was received from a HD broadcast encoder. An HD broadcast encoder is a very expensive machine that is capable of performing a massive amount of computations to arrive at the most optimal decision on the frame type and macroblock modes and motion vectors within each frame. In addition, for broadcast encoding, it must support rapid channel change and trick play, typically requiring I frames to be inserted at ½ second intervals. For 30 fps, a typical group of pictures (GOP) sequence involves inserting an I frame every 15 frames. Because an I frame does not use any temporal compression, they are expensive in terms of bit consumption, consuming a much larger percentage of the available bits than the non-I frames that occur within the sequence.

Within a GOP, there are I, P, and B frames. The predictive (P) frames reference a previously decoded frame while the B frames reference both previous and future decoded frames. Typically a P frame requires a much lower amount of bits to encode than an I frame while a B frame requires an even lower amount. While the B frames typically consume the least amount of bits in a sequence, they add cost and complexity to the encoding process as well. Since they reference future frames, they require frames to be decoded ahead of time introducing additional frame buffering by the encoder and decoder which also introduces additional delay. The added processing, memory, and delay as a result of using B frames provides additional compression efficiency which is desirable for broadcast applications, but is considered to be cost prohibitive for portable devices. It's not just chip cost associated with the additional processing of handling B frames, but it is also added system cost as a result of needing more memory to store these additional reference frames. A typical GOP structure for a broadcast encoder is illustrated in Figure 9.

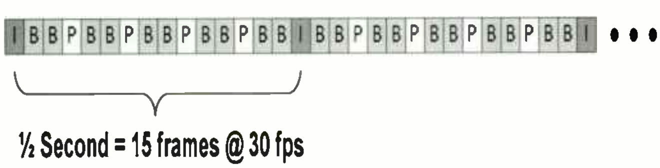


Figure 9: Example Broadcast GOP Structure

In contrast, the cell phone is a consumer device and is very sensitive to added cost from increased memory and unwanted delay. For cell phones, power consumption and memory is at a premium and they often times can only support a much lower profile and level of the codec tools than what is possible for a broadcast application. The reduction in bit rate from the added cost and delay associated with using B frames is not considered justifiable. Since the overall bit rate for portable devices is much lower that that of a broadcast encoder, the added cost of inserting I frames every ½ second is also not justifiable. Also the need for fast channel change and trick play is much lower that what is required for a broadcast STB. Hence, the separation in I frames is increased to many seconds. Often times only inserting an I frame on an as needed basis or at regular intervals of every 10 to 15 seconds, for example. An illustration of the GOP structure for a typical cell phone is provided in Figure 10.

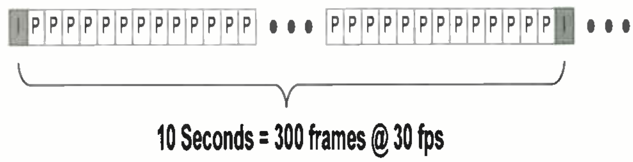


Figure 10: Example Cell Phone GOP Structure

Figure 11 provides a further illustration of this process by showing the replacing of the B frames with P frames and a second I frame with a P frame.

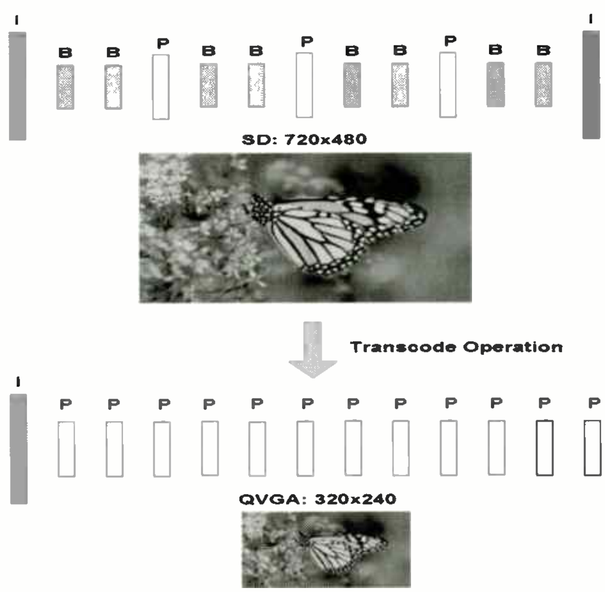


Figure 11: Example of a GOP Structure Change (SD to QVGA)

The top portion of this diagram would be representative of a GOP structure of a standard definition (SD) sequence associated with a broadcast application while the bottom portion is the resulting bit stream and GOP structure of a QVGA sequence as a result of transcoding this for a portable device. Here we can see the replacement of the B frames in the original content with P frames and the removal of the second I frame with a P frame with a corresponding reduction in image size. The reduction in image size and number of I frames will result in a reduced complexity bit stream at a substantially lower bit rate thereby enabling the cell phone to process the newly transcoded content. As a result of transcoding to a lower H.264 profile (i.e., Baseline Profile), error resiliency can be improved in the resulting bit stream by using additional H.264 tools such as FMO, ASO and redundant frames.

During the transcoding process, passing information by the decoder to the encoder as to the type of each frame (I, P, B) and the macroblock modes and motion vectors can be very beneficial. In the case where there is a reduction in frame rate such as going from 30 fps to 15 fps, knowing where the B frames are in the original source can result in substantial reduction in processing. Because B frames are not used as a reference frame for any previous or future frame, the B frames can simply be removed from the bit stream to arrive at the lower frame rate without requiring any additional processing by the encoder.

Also knowing the motion vector of each macroblock in the HD bit stream can result in a huge reduction in processing by the encode process of a transcoder. The motion search process of any encoder is the most computational intensive of all its processes. By knowing the general vicinity of the motion vector for each of the macroblocks from the original source, this substantially reduces the amount of computations required to determine the final motion vector. The motion search of the transcoder becomes more of a refinement about a known reference point, instead of a massive search in a very large image plane. The point here is that an HD broadcast encoder spends a tremendous amount of processing and memory to arrive at an optimal set of parameters for each frame and macroblock in the sequence, often times performing multiple passes of the encode process for each frame. To disregard this information is very unwise and puts a huge amount of additional processing on the transcoder that is otherwise unnecessary and will often times result in increased artifacts and overall reduced image quality. Having the decoder tightly coupled with the encoder reduces the overall complexity and results in a more optimal solution. An example diagram of such a coupled solution is provided in Figure 12.

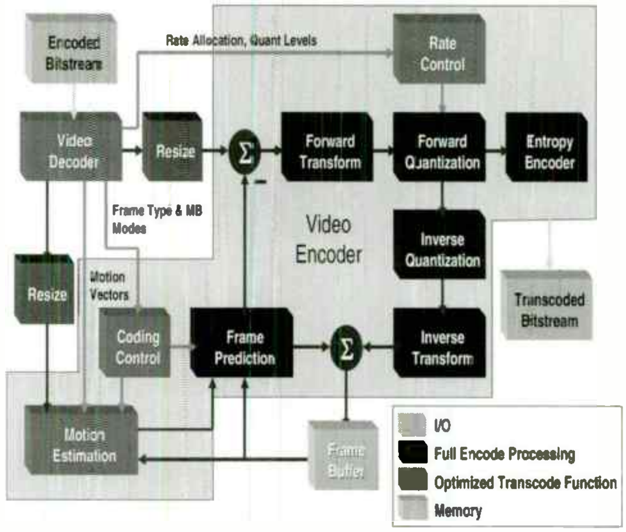


Figure 12: Transcoder Block Diagram

Here we can see the decoder providing frame type, macroblock types and modes, motion vectors, quant levels and bit rate parameters to the encoder in order to help it make better decisions on the transcoded bit stream that it creates. Although the encoder has to perform a full encode on the data, many of the modules have reduced complexity due to the apriori information provided to it by the decoder. This is especially true in the motion estimation process which is more of a refinement than a massive search. In the case where there is an image size change, image scaling can create the proper frame sizes prior to passing the frame off to the encoder and thereby decrease the memory required in the system, further reducing the overall cost of the transcoder.

Since the early days of MPEG-2, much has been spent on trying to find additional methods to improve the image quality and reduce the artifacts associated with compression. Two of the more predominant artifacts associated with using the tools of MPEG-2 are the following:

- mosquito noise - comes from using the DCT on each macroblock with high quantization on sharp edges.
- blockiness - appears throughout the image as a result of tiling the image into a mosaic of macroblocks with high quantization.

These artifacts can be reduced by the introduction of filters after the decode process. This is illustrated in Figure 13.

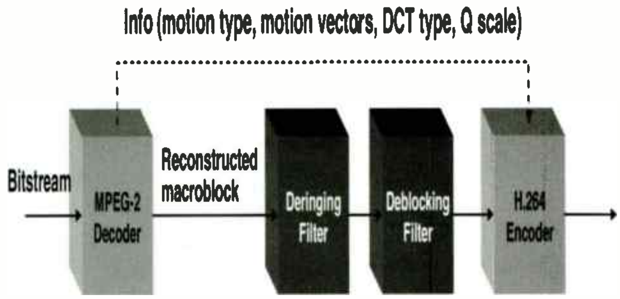


Figure 13: MPEG-2 to H.264 Post/Pre Filtering

By inserting these filters after the decode process but before the encode, especially for an MPEG-2 to H.264 transcode, the overall image quality can be improved with a corresponding reduction in bit rate due to the removal of the unwanted high frequency artifacts that are introduced from the MPEG-2 decode process. This is particularly important when there is high frequency content in the image or the bit rate is lowered due to high quantization levels, because this is when the artifacts are most noticeable and will be magnified and propagated to future frames if they are not removed or suppressed before being processed by the encoder. For example, these MPEG-2 artifacts will translate to inaccuracies of the MV data and intra prediction modes that are determined by the H.264 encoder and as a result will enhance or magnify these artifacts even further. However if they can be suppressed or removed, the end result will be a reduction in bit rate with improved image quality as a result of transcoding from MPEG-2 to H.264

Implementation

Now that the concepts involved in transcoding have been introduced, let us now turn our attention to how to best implement this in an efficient manner in a single device. As mentioned earlier, there are upwards of four different parameters that will need to be modified during a transcode operation: codec type (e.g., MPEG-2, H.264, VC-1, etc), image resolution, video frame rate, and compressed bit rate. The transcoder must be able to adjust any one or all of these parameters depending on the device requesting access to the content. In addition, the resulting compressed video data and corresponding compressed audio data will need to be put into a container format and time synchronized. And finally, the resulting system stream must be packetized and transported or stored for access by the device requesting it. To transcode the content and transport it, three processes need to be implemented:

1.) Transcode the AV content.
2.) Assemble the AV content into a container and time synchronize it.
3.) Transport and/or store the content for access by the device requesting it.

Each of these processes requires a different type and level of processing. Transcoding is by far the most computational intensive, especially for HD content. This task requires a very high end processor that is bit intensive and deterministic while the assembling and transport/storage tasks are more byte and packet oriented and because of the network environment, require a lot of non-deterministic types of event processing, such as interaction with clients and error handling.

Texas Instruments has recently developed an HD transcoder device, the DM6467 SOC, which has been architected specifically for this type of application. Figure 14 provides a block diagram of this device.

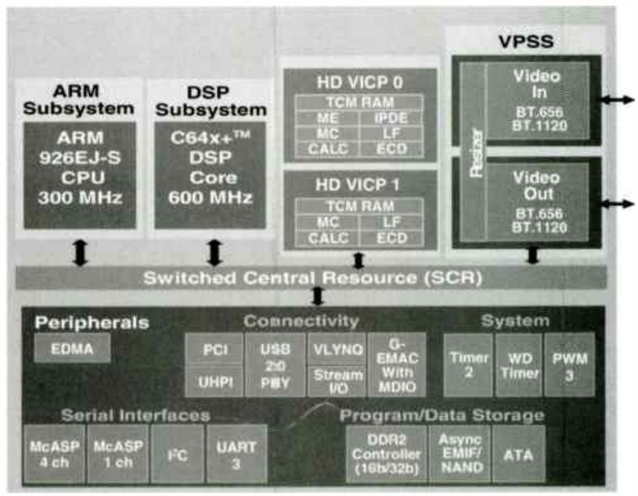


Figure 14: DaVinci HD Block Diagram

The DM6467 has two internal high end processors, an ARM926 and a 64x DSP along with two additional video image coprocessors (VICP), to perform the decode/encode processing of the transcode task. One VICP can perform a full HD decode operation and as mentioned previously hand off the resulting data to the second VICP which performs the HD encode. The 64x DSP performs the rate control, ME decisions, intra/frame/macroblock mode decisions, and error concealment. Working together, the DSP and VICPs can perform an HD transcode as well as any necessary scaling operations along with the audio transcode and provide the compressed data to the ARM926 to complete the remaining tasks.

Serving as the host processor in the system, the ARM926 can perform the assembly task of creating the AV container and providing the necessary time synchronization and packetizing this data for transporting over a network interface. In particular, it can create the UDP multicast, HTTP, or RTSP connection and transport the AV container (TS, PS or MP4) to the requesting device. Alternatively, it can store this resulting data to a hard drive or storage device for access later. Figure 15 provides an illustration of a block diagram of the DaVinci HD SOC being used in this type of application.

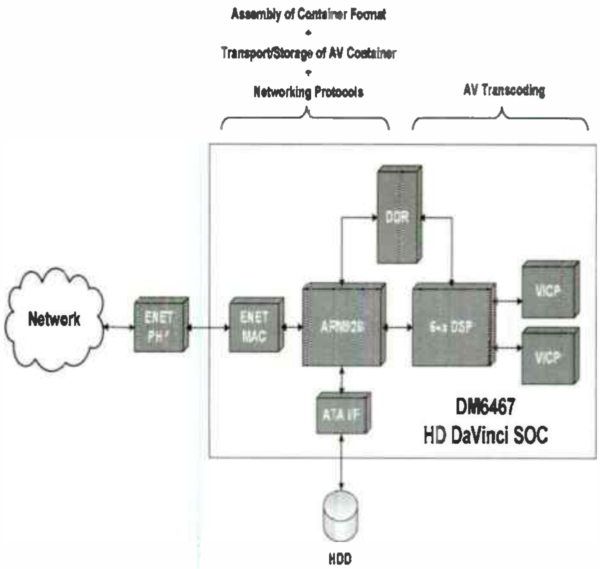


Figure 15: DaVinci HD SOC Block Diagram

Summary

As consumer devices continue to grow, along with a myriad of different applications running on them with a corresponding broad range of capabilities, the need for adaptive media devices with transcoding support will become paramount. This need further increases, especially as HD becomes more prevalent in the home. Most portable devices will not have the processing power, memory, or display capabilities to process the HD content. In order to provide HD content to these portable devices, devices that have access to the HD content, either remotely or in local storage, will need to adapt to the network protocol, container format, and compression format of the less capable portable devices. For the different network protocols and container formats, a programmable host processor can be used to adapt to the devices that wish to access the content, but it will not have the processing power to handle the much needed transcode functionality, especially with regards to converting HD content. For this, a coprocessor is needed to transcode to the many different codecs that are currently available, namely standards based MPEG-2, H.264 and VC-1, but also be able to implement proprietary codecs and future codecs that may become popular in the future such as the already popular On2 for FLASH, China's AVS and further extensions of H.26x. The DaVinci HD architecture is well suited for such a task. Figure 16 provides an example of a DVD recorder with a transcoder receiving compressed content from the STB and providing transcoded content back as well as being able to decode the video to be displayed by the STB hardware using picture in picture (PIP).

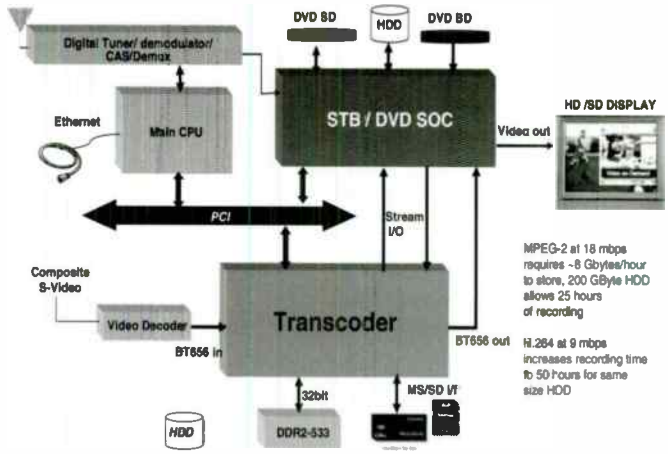


Figure 16: STB + Transcoder Block Diagram

This application primarily needs just the transcode functionality. In the end, a fixed function type of device will not work. There are just too many codecs and too many devices that need to interoperate. For codecs, container and file formats, and networking protocols, there is no one size fits all. It must be adaptive to allow interoperation such that the consumer can access any content, anywhere, at any time, and on any device, both now and in the future. For change is inevitable, and being able to adapt to this change and provide maximum utility to the consumer while doing so will determine which products succeed in this ever changing environment of the home.

References

[1] ISO/IEC 13818-1:2003. Information technology – Generic coding of moving pictures and associated audio. Part 1: Systems.

[2] ISO/IEC 13818-2:2003. Information technology – Generic coding of moving pictures and associated audio. Part 2: Video.

[3] ISO/IEC 14496-10:2003. Information technology – Coding of audio-visual objects. Part 10: Advanced Video Coding.

[4] J. Golston, A. Rao, "Video Compression: System Tradeoffs with H.264, VC-1 and Other Advanced Codecs", Embedded Systems Conference, April 2006.

[5] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, T. Wedi, "Video Coding With H.264/AVC: Tools, Performance and Complexity", IEE Circuits and Systems, Q4/2004.

[6] J. Golston, "HD Transcoding Connects Home Video Applications", www.ti.com

[7] D. Marpe, T. Wiegand, S. Gordon, "H.264/MPEG-4 AVC Fidelity Range Extensions: Tools, Profiles, Performance, and Application Areas", IEEE, 2005.

[8] M. Zhou, "Evaluation and Simplification of H.26L Baseline Coding Tools", JVT of ISO/IEC, Jan 29 – Feb 1, 2002.
[9] M. Zhou, "Benchmark Analysis of H.26L Decoder Functional Blocks", ITU Study Group 16 (VCEG), Sept 2001.
[10] A. Osamoto, "DaVinci HD Simplified Algorithm For MPEG-2 to H.264 Transcoder", TI Systems Lab, Nov 2007.
[11] V. Lappalainen, A. Hallapuro, T. Hamalainen, "Complexity of Optimized H.26L Video Decoder Implementation", IEEE Circuits and Systems, July 2003.
[12] T. Simerly, "HD Transcoding and Developing an Adaptive Media Client/Server Model", SMPTE Technical Conference, Nov 2007.

AUTHOR AND PRESENTER INDEX

Allard, Frédéric 91, 146
Anderson, Timothy B. 167, 182
Andries, Luc 330
Babbitt, Michael, ... 419, 430
Bachofen, Ralph, ... 419, 425
Ball, Talmage 369, 391
Bancroft, David 604
Band, Bob 167, 169
Beckers, Tom 330
Behr. Lawrence 455, 457
Bishop, Mark 45, 62
Bisset, John 157
Briscoe, Paul 419
Caron, Bernard 387
Casaccia, Ed 191
Castro, Michael 45, 51
Chernock, Richard 303, 311, 425
Cho, Injoon 592
Church, Steve 369, 378
Claudy, Lynn 1
Compton, Matthew 565, 567
Converse, Dave 45
Coppa, Gregory 226
Costa, Acacio Luiz 91
Cozad, Kerry 441
Czarnecki, Edward 389
Daniels, Simon 91, 111
Dawson, Benjamin 455
De Geyter, Matthias 330
Deventer, Etienne 446
Driscoll, Melinda 3, 31
Estrada, Luis 303
Etch, T. Goji 193
Evain, Jean-Pierre 191
Fannechere, Nicolas, ... 45, 79
Fanton, Myron D., P.E., ... 257, 271, 391, 405, 441, 443, 451
Felts, Boris 45, 83
Fluker, Steve 283
Fourcher, Fred 303
Francis, Scott 522
Freinwald, Clay 389
Gabriel, Anne 455
Gampell, Guy 167, 176
Gardina, Mick 191
Gauderis, Hugo 446
Gleason, Frederick, Jr. 3
Glenn, Dave 45
Godfrey, John 45
Goldfarb, Robert 191, 253
Grossniklaus, Al 159
Guel, Sid 303, 361
Hahm, Sangjin 565, 592
Hamazumi, Hiroyuki 73
Hardy, Tim 455
Hayashida, Tetsuya 193
Herlein, Greg, ... 522
Heywood, Gareth M. 513, 528
Hisatomi, Kensuke 193
Holt, Tim 391, 393
Holtz, Alex 62
Honda, Rie 557
Howard, John 391
Iijima, Yuichi 557
Illgner, Klaus 387
Isreal, David 157
Ito, Yasuhiro 73
Iwadate, Yuichi 193
Jablonski, Charles 513
Janitschek, Andrew P. 257, 274
Jensen, Laura 3, 31
Jones, Graham 565
Joseph, Wout 446
Katayama, Miwa 193
Kean, John, ... 91, 137, 283, 299, 391, 411
Kelley, Don 157
Kelly, Charles W. 257, 259
Kereakos, Nick 3, 31
Kim, Byungsun 592
Kim, Gyu-Young 129
Kim, Junius 283, 285
Kim, Man-Sik 129
Kim, MunChurl 592
Kim, Sang-Hun, ... 91, 129
Kisten, Shadrach 253
Kline, Gary 157, 455
Knee, Mike 303, 320
Korte, Olaf 257, 266
Kube, Wayne 191
Kyte, Dan 522
Latchman, Ryan 513, 528
Lattie, Tom 303, 344
LeBow, Jerry 389
Lemer, Robert 303
Lennon, Chris 303, 315
Lhermitte, Richard 45, 55

Libin, Louis 441
Liebisch, Gary 455, 469
Locke, Ian 83
Lübcke, Gerold 522
Luplow, Wayne 45
Lyons, Christopher 159, 161
Macher, Jean 45
Maejima, Hironori 557
Mansergh, Daniel 3, 15
Marie, Pascal 303, 348
Martens, Luc 446
Maruyama, Hirotaka 193
Maxson, David 391, 397
Meintel, William 159
Merrill, John 419
Miller, Jerry 67
Mitsuhashi, Seiji 513, 557
Morello, Alberto 387
Narikiyo, Yoshikazu 73
Naylor, John 565, 604
Nobbs, Andy 303
Ogonowski, Greg 91, 99
Okano, Masahiro 45, 73
Otake, Hisashi 557
Overmeire, Luk 303, 330
Pallett, John 45, 47
Park, Changseob 592
Park, Keunsoo 592
Parker, Darryl 389
Paulausky, Larry 283, 293
Paulsen, Karl 191, 199
Pelkey, Gary 91, 93
Peterle, Tony 91, 146
Philippe, Vincent, ... 146
Plets, David 441, 446
Poarch, Derek 389
Porter, Marshall 45, 67
Richards-Smith, Van 455, 502
Richer, Mark 45
Rietschel, Johannes 283, 369
Robidoux, Pierre 369, 371
Roth, Todd 565, 573
Rucktenwald, Tom 3, 5
Sader, Elie 565, 597
Santos, Ernesto 303, 337
Scherer, Chriss 157
Schmid, Philipp 91, 118
Schwab, Dan 299
Schwartz, Robert 1
Seidel, Robert P. 226
Sheffield, Ellyn G. 283, 299
Shibuya, Kazuhiko 73
Shogen, Kazuyoshi 515
Shulins, Paul 3, 167, 455, 464
Shumate, Sid 91, 105
Sidas, Walter 191, 226
Simerly, Tim 565, 612
Smith, Milford 91
Sobue, Shinichi 557
Sockett, Pete 45
Stockhammer, Thomas 67
Suh, Young-Woo 441
Sujikai, Hisashi 513, 515
Suzuki, Yoichi 515
Swanton, Ken 191, 217
Tachino, Junichi 557
Taguchi, Makoto 73
Takada, Masayuki 73
Tanaka, Shoji 515
Tanghe, Emmeric 446
Tanimoto, Kazuo 557
Tawil, Victor 441
Thaler, Larry 303, 356
Thomas, Barry 257
Thomas, Peter 303, 305
Tinkle, Paul 157
Tomiyama, Kimihiro 191, 193
Tonioka, K. 387
Tsuchida, Kenichi 73
Turner, John 303
Van Assche, Steven 330
Verloock, Leen 446
Virag, David 45
Vlaicu, Roland 513, 546
Walden, E. Glynn 455, 474
Wallace, Dennis 159
Warren, Diane 1
Weiss, S. Merrill 441
Whitbread, Colin 387
Wiler, Eric 91, 93
Wilhite, Joel 191, 206
Wilson, Dave 3, 21
Wood, David 387
Yamauchi, Masahito 557
Yamazaki, Junichi 557
Yoshida, Tetsuo 193
Zink, Alexander 283
Zwiebel, Rob, ... 419, 421
Zwing, Rainer 513, 522

COMPANY INDEX

ABC/Disney Television Group....45
Advanced Television Systems Committee....45
American Public Media | Minnesota Public Radio....3, 31
APT....91, 111, 167, 176
Association for Maximum Service Television....441
Audemat....91, 146
Audio Processing Technology (APT)....91, 111, 167, 176
AZCAR Technologies....191, 199
Barix AG....283, 369
BBC....387
Belo Corporation....191
Bird Technologies Group....391, 393
Bitcentral....303
Blue Order Solutions AG....303, 305
Bonneville International....369, 391
British Broadcasting Corporation (BBC)....387
Broadcast Automation Consulting....303, 361
Broadcast Electronics....157
Broadcast Pix....191, 217
Broadcast Signal Lab, LLP....391, 397
BroadComm Inc.....441
Canadian Broadcasting Corporation/ Société Radio-Canada....369, 371
CBS....191, 226
CBS Radio....455, 474
CBS5 KPHO-TV....419
Communications Research Centre....387
Consumer Electronics Association....3, 21
Consumer Electronics Retailers Coalition....1
Cox Radio/Orlando....283
Cumulus Media....157, 455
Current....455
Dielectric Communications....441
Digital Fountain....45, 67
Dolby Laboratories....419, 430, 513, 546
Electronics Research, Inc.....257, 271, 391, 405, 441, 443, 451
ENENSYS Technologies....45, 55, 79
Entercom....389
Envivio, Inc.....45, 83
European Broadcasting Union....191, 387
Federal Communications Commission....389
Flemish Radio and Television Network....446
Fraunhofer IIS....257, 266, 283
Gefen....303
Gennum Corporation, Connectivity Technologies....513, 528
Ghent University, Department of Information Technology....441, 446
Givens & Bell, Inc.....91, 105
Greater Media....3, 91, 157, 167, 283, 293, 455, 464
Harmonic Inc....191, 206, 303, 344, 565, 597
Harris Broadcast Communications....419
Harris Corporation....167, 182, 283, 285, 303, 315, 419, 421
Hatfield & Dawson....455
HD Digital Radio Alliance....1
Hitachi Kokusai Electric Co. Ltd....193
IBM....303
ICU....592
iDirect Technologies....191
Ion Media TV....45
IRT....387
JAXA....557
Jones Radio Networks....91, 93
Kinki University....193
Kochi University....557
Korean Broadcasting System....91, 129, 441, 565, 592
KQED Public Radio....3, 15
LBA Group, Inc.....455, 457
Lincoln Financial Media....257
Lyncole XIT Grounding....391
Meintel, Sgrignoli & Wallace....159
Meisei Electric Co. Ltd.;....557
Merrill Weiss Group LLC....441
Mix TV Network....91
MOG Solutions....303, 337
NAB....1, 565
National Geographic Television....45, 51
Nautel....91, 118, 257, 259, 455, 469
NBC-Universal....191, 253, 303, 356
NDS....3, 5
NEXIO Server Systems, Broadcast Communications Division, Harris Corporation....565, 573
NHK (Japan Broadcasting Corporation)....387 513, 557
NHK Science & Technical Research Laboratories....45, 73, 191, 193, 513, 515
NPR....3, 31

NAB Broadcast Engineering

CONFERENCE PROCEEDINGS 2008

NATIONAL ASSOCIATION OF BROADCASTERS

1771 N Street NW   Washington DC   20036 2800
Phone 202 429 5300

NAB 2008 Proceedings / Book and CD $115.00
3108NB9780893243944