THE COMPLETE
TECHNICAL PAPER PROCEEDINGS
FROM:

NCTA
Technical
Papers

750MHz Power Doubler and Push-Pull CATV Hybrid Modules Using Gallium Arsenide

D. McNamara*, Y. Fukasawa**, Y. Wakabayashi**, Y. Shirakawa**, Y. Kakuta**

*California Eastern Laboratories
Product Development Engineering
4590 Patrick Henry Drive
Santa Clara, CA 95054

**NEC Corporation
Compound Semiconductor Device Division
1753 Shimonumabe, Nakahara-ku
Kawasaki, Kanagawa 211 Japan

Abstract

For the growing channel capacity in the CATV market, a new breed of Push-Pull (Pre Amp) and Power Doubler (Post Amp) amplifiers have been developed that operate to 750MHz. These units incorporate highly reliable GaAs MMICs, surface mount transformers and advanced manufacturing technologies which have resulted in significant improvements in amplifier performance.

INTRODUCTION

In order to accommodate the increasing worldwide demand for broadband services over the television medium, CATV operators are increasing the amount of channels that are being transmitted to their customers. Distribution amplifiers are required to operate to 750MHz, 860MHz and even to 1GHz. To address this increasing demand for greater bandwidth in the hybrid amplifiers, GaAs technology has been selected for its excellent frequency and noise performance. This paper describes newly developed hybrid amplifiers that use GaAs technology for the active devices and all surface mount components in the substrate assembly. The results are high performance and highly repeatable Power Doubler and Push-Pull amplifiers that operate to 750MHz.

SURFACE MOUNT DESIGN

In large volume production, there is significant pressure to minimize the cost of assembly while providing repeatable performance. The primary method of achieving this goal in this development was to automate the entire assembly process and reduce the component count. This requires that all components be packaged for surface mount attachment to the substrate. Based on the components in the existing, traditional hybrids, this necessitated that the active devices and the transformers be re-designed to be compatible with existing automated assembly equipment.

Transformers

The development of the surface mountable transformer presented a significant technical challenge. Not only did the new device need to be in a surface mount style package, it was not to require any tuning after assembly. Many different approaches were tried: thin film, flexible substrates and alumina substrates to name a few. None of these designs had adequate electrical performance. The final design was an auto-transformer configuration using a double ferrite core. This configuration offered a small size, 6.2mmx6.2mm by 3.9mm high, that can be mounted along with other surface mount components. An outline drawing of the transformer is shown in Figure 1.

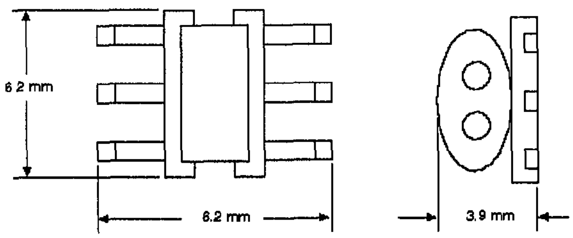


Figure 1. Surface Mount Transformer

Several transformers have been measured in a 50Ω environment for basic characteristics. The insertion loss is less than 2.5dB from 10MHz to 650MHz and is 3dB at 750MHz. The return loss is >19dB across the band. Also, the worst case phase balance is ±3° for the input transformer and ±6° for the output transformer.

GaAs MMIC

The other major design effort focused on making the active devices compatible with surface mounting equipment. This involved the design of both a GaAs MMIC and a carrier for the new IC. The most critical issues for this package were heat dissipation, auto insertion and low cost. A plastic package was not feasible due to the heat dissipation of the IC and a ceramic type package was too expensive. A simple leadframe with potting material placed over the active device was the optimum solution. A similar package is used in the manufacture of cellular phone hybrid amplifiers. The leadframe has 22 pins and is made of a nickel-iron, Ni-Fe, alloy with a copper heatsink that attaches directly to the main heatsink of the amplifier. The carrier is shown in cross-section in Figure 2.

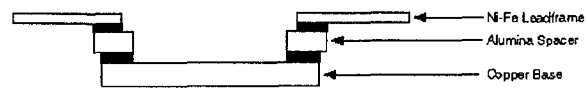


Figure 2. Carrier for GaAs MMIC

Historically, multiple transistors have been individually attached and then wire bonded to the substrate. These new amplifier hybrids only have one active device each which incorporate all the transistors. The MMIC is fabricated using a standard NEC GaAs process with a gate length of 1μm, breakdown voltage (BVgd) of 18V and a maximum current density of 310mA/mm. The CATV MMIC incorporates 6 devices in a push-pull configuration along with biasing resistors. Two of these resistors are used to set the bias of the output stages and dissipate large amounts of DC power. It is advantageous to be able to include these resistors on the MMIC, as it eliminates the need for having high power resistors on the substrate which are large and expensive. The driver stages also have source resistors on the GaAs MMIC. In addition, there are some resistor dividers to set the bias on the gates of the MESFETs (not shown). A simplified schematic for the MMIC is shown in Figure 3.

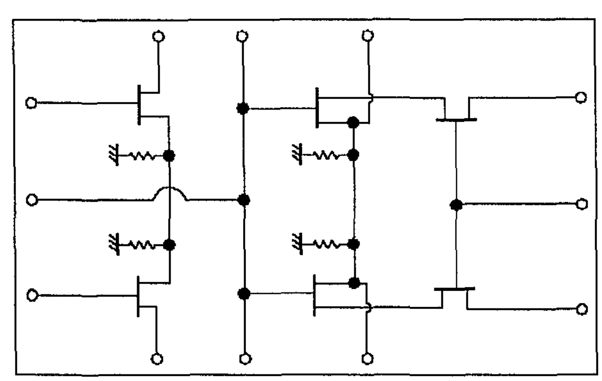


Figure 3. Schematic of CATV MMIC

Assembly

With these changes to the transformers and the active device, the substrate assembly has been fully automated. In fact, these CATV amplifiers are made on the same assembly lines as power amplifiers used in cellular phones. A photograph of a typical amplifier is shown in Figure 4.

The MMIC can be seen in the middle of the substrate, with a transformer at each end. The remaining components are used to provide feedback and biasing for the FETs. The developments detailed thus far have been applied to both the Power Doubler and the Push-Pull amplifiers. In fact, the photograph shown in Figure 4 could easily be of either the Push-Pull or the Power Doubler amplifier because the primary difference is only in the GaAs IC used.

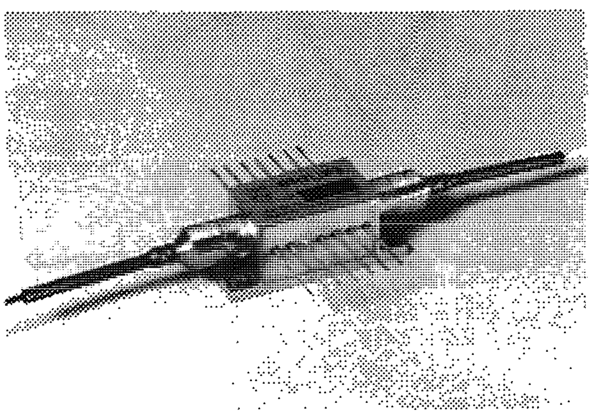


Figure 4. Photograph of substrate

ELECTRICAL PERFORMANCE

With many companies in the process of upgrading the CATV systems to 750MHz, due to the increasing demand for services and competition from DBS, it seemed logical to utilize a technology that is well suited for higher frequency operation. GaAs offers some basic advantages to silicon devices. Primarily, the higher electron mobility of the GaAs devices allows for better frequency performance, especially for noise figure and gain. While the bandwidth of a Si bipolar transistors is being stretched to cover the 750MHz band, the upper frequency range of a GaAs MMIC is limited primarily by the characteristics of the transformers. With the expected expansion of the bandwidth to 860MHz and even to 1000MHz, GaAs seems to be the logical step to service the anticipated capabilities for the next generation of plant expansion. Each of the major electrical parameters are described in the following section.

Gain

One advantage of using a GaAs device is that there is plenty of gain available at 750MHz. The power doubler and the push-pull amplifier have high typical gain levels of about 22dB. The characteristic shape of the gain curve is the same for both units. A plot of the gain response for the Power Doubler amplifier is shown in Figure 5.

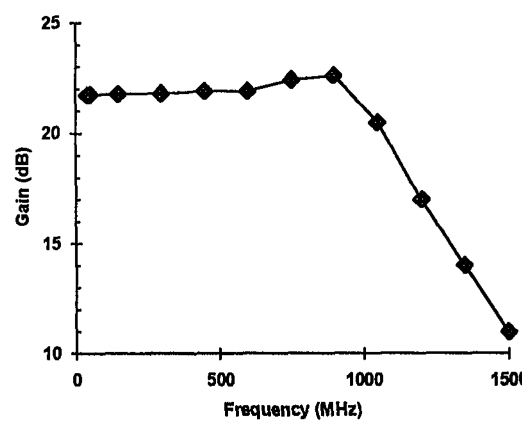


Figure 5. Gain of Power Doubler Amplifier

Noise Figure

Another benefit of the GaAs devices is the low noise figure that is achievable over a broad bandwidth. A characteristic of the Si CATV hybrid amplifiers is that the noise figure increases with frequency. This is due to the fact that the minimum noise figure for a Si bipolar transistor is quadratically related to frequency. The minimum noise figure for a GaAs MESFET is linearly related to frequency. Consequently, it is not surprising to find that the GaAs device has superior noise performance at the higher frequencies[1]. Both the Push-Pull and Power Doubler amplifiers have maximum noise figures of 5dB over the 50-750MHz frequency range. A plot of this data is shown in Figure 6.

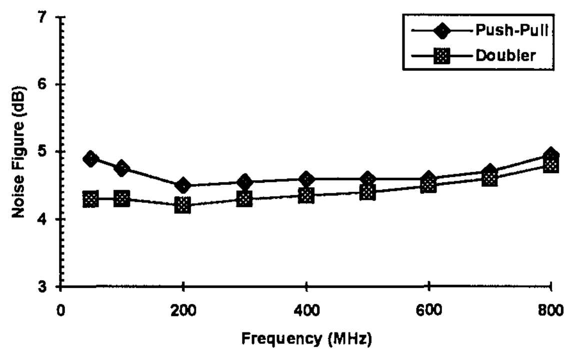


Figure 6. Noise Figure of Push-Pull and Power Doubler Amplifiers

Return Loss

Due to the multiple cascades in amplification along any given cable line, low return loss is critical to minimize reflected power. The transformers contribute significantly to the return loss, with smaller contributions coming from the component values and the input impedance of the active devices. The input and output return losses for the Push-Pull and Power Doubler amplifiers are shown in Figures 7 and 8 respectively. This is an intermediate point in the transformer design. Further development is in process to improve the return loss performance to be a minimum of -18dB across the entire band.

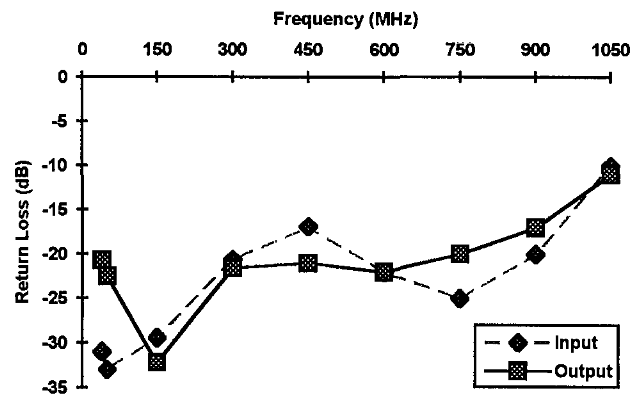


Figure 7. Return Loss of Push-Pull Amplifier

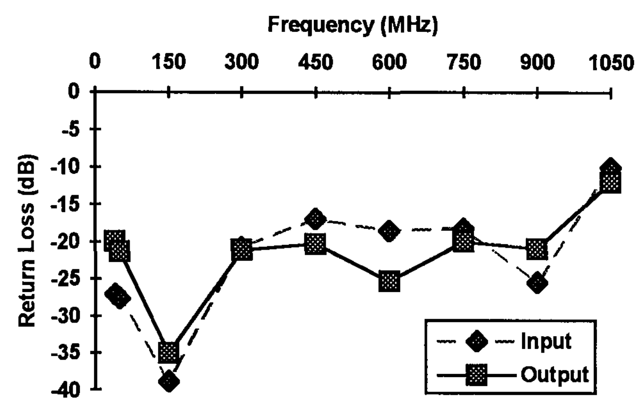


Figure 8. Return Loss of Power Doubler Amplifier

Distortion

This last section of the electrical performance is perhaps the most critical, distortion. It is the primary specification upon which the system performance is judged. The distortion performance of both the Power Doubler and the Push-Pull are extremely good over the entire band. A summary of measured CTB, CSO and XMOD are shown in Table 1.

In particular, the CSO performance stands out as being exceptional. Conventionally, the primary method of canceling second order products is to use a push-pull circuit topology. This design is no different. However, since all the active devices are included on a single IC, not only are the phase and amplitude characteristics of the amplifiers very similar and repeatable, but the interconnect parasitics are minimized. This results in a greatly improved cancellation of the second order products generated by the amplifier.

Table 1. Distortion Performance of Push-Pull and Power Doubler Amplifiers

Parameter	Description	Condition	Push-Pull	Power Doubler	Units
CTB	Composite Triple Beat	Vo=44dBmV 110 channels	-58.1	-63.4	dB
CSO	Composite 2nd Order	Vo=44dBmV 110 channels	-65.9	-70.0	dB
XMOD	Cross Modulation	Vo=44dBmV 110 channels	-61.9	-65.1	dB

[data is at 25°C, 50-750MHz]

Table 2. Electrical Parameters of Push-Pull and Power Doubler Amplifiers

Parameter	Description	Condition	Push-Pull	Power Doubler	Units
SL	Gain Slope		0.6	0.9	dB
FL	Gain Flatness		±0.15	±0.1	dB
$\mathbf{I_{tot}}$	Total Current	Vdd=24V	218	344	mA

[data is at 25°C, 50-750MHz]

Power Consumption

There are a few other parameters that are commonly reported to co⌃nletely characterize amplifiers of this type. These are summarized in the Table 2. Note that there has been an improvement in the power consumption of the Power Doubler unit. Using a novel biasing scheme, the current required is only 344mA typical. This represents approximately a 20% reduction in power consumption of the amplifier over existing units.

PERFORMANCE: GaAs vs Si

Many of the advantages of choosing a GaAs device have been explained previously. To put the reported performance into a familiar context, Tables 3 and 4 compare the performance of the hybrid amplifiers using a GaAs MMIC and hybrid amplifiers using Si bipolar transistors. It should be noted that the specifications reported for the Si devices are known to be conservative. Actual devices could have performance of 3-5dB below the stated value, depending on the vendor.

Regardless, the distortion levels are still improved by using the GaAs MMIC. Of note, for the Power Doubler amplifier, the distortion of the GaAs device is better while providing higher gain than the Si device.

Table 3. Performance Comparison Between GaAs and Si Power Doubler Amplifiers

Parameter	Description	Condition	GaAs	Si	Units
Ga	Gain	f=50MHz	21.9	20	dB
		f=750MHz	22.7	21	dB
CTB	Composite Triple Beat	Vo=44dBmV 110 channels	-63	-57	dB
CSO	Composite 2nd Order	Vo=44dBmV 110 channels	-70	-56	dB
XMOD	Cross Modulation	Vo=44dBmV 110 channels	-65	-61	dB
NF	Noise Figure	f=50MHz	4.7	5.5	dB
		f=750MHz	4.9	9	dB
I_{tot}	Total Current	Vdd=24V	340	415	mA

[data is at 25°C, 50-750MHz unless otherwise noted]

Table 4. Performance Comparison Between GaAs and Si Push-Pull Amplifiers

Parameter	Description	Condition	GaAs	Si	Units
Ga	Gain	f=50MHz	21.7	21.0	dB
		f=750MHz	22.2	21.5	dB
CTB	Composite Triple Beat	Vo=44dBmV 110 channels	-58	-51	dB
CSO	Composite 2nd Order	Vo=44dBmV 110 channels	-66	-50	dB
XMOD	Cross Modulation	Vo=44dBmV 110 channels	-62	-51	dB
NF	Noise Figure	f=50MHz	4.7	6	dB
		f=750MHz	5.0	8	dB
I_{tot}	Total Current	Vdd=24V	218	225	mA

[data is at 25°C, 50-750MHz unless otherwise noted]

REPEATABILITY

Along with the demonstrated improvement in the assembly flow and in the electrical performance, the integration of the transistors on one IC and the use of surface mount transformers are also beneficial in increasing the repeatability in the performance of the amplifiers. Other designs that use discrete transistors and wire wrapped toroidal cores require tuning to compensate for variations in each of the individual components. An advantage to using a MMIC is the uniformity in performance of individual elements on the IC. Consequently, each of the transistors on the CATV MMIC will have the same electrical characteristics. Also, the transformers are self-contained, pre-screened units that do not have any tuning sections. The result are amplifiers that emerge from the assembly line with similar performance, thereby eliminating the need to tune each unit. Quality performance is ensured by constant lot testing of a percentage of the production units based on the distribution of performance measured. This reduces the overall time needed to test the amplifiers, and it also reduces the number of test stations. Capital equipment for CATV testing is expensive and these savings contribute to the reduction of the overall cost of the amplifier. Using these assembly and testing techniques on the pre-production units has yielded very encouraging results.

CONCLUSION

Two new types of amplifiers have been developed that can be used in CATV systems. The incorporation of GaAs technology for the active devices results in high gain, low noise figure and excellent linearity. Manufacturing technology has also been enhanced by reducing the component count and designing all the components to be surface mountable. A summary of the measured performance of the Push-Pull and the Power Doubler amplifiers is shown below in Table 5.

As of the writing of this paper, the prototype stage has been completed for this development and pre-production evaluation is in process. More work is being done on the MMIC to further improve the distortion levels, especially XMOD and CTB. Also, the transformers are being optimized in order to improve the input and output return losses. More details will be available during the presentation of the paper.

Table 5. Summary of Performance

			Push-Pull	**Power Doubler**	
Parameter	**Description**	**Condition**	**Typical**	**Typical**	**Units**
Ga	Gain	f=50MHz	21.7	21.9	dB
		f=750MHz	22.2	22.7	dB
SL	Gain Slope		0.6	0.9	dB
FL	Gain Flatness		±0.15	±0.1	dB
S11	Input Return Loss	f=50-750MHz	≥16.5	≥16.9	dB
S22	Output Return Loss	f=50-750MHz	≥18.3	≥17.2	dB
CTB	Composite Triple Beat	Vo=44dBmV 110 channels	≤-58.1	≤-63.4	dB
CSO	Composite 2nd Order	Vo=44dBmV 110 channels	≤-65.9	≤-70.0	dB
XMOD	Cross Modulation	Vo=44dBmV 110 channels	≤-61.9	≤-65.1	dB
NF	Noise Figure	f=50-750MHz	4.85	4.81	dB
I_{tot}	Total Current	Vdd=24V	218	344	mA

ACKNOWLEDGMENTS

I would like to recognize and thank the managements of CEL and NEC/CSD for their support in the preparation of this paper. Also, thanks to the NEC engineers in Kawasaki and Kansai who contributed their time to assist in the characterization of the amplifiers.

REFERENCE

1. Bahl, Inder and Prakash Bhartia, *Microwave Solid State Circuit Design*, 1st Ed. John Wiley & Sons, Inc., 1988.

A Broadband Interactive Cable Gateway

Michael Adams
Sr. Project Engineer, Advanced Engineering
Time Warner Cable

Abstract

MPEG-2 and ATM technologies each have their own set of advantages and disadvantages in a Full Service Network. We have proposed the design for a Broadband Interactive Cable Gateway (ICG) that bridges the gap between these technologies and allows each to be deployed where it is best-suited in a digital cable system.

The ICG allows initial deployment of broadcast digital television services to be rapidly and cost-effectively upgraded to on-demand services with no impact to existing subscribers.

This paper will discuss the ICG in detail including protocol conversion, signaling, and management functions and provide a set of requirements for prospective ICG vendors. The paper will also provide an overview of the hybrid ATM/MPEG-2 network architecture and explain why it is optimal for the deployment of a mix of broadcast and on-demand digital services.

INTRODUCTION

The Orlando Full Service Network (FSN) was constructed with available technology a little over a year ago. It is an extremely sophisticated system, capable of satisfying Time Warner's stated goals; to understand the technical, operational, marketing and business issues surrounding interactive television systems.

As we move towards large-scale deployment of Full Service Networks, we need to re-examine our options in every aspect of the system in light of technical developments and our implementation experience in Orlando. One important aspect is the digital transport mechanism from the headend to the set-top [1].

Orlando uses Asynchronous Transfer Mode (ATM) all the way to the home. This is a logical decision in a purely interactive system because ATM is the switching technology used to interconnect media servers with the network [2], [3]. However, with the advent of digital satellite broadcast, another technology, MPEG-2 Transport, has been developed [4]. (MPEG-2 Transport is well suited to the delivery of compressed television channels in broadcast networks.) At the same time, mechanisms for mapping switched, interactive channels into MPEG-2 Transport have also been developed and are now close to standardization. Thus MPEG-2 Transport has emerged as the common transport mechanism that can be shared in an integrated broadcast and interactive network.

An Interactive Cable Gateway (ICG) is required to connect ATM-based media servers to MPEG-2 Transport based networks. The ICG allows us to gracefully extend today's digital broadcast architecture into an on-demand architecture that combines ATM and MPEG-2 Transport technologies. MPEG-2 Transport provides a uniform, efficient distribution mechanism to the set-top while ATM provides the necessary switching and load balancing functions required by on-demand services. This paper will discuss the protocol conversion, conditional access, and management requirements of the ICG.

INTEGRATING BROADCAST, INTERACTIVE AND ON-DEMAND SERVICES

The digital network architecture must integrate broadcast, interactive and on-demand services. First we will define what we mean by these terms:

Broadcast

Digital broadcast services are a direct evolution of analog broadcast. The transmission medium is digital and content is compressed to make more efficient use of spectrum, but otherwise it is the same medium.

Examples of digital broadcast services are higher-quality versions of their analog counterparts; for example, HBO provided by a direct-broadcast, digital satellite service.

Interactive

Interactive services require a two-way, real-time connection from the set-top to the headend. The service can become much more responsive, allowing the subscriber to interact with an application program.

Examples of interactive services are the World-Wide-Web, multi-player games, and home-shopping services.

On-Demand Interactive

On-demand services extend the level of service again; the subscriber is given interactive random-access to the entertainment medium itself. The Orlando FSN is testing the marketplace for on-demand services.

Examples of on-demand services are movies-on-demand, news-on-demand, and sports-on-demand.

We believe digital broadcast will be the first of these services because broadcast is a proven paradigm and does not require a change in the subscriber viewing habits.

Interactive services will follow rapidly on the heels of digital broadcast deployment; because the incremental cost of an interactive set-top is relatively small, it makes good business-sense to deploy interactive-capable set-tops.

On-demand services require a media server to supply an on-demand stream to the interactive set-top. The cost of media servers continues to fall rapidly and on-demand services will soon become cost-effective.

The ICG transforms the on-demand stream from a media server so that it is identical to the digital broadcast format; this effectively extends the life of the interactive set-top by allowing it to receive on-demand programming.

DIGITAL COMMUNICATION CHANNELS

Time Warner Cable's digital network architecture defines three digital communication channels in addition to the conventional analog broadcast channel. These are:

- the Forward Application Transport (FAT) Channel
- the Forward Data Channel (FDC)
- the Reverse Data Channel (RDC)

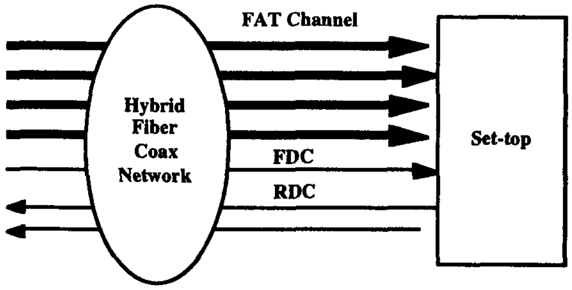


Figure 1 Digital Communication Channels. The arrow thickness is in proportion to the channel bandwidth.

Figure 1 illustrates the three types of channel.

- Forward Application Transport (FAT) channel. The set-top terminal can select any FAT channel by tuning to it.
- Forward Data Channel (FDC). The set-top terminal can always receive the FDC, even while tuned to analog services.
- Reverse Data Channels (RDC). The set-top terminal can only transmit in one RDC. However, more than one RDC may be defined per node for capacity reasons.

All channels are shared by a number of set-top terminals. A FAT channel can carry broadcast digital services in which case it is shared by all set-top terminals, or a FAT channel can carry on-demand services in which case it is shared by relatively few set-top terminals.

Digital Broadcast Network Architecture

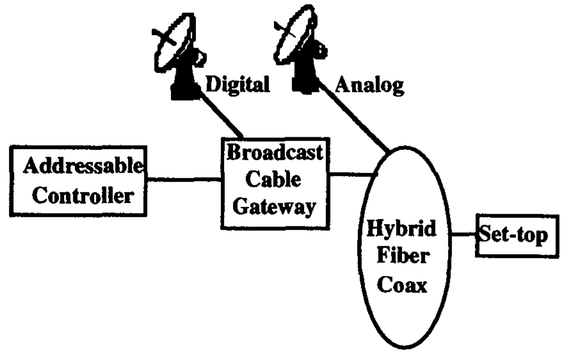


Figure 2 Digital Broadcast Network Architecture

The digital broadcast network architecture is illustrated in Figure 2. The digital services are received from satellite in MPEG-2 transport stream format. The satellite feed is sent to a Broadcast Cable Gateway (BCG) which transforms the signal for distribution over the Hybrid Fiber Coax (HFC) network:

1. The satellite signal is demodulated to recover the MPEG-2 transport stream.
2. The MPEG-2 transport stream payload is decrypted. The MPEG-2 transport stream is demultiplexed into separate program streams.
3. A new MPEG-2 transport stream is created from selected program streams.
4. The MPEG-2 transport stream payload is encrypted to the cable operator's requirements.
5. The MPEG-2 transport stream is modulated into a 6 MHz FAT channel.
6. The FAT channel is combined with other FAT channels, using FDM, and broadcast over the HFC network to all subscribers.

Interactive Network Architecture

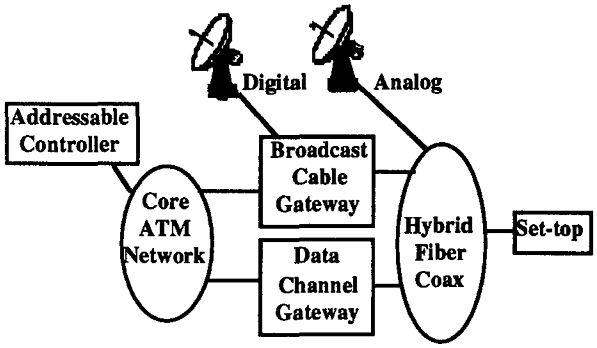


Figure 3 Interactive Digital Architecture

The digital broadcast architecture described so far is merely a one-way, broadcast system. To support interactive services, a two-way, real-time data communications infrastructure is required. Figure 3 illustrates the addition of a Data Channel Gateway (DCG) to support two-way data communications. The DCG supports the Forward Data Channel (FDC) and the Reverse Data Channel (RDC). These channels provide a two-way, Internet Protocol (IP) datagram service between the headend components and the set-top terminal. IP is chosen because it is an open, industry-standard, protocol suite that can support interactive services, as well as management, signaling and application download.

On-Demand, Interactive Digital Architecture

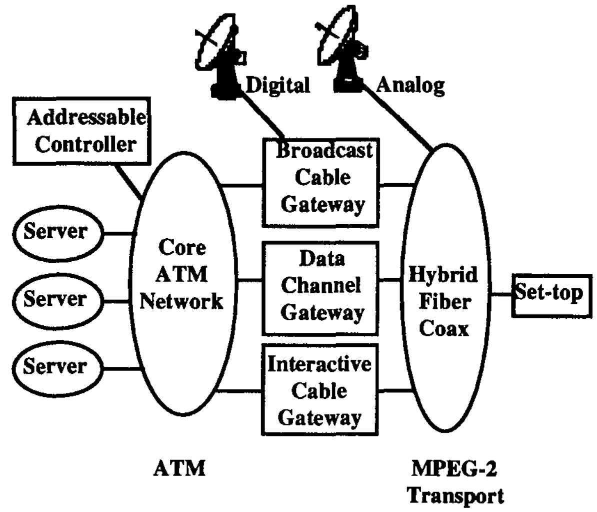


Figure 4 Reference On-Demand, Interactive Digital Architecture

Figure 4 shows the addition of on-demand services. Servers are connected to a core Asynchronous Transfer Mode (ATM) network. An Interactive Cable Gateway (ICG) translates from ATM into MPEG-2 transport. An ATM network is specified for the following reasons:

- Standard physical interfaces are available at very high rates for ATM; 155 or 622 Mbps SONET interfaces can be used to aggregate a number of digital streams into a single physical channel. SONET is an international standard and can be used for local or wide area connections [5].
- ATM can switch an on-demand digital stream from any server to any ICG. This allows new servers to be added incrementally, and service can be offered immediately to all subscribers.
- ATM allows dynamic allocation of bandwidth where it is required (bandwidth-on-demand). For example, a client application that requires 3 Mbps to play a movie can request a 3 Mbps virtual channel and

subsequently deallocate it when it is no longer required.

- ATM is a Wide Area Network (WAN) protocol. This allows the location of servers in the headend or at a distant facility. ATM supports wide-area interconnect over SONET facilities.

- The ATM network also supports data communications between all of the headend components. This avoids the expense of a separate overlay network.

The server encapsulates an MPEG-2 program stream into an ATM Virtual Channel that is routed, by the ATM network, to the ICG. The ICG performs the following transformation of the digital stream:

1. The ATM Virtual Channel is reassembled to recover the MPEG-2 program stream.
2. A new MPEG-2 transport stream is created from selected program streams.
3. The MPEG-2 payload is encrypted.
4. The MPEG-2 transport stream is modulated into a 6 MHz FAT channel.
5. The FAT channel is combined with other FAT channels, using FDM, and narrowcast over the HFC network to a subset of subscribers.

The advantages of a hybrid ATM/MPEG-2 Transport network are:

- Broadcast and on-demand programs are transmitted in an identical format. Therefore, a single, standard, set-top terminal can receive all services.

- Channel bandwidth is used efficiently. MPEG-2 Transport uses 12% less overhead than ATM.

External Device Data Services and Game Applications Reference Architecture

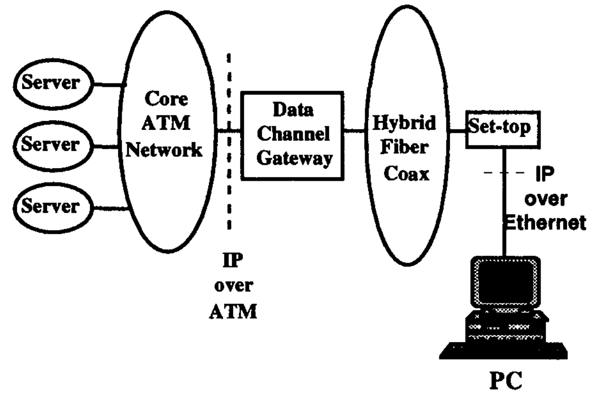


Figure 5 PC Data Services Architecture

The network architecture is also capable of supporting data services and game applications for personal computers and other external electronic devices. Figure 5 illustrates the architecture for personal computer or external consumer device data services. A transparent and secure IP service is provided between an interface in the headend and an Ethernet interface in the home. The service is connectionless and always available to the user. A best-effort, datagram service is provided. Reliable protocols (such as TCP) will be employed at the endpoints to provide re-transmission and sequencing functions when required.

The subscriber sees local area network performance, with high burst data rates and low latency. For example, the Time Inc. New Media LineRunner™ service includes client software for a PC or Macintosh personal computer, and provides access to email, on-line and web-browser services [6].

Game players can also be supported with this same architecture. The Ethernet interface can support two-way communications for multi-player gaming. On-demand download of games is also supported.

INTERACTIVE CABLE GATEWAY

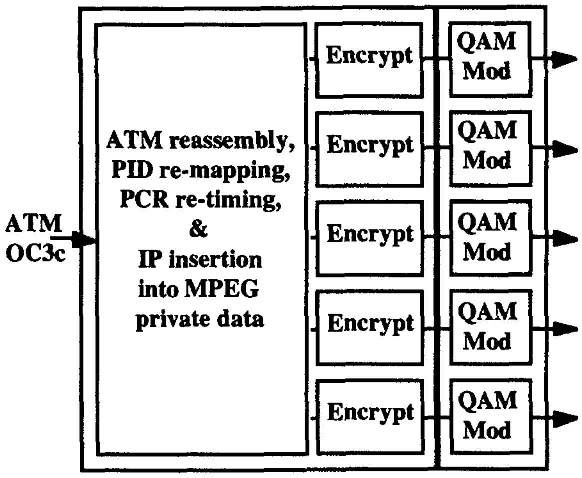


Figure 6 Interactive Cable Gateway

The Interactive Cable Gateway (ICG) transforms the on-demand stream supplied by the media server. Figure 6 is a block diagram of the ICG. The input is a SONET ATM OC3c. The ICG provides up to five FAT channel outputs. The ICG performs the following functions:

- Protocol Conversion
- Rate Conversion
- System Information
- Conditional Access
- Encryption
- Forward Error Correction
- QAM Modulation

Protocol Conversion

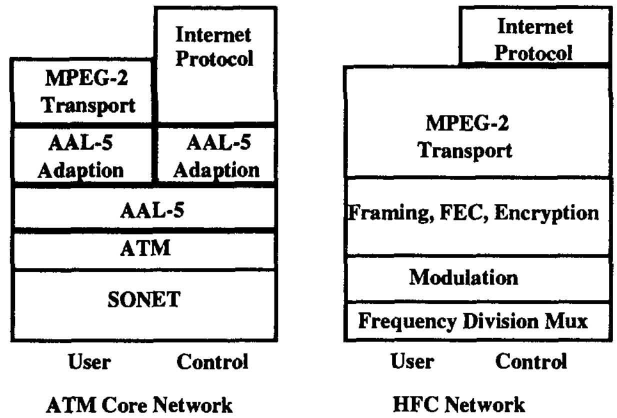


Figure 7 ICG Protocol Conversion

The ICG provides a protocol conversion from ATM Adaptation Layer type 5 (AAL-5) to MPEG-2 transport [7]. The protocol stack of the ATM network and of the HFC network is shown in Figure 7.

User Plane

In the user plane, the ICG receives an ATM Virtual Channel for each single-program stream. The media server assigns fixed Program Identifier (PID) values for program map, video and audio elementary streams. The ICG must re-map these PID values so that they are unique within the new transport stream. The ICG constructs a program-map table and uses this to drive the PID re-mapping function.

1. The ICG terminates and reassembles ATM virtual channels carrying MPEG-2 single-program transport streams.
2. The ICG allocates MPEG-2 PIDs such that each is unique in the output channel.
3. The ICG re-maps the incoming PIDs into their new values.
4. The ICG re-times the MPEG-2 transport streams by adjusting the program clock reference to compensate

for timing errors introduced by the ATM network.

5. A new MPEG-2 transport stream is created from selected program streams.

Control Plane

In the control plane, MPEG-2 Transport provides a high-speed delivery mechanism for Internet Protocol data. The ICG is a suitable point to perform flow-control because the instantaneous output bandwidth of the transport stream is known.

1. The ICG reassembles Internet Protocol (IP) data received on separate ATM connections [8].
2. The ICG inserts IP data into MPEG-2 private data sections [9].
3. The ICG transmits IP data such that it consumes any bandwidth left in the output channel after all video and audio streams have been serviced.
4. The ICG discards IP data if there is insufficient output channel bandwidth to accommodate it.
5. The ICG sets congestion indicators in the reverse direction on ATM virtual channels that experience congestion according to the ATM Forum Available Bit Rate (ABR) specification.

Rate Conversion

The ICG performs a rate-conversion function from SONET OC3c (155 Mbps) to 27 or 36 Mbps payload, suitable for the FAT channel modulator.

System Information

The ICG must generate the program association section (PID 0). Every valid Transport Stream must have a PID 0. The program association section provides a directory of services on the transport stream by listing each by MPEG program number and PID (of each program map section). The program association section also defines the network stream PID.

A program map stream is required for every program stream within the transport stream. The program map stream identifies the PIDs, types, and languages for all elementary streams associated with the program. Each program map stream is supplied by the media server.

The ICG also generates the network stream that may be used to carry various network information [10].

Conditional Access

The ICG must provide a conditional access section (PID 1) for each encrypted transport stream. PID 1 carries the conditional access message that identifies the streams carrying Entitlement Management Messages (EMMs).

EMMs must be uniquely addressed to each set-top and define access rights for each subscriber. EMMs may be sent in-band as part of the Transport Stream or out-of-band on the Forward Data Channel.

Encryption

The MPEG-2 payload from the server is not encrypted. Therefore, it must be encrypted to the cable system requirements before transmission.

1. The ICG encrypts the MPEG-2 payload.
2. The ICG inserts encrypted key information in Entitlement Control Messages (ECM's).

Forward Error Correction

The ICG must apply forward error correction to MPEG-2 transport stream to protect against data errors that occur during transmission over the HFC network.

Modulation

The ICG modulates the payload into a 6 MHz FAT channel.

Management

The ICG must be provisioned with ATM VCI to PID re-mapping information. For conditional access, key information must be provisioned for each encrypted program stream; this is used by the ICG to generate ECMs. A standard SNMP MIB must be defined to allow for centralized management of multi-vendor ICGs in a single system [11].

CONCLUSIONS

The Orlando FSN showed that ATM can be used to deliver interactive multimedia all the way to the home over Hybrid Fiber Coax networks. However, MPEG-2 Transport is a better choice for *integrated* delivery of digital broadcast and interactive services.

ATM is the only available technology that satisfies the *switching* demands of interactive multimedia. Therefore, a combination of ATM and MPEG-2 Transport provides an optimal solution and an Interactive Cable Gateway is required to interconnect the ATM and MPEG-2 Transport domains.

REFERENCES

[1] Communications Technology, "ATM and MPEG-2", M.B.Adams, Dec 1995, Feb 1996.

[2] ATM Forum. ATM User-Network Interface (UNI) Specification Version 3.1. ISBN 0-13-393828-X, Prentice-Hall, Inc., Upper Saddle River, NJ, 07458, September, 1994.

[3] TA-NWT-001113. Asynchronous Transfer Mode (ATM) and ATM Adaptation Layer (AAL) Protocols Generic Requirements. Bellcore, Issue 1, August 1992.

[4] ISO-IEC IS 13818-1, International Standard (1994), MPEG-2 - Systems.

[5] TR-NWT-000235. Synchronous Optical Network (SONET) Transport Systems: Common Generic Criteria. Bellcore, Issue 2, December 1991.

[6] Inter@Active Week, "Time Warner to Offer High-Speed Net Access", S.Vonder Haar, Dec 4, 1995.

[7] ITU-T Document TD-XV111/10. AAL Type 5, Draft Recommendation text for section 6 of I.363. Jan 2, 1993, Geneva.

[8]Internet Draft: Classical IP and ARP over ATM Update (Part Deux). Network Working Group, M.Laubach. 31 Augiust 1995.

[9] ISO-IEC CD 13818-6, International Standard (1994), MPEG-2 - DSMCC.

[10] System Information for Digital Television, Advanced Television System Committee Standard Jan 3, 1996.

11 J.D.Case, M.S. Fedor, M.L. Schoffstall, and J.R. Davin. Simple Network Management Protocol, Request for Comments 1157. DDN Network Information Center, (May, 1990).

A Migration Strategy to High Capacity Return on HFC Networks

Jim Chiddix, Time Warner Cable
Donald Gall, Time Warner Cable
Gerry Shimirak, Raychem

Abstract

This paper explores the potential for the cable industry to use the high end of the RF spectrum for in-bound signals. Experience with such a system is documented, and its performance examined. Approaches to migrating from today's sub-low (5-40 MHz) return systems to the use of high-end return are explored, and a specific proposal, which would minimize or eliminate interruption of existing services during such a transition, is discussed. The conclusion is drawn that this technology is feasible, and has the potential to offer significant benefit as the need arises for high-capacity return in the context of more symmetrical telecommunications and data services over hybrid fiber/coax systems.

INTRODUCTION

The use of cable television transmission systems for two-way communications has been discussed, and occasionally implemented, for at least 25 years. Occasional uses have included return transmission of analog video material originated at schools and other locations remote from the headend, system telemetry, limited data communications, and customer response information in innovative systems such as Warner's QUBE network in the late '70s. More recently, there has been relatively widespread deployment of impulse pay-per-view capability in set-top boxes with non-real-time store-and-forward return transmission capability.

The cable industry has before it today, however, a much wider variety of potentially profitable two-way digital service opportunities, spanning telephony, high speed data communications between residences, businesses, and the Internet, and interactive television. Implementation of these services will require coming to grips with a host of new challenges, notably in the areas of capacity and reliability. Time Warner and other cable companies are shouldering these challenges. This paper will explore a unique approach which Time Warner has pioneered in its Orlando, Florida, interactive television system, which holds significant potential for high volume, highly reliable transmissions over the incoming portion of two-way cable systems.

Historically, cable systems have used the spectrum below channel 2, typically a band from 5 to 30 MHz or, more recently, 5 to 40 MHz, for inbound transmissions – those returning from the subscriber to the headend (Figure 1).

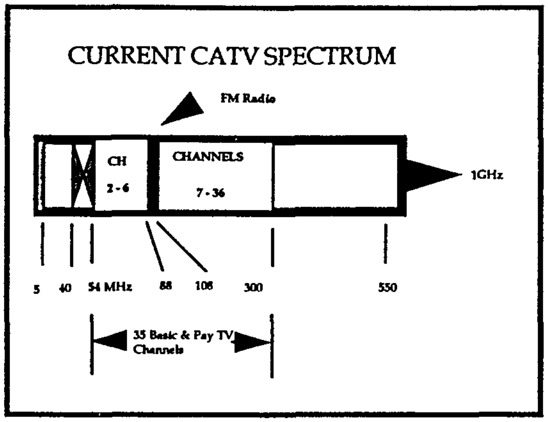


Figure 1

While cable system shielding integrity has improved dramatically in recent years, due in part to the implementation of FCC-mandated leakage monitoring requirements, the physical realities of coaxial plant, both in the distribution system in public rights-of-way and within subscribers' homes, dictate that shielding integrity will never be perfect. This gives rise to ingress of signals from a variety of interfering sources, including discrete carriers from radio sources ranging from short-wave broadcasts to CB and amateur radio transmissions, and from broadband impulse radio sources such as leaking electrical insulators, electric motors, and noisy electrical contacts. These sources are particularly plentiful and strong at low frequencies such as those in the 5 to 40 MHz band.

A second challenge with regard to traditional implementations of return transmission in cable networks arises from tree-and-branch cable architecture (Figure 2).

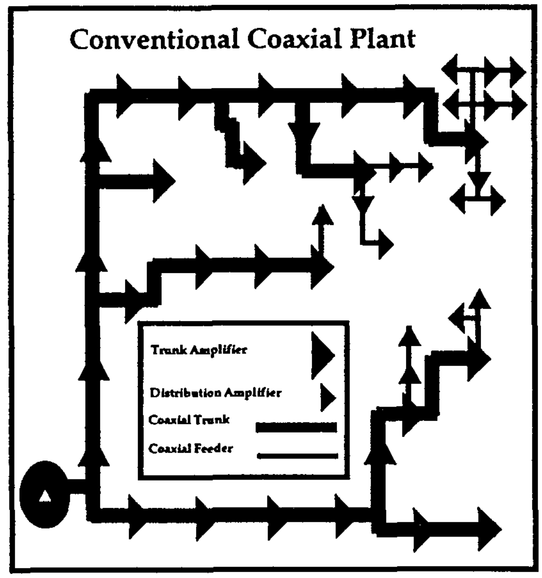


Figure 2

This architecture is a highly cost-effective means of distributing outgoing RF signals, which appear at all end points of the tree-and-branch system, but provides a highly undesirable additive funneling of ingress signals and thermal noise, along with the return transmissions from all points in the network. These factors have forced the adoption of extremely robust, and sometimes complex, schemes for reliable return transmission. Digital modulation schemes used in such systems, for example, generally trade throughput, in terms of bits per hertz, for robustness, using approaches such as QPSK (Quadrature Phase Shift Keying).

Bridger switching, a means of sequentially turning on and then off each distribution leg in the cable plant, and high pass filters on subscriber drops have been tried at various times as a means of making sub-low return transmissions more reliable and manageable. These approaches are, however, expensive and add considerable complexity to the operation of the cable system. Polled "store-and-forward" return schemes are in wide use for non-time-sensitive traffic, and derive robustness from their ability to ask for retransmission of missed data until it finally gets through.

The industry's rapid migration to hybrid fiber/coax ("HFC") architectures mitigates the challenges of utilizing sub-low return frequencies to a large extent by greatly reducing the size of the tree-and-branch sub-network which is funneling ingress, noise, and valid return signals together (Figure 3). In systems utilizing

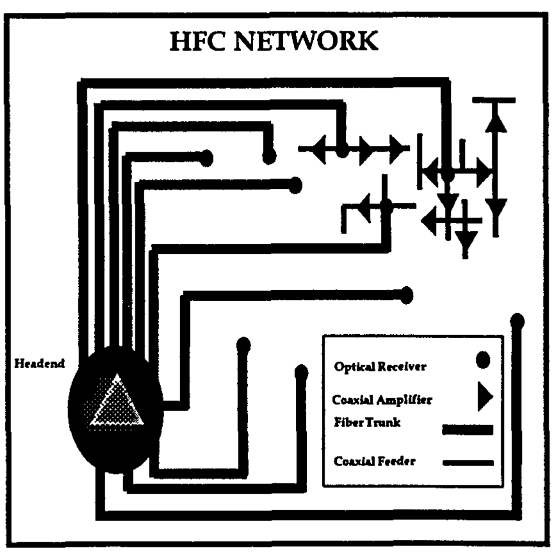


Figure 3

HFC architectures, return digital transmissions in the sub-low band can work quite well. Time Warner has implemented HFC telephony to a large number of customers in Rochester, New York, and has achieved a very high level of reliability. Early trials with cable modems also support this point.

Such systems make use of forward error correction and frequency agility, in addition to robust modulation schemes, to reach a high level of reliability. Nevertheless, HFC systems which utilize sub-low frequencies for the return have limited capacity. Future digital services which require a more symmetrical outbound and inbound transmission may, as penetrations grow, run up against capacity constraints. Such services include extensions of the HFC telephony, cable modem, and interactive television services which are in their early phase at present. The evolution of desktop video conferencing and video telephony, work-at-home applications, and web sites in homes and small businesses may bring about an increase in symmetry and demand for return bandwidth.

In addition to sub-segmenting HFC neighborhoods, the cable industry has available to it a source of a great deal of high-quality return bandwidth. This was demonstrated by Time Warner in its Orlando interactive television system, which has been connected to more than 4,000 homes and involves the use of the highest frequencies in the coaxial spectrum rather than the lowest. The results of the Orlando experience have been very promising, and the high-end return scheme has worked well. There is, however, a challenging migration problem, if and when the industry needs to take advantage of this additional spectrum. The industry has a rapidly growing base of deployed equipment in subscribers' homes which makes use of the sub-low frequencies. Obsoleting this equipment by switching to the high end of the spectrum for return transmissions is unlikely to be economically feasible. It will therefore be necessary, in order to tap the benefits of high-end return, to maintain sub-low transmissions at the same time. This paper will outline ways in which this can be accomplished.

Orlando System UHF Return

A UHF return system is used in the Full Service Network interactive television system in Orlando. There are 20 fiber nodes, serving approximately 100 miles of system configured with the high-return amplifiers. The forward system uses an HFC architecture that consists of a node receiver and up to 6 active devices in cascade. The forward bandwidth is from 54 to 735 MHz and the return bandwidth is from 900 MHz to 1 GHz, with the diplex filter crossover between 735 and 900 MHz. There is no sub-low return path (Figure 4).

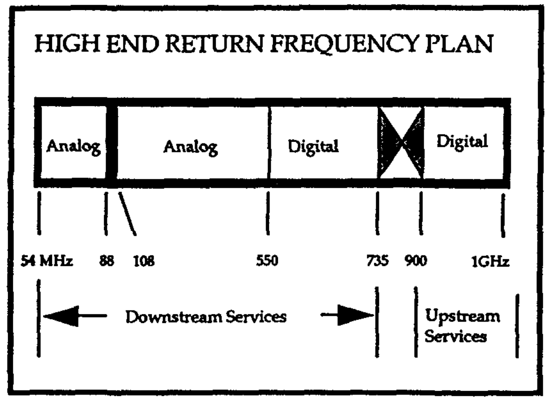


Figure 4

The high-return amplifier module has an operational gain of 36 dB, a third-order intercept point of 24 dBM, and a noise figure of 11.5. System specifications call for setting the lowest return-channel input level to a given amplifier at +1 dBmv. This fixes the system carrier-to-noise ratio ("C/N") at between 32 and 34 dB. This is sufficient to support high-level modulation techniques. Unlike sub-low return, the UHF bandwidth is very clean with respect to ingress and other interference sources.

The current version of the Home Communications Terminal ("HCT") used in Orlando has a frequency-agile transmitter utilizing QPSK modulation, carrying a DS1 (1.544 Mb/s) data signal with a variable RF output of 30 to 55 dBmv. Each HCT is assigned a frequency and a time slot when it is initialized. Current channel loading for the high return is 15 DS1 carriers between 906 and 937

MHz. This yields a peak loading of 375 HCTs per node. Expansion plans call for up to 24 DS1 circuits per node, or up to 600 terminals per node. All planned usage would occupy only slightly more than half the available return bandwidth for interactive TV, leaving the remainder for other services.

Diplex Filter and Gain Considerations

The diplex filter in the current high-return amplifier has a crossover bandwidth of 165 MHz. This was done as a matter of expediency for a prototype product and can be improved upon; significantly better crossover efficiencies should be readily achievable. A UHF return module selected at random and swept for frequency response had 1 dB roll-off at 750 MHz in the forward direction and 860 MHz in the reverse (Figure 5).

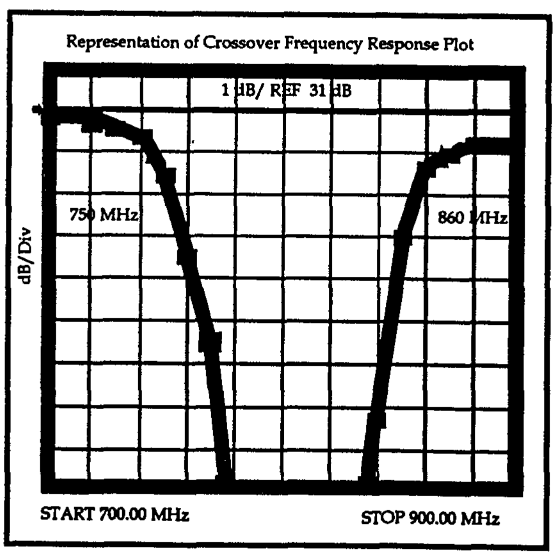


Figure 5

In the Orlando test there is 36 dB operational gain in the active devices (i.e., trunk and line extenders). This capability more than compensates for the losses between amplifiers.

The input to the first active device is set to +1 dBmv. This fixes the maximum loss between the home and amplifier at 54 dB (Figure 6).

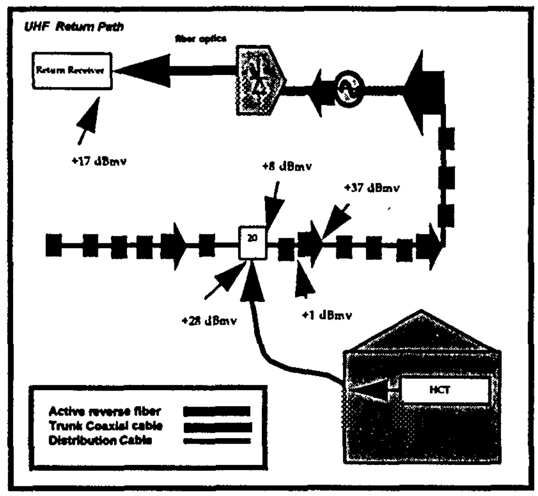


Figure 6

To achieve this maximum loss number, the system has to overcome the drop, coaxial, and passive losses. This can be a problem when a long drop and multiple splitters have been installed, but could be eliminated by giving the HCT a somewhat higher output. Another device that would be helpful for the longer drops and more-than-average number of outlets would be a house amp with gain in the reverse (UHF) direction.

The overall system C/N is between 32 and 34 dB with proper alignment. The return system currently carries 15 DS1 channels that occupy approximately 1 MHz of bandwidth each and can operate without error correction down to a C/N of 25 dB, leaving a margin of 7 to 9 dB. In a more widely deployed system, a somewhat greater margin would be desirable.

The UHF frequency band has so few problems with ingress and all the carriers fall within one octave, interference from ingress and distortions is almost nonexistent.

METHODS OF IMPLEMENTATION FOR NEW AND EXISTING PLANT

In new HFC plant, the implementation of high-end return or of both high and low-end return is relatively straightforward. Amplifiers may be developed by industry suppliers with built-in triplex filtering (Figure 7).

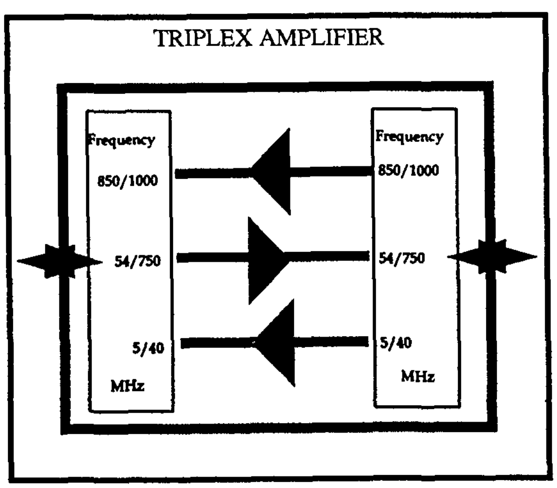


Figure 7

Within a single station, filtering would route sub-low and high-end return signals to their respective return amplifiers, while routing downstream signals, probably in the 50 to 750 MHz band, through a forward amplifier. This approach does include some engineering challenges, as a high degree of isolation must be maintained between these 3 signal paths in order to avoid the establishment of positive loop gain at certain frequencies, with resulting parasitic oscillations. This is a matter of filter isolation, as well as shielding isolation within the amplifier, but may well be achievable with careful design. Such amplifiers, when installed in new or rebuilt plant, would allow the use of customer premises equipment with either a sub-low or a high-end return transmitter.

A more difficult challenge arises in a system which has already been upgraded to 750 MHz downstream capacity, with a sub-low return, where it is desired to add high-end return as well.

Changing out all amplifiers in such a system is economically unattractive, but there are approaches which would allow the addition of high-end return equipment without necessitating the removal of existing active devices. One method is shown in Figure 8. In this approach, outboard high-split filters have been spliced into the cable at the input and output of each amplifier.

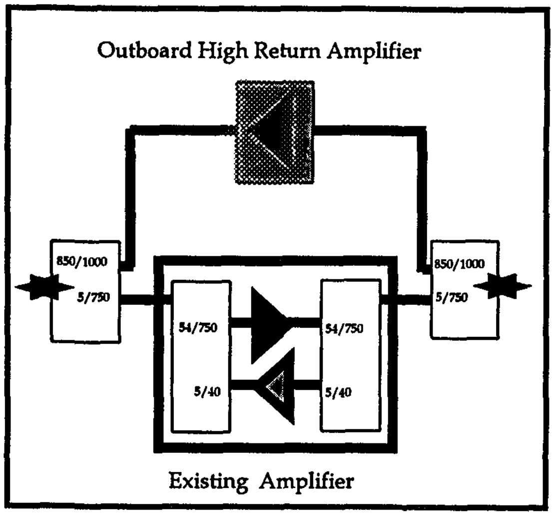


Figure 8

The sub-low diplex filters within the amplifier have been modified to accomplish a low-pass roll-off at 750 MHz, and an outboard housing with a new high-end return amplifier has been added. This approach preserves the investment in active equipment, and also eliminates the need for the high degree of shielding which would be required in the triplex amplifier outlined above. The only factor affecting isolation, and preventing loop-gain oscillation problems, lies in the directivity and isolation of the filters being added to the system. In this approach, there is also an opportunity to couple power from the center conductor of the coaxial cable to the new high-end return amplifier through the new filter.

There is, however, a significant drawback to this approach. It requires interruption of the system while new devices are cut into the cable at the input and output of every active device. This is of particular concern in systems which have implemented HFC telephony, with its need for a high level of reliability and continuity in order to maintain customer satisfaction in a competitive market. There is another approach to adding high-end return to an existing plant which avoids this problem.

Bypass Tap

The bypass tap is designed to route specific RF signals around existing working hardware (e.g., amplifiers) while causing no service interruptions (Figure 9). The bypass tap is modular and consists of 3 pieces: the clampshell, the tap, and the tap insert module. The tap provides an efficient means to clamp onto common coaxial cable sizes while at the same time providing an environmental seal and RF shielding. To install the bypass tap, a special cordless tool attaches to the clampshell

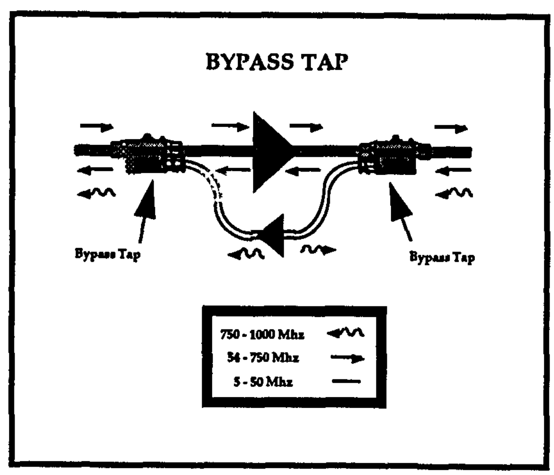


Figure 9

allowing access to the center core of the coax by cutting away part of the coax sheath and dielectric material.

The method used for the RF signal extraction from the distribution coaxial cable is innovative and unlike the traditional tap methods. Signals are extracted from the center conductor without physical contact, through capacitive and inductive coupling. Each tap has isolation of approximately 35 dB. Pairing two of the bypass taps allows the routing of specific RF signals around existing hardware without interrupting service. If power is required, it can be provided by changing to a power-passing insert. In the power-passing path, an RF choke is used to control return loss to avoid a second RF path interference (Figure 10). In both power passing and

FEATURES	SPECIFICATION
Frequency Response	.7 to 1.0 GHZ ± 1.0 dB
Tap Values	10 to 30 dB
Insertion Loss	≤ 1.2 dB
Return Loss	≥ 18.0 dB
Isolation	≥ 35 dB
Hum Modulation	≤ -80 dB
Bypass Power Capacity	2.0 Amps
RF Shielding	≥ 100 dB
Cable Connector	Bypass Port Only
Signal Interruption - Installation	NO

Figure 10

non-power passing versions, the existing signals and electrical power in the coax are not interrupted during installation.

Relative Costs

In 1994 Time Warner completed an informal survey concerning the relative costs of manufacturing amplifiers with either a UHF return or a tri-band device that used both the sub-low and a UHF return. The cost premium for a high-return amplifier over a standard sub-low split was approximately $100. This cost differential would diminish with volume. The triplex filter version is technically more challenging and would require complete redesign of the majority of vendors' product lines. It would also, in many cases, be incompatible with existing amplifier housings.

An interesting alternative may be the bypass tap concept referenced earlier in this paper. This method could be retrofitted into existing plant without replacing existing equipment. The high-return amplifier would be mounted externally to the existing amplifier. Any additional isolation not furnished by the bypass taps could be added by modifying the diplexers in the sub-low split station. The estimated cost of the bypass tap approach described is about $50.

CONCLUSIONS

We have explored the nature of the return transmission challenge: the need for return capacity, and the need to find an evolutionary approach to increasing that capacity. We have explored a means of implementing a transition with minimal disruption of existing services and capital investment. The economics of this transition are probably attractive in the context of the kind of revenues which the need for the additional capacity implies.

In summary, high-end return is an excellent approach to high-capacity return transmission in cable systems. Like all technology introductions, however, its implementation is much more practical if it can be effected in an incremental fashion. We have described such an approach, and believe that it will prove to be practical and cost-effective. As the provision of two-way services over cable television plant grows in importance, the cable industry can rest assured that an increasing need for two-way capacity can be met with practical, evolutionary approaches such as those outlined in this paper.

ANALYSIS OF TWO-WAY CABLE SYSTEM TRANSIENT IMPAIRMENTS

Richard S. Prodan, Ph.D.
Majid Chelehmal, Ph.D.
Tom Williams

CableLabs
Louisville, Colorado

Abstract

This paper presents an analysis of the effects and potential sources of transient impairments found in field testing of two-way cable systems currently supporting trials of digital transmission equipment for two-way services. Data on the characteristics and statistics of impairments that have been captured over extended time periods using a device known as the CW Tester™ are presented. This device demodulates a test carrier wave into in-phase (I) and quadrature (Q) baseband signals. When a transient impairment is sensed, the device triggers a digital oscilloscope to acquire the demodulated impairment. The impairment data is stored on a computer for subsequent analysis using a data browsing software analysis tool for detailed inspection of individual impairments, as well as for longer term statistical analysis and transmission performance characterization. The results of laboratory experiments designed to replicate and thus identify the source of some impairments are compared with similar characteristics observed in the field.

INTRODUCTION

Successful deployment of interactive two-way services on the existing HFC cable systems requires an in-depth characterization of the return transmission medium[1]. Such a characterization seeks to determine the nature of transmission channel impairments as well as limitations imposed on transmission capacity[2]. Characterization should be done for both stationary impairments (e.g., thermal noise, intermodulation distortion, frequency response) and for transient impairments (e.g. impulse noise, RF ingress, signal clipping). An extensive study of cable channel characteristics for stationary impairments has been previously reported [3]. This paper focuses only on the transient impairments for the cable return channel due to the predominance of transient disturbances in this portion of two-way cable systems.

Diagnostic hardware and software tools to facilitate cable system characterization are under development. These tools are useful for plant certification, real time status monitoring, deduction of impairment sources, and prediction of transmission system performance of vendor equipment in the two-way cable environment.

TEST SYSTEM DESCRIPTION

A transient impairment capturing system

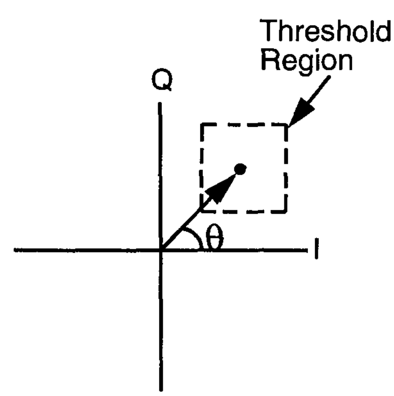


Figure 1. I-Q Constellation of CW Tester™

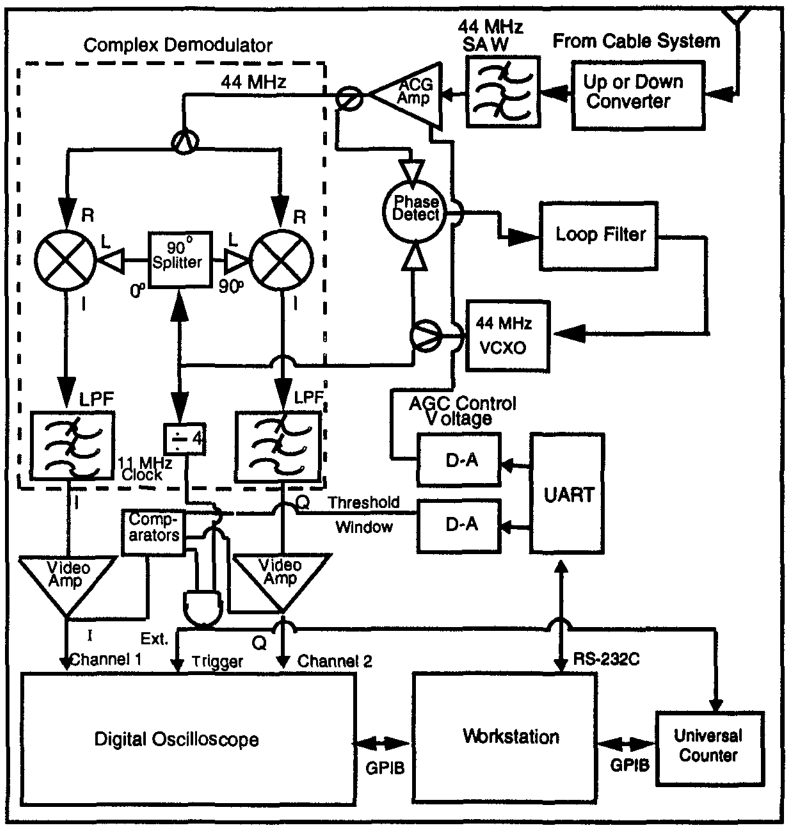


Figure 2. CW Tester™ block diagram

called the CW Tester™ has been developed by CableLabs. This system currently allows for testing of a cable channel in up to a 6 MHz bandwidth. This tester works by placing a carrier wave (CW) in a vacant channel, and may be used on either the forward or return cable system.

Figure 1 is a constellation diagram of a CW carrier. By way of comparison, a 16-QAM signal demodulated into I and Q components and sampled each data symbol interval would produce a 16 point constellation. An unmodulated CW carrier has a single point constellation with some static magnitude and phase angle. A threshold region is established around the static carrier. The power level of the carrier and the size of the threshold region are chosen for the appropriate bandwidth and transmission modulation format respectively. The results reported here assume a 6 MHz QPSK carrier. If the carrier trajectory wanders outside of the threshold region, the oscilloscope is triggered and the demodulated carrier components are downloaded from the oscilloscope into the computer. The amount of time the carrier spends outside of the threshold region is also measured with a gated counter.

A block diagram of the CW Tester™ is shown in Figure 2. A CW carrier is transmitted through the system under test, and either up or down converted to an IF frequency of 44 MHz at the receiver input of the CW Tester. After IF filtering by a SAW filter, automatic gain control (AGC) is applied to the received carrier under computer control to monitor and compensate for cable plant gain fluctuations. A phase-locked loop (PLL) tracks the signal using a voltage-controlled crystal oscillator (VCXO). The VCXO output is divided with a quadrature splitter and applied to a complex demodulator to provide in-phase (I) and quadrature (Q) outputs. The I and Q signals are amplified and supplied to a digital oscilloscope. The I and Q outputs are also provided to comparators which drive an "and" gate. The "and" gate's other input is an 11MHz clock. The output of the "and" gate drives a universal counter and triggers the digital oscilloscope. Once an event triggers the oscilloscope, 5000 I and Q channel samples are collected in about 0.5 milisecond, and the oscilloscope memory contents and the digital counter value are transferred to the workstation. About 2 second are required for this operation. The purpose of the counter is to provide an estimate of the amount of time that the carrier spends outside the threshold region, independent of the oscilloscope signal acquisition. This estimate gives an indication of whether the event captured by the oscilloscope was of a short or long duration. The data captured by the oscilloscope are useful for evaluating the nature and severity of the impairment that triggered the data capture, and can also be Fourier transformed to provide information about the spectral content of the interferer.

Impairments are stored on a computer disk for later data analysis. A data reduction algorithm regulating the interval between event samples from a minimum of two seconds to an arbitrarily long duration will prevent a continuous impairment from overflowing available disk storage with redundant oscilloscope data. However, the counter continuously monitors the error inducing effects of all impairments during any interval.

Figure 3 shows a high level flow diagram explaining the control software and data acquisition process. Three daily files are created each day and are used to record data. The first file stores the bulk of the data. The second file is a separate log file used to store a summary of each event. The third daily file is an AGC file to

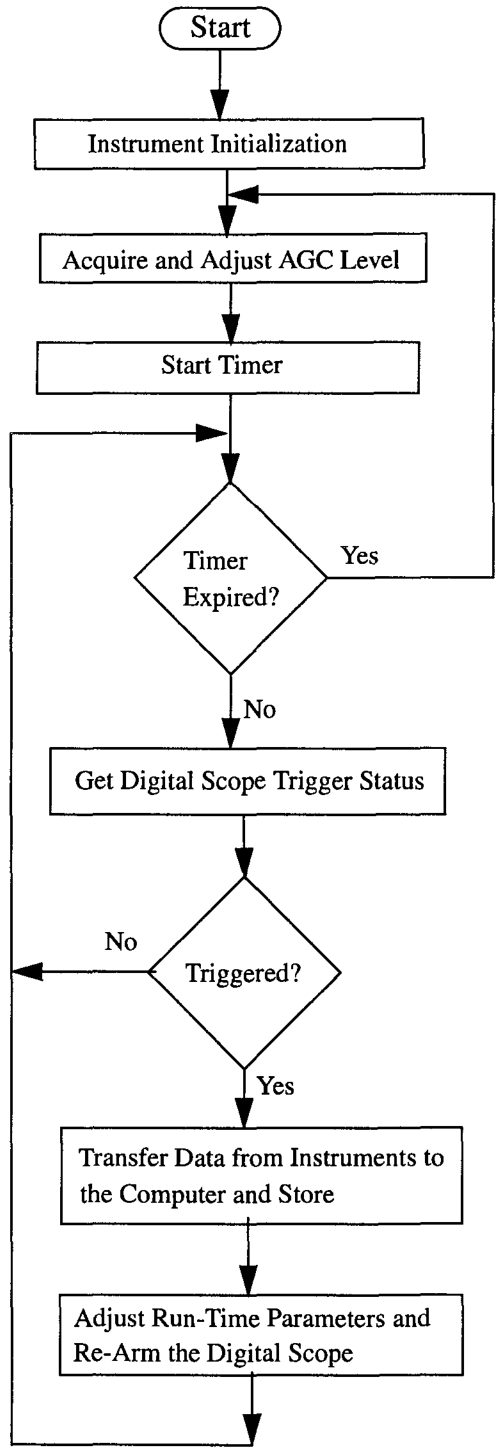


Figure 3. CW Tester™ software control program high level flow diagram

record signal level variations during the day on a minute-by-minute basis. In addition to AGC level, carrier-to-noise ratio is measured each minute and recorded in this file. A unique file format for each daily file identifies date, time, counter value, count interval, and instrument settings so that the captured impairment waveforms can be correctly interpreted.

Besides using the computer for data acquisition control and storage, the control program will initialize various instruments and monitor for abnormalities. This is accomplished through fault tolerant and error recovery software routines included in the control program. Furthermore, a number of flexible features are added to allow for reconfiguring the control environment by reading an input parameter file. The daily files contain actual instrument settings in case the operator changes the default values in the input parameter file. These parameter changes will become effective immediately upon restarting the control program.

DATA BROWSING TOOL

The CW Tester™ can capture a large amount of data making individual observation of each captured event impractical. The data browsing tool developed at CableLabs simplifies the analysis of captured impairment statistics and provides overall channel performance metrics while identifying important and interesting impairment characteristics in detail. This tool provides both long and short term data analysis. A summary of several important digital transmission performance characteristics with a hierarchical structure that can quickly point the user to the individual events of interest for direct inspection.

The LabVIEW software package from National Instruments implements the tool with an interactive graphical user interface (GUI). LabVIEW provides a rich class of ready to use graphical modules for a variety of applications, including tools for signal processing, statistical analysis, and data manipulation. In addition, waveform plotting tools (widgets) are supported for presentation of individual impairment events.

The data browsing tool calculates several layers of performance characteristics pertinent to digital transmission systems providing more data reduction and performance summarization at each level. Thousands of channel impairment events are analyzed with data reduction algorithms providing statistics on captured impairments on a second by second basis, a longer interval of interest (e.g. 1 minute, 15 minutes, 1 hour etc.), a daily basis, and global performance statistics over the entire duration of testing.

The CW Tester™ captures a half millisecond demodulated sample of an impairment that exceeds the trigger threshold in addition to a gated two second count for each time stamped transient event. The duration the impairment forced the test carrier outside of the threshold region is determined for each demodulated sample, as well as for the longer count interval. This data is used to calculate several channel characteristics and performance metrics.

The time that the CW carrier spends outside the threshold region in the demodulated sample determines the *burst duration* of the transient impulsive noise burst. The counter value determines the number of such burst events in the two second (or longer) interval, which in turn provides a short term estimate of the average time separation between impulsive disturbances or *interarrival time*. The proportion of time the CW carrier spends outside of the threshold region in the two second count interval defines the *symbol error rate*. Each second containing symbol errors defines an *errored second* at the symbol error rate so determined. Thus an *unavailable second* at a specified symbol error rate exceeds that error rate during that second. The percentage of errored seconds above a specified symbol error rate in a given time interval (e.g. 15 minutes) defines the *percentage unavailability* at that symbol error rate during that time interval. The unavailability at any error rate over any time interval for one minute, one hour, one day, or the entire test epoch could be

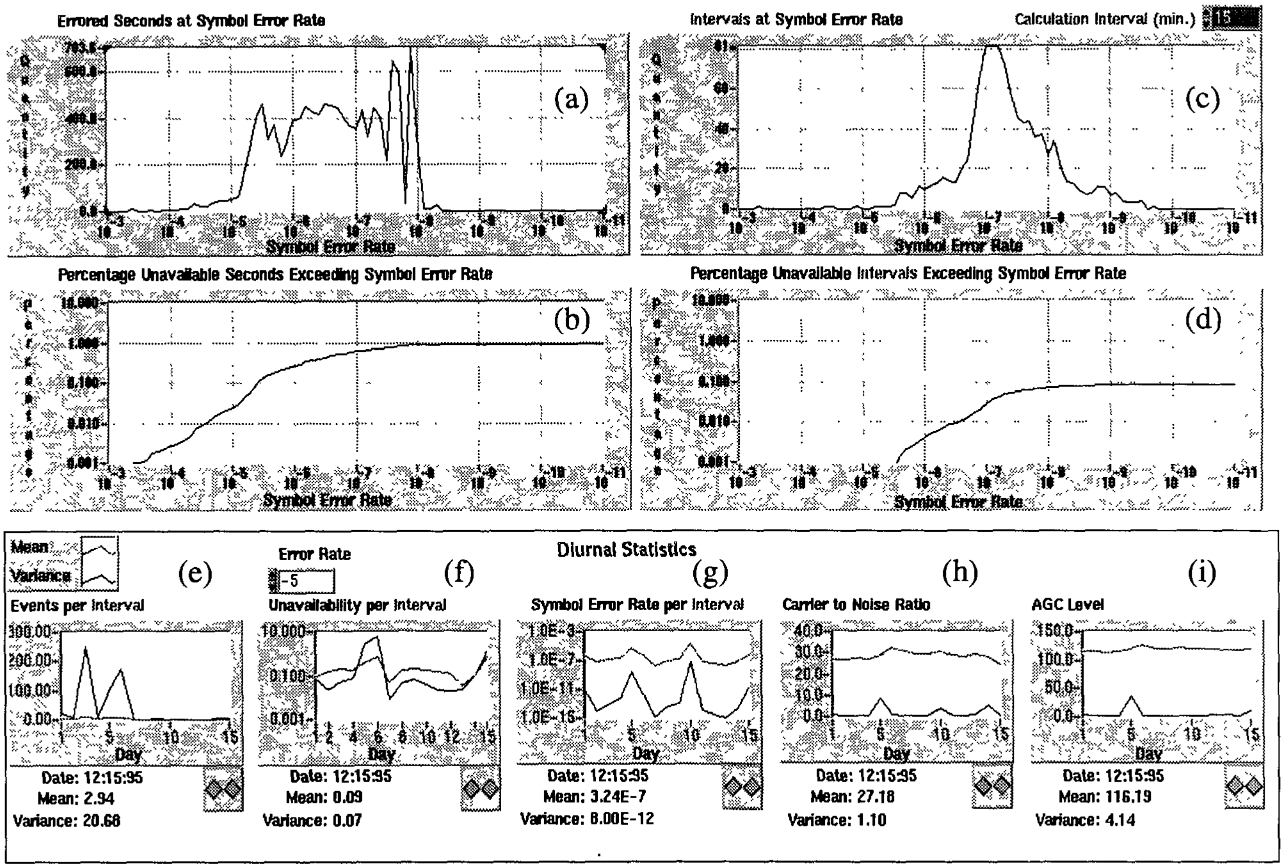


Figure 4.

calculated from the errored seconds statistics.

The CW Tester™ data browsing tool performs a hierarchical analysis of the CW Tester™ daily data files for an extended period of time. The features of the data browsing tool are illustrated in the presentation of field test results obtained from a two-way cable television system in the following section.

PRESENTATION OF RESULTS

The architecture for the field tested system consisted of a two-way express feeder upgrade with a maximum cascade of three amplifiers passing 600 subscribers from the node. The node tested consisted of all newly connected drops.

Figure 4 summarizes the return path performance of the node appropriate for QPSK transmission in a 6 MHz bandwidth centered at 26.75 MHz. The data for this analysis was acquired over approximately a two week test period. Figures 4a through 4d plot several histograms of statistics as a function of the symbol error rate. Figures 4e through 4i show the daily mean and variance of several statistics.

Figure 4a shows a histogram of the number of errored seconds as a function of symbol error rate for the entire test period. Figure 4b is a cumulative distribution of the percentage of unavailable seconds exceeding a given error rate, calculated as the normalized integral of figure 4a. For example, 1.0 percent of all seconds during the entire two week test period exceeded an error rate of 10^{-8}, and less than 0.1 percent exceeded 10^{-5}.

Figure 4c shows the number of 15 minute intervals as a function of symbol error rate. The

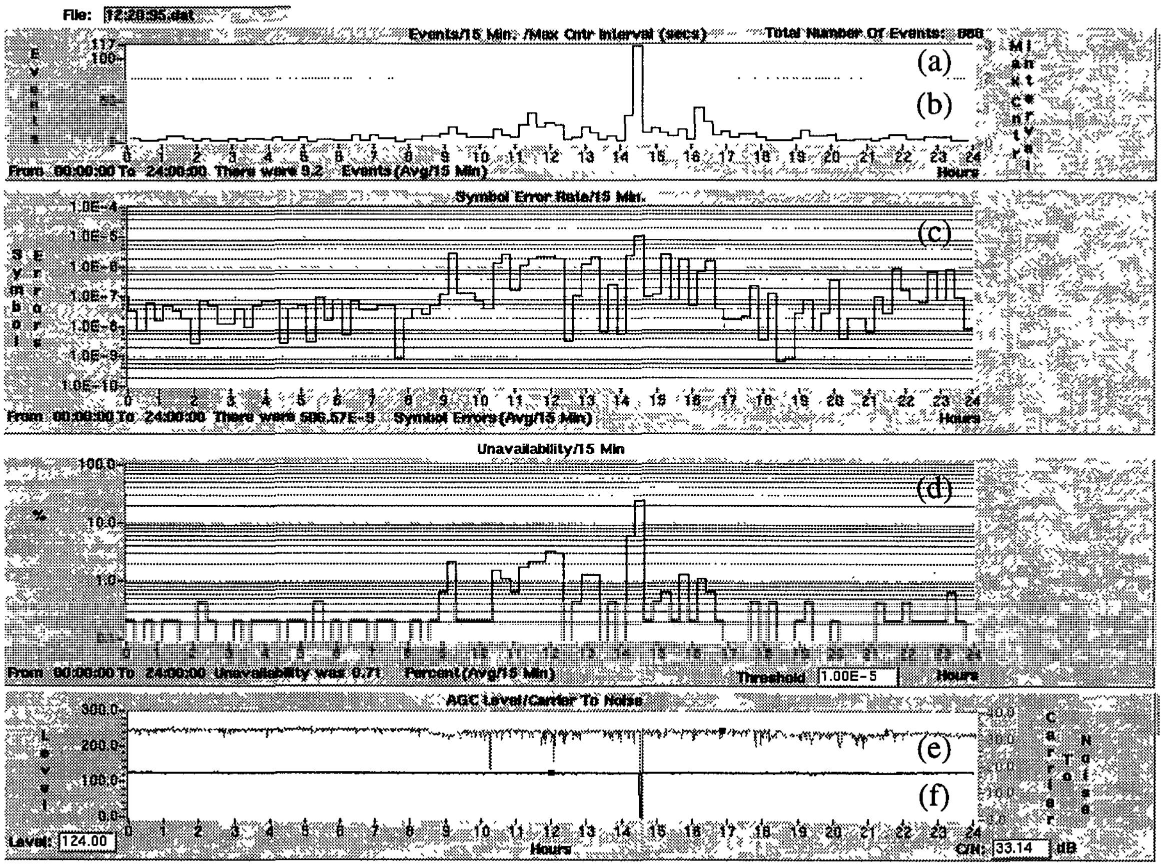


Figure 5.

interval over which errors are accumulated is increased to every 15 minutes in figure 4c from every second in figure 4a. The interval can be varied from 1 minute to 24 hours. Figure 4d is a cumulative distribution of the percentage of unavailable 15 minute intervals exceeding a given error rate, calculated as the normalized integral of figure 4c. For example, 0.1 percent of all 15 minute intervals during the entire two week test period exceeded an error rate of 10^{-8}, and less than 0.001 percent exceeded 10^{-5}.

Figures 4e, 4f, 4g, 4h, and 4i display the daily mean and variance of several interesting statistics. Figure 4e shows the daily mean and variance of the number of burst impairment events per 15 minute interval. Figure 4f shows the daily mean and variance of the percent unavailable 15 minute intervals exceeding 10^{-5} error rate. Figure 4g shows the daily mean and variance of the symbol error rate per 15 minute interval. Figure 4h shows the daily mean and variance of the carrier-to-noise ratio sampled every minute. Figure 4i shows the daily mean and variance of the received CW carrier AGC level, which is proportional to the received carrier level, sampled every minute.

A fluctuation in the mean or a large variance indicates significant time variability in the measured statistic during a 24 hour period. Thus the individual days with significant variation in plant performance can be readily identified for analysis over a single day from the top level of the browsing tool.

Figure 5 illustrates the next lower level of the browsing tool, which shows a representative output of a 24 hour analysis selected from the top level of the browsing tool. A comprehensive report of the transmission performance during

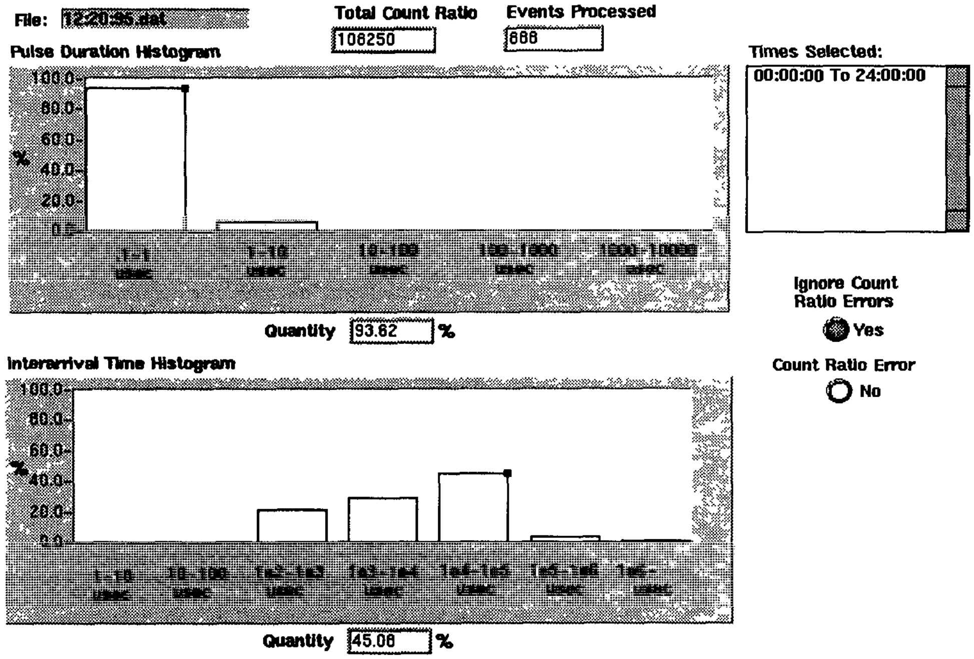


Figure 6.

each 15-minute period in a single day is displayed using several graphs. Figure 5b is a plot of the number of impairment events that triggered the oscilloscope in each 15 minute period. Since a minimum of 2 seconds is required to reset the oscilloscope for another trigger, this number is limited to a maximum of about 450 per 15 minute interval. However, if the impairments are continuous, the system has a data reduction algorithm that increases the minimum time interval between event samples preventing overflow of available disk storage. Figure 5a shows the maximum count interval time in seconds during each 15 minute period.

The counter value is used to estimate the relative amount of time the carrier spent outside of the threshold region. This can be expressed as a symbol error ratio and is shown as figure 5c. The percentage of unavailability for a given symbol error rate over a 15 minute period may be read on figure 5d. In this example, a symbol error rate of 10^{-5} is shown. Other functions that are available are the carrier-to-noise ratio and the AGC level. The carrier-to-noise ratio is obtained by periodically triggering the oscilloscope every minute and computing the ratio. The carrier-to-noise ratio is shown as figure 5e. Since an 8-bit A-D converter is used, the quantization noise limits the measurement range to about 40 dB. The AGC level, which is updated once per minute, is an indication of carrier level drift or an outage if the value is zero. The AGC level is shown in figure 5f.

The burst error duration and interarrival time statistics for the impairments for the entire 24 hour period is shown in figure 6. The predominant number of bursts occur with a duration of less than 1 microsecond, and average interarrival time of about 10 milliseconds. This is typical for most days. A portion of the day may also be selected for analysis of burst error statistics.

The lowest level of the browsing tool allows an operator to set cursors selecting interesting periods of time and inspect the individual impairments that caused the errors. For example, the cursors select 14:30 hours and the impairment shown in figure 7 is a typical scope trace. The I and Q voltages are shown on the

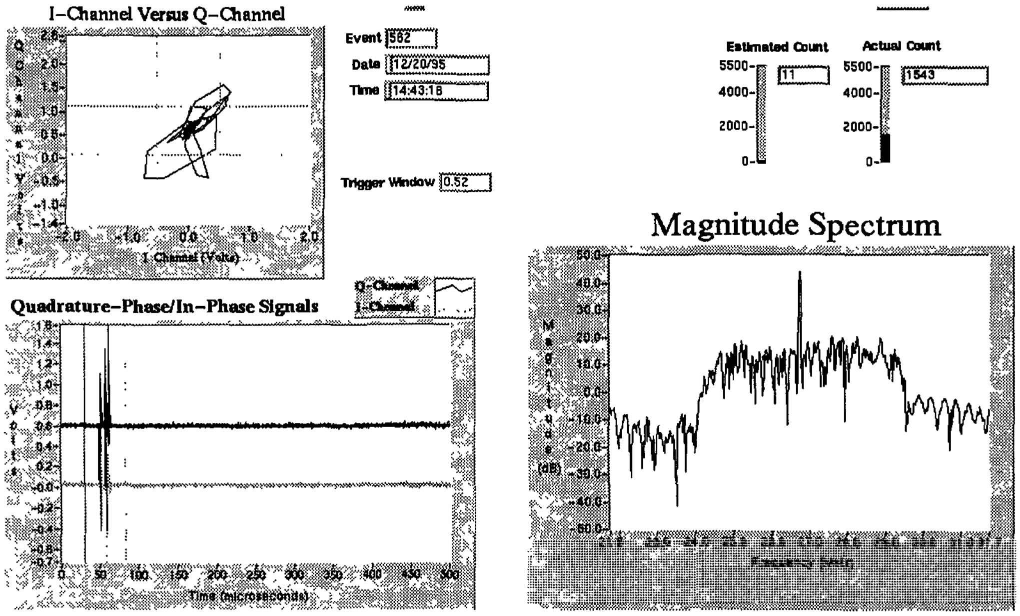


Figure 7.

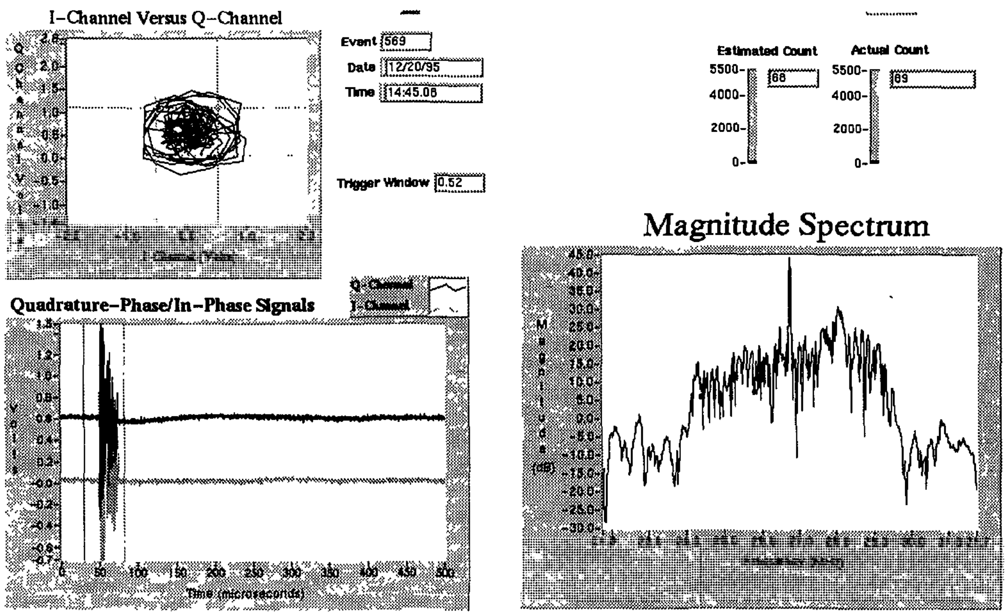


Figure 8.

lower left graph. The vector I-Q voltage is shown on the upper left graph, and a Fourier transform of the time domain data is displayed in the lower right graph. Note that the impulse is brief. This is collaborated in the estimated count box which shows a value of 11. This value is obtained by computing what counter value would have been obtained during the 500 microsecond scope trace time. The total time for this count is 1 microsecond for an 11 MHz clock. However, the counter totaled for 2 seconds, and a count of 1543 was actually received. This strongly suggests that this scope impairment was only one of many that were received during the 2 second period. If the estimated count is close in value to the actual count, the entire impairment was captured by the scope within the 500 microsecond scope trace as

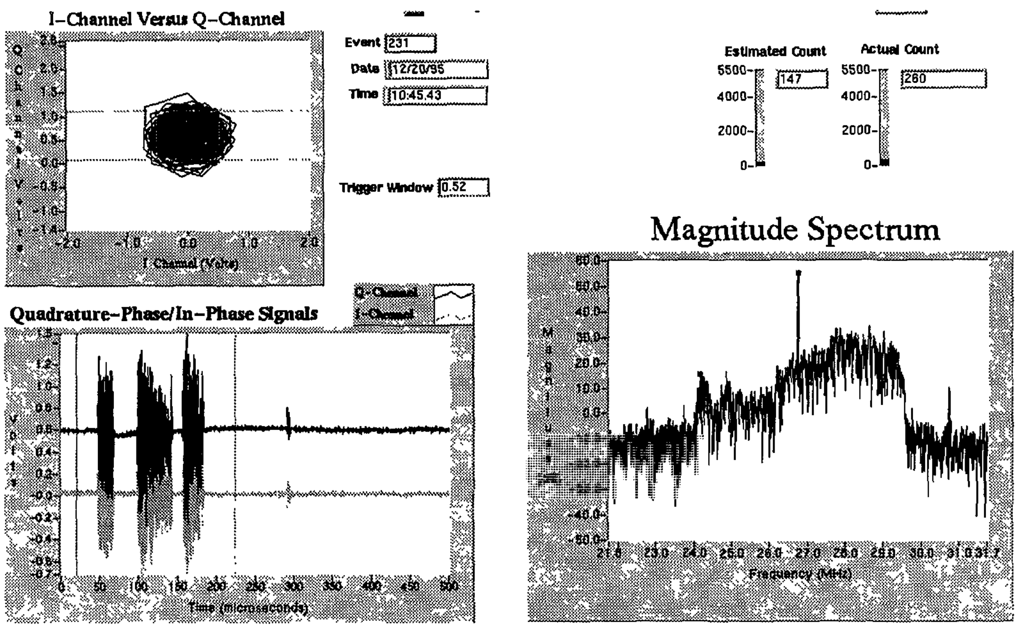


Figure 9.

shown in figure 8. The total time for this isolated event with a count of 68 is 6 microseconds.

LABORATORY VERIFICATION

Some of the impairments captured in the field possessed distinctive characteristics which were repeated in several test sites where the CW Tester™ system was installed. These observations revealed some patterns that suggested experimental replication of these impairments by duplicating the impairment sources and mechanisms in the laboratory. An inspection of the oscilloscope traces from field data inspired two laboratory experiments to determine the probable cause of several observed impairments.

One very prevalent burst type impairment known as impulse noise is similar in appearance to the data shown in figure 9. Switching of electrical devices in homes such as TV sets, VCRs, hair dryers, vacuum cleaners, washers and dryers, etc. could be the source of these types of impairments. Furthermore, the majority of impulse noise ingress originating from inside of the subscriber home enters the cable network through the drop cable connection. The question arises as to the mechanism by which this impairment enters the cable plant. One possible mechanism injects this noise from the public utility neutral wire into the cable system at the common bonding point at the side-of-home ground block.

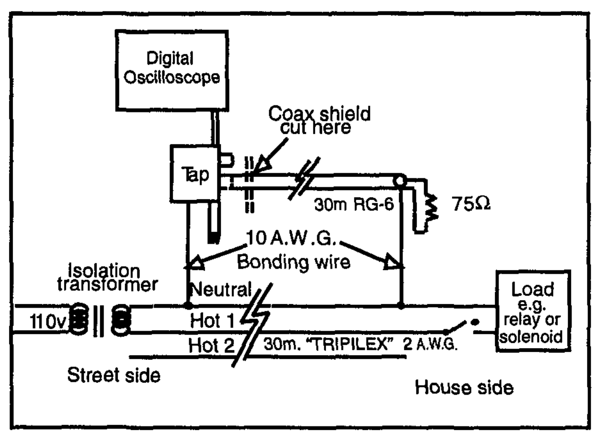


Figure 10. Lab experiment 1 simulating home wiring

Figure 10 is a wiring diagram of the lab experiment attempting to recreate this situation. The point of the experiment was to determine if devices in the home, when they are switched on and off, can create impulsive energy in the return band, if there is a mechanism to conduct the

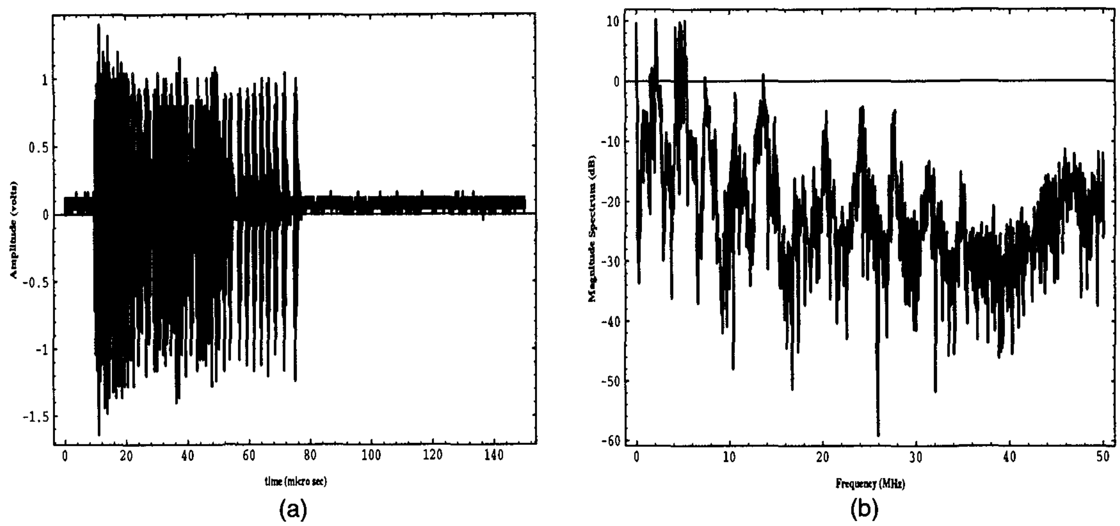


Figure 11. Lab unshielded drop cable time and frequency responses.

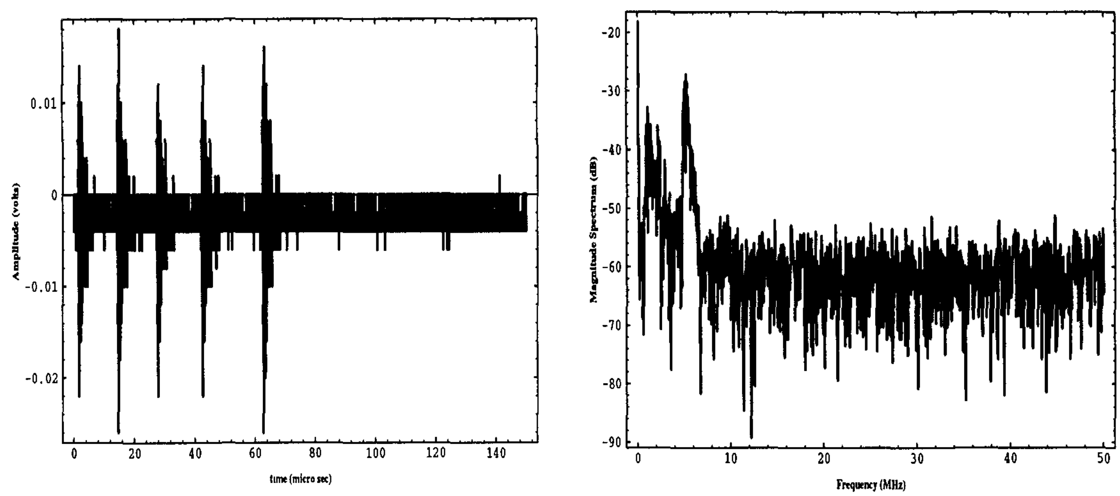


Figure 12. Shielded humidifier load time and frequency responses.

energy from the AC power lines into the cable plant, and if the impulsive energy had similar characteristics to what was recorded in the field. About 30 meters of 2 AWG "triplex" aerial house wiring was bonded on the neutral wire at both ends to a 30 meter piece of RG-6 coaxial cable. The RG-6 coax was applied to an 8 dB tap port, and the output of the tap was connected to an oscilloscope and to a CW Tester™. An isolation transformer provided power to the home wiring at the "street side" and in the "home" a motor starter relay was powered on and off by means of a switch.

A small amount of impulsive energy was coupled into the cable return path. However, when the shield of the coaxial cable was cut on either end, the magnitude of the impulsive energy increased dramatically (about 20 to 30 dB) as shown in figure 11a. A Fourier transform

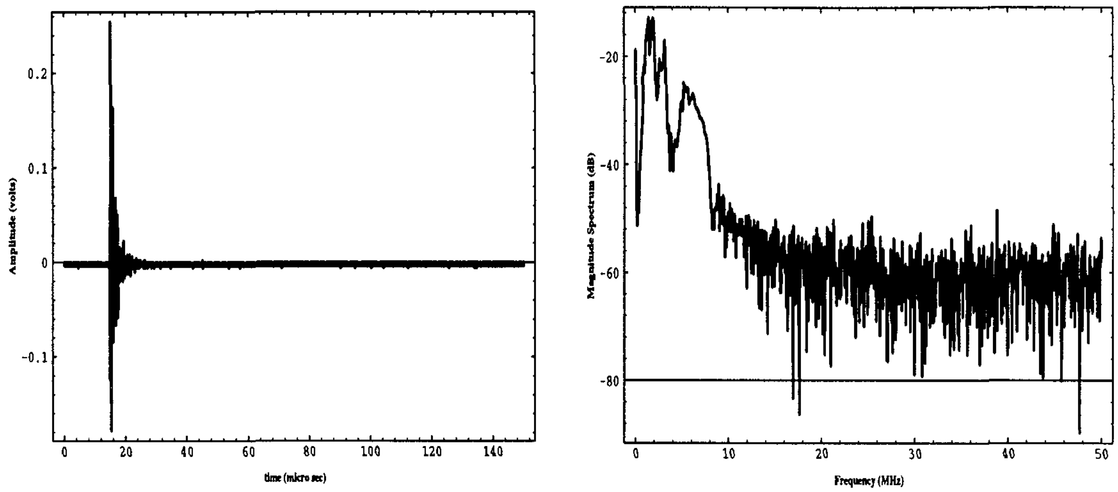


Figure 13. Shielded static electricity discharge time and frequency responses.

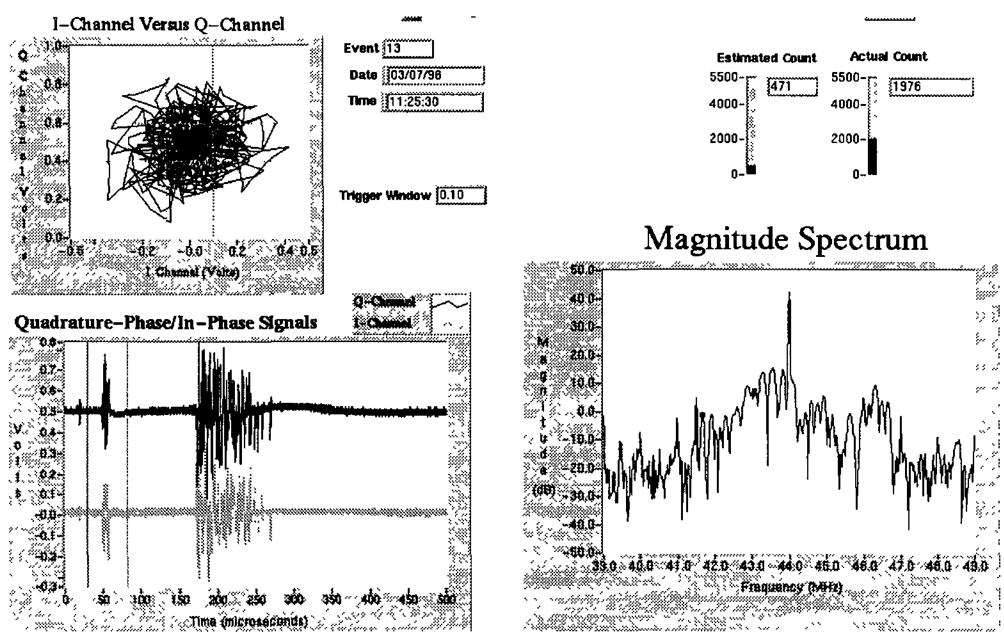


Figure 14.

of the signal indicated that significant energy is available in the entire return band as shown by the spectrum plot in figure 11b.

Similar impulsive events were observed from a randomly chosen home with underground utilities. In this home, the cable was grounded with a ground rod at the entry point, not directly connected to the power ground. The impairment in figure 12 was caused by a humidifier load. Another observation was that static electricity from walking across a carpet being discharged on a wall switch plate could also cause an impairment as shown in figure 13. This home proved to be quite capable of generating impairments with its own appliances without being supplied special loads and switches from CableLabs. When the drop coax shield was

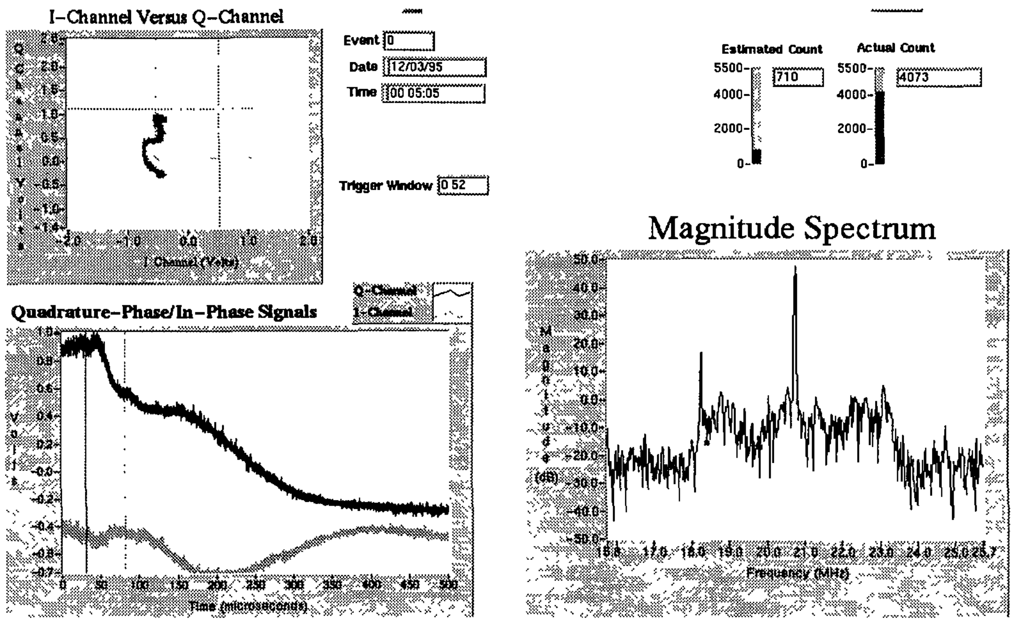


Figure 15.

severed at the tap, the impairments increased in amplitude by 20 to 30 dB.

The spectrum of the impairment indicated that much of the high energy from the house was attenuated relative to the lower frequency energy. This characteristic was not as evident in the lab results. Apparently, high frequency roll-off of the energy may be a general trend of house wiring.

The CW carrier transmitted through the home simulated wiring was connected to the CW Tester™. The result is shown in figure 14. Note the burst characteristics similar to the field results previously shown in figures 7 through 9.

Another type of impairment that was found in the field with the CW Tester™ is illustrated in figure 15. This impairment was much less common than the impulsive impairments and was most easily found in plant with no drops attached. The characteristics of this impairment include a low frequency disturbance of the demodulated carrier. This impairment was also observed in downstream channels. One possible cause could be that the ferrites in the cable plant were being saturated by current from the power grid.

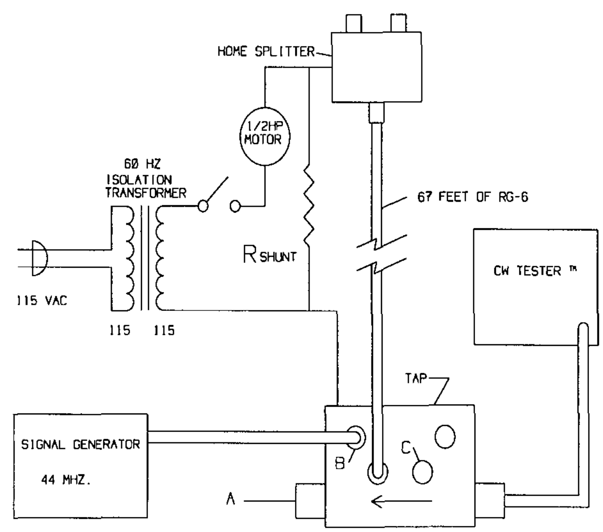


Figure 16. Lab experiment 2 simulating ferrite saturation in the tap.

Figure 16 is the wiring diagram of a lab experiment to replicate this phenomenon. A tap without a blocking capacitor on the tap port was connected through a long piece of RG-6 cable. The cable was connected to a splitter that provided a path between the shield and the center conductor. The coax was used as a

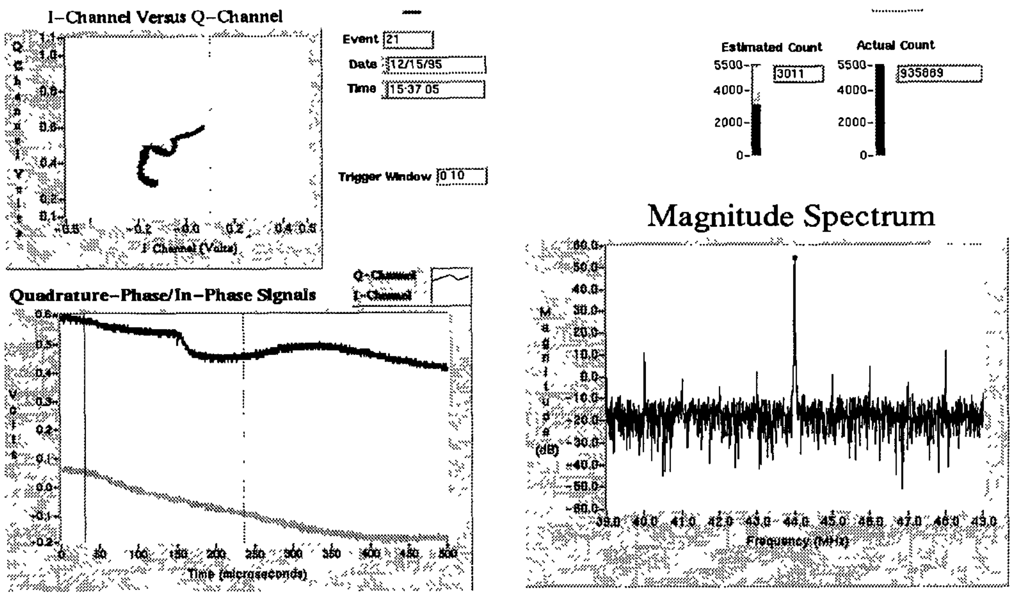


Figure 17.

conductor for a 1/2 h.p. motor. The current was shared between the center conductor and the shield. When the motor was started, a surge of about 36 amps flowed for several hundred milliseconds, after that it went to a steady-state value of 8 amps. The signal for the CW Tester™ was passed through an adjacent port and was attenuated by ferrite saturation as shown in figure 17. Although it is difficult to state if the impairment that was seen in the field was caused by ferrite saturation, the results are similar.

CONCLUSIONS

The CW Tester™ transient impairment data acquisition system and the companion data browsing software analysis tool provides a comprehensive transmission channel performance evaluation system for two-way cable systems. The utility of the hierarchical data analysis features allowing overall global performance evaluation of a two-way cable system (primarily the return path) over a significantly long duration has been described.

The results provided by the CW Tester™ channel characterization and transmission performance evaluation system presented here allows a transmission system designer to evaluate the requirements for modem performance at any desired level of service availability defined in terms of both short and long term channel error rate and burst error characteristics.

The evaluation of transient impairment data acquired over a two week period in the field on a typical 600 home passed node was presented. The error rate measured over successive 15 minute intervals was observed to vary significantly (many orders of magnitude) over a 24 hour period. Some days showed a higher average error rate as well as more variation in error performance.

Several characteristics of typical impairments were illustrated. A potential cause of these impairments was hypothesized. Laboratory and field experiments and results were presented to verify the sources and effects of these transient impairments. Common bonding of the cable plant with power lines apparently provides an entry point for electrical impulse noise generated by electrical appliances in homes.

As demonstrated in the lab and verified in the field, maintaining the integrity of cable shielding is extremely important. A compromised shield on a single drop could produce a hundred-fold increase in interference level from a single home connected to the common return path serving several hundred subscribers. In this situation, locating the individual drop responsible can be extremely difficult due to the time-varying nature of the impairment source.

The two-way cable system return path requires consistent attention to properly maintain a high level of performance. Diagnostic tools for cable plant proof of performance verification, problem isolation, and real-time status monitoring are needed for the operational success of two-way cable services. The CableLabs CW Tester™ system is an embodiment of one such tool for assisting cable operators in the successful deployment and operation of these services.

ACKNOWLEDGMENTS

The authors wish to extend their appreciation for the efforts of David Eng, Craig Owen, Alex Mends-Cole, and Moisey Goldshteyn at CableLabs for their contributions to the successful development and deployment of this project.

We would also like to thank Tony Werner of TCI and his staff and field personnel for their collaboration and assistance in making possible the field studies on several cable systems.

REFERENCES

[1] R. Prodan, M. Chelehmal, T. Williams, and C. Chamberlain, *Cable System Transient Impairment Characterization*, NCTA Proceedings, 1994.

[2] R. Prodan, M. Chelehmal, T. Williams, and C. Chamberlain, *The Cable System Return Channel Transmission Environment*, NCTA Proceedings, 1995.

[3] *Digital Transmission Characterization of Cable Television Systems*, Cable Television Laboratories, November 1994.

ARCHITECTURAL CONSIDERATIONS FOR OVERLAYING FAST-PACKET DIGITAL NETWORKS ON HYBRID FIBER/COAX SYSTEMS

Gaylord A. Hart
XEL Communications, Inc.
17101 East Ohio Drive
Aurora, CO 80017-3878

ABSTRACT

This paper describes basic fast-packet switching network attributes, providing background about the operation and advantages of fast-packet technologies. Fast-packet networks achieve their high throughput by lowering the processing overhead required for transport and shifting this responsibility to the terminal equipment at either end of the network. At the same time, these networks achieve great efficiency by statistically multiplexing data from several users onto a single transport path. Frame and cell relay technologies are discussed, with a brief overview of the relative merits of each.

This paper primarily focuses on architectural issues unique to providing fast-packet transport over hybrid fiber/coax (HFC) networks. Various transport topographies and multiplexing methods over the HFC network are described and discussed. Particular attention is given to the return path, which presents unique challenges for achieving high throughput efficiency just where spectrum is currently most limited. Congestion considerations and the effects of BER on system throughput are also discussed.

INTRODUCTION

Traditional telephony network architectures are circuit switched. In this case, an actual circuit or channel of fixed bandwidth is assigned to connect two endpoints in the network. This connection may be set up on a permanent basis (as with leased lines) or on a real-time basis (as with a telephone call). The circuit is dedicated to these endpoints for the duration of the connection, even if no traffic is flowing between them. This requires that additional channels be provided in the network for all other active users, one for each connection.

If the demand for connections exceeds the network's channel capacity, no further connections can be made until a channel is released. This condition is referred to as blocking because access to transport is blocked. To eliminate blocking, enough channels have to be provided for every possible connection in the network under peak load conditions, but this is not economical since average traffic loads are typically far below peak loads. Most switched circuit networks are designed to minimize blocking, but not to eliminate it. Queuing theory is used to statistically derive how many channels are required to maintain a certain probability that blocking will not occur or will occur less than a certain percentage of the time.

The fixed channel bandwidth restriction of circuit switched networks is a disadvantage if data or other services with mixed or varying bandwidth requirements are being transported. The end user may wind up spending more for a large, yet mostly under-utilized transport channel just to meet his peak demand transmission requirements. On the other hand, another user may end up spending less for a channel which is utilized very efficiently most of the time, but which fails to meet peak demand requirements because the transport channel bandwidth is too small. Since switched circuit networks cannot provide bandwidth on demand they are not very flexible in this regard.

For voice applications, switched circuit networks are efficient. Voice traffic tends to utilize a channel 100% of the connection time, and the bandwidth requirements do not change over time. However, this will not be the case where different services may be offered over the same pipe at different times (each with its own bandwidth requirements) or with data transmission (which tends to be bursty). Channel utilization will likely be less than 100%, and different channel bandwidths will be required depending on the application and the peak to average transmission requirements of that application.

For modern multimedia networks, which must carry voice, video, and data, circuit switched topologies are highly inflexible and inefficient. In a purely digital network, voice, video, and data applications each have unique transport requirements which demand flexibility from the network. Ideally, a single digital pipe to the home should provide all these services. An advanced digital network must somehow provide enough built-in flexibility to provide the unique transport requirements of each of these applications while at the same time using a common transport mechanism for all. Also, an advanced digital network must provide for efficient and graceful evolution of the network and network applications. Fast-packet switched networks will provide the flexibility, efficiency, and low cost necessary to accomplish these objectives.

FAST-PACKET NETWORKS

Packet switched networks have been around for quite some time, typified by protocols such as X.25. Early packet transmission protocols were developed when transport speeds were relatively low and transport errors relatively high. These protocols were developed with an emphasis on error-free transmission, not speed, and typically require an elaborate handshaking whereby each network link has to verify or repeat transmission until error free reception is achieved before a packet can be sent on to the next link, where the process is repeated once again. This handshaking demands considerable link overhead and thus makes transport more difficult at the speeds required today.

New packet protocols (referred to as fast-packet protocols) have been developed which eliminate most of the transport overhead and error processing in the transmission path itself, thus allowing higher transmission throughput. These fast-packet protocols have been developed to take advantage of the virtually error-free transmission of today's networks and the use of low-cost, intelligent terminals at the network endpoints.

When transmission errors are rare, it makes no sense to provide elaborate mechanisms for detecting and correcting errors at each link in the network. Consequently, most fast-packet networks perform little or no error processing at the link layer (OSI transport model layer 2). In many cases, if this processing exists at all within the network, it consists only of detecting errored packets and discarding them. After all, it makes no sense to send an errored packet any further and risk congesting the network.

Nevertheless, providing reliable service requires that error checking and correction be done somewhere. In fast-packet networks this function is typically performed at the transport layer (OSI model layer 4) by intelligent endpoint terminals. Processing errors at the end points eliminates transport overhead and allows the network to transport data quickly and efficiently. The higher throughput of fast-packet networks is accomplished to a high degree by eliminating lower level layers from the OSI transport model or by transferring the functional responsibility for these lower level layers to higher level layers in the model. Errored packets are simply thrown out by the network without any attempt to correct the error or notify anyone of the error. The end terminals must themselves determine if packets have been received with errors or have been lost and then provide for retransmission.

Packet switched networks further increase throughput by statistically multiplexing several users' digital information onto a single transport channel. This information may consist of voice, video, or data applications, or a combination of these. In fact, the packet payload may itself be packets using a different transport protocol. Unused channel capacity from one user is then allocated to another user on a real-time basis, thus taking advantage of the dynamic nature of channel capacity requirements. Fewer transport channels are now required in the network since each individual channel is more fully utilized. This results in more efficient use of network resources and lower transport costs.

Multiplexing is carried out by partitioning each user's data into small packages, or packets, for transport. Each packet has a strictly defined structure, determined by the particular packet technology in use, containing the user data and additional overhead for performing other critical transport functions: 1) packet boundary delineation, 2) packet routing to the intended destination, 3) congestion control, and 4) error detection. The packet overhead is kept to a minimum since it consumes transport capacity.

Individual packets from each user are buffered at nodes in the network, then time-division multiplexed into the network when transport capacity becomes available. Packets are transported and routed within the network as indivisible units. At the far end, the packet overhead is stripped off, and the payload data bits are reassembled in the correct order before handing off to the end user or terminating application.

Unlike circuit-oriented connections which have fixed data rates, packet-switched transmission has the unique advantage that packets do not inherently have any data rate associated with them. Providing that the transport channel has sufficient bandwidth, the actual channel speed is transparent to the end users. Since packet transport is not locked to a fixed data rate, bandwidth may be allocated to users on demand and without hardware changes. Higher bandwidth is simply allocated by allowing a given user to transport more packets in a given period of time.

Unlike switched-circuit networks which may have connections blocked while circuits are unavailable during heavy traffic loads, packet networks can still accept packets under heavy load conditions. When congestion occurs, a packet network will simply experience greater transmission delay. The nodes in a packet network are connected by fixed size transmission pipes. Under normal conditions, the number of packets entering one of these pipes is not enough to fill the pipe, and packets are allowed to enter the pipe as soon as they are available.

Congestion occurs when more packets are trying to enter the pipe than the pipe has capacity for. When this occurs, packets must be buffered at each node while awaiting a slot in the pipe for transmission. The buffers thus serve to mitigate peak demand by spreading it out over a longer period of time. As the buffers begin to fill up with packets, the delay increases for each packet before transmission to the next node. However, the buffers are also finite in size. If the network experiences extreme congestion (i.e., the packet buffers overflow), significant delays will result because those packets lost in the buffer overflows must now be transmitted again, thus adding more traffic when it is least desired. Once congestion begins to occur in a packet network, performance tends to degrade rapidly.

For data services, delays may be unimportant. But for video and voice services any such delay will likely be intolerable. However, techniques exist for mitigating congestion and system delay. Dynamic routing can help reduce congestion by balancing the transmission loads of the various links in the network. Packet networks usually also allow packets to be prioritized, which means that more important services or services requiring real-time transport (e.g., television signals) can be given greater access to the network when desired. An alternative approach allows the network to discard less important packets when congestion occurs. Careful network planning is required to control and minimize congestion.

Packet transport may be further characterized as being connection-oriented or connectionless. Regardless of whether a network is connection or connectionless based, switching functions must be provided at each node in the network for routing packets on to the next node.

As the name implies, connection-oriented networks require that a logical connection be established between two endpoints before data may be transferred between them. These connections are made via virtual circuits and require setup operations to establish each connection and its routing path. Virtual circuits use a pre-defined routing path for all packets traveling between two network endpoints. These routing paths are defined by logical and physical paths through the network from one endpoint to the other. Since all packets follow the same path through the network, they also arrive at the end node in the same sequence as originally transmitted (providing no packets have been discarded due to errors).

The term 'virtual circuit' has been coined because such a channel appears to the end user very much like that provided by switched circuit networks. Unlike switched circuit networks, however, several virtual circuits (and their packets) can share the same physical channel between any two nodes internal to the network. Virtual circuits may be further characterized as switched virtual circuits (SVC's) or permanent virtual circuits (PVC's). SVC's are analogous to switched circuit network dial-up connections in that the setup and teardown of the connection is done on a demand basis. PVC's are analogous to leased line connections and must similarly be provisioned to establish the connection..

For connectionless networks, no previously established connection between endpoints is required, and thus no pre-defined data path exists through the network. Packets are simply put into the network with a final destination address inserted into the packet at the originating node. Datagram transmission is used within the network, whereby each individual packet (or datagram) is routed through the network independently of any preceding or following packets.

Each successive receiving node examines the packet address to determine if this node is the final destination. If not the final destination, the node takes into account its position in the network and the possible paths to the destination and then routes the packet on to the next node. Each packet may take a different path through the network and may actually arrive at the far end out of sequence. In this case, the receiving terminal is responsible for sequencing the packets in the correct order. Datagram networks allow dynamic fault recovery by routing packets around damaged links and on-the-fly congestion control by sending packets over lower utilized routes in the network.

Fast-packet transport is typified by two technologies: frame relay and cell relay. The primary difference between frame relay and cell relay is that frame relay uses variable length frames whereas cell relay uses small, fixed length cells. Both have fixed transport overhead associated with each packet. For frame relay, five octets of overhead are required per frame, but the payload may consist of one to 4096 octets. Frame relay typically incurs a lower transport penalty for its overhead since more user data may be transported per packet than with cell relay. On the other hand, variable length frames are more difficult to process because of their variable length. Longer frames, simply because of their size, are also more likely to take errors, which means that a larger frame must then be retransmitted for error recovery.

Cell relay is best typified by ATM, which uses a 53 octet transmission cell comprised of five octets of overhead and 48 octets of payload. Fixed size cells are inherently easier to process because the location of each component in the cell is always the same, which readily allows direct implementation of cell processing in silicon. Because fixed size cells allow greater control and predictability of transmission timing and delays, cell relay is more easily optimized for low delay, high bandwidth applications and some versions are suitable for voice, video, and data.

Other differences between frame relay and cell relay exist as well, with each technology providing unique advantages and disadvantages. Both frame relay and ATM are connection oriented fast-packet technologies supported and defined by industry-wide standards.

HFC PACKET NETWORK ARCHITECTURES

Traditionally, CATV systems have existed as isolated islands receiving signals via satellite. It is common for a large metropolitan area to have several CATV systems operated by different MSO's, each system providing service to a particular geographic section. Because of the broadcast nature of traditional CATV television services, little need existed for building large interconnected CATV networks between these systems. Even within systems, individual headends typically only needed to be connected to transport video signals between them. This was accomplished with AML links or more recently with fiber.

All of this has changed with the introduction of new services, deregulation, and competition to provide existing and new services. Today's HFC network is envisioned to provide in the near future analog television, compressed digital television, HDTV, telephony, data services, internet access, and numerous other interactive and multimedia services. Many of these services cannot be provided without being connected to a larger universe of other networks, indeed, both national and international networks. It will no longer be possible to operate systems as islands.

Nor will it be possible in the long run to operate HFC networks as providing disparate services, each sharing the HFC transport path but essentially using different transmission formats and signaling schemes. The first stage of this evolution, replacing analog signals with digital signals, is already well along. Many operators already simultaneously carry FM radio signals and CD quality digital audio music services. In a short time, compressed digital NTSC services will be common. The benefits of digital technology, including better signal quality and bandwidth efficiency, will accelerate this change.

As more signals on the HFC network go digital, there will be greater economic incentive to process and transport these signals through common channels and equipment. After all, bits are bits, whether they encode voice, video, or data. Such an integrated network will be capable of delivering all these services over a single data stream with unique content and connections for each home. This network can best be built using modern fast-packet technologies.

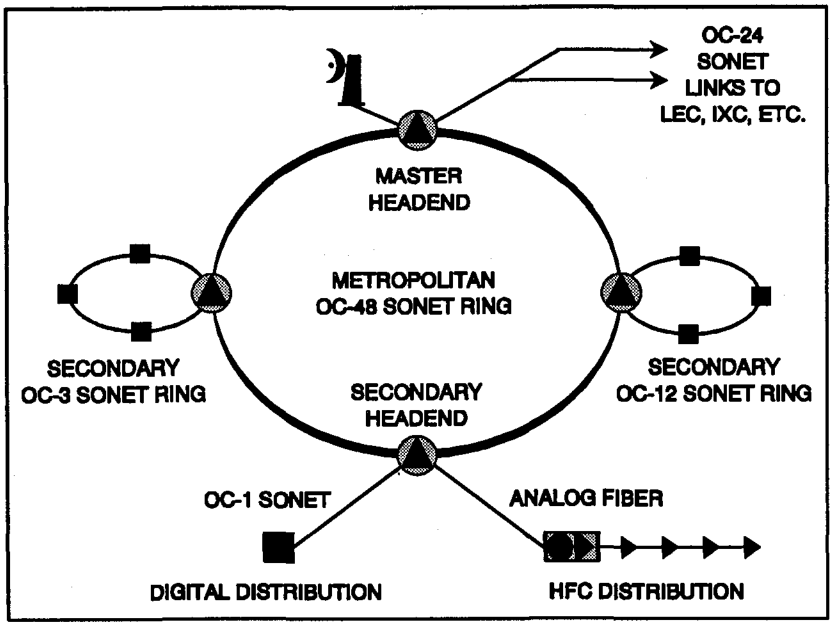


Figure 1--Advanced HFC Metropolitan Network

Advanced Metropolitan HFC Networks

Figure 1 shows a large metropolitan HFC network which consists of a master headend for gathering and processing signals and several secondary headends. All of these headends are connected by a SONET ring which is used to transport digital signals between the headends. Secondary SONET rings may also be present at each headend for local transport of digital signals to businesses or secondary hubs attached to each headend. Each headend will in turn have several HFC nodes to support local distribution of services to residential and small business areas. Point-to-point SONET links may also be used to distribute digital signals to individual businesses or other sites. The overall metropolitan network has additional SONET links which connect to the larger network universe: LEC's, IEC's, and other CATV systems.

Though not ubiquitous, networks of this type have already been built. From a functional standpoint, the overall SONET ring architecture has been used to transport digital television and audio signals from one headend to another, much as with earlier AML systems, and for local ad insertion and distribution. The ring has also been used to support parallel alternate access telephony services and data transport services, though these have typically been operated separately from the CATV network itself. More recently, some operators have been installing telephone switches in headends to accommodate basic telephony services as well, and these switches use the SONET ring for connecting to customers and other carriers.

In most cases, however, these networks are primarily characterized by multiplexing equipment and have been deployed using traditional circuit based technologies. The various services are multiplexed onto and off of the SONET ring using dedicated channels. As these networks migrate toward packet based transmission systems, the overall metropolitan architecture will not change, but more of the SONET transport capacity will be used as the physical transport layer to carry packets between switching nodes. The fundamental changes will take place at the switching nodes, which will be comprised of the regional headends and other hubs where signal distribution takes place. Here packet switches will be added after the SONET multiplexers to route packets onto and off of the SONET ring and into and out of the local distribution system.

The functions described above already being performed by the circuit based version of this architecture will continue to be performed, but through virtual circuits in the packet network. Television and audio signals will still be distributed between headends, but as packets. Imagine how easy it will be with a packet switched network to distribute local adds to particular headends or even to a particular node on a headend. In the long run, ads may literally be targeted to individual homes.

New services will also be made possible or economical by packet networks, and these too will be integrated into the network as it evolves. For example, video on demand can readily be implemented since video packets may be switched and routed between any two points on the network. Ultimately, this switching capability will extend to the individual subscriber, allowing true virtual channels. Additional services such as basic telephony and internet access will at some time also be carried by the packet network. Since both of these services depend heavily upon switching, the inherent switching capability of the packet network provides an efficient, integrated approach to providing these services.

<u>HFC Headend with Packet Network Overlay</u>

Figure 2 is a block diagram for a headend in the packet switched metropolitan HFC network, showing the additional digital and packet transport components. The traditional analog television and audio distribution components have been left out for simplification, but these would also feed each of the HFC nodes. The SONET multiplexers used to interconnect this headend with other headends and secondary sites in the packet network are shown on the left. These multiplexers allow signals on the SONET ring to be dropped off at this headend and signals originating from this headend to be added onto the ring.

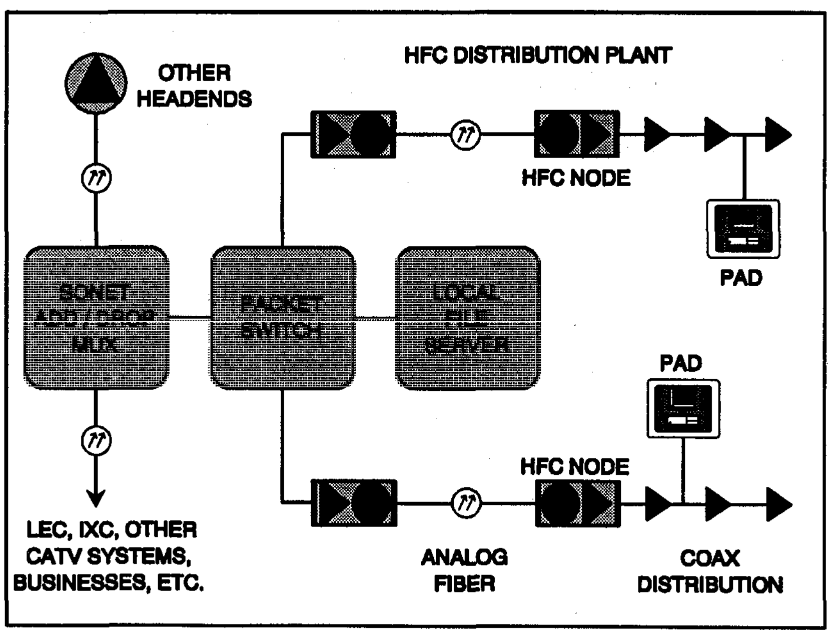


Figure 2 -- HFC Headend with Packet Switch

Since the headend is a major routing node in the network, a packet switch is central to the headend and is shown here with ports feeding the SONET multiplexer, a local file server, and all the HFC nodes on the network. This switch will examine every packet entering one of its ports and route each packet out the appropriate port towards its final destination. For packets being sent out to HFC nodes, this will typically be the last switch the packet sees in the network. For packets coming in from an HFC node, this may be one of many switches the packet will pass through before reaching its final destination.

The file server shown in this headend might support any one of a number of applications. This could be an internet file server or a video file server for video on demand services. In all likelihood, a headend will contain many file servers, each designed to support a specific application. But each of these servers will be connected to the network through the packet switch.

PAD's, otherwise known as Packet Access Devices or Packet Assembler/Disassemblers, are shown on the right hand side of this diagram. These would be located in the subscriber premises, whether residential or business, and would be served by a drop from the coaxial distribution plant just as any other service. For the purpose of this paper, a PAD may be considered any device providing a packet network service to an end user. This could be an internet access port for a computer, an ethernet port connected at the other end to an office for telecommuting, a set-top box providing compressed digital NTSC to a TV, or even a telephone.

HFC Node with Packet Network Overlay

Figure 3 shows the HFC node with packet network overlay in greater detail, though still greatly simplified. A standard HFC distribution architecture is presented. Here you can see the actual forward and return RF paths of the node.

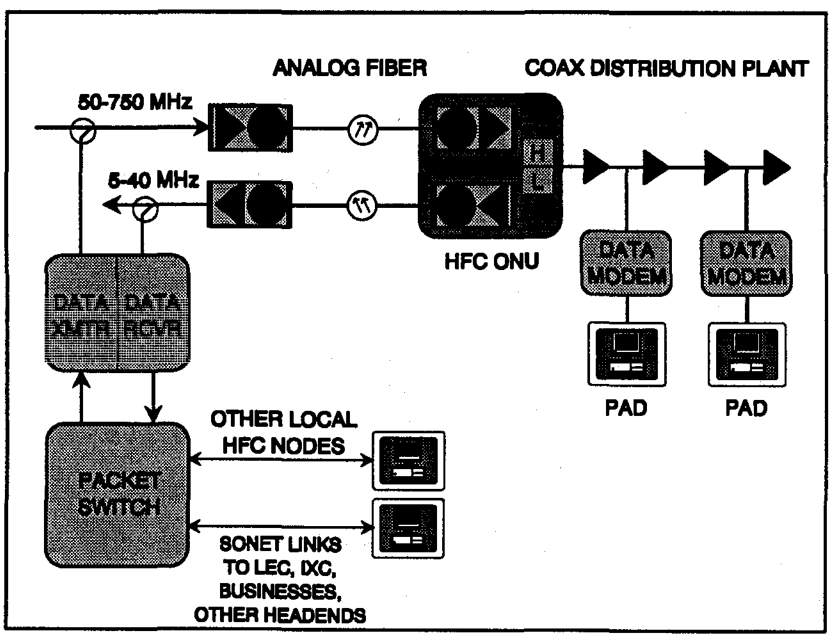


Figure 3 -- HFC Node with Packet Network Overlay

Typical RF frequencies for these paths are used (50-750 MHz downstream; 5-40 MHz upstream). An RF packet transmitter is shown at the headend, where packets to be sent downstream are coupled with other RF signals prior to being fed to the downstream analog laser. An RF packet receiver is also shown at the headend. Upstream RF signals are coupled into this receiver from the analog RF optical receiver in the headend.

Since the packet transmitter and receiver represent different physical paths in the network, each is connected to a separate port on the packet switch. In reality, several headend RF packet transceivers will probably be connected to each node, one for each RF packet channel on the node. Each of these will also be connected to the packet switch. The packet switch shown here is the same as in Figure 2.

Two subscribers are shown connected to this node. At each subscriber's end, an RF data modem serves as the interface between the PAD and the coaxial distribution plant. In all likelihood, this modem will be integral to the PAD itself, and the two together will be considered a subscriber network unit (SNU). A subscriber may have several service specific SNU's in his home, one for each service being taken. One SNU could support telephony, while another might support internet access or compressed digital NTSC television service. Over time, the SNU may actually be a packet gateway to the home, providing several services from a single device. As with other fast-packet networks, the SNU has the responsibility for making sure that errored or lost packets get retransmitted.

If the two subscribers shown in Figure 3 have SNU's that support data connections and they wish to communicate directly via computer, first they must establish a connection. Establishing this connection will consist of defining type of service, gaining access to transport capacity, negotiating bandwidth and quality of service, and setting up the routing path. Once this is accomplished, data packets will flow upstream from one subscriber, enter the packet switch, and then be routed downstream on the same node to the other subscriber. Of course, connections could similarly be made with other terminals anywhere on the network, either on other nodes or nodes served by other headends. Similarly, other services can be supported via the same network.

HFC NODE CONSIDERATIONS

Given the flexibility and transparent transport capability of the HFC architecture, overlaying a basic fast-packet network on the HFC network is largely a matter of installing the appropriate terminal equipment at the headend (switches, file servers, and RF data modems) and customer premises (RF modems, routers, and PAD's). However, deploying advanced, fully integrated packet networks capable of delivering voice, video, and data over the HFC architecture may not be so easy.

The downstream path of the HFC network is well understood and should provide adequate bandwidth and RF transport performance for advanced packet services. The upstream path is less understood, and many technical questions and challenges remain in this direction. The two primary concerns are error performance and transport capacity. Packet networks rely on excellent transmission error performance to help avoid congestion, and advanced integrated packet networks will require significant return spectrum to support the many services envisioned. In the long run, changes to the basic HFC system may be required to ensure adequate error performance and available spectrum in the return path.

Since businesses and homes will use this network, and advanced data services are envisioned which will involve financial transactions or other exchange of confidential information, security is a critical issue. By virtue of the HFC architecture, all downstream packets are transmitted in a broadcast mode over a given node, and all downstream drops have equal access to this RF data channel. This means confidential information from other users, albeit difficult to access in a meaningful form due to the complexity of the RF and packet transport media, will be present in every business or home on the node.

The return path is less troublesome from a security standpoint. Return path signals are considerably isolated from individual drops by the directional couplers used in taps, and due to the reverse tree and branch structure in the return path not all return signals pass every drop. However, steps must be taken to ensure each subscriber's privacy both in the upstream and downstream direction, whether that subscriber is at a business or a home. Encryption and decryption will be required for some services, if not for all. Encryption can take place at the transport level or at the applications level.

Downstream Path

To realize the full benefits packet switched networks have to offer, efficient statistical multiplexing of packets is necessary. In the downstream direction, multiplexing data packets from several users or services onto a single, high-speed RF data channel at the headend is relatively simple and inexpensive to accomplish. All downstream SNU's then simply monitor this data stream for their intended packets and then extract them.

In most cases, the RF data channel and the packets sent down it will be unique to each HFC node, and each RF data channel will be associated with a specific port, and hence routing path, on the packet switch in the headend. As traffic requirements increase, more RF data channels are simply connected to other ports on the packet switch and added to the HFC node via frequency division multiplexing.

On the other hand, increased traffic demand may be dealt with via fiber division multiplexing by subdividing the HFC node into two smaller optical nodes. In this case, a given RF data channel would then support roughly half the number of subscribers as before, cutting by half the traffic demand on the channel. Of course, a new RF data modem and packet switch port would still be required to support the new node generated by the subdivision. The flexibility of the HFC architecture and the packet network itself allows economical and efficient evolution of the network.

The second requirement to realize the benefits of fast-packet networks is that a low bit error rate (BER) must be maintained to keep the network from filling up with retransmitted packets. The downstream carrier-to-noise ratios and transmission performance of the HFC network are such that very high order modulation schemes may be used to achieve excellent spectral efficiency with very little sacrifice in BER performance. If necessary, tradeoffs may be made between BER performance and modulation spectral efficiency. Of course, forward error correction (FEC) may be used in the RF transport path to improve BER performance for any modulation scheme, but this requires additional overhead bits in the data, which in turn reduces spectral efficiency. In many cases, good BER performance may be achieved without requiring FEC at all in the downstream path.

Upstream Path

The upstream path in an HFC system is subject to noise funneling brought about by the noise summation of all the return legs on a node. This noise is a potential problem for maintaining acceptable bit error rates in the return path. Fortunately, this effect is minimized by modern fiber architectures, which allow the use of smaller node sizes, and hence less noise to be summed. The return path can also be susceptible to ingress (short-wave signals, CB and ham transmitters, spurs and LO products from consumer devices connected in the home) and impulse noise (appliances, lightning strikes, etc.). Studies have shown that most of these secondary problems occur not in the hard-line portion of the plant, but in the drop or customer premises wiring. These parts of the plant must be brought under control, either by hardening the drop or using filters to limit unwanted signals from entering the plant.

A fast-packet network achieves considerable efficiency by unloading most of the lower level transport error processing and its overhead. Virtually error-free transmission in optical networks has made this possible. When errors are rare, discarding an occasional errored packet and retransmitting it does not consume much bandwidth. But if errors are frequent, retransmitting packets can consume considerable transport capacity. A single bit error requires an entire packet to be retransmitted. HFC networks must provide excellent BER performance if fast-packet networks are to be operated over them.

If necessary, forward error correction (FEC) may be used at the physical layer to reduce the loss of transport capacity due to errored transmissions. However, FEC also incurs a bandwidth penalty as extra bits must be transmitted to accomplish FEC at the receiving terminal. A trade-off exists between bandwidth consumed by retransmitted packets vs. bandwidth consumed by additional FEC bits. The choice of modulation schemes will also make a difference in BER performance. Lower order modulation methods are more robust when faced with poor transmission channels, but these methods are also less spectrally efficient.

One approach to overcoming noise and interference problems is simply to increase transmitter power to a level where good BER performance is achieved. Within reasonable limits, this is perhaps the most economical and spectrally efficient approach, but requires return amplifiers and lasers with very good linearity to withstand the higher signal levels. In any case, the return plant will have to be designed and balanced with as much care and attention paid to performance as with the downstream path. Careful analysis must be done in the system design to optimize BER performance with the least impact on spectral efficiency. Higher overall throughput is the primary goal.

As also pointed out, packet networks derive increased throughput by statistically multiplexing several users onto a single transmission channel. Multiplexing packets efficiently from separate SNU's on the same channel in the upstream direction on an HFC fiber node is not easy to accomplish. Normal multiplexers operate by taking in several data streams and combining them into a single high speed stream. In the upstream direction on an HFC network, each SNU functions independently, but typically transmits on a common return channel with other SNU's. No single device performs the multiplexing function, and in essence all the SNU's on the RF return channel comprise a distributed multiplexer of sorts.

Statistical multiplexing can only be achieved in the HFC return path by providing a mechanism for controlling each SNU's access to the upstream RF channel to prevent collisions which will occur if two or more transmitters become active at the same time. This mechanism must also minimize any overhead penalty imposed by this coordination and minimize any periods of inaccessibility to the channel. Otherwise, channel throughput will be limited.

Unfortunately, due to the HFC architecture, an SNU cannot monitor or coordinate the return path transmission activity of other SNU's on its RF channel. One simple solution is to provide an independent return RF channel for each SNU (frequency division multiplexing). Unfortunately, this approach is the same as the dedicated channel approach used by switched circuit networks, and none of the benefits of statistical multiplexing are achieved. Dedicated channels are inefficient from a spectral perspective and costly from a hardware perspective. Another mechanism must be used for statistically multiplexing several SNU's onto individual RF channels to realize transmission efficiency and bandwidth on demand.

Another approach is to use CDMA transmission (Code Division Multiple Access, a form of spread spectrum), which allows several users' transmission spectrums to overlap simultaneously. CDMA offers the benefit of noise immunity when high processing gains are used, a desirable characteristic for return path transmission. But CDMA has limitations for this application, especially if any services or users require large data throughput, which is likely to be the case for advanced services.

The return path's available spectrum is too small to allow effective spreading, and hence processing gain, for anything but the narrowest of return data channels. Even so, the entire return path would likely have top be used for CDMA, which would rule out sharing this spectrum with other types of services. Any attempt to allocate a smaller subset of the return spectrum for CDMA signals alone will only make good processing gains more difficult to achieve. Finally, CDMA is still a channel oriented transmission method, and another means would have to be used along with CDMA for statistically multiplexing several SNU's onto individual CDMA channels.

There are four fundamental ways to coordinate multiple SNU transmission access on a single upstream RF channel: time division multiplexing, polling, token passing, and collision detection. Combinations of these methods may be used as well.

With time division multiplexing, each SNU upstream transmitter has ownership of a fixed time-slot in which to transmit. Reference timing and slot assignment are provided by the downstream data path for the service link. The subscriber buffers data until his time-slot becomes available, then bursts this data back to the headend within the allocated time. Non-transmission guardbands must be provided before and after each time slot to ensure no two transmissions overlap, and these guardbands consume some channel transport capacity. The primary disadvantage to this approach, however, is that the subscriber gets a time-slot even if he has no data to transmit, and thus bandwidth may go unused.

The benefits of statistical multiplexing are not possible under this scheme unless dynamic access to all time slots is provided to all SNU's. This would allow slots to be assigned only to SNU's that need transport capacity, and multiple slots could be assigned to SNU's needing extra capacity. This implies that a central, intelligent channel manager is coordinating slot assignment and that this manager has a means of determining each SNU's requirements. This can only be accomplished through some form of polling, which entails additional transmission overhead as the time slot manager queries SNU's for their needs and assigns resources accordingly.

The CATV industry has used device polling for years to coordinate upstream set top converter communications. An upstream channel manager, or host terminal, located in the headend informs each SNU in turn that transmission access has been granted to the upstream RF channel. If the remote SNU has no data to transmit, it informs the host terminal of this fact with a very short message or by failing to transmit within a pre-defined time span. If the SNU has data to send, it then seizes the transmission channel until all its data has been sent or until a defined limit is reached, thus preventing any single SNU from hogging bandwidth in the return path. Polling itself may require a significant amount of transmission path overhead, depending on the complexity of the polling mechanisms. Due to the nature of polling, transport capacity is consumed by overhead in both the downstream and upstream paths.

Token passing is another means of coordinating individual SNU access to the return transmission channel. In this case, only the current holder of a token (a software authorization to transmit) is allowed to transmit in the return path. Only one token is allowed to be shared among all the SNU's using a given return path channel, and care must be taken to ensure channel recovery if the token is somehow lost. Typically this token is passed from terminal to terminal on a bus, usually in a ring architecture. However, direct communication between SNU's is not possible in an HFC network, so any token passing has to take place through the packet switch in the headend. This implies that the token must either be sent to a channel manager for forwarding to the next SNU or that each SNU has knowledge of the other SNU's on the channel and knows which SNU is to receive the token next. This type of system is complex to implement on an HFC network and will create considerable delays as the return path remains idle while the token is passed. Since the token must be passed up to and down from the headend packet switch, polling may be the more attractive alternative.

Collision detection is yet another mechanism of regulating multiple access to a channel (ethernet is a good example of collision detection). In this case, each terminal on a bus monitors the channel for inactivity. Once silence is detected, any terminal that wishes to transmit may do so. The channel is then monitored for transmission errors to make sure that only one terminal seized the channel. If more than one terminal did transmit, all terminals cease transmission. Each terminal then generates a short, but random time-out period in which it cannot transmit. Usually, one terminal will emerge from its time-out period before the others, and this terminal will then attempt to seize the channel again, and the process repeats until a single terminal gains transmission access. Once a terminal gains this access, it transmits until it has no more data to send or until a defined limit for channel access is reached.

Collision detection is not very efficient for HFC networks. Due to the HFC architecture, collisions cannot be directly detected on the return path by an SNU. The return path can only be monitored for collisions at the headend. This means that collision detection must be mediated through a headend channel manager and communicated back downstream, thus adding mediation overhead and delays to transport access. Because collisions are unpredictable, access delays may be highly variable, and this is not desirable for isochronous applications such as voice or video. The complexity of a collision detection system for an HFC network is also not desirable. Again, some form of polling appears to be a more attractive alternative.

No simple solution exists for multiplexing several remote terminals onto the same return channel with minimal system delay or overhead. But with careful design, channel overhead may be reduced, and delays may be minimized and made acceptable for service transport.

Bandwidth capacity is the final and perhaps greatest concern with the HFC return path. Typical HFC networks have been built with 5-30 MHz return paths. More recently, networks have been built using a 5-40 MHz return. In either case, this provides little spectrum for advanced services, and this may be even more limited since recent studies have shown that the region below 10-15 MHz may be difficult to use due to noise and ingress. In all likelihood, this lack of bandwidth is the greatest barrier to deploying full service fast-packet networks on HFC systems. Five solutions to this bottleneck are readily apparent, but other solutions exist, as well.

First, mid-split systems may be deployed rather than the current sub-split design. As an example, the upstream band might occupy 5-150 MHz, and the downstream band 200-750 MHz. A 50 MHz guardband would be provided between these. Of course, this takes away from downstream capacity, but with the use of compressed digital NTSC television signals instead of analog NTSC, considerable downstream bandwidth can be freed up.

Second, the coaxial legs coming into the optical network unit at each node may have their individual sub-split return paths kept separate. These would then individually be block converted to a unique frequency and then combined for transport over a single fiber back to the headend. Of course, one of these coaxial legs need not be block converted since its return signals may be transmitted on their original frequencies. For example, an optical node with four coaxial arms radiating from it may block convert one 5-40 MHz return leg to 65-100 MHz, a second leg to 125-160 MHz, and a third leg to 185-220 MHz. One leg would continue to occupy 5-40 MHz. When combined for transmission back to the headend over a single laser, a 25 MHz guardband is provided between each frequency grouping for filtering purposes.

Third, the return path spectrum can be placed above the downstream path spectrum, perhaps occupying 850-1000 MHz. At these frequencies, the necessary diplex filtering between the forward and return paths is more difficult to accomplish, so additional guardband between the two will be required. On the other hand, the noise and ingress performance at these frequencies is much better than in the 5-40 MHz range.

Fourth, a completely separate return path may be added to the coaxial portion of the HFC network (the return path is already separate on the fiber side of the node). In this case, a dual cable approach is taken, one cable for the forward spectrum, one for the reverse. This eliminates the need for diplex filters and guardbands between the forward and reverse paths and allows symmetrical or asymmetrical spectral capacity to be provided. Adding bandwidth to either direction is also easier to accomplish if network demands require it. While attractive from a technical perspective, this approach may not be cost effective.

Fifth, an optical node may be subdivided into two or more optical nodes when either the forward or return path runs out of spectrum. Of course, each new node will still operate over the same frequencies as the original node, but now fewer subscribers will be sharing this spectrum, thus reducing the demand for capacity.

SUMMARY AND CONCLUSIONS

Fast-packet technologies are capable of efficiently delivering advanced digital services, including voice, video, and data, over an integrated transport network. This is made possible by reducing the transport overhead normally associated with error processing and by statistically multiplexing several services and users' data over single transport paths within the network. Such a network can most cost-effectively be implemented using HFC architectures. To support a fast-packet network capable of delivering voice, video, and data, an HFC system must provide significant bandwidth and excellent bit error rate performance, both of which are possible.

REFERENCES

Black, Uyless, *Data Link Protocols*, PTR Prentice Hall, Englewood Cliffs, NJ, 1993.

Black, Uyless, *Emerging Communications Technologies*, PTR Prentice Hall, Englewood Cliffs, NJ, 1994.

Kessler, Gary C., *ISDN: Concepts, Facilities, and Services*, 2nd Edition, McGraw-Hill, Inc., New York, NY, 1993.

Pecar, Joseph A., Roger J. O'Connor, and David A. Garbin, *The MGraw-Hill Telecommunications Factbook*, McGraw-Hill, Inc., New York, NY, 1993.

Spohn, Darren L., *Data Network Design*, McGraw-Hill, Inc., New York, NY, 1993.

Stallings, William, *ISDN and Broadband ISDN with Frame Relay and ATM*, 3rd Edition, Prentice Hall, Englewood Cliffs, NJ, 1995.

BACK TO 1310 nm -- WITH SEMICONDUCTOR OPTICAL AMPLIFIERS

Venkatesh G. Mutalik
Fiberoptics Engineer
Philips Broadband Networks Inc.

Luuk F. Tiemeijer
Research Scientist
Philips Optoelectronics Centre

Abstract

Semiconductor optical amplifiers (SOAs) provide a cost-effective optical amplification scheme at 1310 nm. Although they have already been used in enhancing the performance of high-bit-rate digital systems, their use in CATV systems has been limited so far because of their distortion performance. However, recent research suggests that this can be overcome by using a new technique called 'gain-clamping'. In this paper, we will review the performance of presently available SOAs and discuss their performance enhancements using gain-clamping. We will then indicate potential applications of SOAs along with some applicable system tradeoffs with special emphasis on analog networks.

INTRODUCTION

Optical amplification schemes at 1550 nm using Erbium Doped Fiber Amplifiers (EDFAs) have demonstrated their superiority over conventional optical repeaters in terms of noise and distortion performance. They are already being used in connecting headends and in other analog supertrunk applications. However, their use in dense CATV systems has been rather slow because the externally modulated transmitter that is required to combat dispersive effects at the 1550 nm low loss window is very expensive. Semiconductor optical amplifiers (SOAs), on the other hand provide amplification at the low dispersion, 1310 nm optical window. Furthermore, the electronic control of these amplifiers is similar to that of present-day DFB lasers, so they help preserve present system architectures. SOAs are now available as a fully packaged product and have already been used to enhance the performance of digital systems [1, 2, 3]. Their use in CATV applications has been limited because their optical gain varies with instantaneous changes in input signal level. This generates distortions that cannot be tolerated by conventional AM-VSB signals.

Recent research suggests that this drawback can be overcome using a new technique called *'gain-clamping'*[4]. For gain-clamped SOAs, output powers in the range of 10-15 dBm, and IMD products approaching that of linear DFB lasers, are predicted to be possible. This paper reviews the performance of presently available SOAs and discusses the prospects for linearizing their performance using the gain-clamping technique. Finally, we will review their potential applications in conventional AM-VSB systems with respect to optical budget, cost and flexibility.

SEMICONDUCTOR OPTICAL AMPLIFIERS

Semiconductor optical amplifiers are now available as a fully packaged and pigtailed product, with a size and power consumption comparable to that of the well-known 1480 nm pump lasers. The only difference is an additional input-optical fiber. The semiconductor optical amplifier chip inside this package is based on a high-performance multiple quantum well semiconductor laser structure of proven design. The chip employs a 10° angle stripe and high-quality AR coatings to suppress cavity effects and to obtain the optimum performance. Their fabrication technology is identical to that of semiconductor lasers.

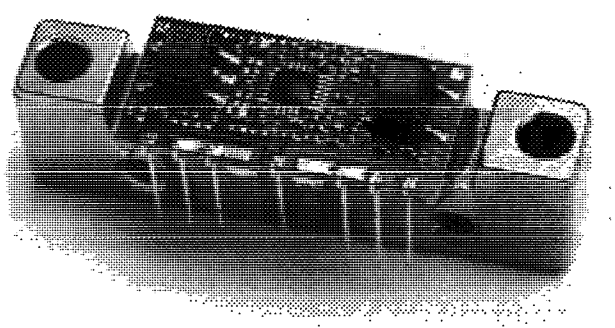


Fig. 1. Photograph of a fully packaged Semiconductor Optical Amplifier

Unfortunately, so far this optical amplifier has been considered unsuitable for use in cable TV fiber optic transmission, mainly because of its nonlinear transfer characteristics. The dashed line in Fig. 2 shows the characteristic output power versus input power of a conventional optical amplifier. The decrease in output power for high levels of input power is caused by gain saturation in the active medium.

Although an EDFA has a transfer function similar to the one shown, the gain mechanism in the EDFA is very slow (about 10 ms) compared to the typical modulation frequencies applied in the CATV transmission. Hence, the above transfer function applies only to average signal levels rather than to instantaneous changes in input signal levels. Therefore, the EDFA can amplify the signals in the CATV frequency band undistorted.

The major obstacle towards using semiconductor optical amplifiers for the same type of application is the fact that their gain mechanism is very fast (about 0.2 ns). This means that even instantaneous changes in input signal levels will modulate their gain. This gain-modulation eventually results in the generation of generally unacceptable levels of second and third-order distortion products.

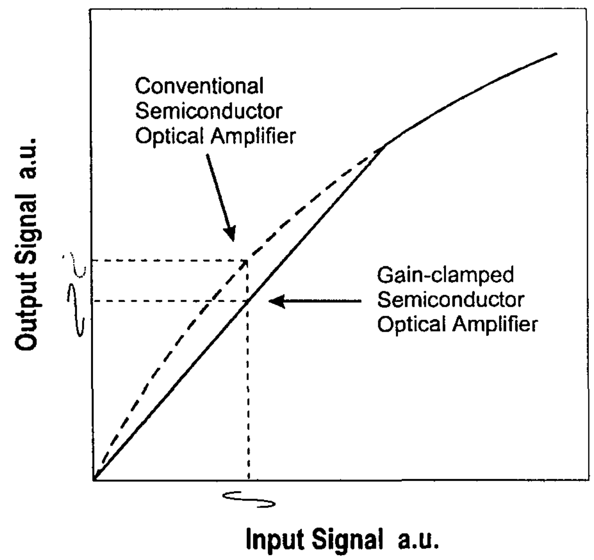


Fig. 2. Diagram illustrating the transfer characteristics of conventional and gain-clamped SOAs

Recently, we demonstrated that *'gain-clamping'* is an effective technique to overcome the problem of gain modulation in the SOA. With this technique, laser action is initiated at an out-of-band wavelength of 1290 nm by a weak 1 - 5 % reflective Bragg grating, which is monolithically integrated with the amplifier chip. This grating is transparent for the 1310 nm CATV signal. Lasing at 1290 nm clamps the carrier density to the laser threshold level, and thus stabilizes the optical gain efficiently. This is illustrated by the solid curve in Fig. 2.

When this gain-clamped amplifier is operated well above the lasing threshold for

1290 nm emission and a fiber Bragg grating filter is used to block the 1290 nm emission, a highly linear 10 to 15 dB fiber-to-fiber gain around 1310 nm is obtained.

Results based on preliminary analysis indicate that after design optimization second-order IMD products in the range of -69 dBc are possible. This is for output powers in the range of +10 dBm using a two-tone modulation frequency pattern of 125 MHz and 226 MHz with a modulation depth of 35%. For the same system, third-order IMD products in the range of -100 dBc are possible. These numbers are better than the corresponding numbers for a DM DFB laser.

SOA PERFORMANCE EVALUATION

Digital

SOAs have already been used to enhance digital performance at 1310 nm. In one experiment, a 1310 nm polarization insensitive semiconductor optical pre-amplifier enabled a power budget of close to 40 dB, at a bitrate of 10 GB/s. This is sufficient to bridge 102 Km of standard singlemode fiber. In another experiment, a cascade of four polarization insensitive amplifiers transmitted a 20 GB/s soliton signal over 200 Km of standard singlemode fiber.

Digital systems are very forgiving of noise and distortions and can tolerate gain-modulation effects inherent in SOAs. Also, the average power levels for these systems is quite low, where the SOA distortions are considerably weaker. These examples prove that SOAs work well in digital systems.

Analog AM-VSB

For conventional analog systems the power at any optical node is roughly 0 dBm. This requirement puts a premium on the output power of the SOA. Presented below is the Pin versus Pout performance of an SOA for varying bias currents.

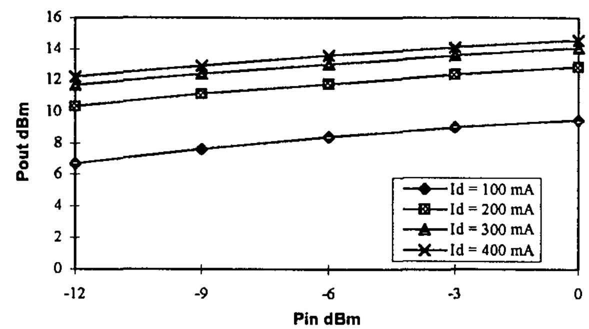


Fig. 3. Pin Vs. Pout graph for an SOA for different bias currents.

The SOA output power saturates at higher bias currents. This graph demonstrates that SOAs can provide enough power gain for CATV level inputs.

However, the requirements on distortion and noise performance for conventional CATV systems are far more aggressive. Shown below is a test set-up for SOA testing.

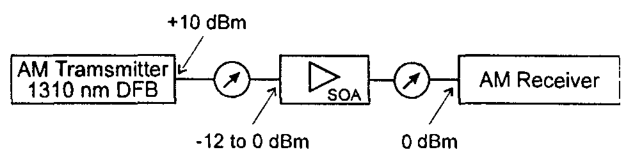


Fig. 4. Test set-up for measuring AM performance of an SOA

Based on tests conducted on the SOAs, using the above set-up, and a 40 channel loading (300- 550 MHz), typical CNR of 40 dB and CTB of 40 dB were measured with presently available SOAs. This clearly indicates that the noise and distortion performance of SOAs in the present form is unsuitable for conventional CATV systems. The gain-clamped version should help in this case.

FM Transmission

FM transmission on fiber closely imitates digital transmission. Shown below is the test set-up for a 16 channel FM test. The output of the transmitter is a set of 16 unmodulated CW carriers which are incident on the FM receiver at an optical power of -10 dBm, as consistent with FM requirements.

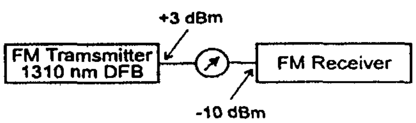


Fig. 5a. FM test set-up

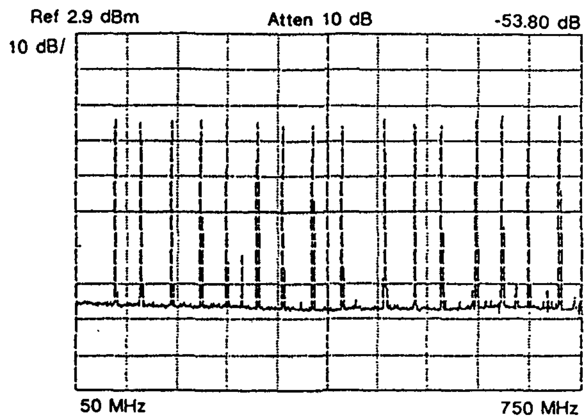


Fig. 5b. FM response with -10 dBm input to the receiver (RBW 300 KHz, VBW 3 KHz)

Now, the SOA is inserted between the transmitter and the receiver. As is clear from the diagram below, the SOA has an input of -10 dBm and gives an output of +10 dBm. Then, as above, -10 dBm is incident on the receiver.

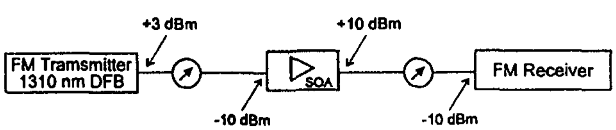


Fig. 6a. Test set-up for measuring FM SOA performance

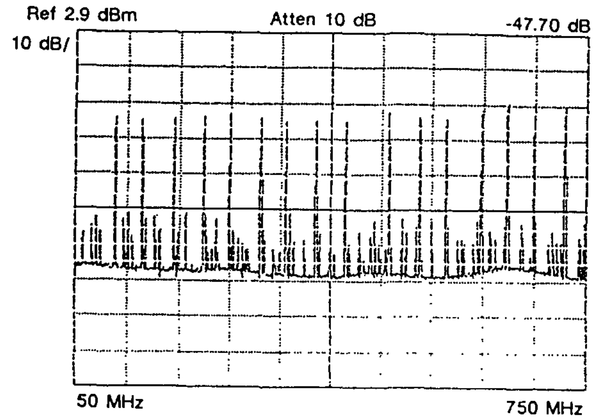


Fig. 6b. Response with SOA with -10 dBm at the FM receiver (RBW 300 KHz, VBW 3 KHz)

A comparison of the two responses indicates that the noise floor has degraded by 6 dB, allowing a CNR of 47 dB and that the distortions are limited to about -35 dBc. The addition of the SOA has also resulted in an enhancement of the link budget by 20 dB, roughly corresponding to 60 km of fiber length.

This is just the preliminary result, and we expect that with optimization of the SOA parameters and optical filtering of the amplifier spontaneous emission noise, cascading of several SOAs while maintaining low noise and distortions is possible.

Return Systems

SOA testing for a return system was carried out using 20 - 80 channels of QPSK at T1 rates in a fiber. The results involving SOAs are inconclusive. This is because the laser used to transmit the QPSK channel is an FP laser with no isolator protecting it from back-reflections. This is not a major concern for normal systems. However, when an SOA is inserted in the signal path, the amplified spontaneous emission (ASE) of the laser along with the Rayleigh back-scatter feed-back into the FP laser, thus increasing the laser noise and affecting its performance. Further work needs to be done to identify this problem precisely and to propose a cost effective solution.

DISCUSSION

Development of viable optical amplifiers at the 1310 nm window is a great opportunity for the CATV industry to protect its investment in presently deployed 1310 nm systems [5]. Clearly SOAs in the present form are not suitable for use in analog CATV networks, other than for FM systems. However, improvements in SOA technology will in all probability lead to better devices that can handle analog CATV signals.

To be fair, we must mention two other competing amplification schemes that exist at 1310 nm. They are the Praseodymium Doped Fiber Amplifiers (PDFAs) [6] and the Raman Fiber Amplifiers [7]. Both these amplification alternatives are bulky and expensive. Since these amplifiers use multiple high-power pumps and special types of doped fibers, their reliability with respect to continuous use has not yet been confirmed. The manufacturing process of building SOAs, on the other hand, is the same as that for DFB lasers. This mature technology and the associated high reliability, at prices competitive to DFBs, make the future of SOAs secure. With this in mind, we need to compare SOAs to conventional repeaters and to 1550 nm Externally Modulated systems.

The superiority of a fully functional gain-clamped SOA over a conventional optical-RF-optical repeater is considerable. SOAs will help in eliminating cost and complexity associated with maintaining a separate receiver and transmitter at each location. This would reduce congestion in dense headend locations. SOAs would also be able to provide optical budgets for bridging medium distances and for multiple splitting options. And they do all of this while maintaining the present system architectures at competitive prices.

In supertrunk applications, especially in providing headend-to-headend connection over long distances (in the range of 80 - 100 km), EDFAs have some obvious advantages. Fiber attenuation at 1550 nm is almost half that of 1310 nm, while the passive split losses are the same at the two wavelengths. In all likelihood, EDFAs will continue to have an upper hand in this sector. However, these systems involve at least one externally modulated 1550 nm transmitter and a combination of EDFAs depending on the specific application.

The design of good external modulators to compensate for Stimulated Brillion Scattering (SBS), Double Rayleigh Backscatter (DRBS) and fiber dispersion is expensive and non-trivial. Maintaining low noise figure (NF) in EDFAs with high reliability is non-trivial as well. The resulting system, while it works well when properly designed, is expensive and complex.

In high-split applications, an additional amplifier is provided at the last headend. The output of the amplifier is split into multiple fiber lines and distributed directly to the hybrid fiber coax network. Here, both amplification schemes have equal loss budgets since splitter losses are the same at both wavelengths.

Hence, if the system configuration calls for short hops followed by multiple splits, SOAs have the distinct advantage. This is because total compatibility of 1310 nm wavelength is maintained. The full potential of narrow casting or broadcasting can now be realized, while preserving all existing 1310 nm networks.

CONCLUSIONS

In this paper, presently available semiconductor optical amplifier technology was reviewed. The basic architecture of these amplifiers including gain-clamping of SOAs for linearizing their performance was briefly explained. Performance of presently available SOAs in different transmission schemes were then discussed in detail. Finally, potential applications of 1310 nm SOAs and how they compare to currently used solution were indicated. The chief advantages of 1310 nm SOA technology are: flexibility, scalability, dispersion immunity, compatibility with existing systems, and low cost. Now, for the first time, both power budget and dispersion are non-factors in transmitting optical signals.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the help of Eric Schnettler in testing SOAs. One of the authors (VM) would like to thank Kerry LaViolette for helpful discussions on systems aspects of SOAs.

REFERENCES

1. L. F. Tiemeijer, "High performance MQW laser amplifiers for transmission systems operating in the 1310 nm window at bit-rates of 10 Gbit/s and beyond", Proc. of ECOC-95, paper Tu.B.2.1, 1995

2. H. de Waardt et al., "2 x 10 Gbit/s WDM 1310 nm optical transmission over 63.5 km standard single-mode fiber using optical preamplifiers", IEEE Photonics Technology Letters, 104-107, Jan 1995

3. L.F. Tiemeijer, "33 dB fiber-to-fiber gain +13 dBm fiber saturation power polarization independent 1310 nm MQW laser amplifiers", PD1-2, Optical Amplifiers and Their Applications, June 1995

4. L.F. Tiemeijer et al., "Reduced intermodulation distortion in 1300 nm gain clamped MQW laser amplifiers", IEEE Photonics Technology Letters, 284-286, March 1995

5. L. F. Tiemeijer and V. G. Mutalik, "Applications of semiconductor optical amplifiers in CATV networks", SCTE Emerging Technologies conference, Jan. 1996

6. Y. Ohishi, et al., "Recent progress in 1.3 um fiber amplifiers", TuG1, OFC, Feb/Mar 1996

7. A. J. Stenz, et al., "Raman ring amplifier at 1.3 um with analog-grade noise performance and an output power of 23 dBm", PD17, OFC, Feb/Mar 1996

Bottleneck of Data over Cable

David M. Fellows
Continental Cablevision

Abstract

Advanced modulation techniques, combined with protocols such as the MPEG 2 transport layer, have the ability to turn cable systems into broadband digital pipes to the home. However, a growing issue will be the bottleneck of data in the reverse direction, especially as the number of services supported over cable networks grows beyond Impulse Pay Per View and traditional status monitoring. Not only does the reverse path offer at best one-twentieth the bandwidth of the forward direction, but the noise and other impairments limit the aggressiveness of the modulation possible. Even though most applications still will use more bandwidth to the home than from it, the reverse path represents the most severe bottleneck in future systems. Only by careful utilization of this resource, and agreement on common protocols across services, will a cable network support all the services that will be required of it.

CABLE ARCHITECTURES

Fiber-to-the Serving Area

For several years, cable systems have been using fiber to reduce amplifier cascades and increase bandwidth. The use of fiber to small pockets of homes has the added benefit of increasing reliability, improving the reverse path, and creating a star architecture out of the headend, so bandwidth can be reused from pocket to pocket. This is essential for different voice and data calls to go to different pockets of homes. These pockets are at most 2000 homes-passed, and more often are 500 homes.

Forward Path

Most cable systems are being rebuild to either 550 MHz or 750 MHz bandwidths, with roughly 500 MHz or 700 MHz dedicated to the forward direction and 25 to 35 MHz dedicated to the reverse direction. These systems have carrier-to-noise figures in the high 40's, and tend to have diplex filters (which contribute group delay) only at the very bottom of the frequency range (where an analog channel 2 sits). The MPEG 2 digital video systems discussed in this session have concentrated on 64 QAM or 256 QAM modulation. Taking 30 Mb/s in every 6 MHz, a system with 700 MHz of forward downstream capacity could, in theory, carry 3.5 Gb/s of data on the cable! The major impairment to digital transmission is micro-reflections due to impedance mismatch and unterminated connections.

Historically, multiple services in the forward direction have been handled primarily by Frequency Division Multiplexing (FDM) - every channel is given a different frequency. Within one MPEG frequency slot, multiple channels are handled by Time Division Multiplexing (TDM). Because all services originate from the same place (e.g. the headend), FDM or TDM is possible, and results in efficient packing of data.

Reverse Path

The reverse direction, however, is everything the forward path is not. It has at most 35 MHz of bandwidth, and suffers from the noise funnel effect, group delay from aggressive diplex filters (in 40 MHz split systems), and lots of high-power interferers (Ham, International AM, and CB radio). Most of the reverse systems being proposed achieve about a 1 bit/second/Hertz throughput (the modulation is higher, but overhead and other inefficiencies discussed below reduce the usable bit rate), meaning that 35 MHz yields 35 Mb/s - one one hundredth of the forward amount.

Different services in the reverse direction start out using FDM; for example, IPPV uses a different frequency than status monitoring. For a given service, within a single frequency, TDM can also be used. But in the reverse direction, data is originated from different sources, and unpredictably so for many services, so a multiple access protocol must be used, further reducing efficiency. In any event, guard bands or guard times must be inserted between channels to separate them from each other (more in the next section).

DIFFERING SERVICE REQUIREMENTS

Services

Cable systems will be called upon to carry many differing types of services in the near future. Going from the existing to the very near future, the list for the reverse direction includes Impulse Pay-Per-View (IPPV), status monitoring, high speed data return, telephony return, interactive set-top real-time return, energy management/UCS/telemetry services, PCS return (rads/rasp), and so forth. Even though many of these applications do not require much bandwidth, the fact that they "talk" different languages means that they need different frequency slots, with the waste of guard bands between the channel frequencies.

Multiple Access

To make matters worse, within each service type, there will be multiple people trying to simultaneously access or use the service. If there were enough channels to go around, then the multiple access is set-up at installation time - each user is assigned their own channel. But this is too inefficient for long - must use concentration. that is, there will be more users than channels or bandwidth, and a channel or bandwidth must be assigned upon demand

Multiple Access Protocols

Many customers vying for and sharing a limited number of channels is a problem attacked for satellite service between Hawaii and the mainland. In a protocol called "Aloha", anyone wanting service just transmitted. A variant of this was slotted Aloha, where the beginning point of customers was synchronized (and resulted in a factor of two increase in loading efficiency). To use ethernet as an example, the transmitter first looks at the channel to make sure no one else is using it (carrier-sense multiple access - CSMA), then transmits its data, making sure that no one else transmits on top of it (because of delays in most channels, the fact that a channel is clear at the start of transmission does not mean that another transmitter has not also seen a clear channel and started transmitting - but that transmission signal has not reached this location yet - called collision detection - CD) Put together, this is called CSMA-CD, and if a collision is detected, both transmitters back off a random amount of time, then try again. Further refinements include central reservation or distributed reservation via signals called tokens. Today, multiple access protocols fall into three categories: Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA), and Code Division Multiple Access (CDMA). These protocols have been further enhanced by reservation or assignment techniques, where channels are assigned centrally after contention on a reservation channel.

With FDMA, channels are assigned different carriers, and users contend for these channels. In TDMA, a single frequency carries a data stream that is logically broken into sub-channels or timeslots, and users vie for these timeslots. CDMA uses the spread spectrum technique of orthogonal codes, and each code represents a different channel, for which users contend. In all of these, a separate reservation channel may be used to centrally assign channels (whether frequencies, time slots, or codes).

Between many services and many users of each service, there is the real possibility that there is not enough reverse bandwidth for every subscriber and every application.

SYSTEM LEVEL SOLUTIONS

More Reverse Bandwidth

Standard systems in the USA are built with sub-split reverse, which used to be 5 MHz to 30 MHz and now typically is 5 MHz to 40 MHz, a 40 percent increase. Since the bandwidth below 10 MHz is tough to use, this results in an even greater increase. Many other splits are available, and are used in foreign countries. In the USA, the must-carry provisions of the 1992 Cable Act makes it very difficult to carry Channel 2 anywhere other than 54 MHz, so that wider bandwidth at the low end is not available. In Orlando, amplifiers were build that allowed more reverse above the forward spectrum. This proved that more reverse bandwidth is technically available, but this is not thought to be an easy answer.

Asymmetric Services

Another answer is not to waste what bandwidth is available. If a service is asymmetric, then the channels in the downstream and upstream direction should be different, and sized for the service. This is especially true if the demand is for more downstream than upstream bandwidth. The high speed data service is the most obvious example where a smaller (although adequate) upstream channel can save spectrum.

Common Protocols

In the vein of not wasting bandwidth is the concept of having multiple services use the same protocol. In Australia, Optus Vision is using the overhead channel in their telephony product to carry the telemetry information for the utility communication service (load shedding, meter reading, etc.). In the future, use of Asynchronous Transfer Mode (ATM) or other cell-based protocol may allow constant and variable bit rate services, connection and connectionless services, and different bit rates to all use a common language.

A simple first step toward this goal is a MIB developed by Time Warner which allows a management system to move services around in frequency. This means that a system can evolve to new protocols over time by shifting existing services around.

Smaller Nodes

When the problem is too many users (subscribers) of a service, and more bandwidth is not available, then fiber nodes can be split into smaller nodes. This splits the number of users into smaller groups, and frees up capacity. Most systems have installed extra fibers to allow node splitting.

Since the bottlenecks in the reverse and forward direction are different, some approaches treat the two paths differently. For example, a fiber node could have one forward path and four bridger outputs. Each output port could have a different reverse path - either four return lasers, or the four reverse paths could be stacked in frequency and sent back on one fiber.

Cheating

It should be noted that not all services that look interactive in fact are. and many do not use any return bandwidth. Examples include the SEGA channel or Electronic Program Guides, both of which have an interactive feel to the customer, but do not require return channels. Another example is the launch of Enhanced Pay-Per-View in place of a Video-On-Demand - not the same look and feel, but much less capacity-hungry. Obviously, interactive services demanded by subscribers must be delivered, but there are fall-back services that may achieve 80 percent of the goal.

Conclusion

While downstream data capacity looks very promising, there are issues with the upstream direction, which could result in a data bottleneck. Multiple access protocols have been developed which allow many users to share limited capacity and nodes can be split to free up capacity. But, for all the services envisioned on future cable systems, protocols that address multiple services must be developed.

Building Resilient Transmission Schemes Over HFC Networks for Telephony

Bouchung Lin, Ramin Borazjani, Scott Hardin

Scientific-Atlanta, Inc.

P.O. Box 6850, ATL 52-D

Norcross, GA 30091-6850

ABSTRACT

Past studies have quantified the behavior of interference and ingress in the reverse path (e.g., 5 - 40 MHz) of Hybrid Fiber Coaxial cable (HFC) plants. Unfortunately, the speculation of a hostile environment has been confirmed. The studies show the detrimental impact of ingress and impulse noise on the transmission quality, and confirm its unpredictable nature.

Preventive measures can be applied to improve cable plants' quality for digital services, but the real challenge is to design a network architecture that survives adverse conditions. Achieving high signal quality is even more important when telephony service is added because it demands high system reliability and fault tolerance.

This paper focuses on the selection of various transmission schemes for this potentially hostile environment. It shows the advantages FDMA (Frequency Division Multiple Access) has over the wideband burst mode TDMA (Time Division Multiple Access). The requirements and mechanism for providing reliable telephone service will be discussed and analyzed. A reliable RF channel protection algorithm and family of channel hopping algorithms that achieve acceptable telephone service will be discussed. System capacity can be estimated given a desired service quality as measured by blocking probability.

INTRODUCTION

As HFC network operators choose to deploy two-way digital services (e.g., telephone service) over HFC, they may consider upgrades to their networks. The extent of these upgrades will be dictated by the digital service equipment's resiliency to ingress. Some vendors' equipment may require a network free of ingress. This may require installation of passive filters at each customer's premises. Other vendors' equipment may employ schemes that readily adapt to ingress that would discount the need for filters at each subscriber's premises. One important factor that affects telephony equipment's ability to adapt to increase in the upstream path is its multiple access method.

Selecting a multiple-access method for upstream resource allocation (a channel between the subscriber and headend equipment) is a most controversial and important decision in building a cable telephony system. Two of the most popular approaches include FDMA and TDMA. FDMA divides the available frequency band into a group of small channels, each dedicated to a telephone line. TDMA has a single carrier that supports several telephone line (typically 24 to 32 lines), with each line assigned a different time slot. FDMA techniques are often used with TDMA systems to juxtapose several TDMA carriers on the same upstream path. The two methods are shown in Figure 1.

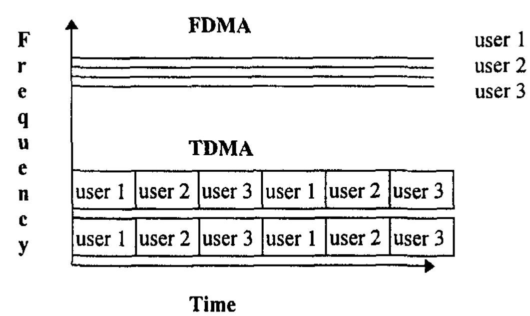


Figure 1: FDMA versus TDMA

FDMA is hypothesized to be more resilient to ingress than TDMA because it can more efficiently be isolated from the contaminated band of spectrum. The first part of this paper quantifies this advantage FDMA systems have over TDMA systems.

In order to help guarantee wireline telephony grade of voice quality, the performance of each upstream channel is monitored by the headend equipment. The system capacity is measured in terms of the number of channels (DS0s) for a given RF spectrum band. A set of channel monitoring algorithms are described in the second part of this paper. A channel hop will be triggered once a corrupted channel is identified. The hopping operation must be brief, and ideally transparent to the end user. Each channel hopping strategy will uniquely impact system capacity, and has a different implementation complexity and algorithm efficiency.

FDMA's ability to monitor individual RF channels (or DS0s) provides an effective mechanism for characterizing and efficiently managing ingress in the upstream RF plant. By dynamically performing upstream spectrum surveillance on narrowband carriers, an HFC network manager can manage applications spectrum usage based on traffic requirements and ingress levels to provide high utilization of the upstream spectrum.

WHY USE A NARROW BAND CARRIER?

One TDMA carrier typcially occupies 40 - 50 times the RF spectrum compared to one FDMA carrier; however, TDMA's channel carrying capacity is significantly less than multiple FDMA carriers occupying the same bandwidth. This is caused by the additional overhead required by burst mode TDMA. Intuitively, an FDMA narrow band system is more efficient in isolating the trouble spots, and efficient at filling gaps between noisy bands due to its smaller channel size. The question remains how much is it better than the TDMA approach? Quantifying this is not an easy task.

Rogers Engineering of Rogers Cablesystems Limited, Ontario, Canada, measured ingress on an HFC reverse plant representative of a fiber to the serving area architecture [3]. Measurements were taken over a period of three days. A review of these field measurements demonstrated that ingress noise typically consumes between 150 and 200 KHz of reverse bandwidth between 5 and 15 MHz. The width of the ingress is significantly smaller than TDMA carrier widths, typically 2 MHz, but is wider than the channel size of the FDMA system, typically 50 KHz or less.

Traditionally, information theory quantifies the quality of the transmission channel by its Signal to Noise Ratio (SNR) under the two conditions: (1) SNR vs. BER (Bit Error Ratio) which is meaningful when considering the same modulation techniques, and (2) as a white noise system. Without dispute, the first assumption is satisfied by assuming QPSK (Quadrature Phase-Shift Keying) modulation for both access methods. The white noise assumption is not satisfied by the ingress patterns observed by Rogers. Becasuse the relative performance is the focus, it is probably not critical to assume white noise for both systems in order to make the comparison possible. However, we will revisit the assumption later.

We segment the return path bandwidth (5 to 35 MHz) into either 600 x 50 KHz (1.5 bits per Hertz to form a 64 Kbps DS0 plus additional overhead) channels for the FDMA system, and into 15 x 2 MHz channels for a TDMA systems as shown in Figure 2. We also assume that the TDMA system uses a signal carrier power 40 times (16dB) stronger than the one used by the FDMA system since we assume that TDMA channel is 40 times wider than FDMA channel.

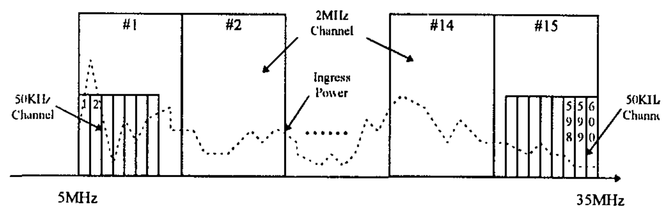


Figure 2: Reverse RF Bandwidth Allocation

Clearly, every channel suffers from different levels of ingress that consequently causes SNR values to vary from channel to channel. Note that the noise power is the total noise power measured within the channel bandwidth. The SNR for each channel is calculated. A distribution of SNR in each system is then defined as $P[SNR_i \leq S]$, which is the probability of the channel having a SNR less than S. The probability distribution functions for both systems is presented in Figure 3.

If the minimum requirements (BER of 10^{-3}) of SNR for QPSK that provides acceptable performance for telephone service is 13 dB, the FDMA system can provide 94.3% of its 600 x 50 KHz channels, or 565.5 channels (one channel carries one telephone line). For the same requirement, the wide channel system can only offer 78% of its 15 x 2 MHz channels, which is 11.7 channels (one channel carries 24 to 32 telephone lines). Therefore, for TDMA these 11.7 channels will provide only 280 to 374 DS0 lines. Thus, for the likely experienced SNR range of 10 dB to 25 dB, the actual efficiency advantage by using the FDMA approach over a TDMA system will be between 50% and 100%.

At low SNRs, the wide channel system delivers slightly better performance than the FDMA system. But QPSK is not normally expected to be deployed in environments with SNR less than 10dB. Even if it is, the difference is very small. As shown in Figure 3, if the reverse channel is exceptionally free of ingress, or when high SNRs are present, there is little difference between these two approaches.

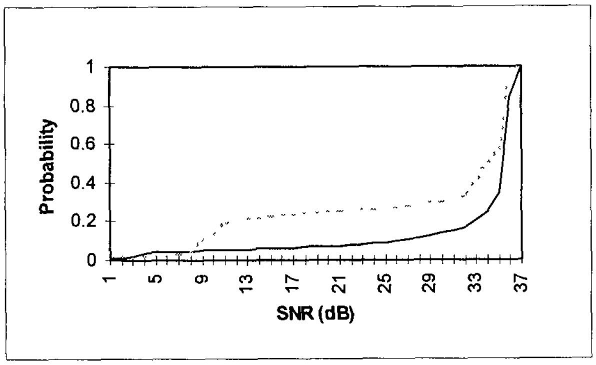


Figure 3: Probability Distribution of Channel SNR for Narrow Band (dashed line) and Wide Band (solid line) Band Carriers over 5 to 35 MHz.

The above comparison is based on the assumption that the transmission performance (bit error ratio) would be the same if the SNR of two band-limited signals of different sizes are identical. This statement is absolutely true only if the noise in these two systems is identical (e.g., white noise).

The data collected by the Rogers engineers demonstrates that the noise characteristics of the same channel change over time, as shown in Figure 4 and Figure 5. The two curves represent the observations at intervals which are three hours apart. In Figure 4, the characteristics of the ingress change very little between the two samples. In Figure 5, taken at different frequencies, the ingress changes dramatically between the two observation times.

These examples also show that ingress either occurs at one or two frequencies, but is unlikely to occur within three or more different frequencies inside of a 2 MHz window. Unfortunately, the resolution of the measurement is not fine enough to show the ingress activity within a 50 KHz band. One can imagine that the ingress is likely to appear as a single interferer, if it is present.

Because the channel noise does not appear to be identical in the two bands (the white noise assumption is violated), the above results are in doubt. Validation of the results is provided in the next section to address this concern.

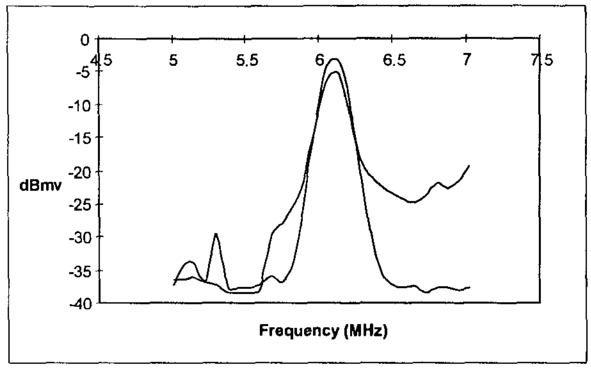


Figure 4: Two samples of Ingress between 5 MHz and 7 MHz

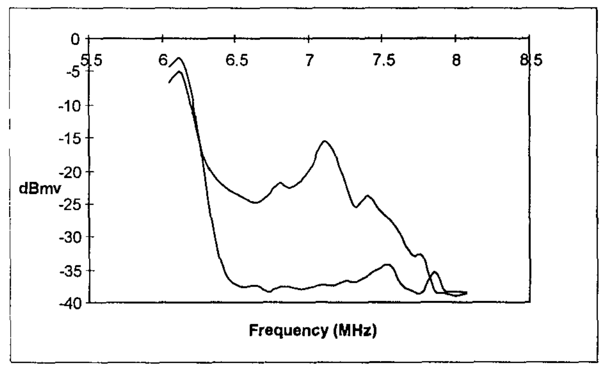


Figure 5: Two Samples of Ingress between 6 MHz and 8 MHz

VALIDATION OF THE RESULTS

A simulation technique is chosen to verify the noise assumption. The simulation model is shown in Figure 6. A stream of digital data is encoded and passed through the QPSK modulator. The modulated signals would be contaminated by the injected noise in the transmission medium. The combined signals (the original signals and the noise) is demodulated. The voltage difference between the output signals and the input signals are collected and compared to the power ratio of the carrier and noise.

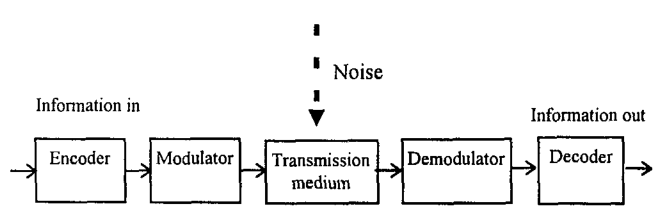


Figure 6: Ingress Simulation Model

Because the objective is to study the QPSK performance under different ingress conditions, we focus on the following two ingress types:

- Ingress, of identical total power, consisting of several different numbers of sinusoidal signals.
- Using the same number of sinusoidal signals located at different frequencies within the QPSK carrier band.

For this purpose, ingress is constructed within a 50 KHz band in four forms as shown in Figure 7. The ingress appears as sinusoidal waves at varying positions from the central frequency: (a) ingress composed of four identical sinusoidal signals at ±1 KHz, ±2 KHz from the carrier's central frequency, (b) ingress composed of four sinusoidal waves at ±16 KHz, ±17 KHz from the central frequency, (c) ingress composed of ten identical sinusoidal waves at ±1, ±2, ±3, ±4, ±5 KHz, (d) ingress composed of ten sinusoidal waves split at ±1, ±2, ±3, ±16K, ±17 KHz, and (e) white noise. Note that the SNRs of each case is fixed, and identical. Even though these waves are modelled in 50 KHz bands, the results will be the same if the frequency is scaled to a 2 MHz band.

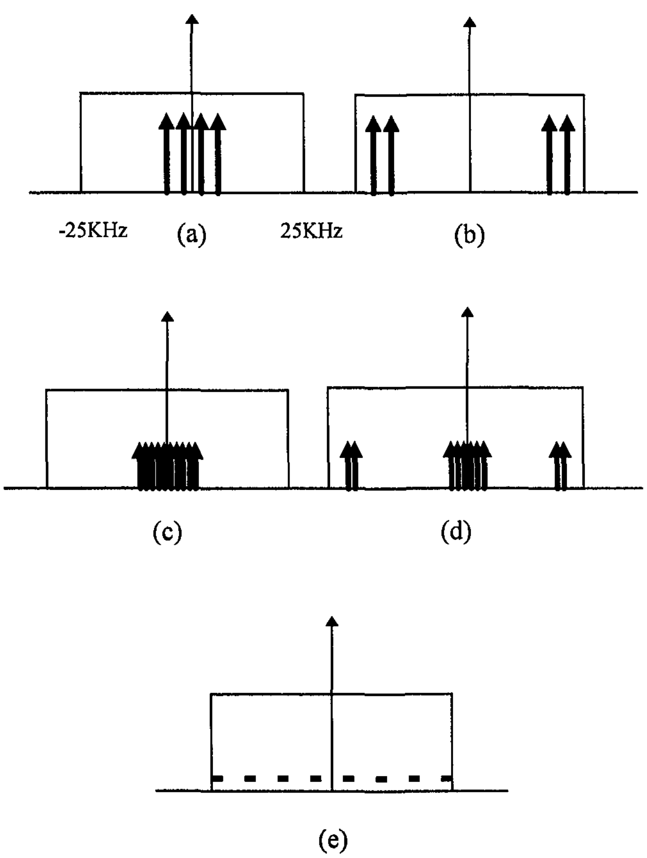


Figure 7: Ingress Models: (a) ±1K and ±2K Hz (b) ±16K and ±17K Hz (c) ±1 to ±5K Hz (d) ±1 to ±3, ±16, ±17 (e) White Noise.

The probability distribution function of each case is presented in Figure 8. The Gaussian white noise has the smallest tail probability (i.e., narrow distribution). Longer tails contribute to higher bit error ratios. The other cases show significantly larger tails, and thus higher BERs compared to white noise. Case d, two groups of ten sinusoidal waves, has the greatest detrimental impact on the performance. Even though the sample space (1888 samples) is not large enough to show the bit error probability, one still can conclude that when the ingress is split into two groups it will generate a longer tail than if it appears as a single interferer.

In all the cases, the total power of the ingress is identical. That is, the power per sinusoidal wave of cases a and b is greater than that of cases c and d. Based on this analysis, several conclusions can be drawn:

- Ingress in two bands within a carrier will increase the error rate compared to a single ingress band.
- Five sinusoidal waves of ingress will cause more errors than two sinusoidal waves.
- Bit error caused by the ingress adjacent to the central frequency is worse than if it occurs away from the central frequency.

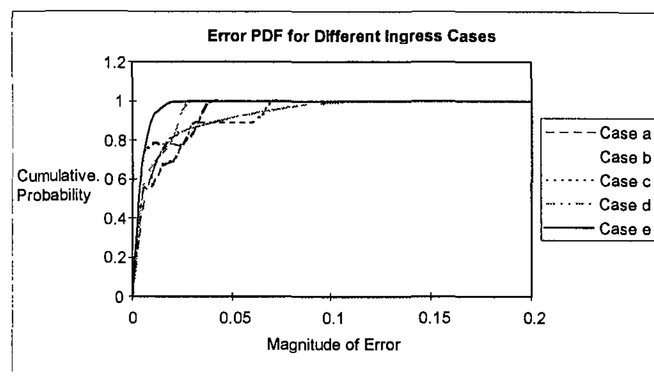


Figure 8: Probability distribution functions of Errors caused by different forms of ingress.

Most importantly, the variation in the magnitude of error is subtle from case to case. This allows us to conclude that with the same SNR, the narrow channel would perform better or at least as well as the wide channel approach. Given that the entire 5 to 35 MHz range is included in the analysis, providing a wide range of ingress scenarios, the approach taken is conservative from an engineering perspective.

RF PROTECTION ALGORITHM

The RF protection algorithm is influenced by the minimum RF channel qualifications for acceptable telephone service [4]. The algorithm includes the thresholds that trigger the frequency hopping mechanism in order to maintain acceptable service. The key prospect of this algorithm is the responsiveness of the mechanism under a given noisy environment.

First, a generic upstream framing structure is presented. The framing is tightly coupled with the RF protection algorithm. The upstream is divided into a sequence of continuous fixed-size frames (or slots). A frame consists a data field carrying DS0 data and a header that consists of an Upstream Frame Alignment (UFA) field, a Cyclic Redundancy Check (CRC) field, and necessary signaling and control information.

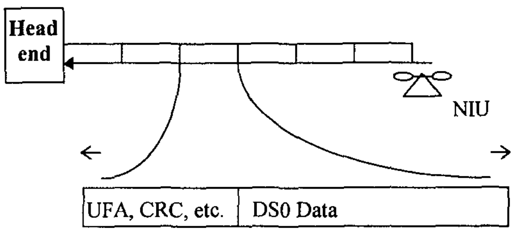


Figure 9: Upstream framing structure

UFA is a fixed bit pattern indicating the beginning of a frame. The CRC field is used to protect the information in the data field at the previous frame.

CHANNEL STATISTICS AND THRESHOLDS

The statistics of the following possible events can be constantly monitored at the headend for each RF channel to measure its transmission quality:

- Code Violation (CV) occurs when a CRC failure is detected.
- Errored Second (ES) is a second in which one or more Code Violations occur [4].
- Out Of Frame (OOF) is the event when headend can not locate the UFA.
- Loss Of Frame (LOF) occurs when OOF persists for a fixed number of times, say L.

There are potentially more events can be defined if implementation is permitted.

The threshold crossing indicates that performance of a channel has reached an unacceptable level. The appropriate threshold values for the above events should be derived from service requirements of POTS and ISDN.

Examples of possible channel-hopping criteria can be:

1. Occurrence of the LOF event, or
2. Occurrence of M CVs in one second.

EXPECTED DETECTION TIME

Given the event definitions, the threshold values, and the system parameters, the probability of these two criteria can be derived.

The derivation of the probability of LOF event *P[LOF]*, for example, starts with the event of OOF. The probability of an OOF is the probability of at least one bit in the UFA field is in error. Given an UFA consists of b_{UFA} bits, the probability *P[LOF]* is

$$P[LOF] = OOF^{L} = (1-(BER)^{b_{UFA}})^{L} \qquad \textbf{EQ. 1}$$

FER (Frame Error Rate), used in the following equations, is defined as the probability of a frame of b_{FRAME} bits containing at least one bit error (a code violation), and is expressed as $FER = 1-(1-BER)^{b_{FRAME}}$.

For the event of more than M code violations in one second , the probability can be expressed by the following equation given there are f frames in one second:

$$P[M\,CV\ in\ one\ second] = \sum_{i=M}^{f} \binom{f}{i} FER^{i}(1-FER)^{f-i} \qquad \text{EQ. 2}$$

The LOF event is designed to deal with the catastrophic condition such as a malfunctioning transmitter. The LOF event would not be expected to occur under typical RF degradation (e.g., thermal noise).

The graphical demonstration of the expected detection times versus BER is given in Figure 10. As the figure suggests, the algorithm will not allow channel hopping to take place if the channel is relatively clean. Using different *M*, the algorithm would activate the frequency hopping mechanism in one second when BER of the transmission environment exceeds a desired value.

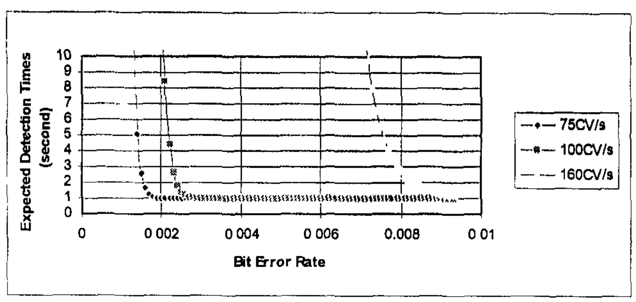


Figure 10: Expected Channel Detection Times

The minimum BER requirements for modem, fax, and telephone modem are not clear at this moment and they may not be as tolerant as the voice application. If a lower

BER is required, the algorithm can be tuned by modifying the parameters.

UPSTREAM CHANNEL ALLOCATION ALGORITHMS

So far, we defined the channel qualification and the channel hopping criteria. Consequently, a channel allocation algorithm is needed in order to choose a new channel when the current channel is declared unusable. The variations of such operation will dramatically affect the hardware and software complexity. A system may choose to implement a subset of the following functionality which may result in different system efficiency. The possible system features are:

1. *Assignment of RF channels to subscribers*: Fixed or Dynamic. *Fixed* means that a channel is assigned to a telephone line permanently. The relationship still exists even when the telephone set is in an on-hook state. *Dynamic* implies that the RF channel is released to a channel pool once the telephone set is on-hook. An off-hook event will cause the headend equipment to assign a channel. The mechanism the Network Interface Unit (NIU) uses to request a channel upon a off-hook event is beyond the scope of this paper, and will not be discussed.
2. Assignment of an NIU to a Demodulator / Receiver: Fixed or Variable. In practice, several Demodulators / Receivers can be grouped together. The NIU may be able to tune to only one group (fixed), or more than one group (variable).
3. Channel Performance Selection Classification: None or Queue. Once a channel is determined to be released due to the ingress noise, an abnormal release, it can be treated differently. A Queue, or linked list, can be maintained at headend such that a normally released channel can be enqueued at the top of the list and a abnormally released channel is enqueued at the end of the list. New calls would be established with preference for the channels that were released normally.

Different combinations of above functionality would result in different system performance with respect to the call blocking probability. For a given requirement, the results would be used to determine the maximum RF channel concentration ratio, and the system capacity.

SYSTEM QUALITY OF SERVICES (BLOCKING PROBABILITY) STUDY

The performance of a telephone system is traditionally described by the blocking probability of a call attempt [1]. To demonstrate how system capacity can vary due to different allocation schemes, we briefly review the queueing models commonly used in the telephony system.

THE EXISTING SYSTEM

A telephone system with m DS0 circuits is usually described either as an M/M/m or an M/M/m/n system, as shown in Figure 1, depending on whether calls are generated from an infinite number of phone lines or a finite number, n, of phone lines. The arrival pattern is a Poisson process. The length of call holding time is exponentially distributed with mean time, t.

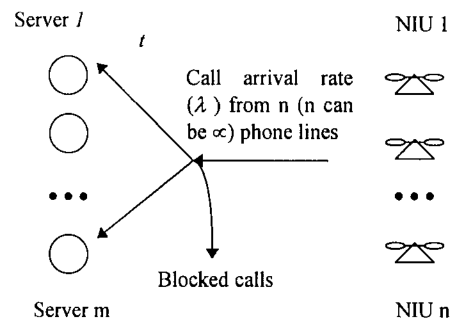


Figure 11: M/M/m/n Queueing Model

The blocking probability of a M/M/m system is known as the *Erlang's loss formula or Erlang's B formula:*

$$B_{infinite}(m,\lambda t) = \frac{(\lambda t)^m / m!}{\sum_{k=0}^{m} (\lambda t)^k / k!} \qquad \textbf{EQ. 3}$$

where m is the number of servers, λ is the average call arrival rate, and t is the mean holding time [2]. In a real system, the assumption of infinite customers is not realistic if the RF concentration factor is not greater than five. The above system has been further modified such that calls are generated by a finite number of telephone lines and the blocked customers will not generate another request immediately, but will after an exponentially distributed idle time. The blocking probability for such system is:

$$B_{finite}(n,m,\lambda t) = \frac{\binom{n-1}{m}(\lambda' t)^m}{\sum_{i=0}^{m}\binom{n-1}{i}(\lambda' t)^i} \qquad \textbf{EQ. 4}$$

where n is the number of customers, m is the number of servers and λ' is the call generating rate for one idle telephone line, and t is the average holding time for a call [2]. One can argue that this is not realistic either because a blocked customer may very well generate another request right after being denied by the system. Nevertheless, Equation 4 can be viewed as the upper bound, and

Equation 5 the lower bound of a real system. The difference of these two systems will not be significant once the concentration factor is greater than five.

OUR MODEL

In a HFC environment, the channel impairment can temporarily disable a channel for a undetermined period of time that makes the number of servers become a stochastic random variable.

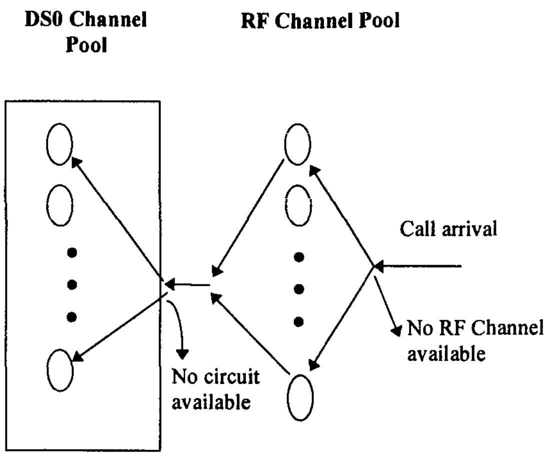


Figure 12: Queueing Model

A call can be rejected by either the DS0 circuit being unavailable or RF channel being unavailable, but not because of both events. One can conclude that the overall system blocking probability would be determined by the more constraining event. Since the two blocking events are mutually exclusive, one can further decompose the system into two parts: one consisting of the DS0 circuit pool (limited by switch ports) as the servers, and the RF channel pool (limited by the available RF channels) as the servers.

Recall the observation from the previous section, a HFC plant may lose a number of RF channels due to the RF impairment. But the stochastic behavior of the effective number of RF channels is not well known, and probably varies from plant to plant. It is not practical to manage the system capacity based on reliable assumption. The conservative approach is to provision a number of spare channels to accommodate channel impairments to maintain the desired service quality based on the worst scenario. The rest of this section will focus on the channel assignment algorithm.

CALL BLOCKING PROBABILITY

Two channel hopping mechanisms are used to demonstrate how the system capacity can be affected. The option of channel selection using a linked list is not included in the study.

OPTION 1

Under Option 1, each NIU transmitter is assigned to one channel. The association will not be released even the NIU is on-hook. If a channel is impaired, the headend equipment will direct the NIU to a spare channel within the same receiver. Therefore, the incoming calls can only be generated from a finite number of customers which is equal to or smaller than the number of channels that a receiver can support. Since the RF channel and NIU transmitter are tied together, the blocking model is slightly different from the M/M/m/n model, but we still can adapt the results to our system.

One observation of Option 1 is that once a NIU is assigned to a RF channel, its quality of service is somewhat determined. A NIU associated with a clean RF channel would never be blocked due to the RF channel unavailability. Blocking only applies to those whose RF channels happen to be in the area that ingress or other sources of impairment are strong. Blocking occurs when the number of impaired channels is greater than the number of spare channels. The blocking probability is no longer identical among the telephone lines. An average blocking probability would not appropriately represent the system blocking characteristics. The customers that happen to be assigned to the impaired channels must be satisfied.

The problem can be stated as follows: given a number of impaired RF channels, how many spare channels are needed such that the traffic generated by the telephone lines assigned to these RF channels will experience a desired blocking probability. Such system can be described by a M/M/m/n model, where m is the number of impaired channels and n is the number of spare channels.

The three-dimensional diagram in Figure 13 is based on a 24 channel receiver, and demonstrates the relationships among the following parameters: the number of impaired channels, the number of spare channels, and traffic density. Four different traffic levels, CCS equal to 3, 6, 9, 12, are considered. Note that the z axis is inverted for readability. The results confirm that one spare channel is required for every expected impaired channel when the total number of impaired RF channels is less than or equal to five channels in order to maintain the requirement of 1% blocking probability. If the HFC plant is likely to have more than five channels impaired, fewer than one for one spare channels will be needed because of the unlikelihood of all spare channels being needed simultaneously.

for one spare channels will be needed because of the unlikelihood of all spare channels being needed simultaneously.

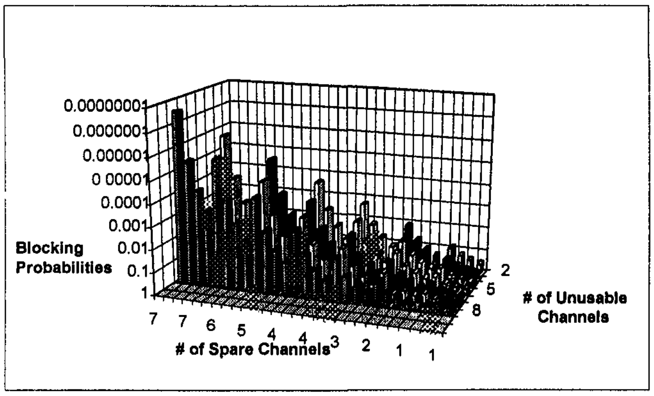


Figure 13: Blocking Probability for a Non-Concentrating System

The service providers need to characterize their reverse channel characteristics in order to determine the number of spare channels if Option 1 is implemented.

Using the results from the previous section, 6% of RF channels will be impaired, two out of 24 will be set aside. Therefore, 550 phone lines can be supported for 5 MHz to 35 MHz bandwidth.

OPTION 2

The above system can be improved by dynamically assigning RF channels between the NIU and headend equipment. An RF channel is assigned to an NIU when it is needed. Each channel will be returned to the channel pool after the call is terminated. Therefore, a telephone line can be assigned to any RF channel within the same receiver. No designated spare channels are needed. This scheme allows an impaired user to use any free channel to improve the channel utilization.

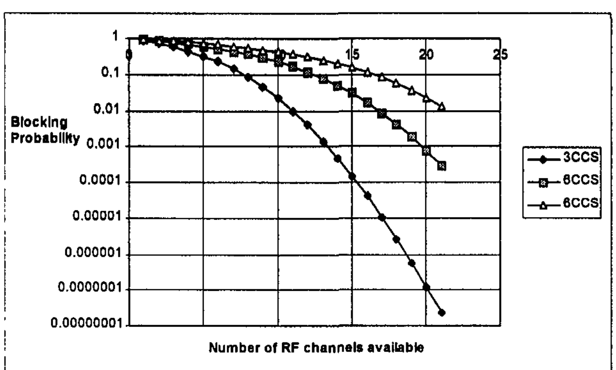


Figure 14: Blocking Probability of a 70 phone-line system

Clearly, the system blocking probability is determined by the HFC system characteristics. Given a finite number of phone lines the blocking probability versus the number of occupied channels can be provided based on Equation 5, as shown in Figure 14. Using the unavailability rate of 4%, a 24 channel receiver can support up to 70 phone lines. Comparing this result to the 20 phone lines that the same capacity receiver supports under Option 1, a more than 200% gain is achieved.

NETWORK MANAGEMENT IMPLICATIONS

If network operators implement systems that use narrow upstream carriers (50 KHz), the benefits can be extended with HFC network managers that perform upstream spectrum surveillance and management. The network manager can use the processed performance monitoring information provided by the receiver / demodulators to maximize equipments utilization of the RF spectrum.

The agent consists of equipment located in each headend that measures , at 50 KHz intervals, RF power levels, or Bit Error Rates of signals generated by transmitters located deep in the HFC subnetworks. Initially, the information provided by the network manager could be used to provide assistance in optimally assigning upstream frequencies to data and telephony applications. Later, a standard interface will allow operators to further optimize the frequency assignment by dynamically allocating bandwidth as dictated by traffic. Using a spectrum manager and agent to dynamically measure and assign bandwidth for each upstream HFC subnetwork provides the following benefits.

- Increases service reliability achieved through optimal frequency utilization
- Maximizes digital services' upstream spectrum utilization for each HFC subnetwork
- Allows performance monitoring of the entire upstream spectrum to help identify marginal HFC subnetworks
- Expediently identifies ingress and isolates it to a specific upstream HFC subnetwork

References

1. "Digital Telephony and Network Integration", Bernhard E. Keiser and Eugene Strange, 2nd edition, International Thomson Publishing Inc.
2. "Queueing System Vol. 1 Theory", Leonard Kleinrock, John Wiley & Sons, Inc.
3. "Two-Way Cable Television System Characterization", CableLabs, Inc., April 12, 1995.
4. TR-NWT-000499, Nov. 1991.

Cable Modem: Old Protocols for a new Paradigm

Frank R. Koperda
Scientific-Atlanta, Inc.

ABSTRACT

Cable modems should take advantage of the lessons learned from the data communications and the cable video communications environments. The resulting system should embody the robust protocols created for the digital video cable delivery system and provide the advanced features of the protocols in today's data communications networks.

*The merging of these environments will require changes to existing data and cable protocols to make them useful. ATM is emerging as a low cost delivery solution optimizedfor a point-to-point delivery system. ATM must be **modified** in several ways to **fit** into the point-to-multipoint topology. Similarly, the digital video systems deployed have substantial delay due to interleaving the data of many programs to keep a noise burst from corrupting any single channel.*

This paper will explore the systems issues involved in building an optimal data delivery system for cable systems. We will examine changes to standard protocols and how these will work in the new cable paradigm.

INTRODUCTION

Data communications is changing the types of possible applications for the home just as it has changed the applications in the business environment. Analog television is giving way to digital television with the advent of MPEG. The evolution of computing in the home environment is continuing with the emergence of the cable modem as a new product which will need to merge the video and data services. To operate in this new situation, existing protocols from both services will need to be altered to create a functional system to meet the needs of a new market.

The first task is to choose a topology that will lead to easy operations and the most efficient use of investment. Figure 1 represents a likely network topology for the delivery of data services. The network consists of three distinct physical blocks. They include the Master Headend (MHE) equipment, Distribution Hub (Headend) equipment and the Home unit. The topology of this network enables the operator to centralize the operations and the location of information serving computers, thus lowering operational and capital costs. Among the functions likely to be performed at the MHE are billing, level-of-service authorization, and network operations control. The MHE supports several Distribution Hubs and therefore is a logical point to perform inter-hub switching or routing, as well as serving as a gateway to external networks.

COMMUNICATIONS PROTOCOLS

There are several communications protocols that will be important in the new paradigm. Of interest are TCP/IP, ATM, MPEG, X.25, Frame Relay, DAVIC, IEEE 802.14, and IEEE 802.2 which includes Ethernet and Token Ring. Figure 2 illustrates their usage in the various parts of the cable network.

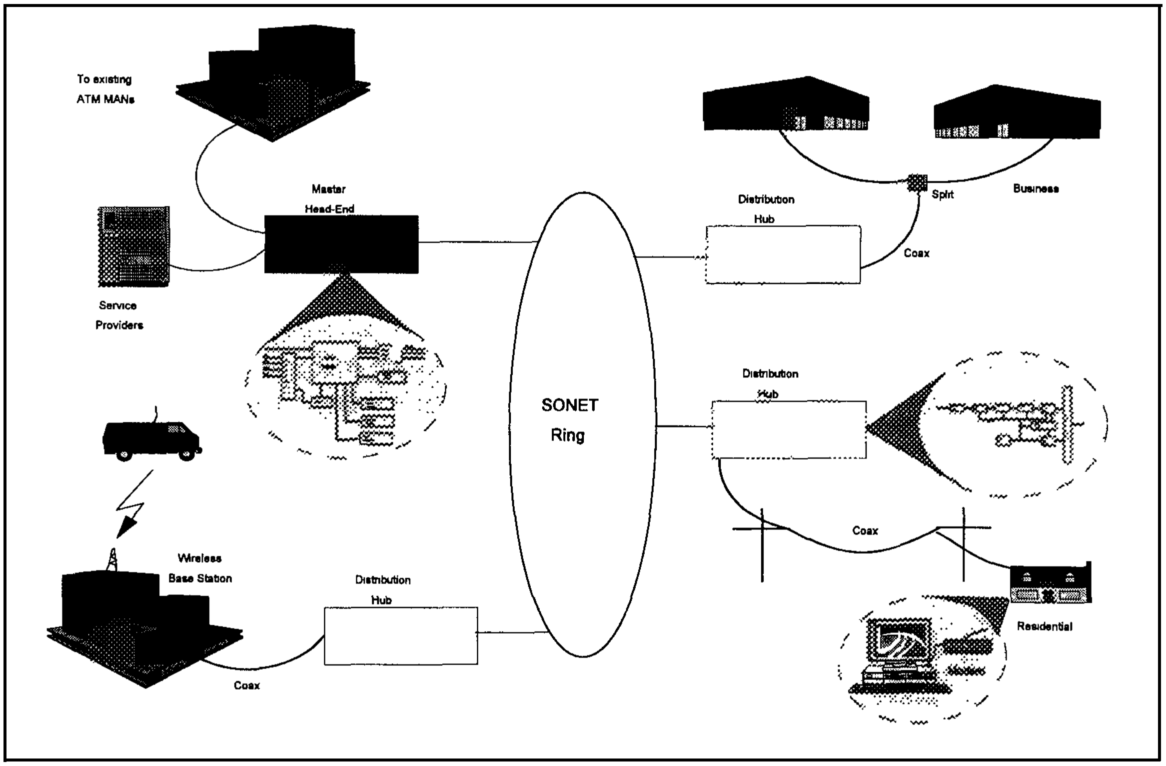


Figure 1. Network Topology

The cable modem will likely exist as a stand-alone box during the first few years. Two primary reasons for it to remain separate from the computer are potential liability issues and obsolescence. Operators will be reluctant to open up the PC, install a card and incur the responsibilities and liabilities of making such a device operational. A customer is unlikely to buy such without knowing that it will be usable over their cable system now and in the event of moving to another cable system. As new standards emerge, the operator will specify compliance to a standard and the ownership of an internal modem will be possible by the consumer. This should then allow a reduction in the price of the modem.

The PC will connect to the cable modem using a LAN card of the type found in many office environments. The cards are relatively inexpensive, are available on most hardware platforms, and are supported by all major operating systems. Ethernet using a 10BaseT interface should meet the various requirements for this situation. As an evolutionary course, the 25 Mb/s ATM might also emerge in several years in response to a need for greater speeds.

The cable modem is connected to the headend using an existing cable plant that has had the reverse path activated. The RF modulation technique will probably be QAM-64 that allows 30 Mb/s of data to be transmitted in the 6 MHz channel typically used for a television channel. On reverse channel, the RF modulation technique used will be some form of QPSK. Proposed data rates of 256Kb/s, 1,544 Mb/s and 2.048 Mb/s have been mentioned in the various standards groups. As seen, this implies that highly asymmetrical data rates will occur, which affects existing protocols that have been optimized assuming symmetrical data transfer rates.

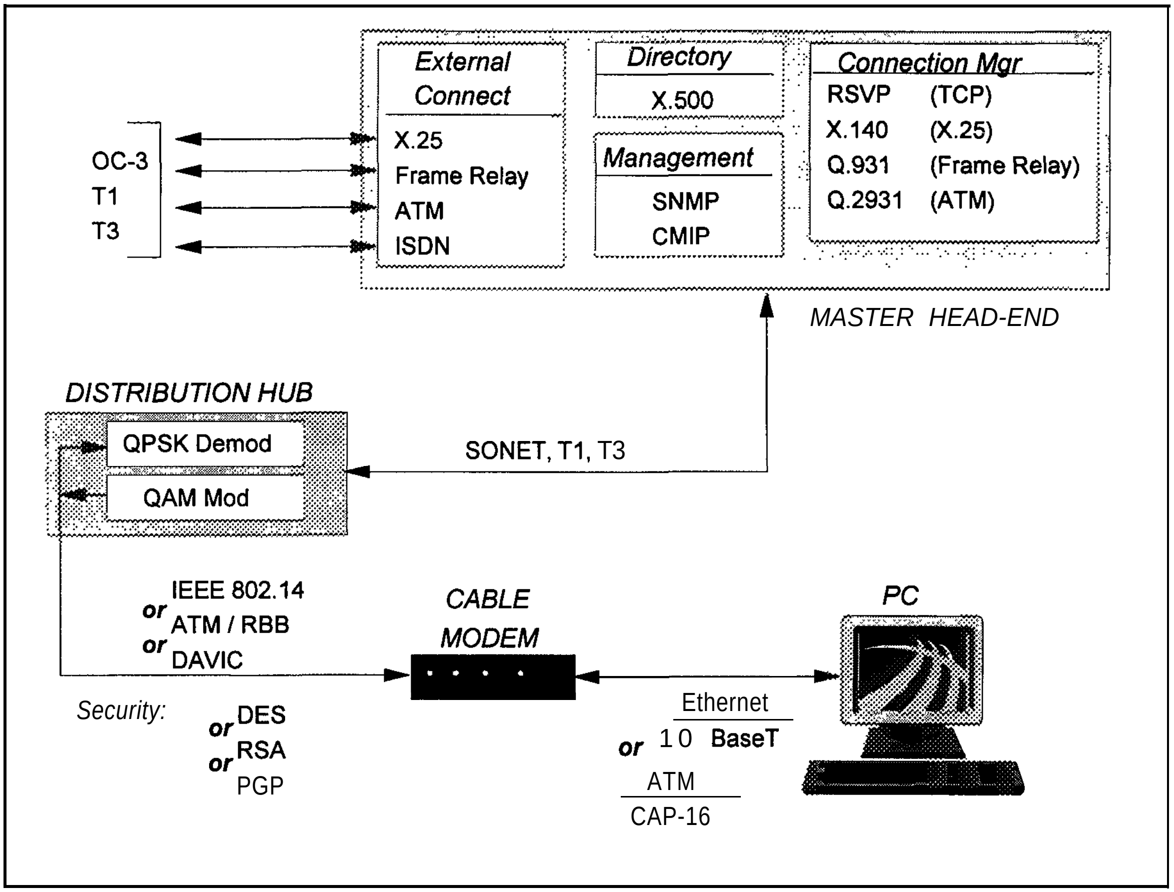


Figure 2. Protocols used within a cable modem network.

Another characteristic differentiating this link from many other link protocols is that it has a point-to-multipoint topology downstream, but a multipoint-to-point topology upstream. Because a single entity controls the downstream, multi-access techniques are not required, but security for user data becomes important because many stations are listening. Upstream, it is difficult for other stations to hear each other because of uni-directional taps, but users of the system perceive their data may be in danger of interception and therefore will desire security. Potential encryption algorithms include DES, RSA, and PGP.

Having many stations sharing this upstream link will require new implementations to control which stations have access to the upstream bandwidth resource. Since bandwidth is a scarce resource, one user should not be allowed to monopolize the link as is possible over existing LANs today. Also, a user may need to be billed based on the level of service being provided. Having a 64 Kb/s tier, 384 Kb/s tier and a burst 1.5 Mb/s tier for example will require a network to guarantee a station can achieve that performance. It must also ensure that the usage for a station does exceed that performance for a given tier.

The link from the modem to the headend is being 'standardized' by several groups. The Digital Audio-Visual Council (DAVIC) is targeting its 1 .1 specification to address data transport issues. The IEEE 802.14 committee is also drafting a standard that defines both the link protocols and system connectivity issues. The Residential BroadBand (RBB) group of the ATMForum has indicated interest in this link but may adopt the IEEE 802.14

recommendations for HFC (Hybrid Fiber Coax) systems.

The benefits of adopting products based on well known standards should allow rapid growth, low prices, and additional functions and features from manufacturers.

From the headend, some information that originates within the area will flow back downstream and some will flow to the MHE. Typical physical connections will be provided by T1, T3, OC-3, OC-12 or even higher speeds. The logical protocols may be ATM or Fast Ethernet. Data from the headend may go to local servers at the MHE, out to other headends or out to the public network.

Because there may be connectivity beyond the local network, the protocols within the network are affected. For example, many network protocols such as TCP have local naming and addressing. This option cannot be used external to the network; hence, resolving names requires connection into other Domain Name Servers (DNS) and the local operator must deal with getting blocks of addresses. The X.500 protocol is a popular choice to map the number addresses to user names.

Security protocols are also affected because it is impossible to ensure that all destinations outside the cable network use the same type of security. For this reason, the HFC link should have local encryption and decryption. Application layer encryption is also required for end-to-end protection. However this encryption is beyond the operator's control and should be left to browsers, shopping, and banking applications that can ensure protection using many different methods appropriate to the expected level of attack.

The MHE also has a number of other system elements to ensure a functional network. Network Management is spread throughout the system with the central control intelligence located at the headend. Common platforms for this are NetView® by IBM and OpenView® by HP. They use the CMIP and SNMP protocols to communicate with various units in the network to monitor performance and to determine the location of problems.

Billing systems need to operate with the equipment to authenticate devices that wish service, determine the amount and type of service permitted and gather connection statistics. These statistics may consist of the length of the connection, number of packets sent and received, or identification of servers contacted. These types of statistics may be saved away so that the user can be billed for the amount of network and system resources that were used. These systems are currently varied and often unique.

As mentioned earlier, the quality of the connection will likely be very important to the user. For this reason the parameters for the type of connection desired could be requested from the network. Several different Quality of Service (QoS) standards exist that can communicate the connection parameters from the user to the system. These QoS protocols are linked to specific protocols. The X. 140 standard is used with the X.25 protocol. TCP uses RSVP, Frame Relay uses Q.931-Q.933 and ATM uses Q.293 1. Other protocols may also reuse some of these QoS standards.

These QoS standards have similar types of parameters. Requested bandwidth for forward and reverse paths, maximum delay, amount of packet delay variations, maximum length of a packet, number of packets in a transmission burst and maximum error rates are examples of these parameters. It is very important that the network keep its QoS contract with a user's connection to ensure proper operation of the applications. Mechanisms that the network uses to implement this contract are planning routes through the modulators, demodulators, switches and links; monitoring congestion; putting packet admittance procedures in place as data enters the network; and delaying data that is not time critical.

Moving the discussion from the network to the protocols within the PC, Figure 3 provides a useful reference. The LAN adapter within the PC may be one of the better ways to tie the PC into the network. The 10BaseT adapter can often be purchased for $40-$60 and supports the IEEE 802.2 Link Layer Control (LLC) for packet transfer. Drivers are required to support the Upper Layer Protocols (ULP) such as TCP/IP, NetBEUI, NetBIOS, Appletalk, IPX/SPX (Netware) and SNA. These drivers exist for various hardware platforms and operating systems such as Windows, UNIX, and OS/2.

Effective network management will be

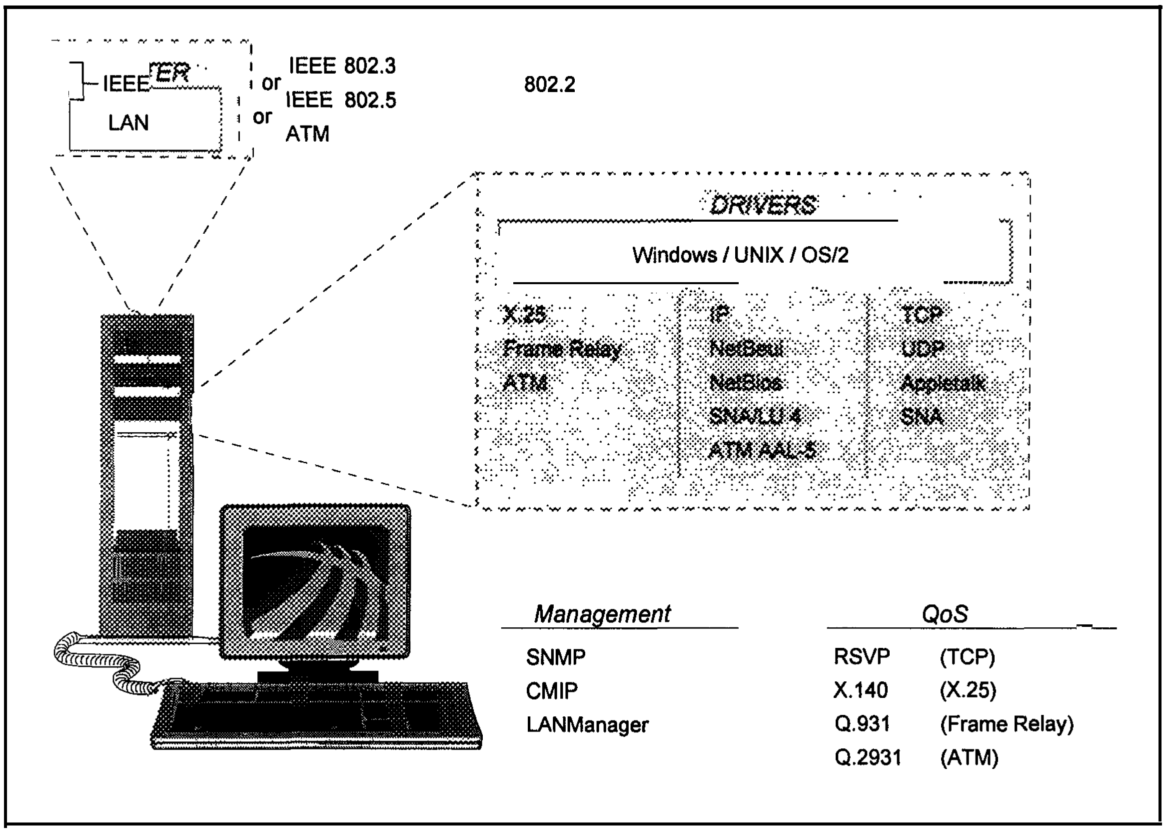


Figure 3. Protocol stack within the PC

Using the serial or parallel ports has some advantages and disadvantages to connect the computer to the cable modem. One advantage is that they are often provided with the system but a disadvantage is there is little existing communications software that make use of them. The serial port also does not have the required speed of several Mb/s, and the parallel port is often used for the printer and often goes to 800 Kb/s.

important for the broadband network because of the cost incurred by an operator for on-site support. It may become a requirement that the modem run an SNMP or CMIP agent that can gather statistics about the connection, and feed it to the network operator automatically to help diagnose the source of problems. The user may need to become aware of these features to participate in a conversation with a network administrator in trying to track down faulty equipment. The system should be as transparent to the user as possible and ideally enable the operator to correct problems before they degrade to the point of being noticeable to customers.

Quality of Service will be increasingly important to customers as more intelligent applications are created. Videoconferencing is an application that will need to communicate with the network to establish a 384 Kb/s link with delays under 50 ms. Efforts are underway to establish the Application Programming Interfaces (API) that will allow the user's applications to convey its requirements to the various network elements required to implement the request.

The above discussions demonstrate only one of several valid configurations that may emerge. A variant of this is to have the downstream cable supplying data, but have the upstream data sent via a standard teleco modem. This will have a useful life for a few years until more of the reverse cable plants are activated. Another alternative is to have a less powerful computer for the user connected to centralized servers to perform many of the tasks. This will allow for cheaper home units, more economical maintenance of software which is done at the servers, and a sharing of the cost of computing. The type of protocols to support this system are somewhat different from those illustrated.

OLD ATM FOR A NEW PARADIGM

ATM may become an important protocol for use in the cable modem systems. Internet access will be a key application and is implemented using TCP over IP. For low capacity nodes, routers are a good option. As the network grows, switching becomes a better option. Routers work by looking inside a packet to understand what functions need to be performed working at the OSI layer 3. Packet switches work at layer 2 of the OSI model and operate by looking at the destination address. They do not have any understanding of the protocol nor the contents of the packets. For a given capacity, switching is about 1/3 of the cost of a router. A useful axiom in network design is to 'switch when you can, route when you must'.

As multimedia applications continue to develop, it is important to have a network to support the applications by guaranteeing the performance of the data transfer. Two important reasons to guarantee the performance are to minimize buffering costs and reduce the delays between the end stations. Applications without proper network guarantees will see their videoconference images become very jerky and excessive delay hinders communication when responses are delayed by seconds.

An important method of guaranteeing performance is to use fixed length cells instead of variable length packets. Packets make it

Attribute	Current ATM	Future Cable ATM	Motivation
Data transfer	Symmetrical	Asymmetric	Ingress noise
Speeds (Mb/s)	25.6, 52, T3 (45)	1.5 (up), 27 (down)	RF modulators
Access Methods	Point-to-Point	Point-to-Multipoint (down) Multipoint-to-Point (up)	Cable topology
Security	optional	mandatory	Multiple access
Data Pacing	OAM cells	leaky bucket	Asymmetric bandwidth
E. 164 Addresses	Fixed per port	Movable between ports	Alternate routes

Table 1. Migration of ATM for cable

difficult to control the variation in arrival time since multiple packets of indeterminate length can get in front of critical packets. ATM appears to offer several advantages in the cable environment by reusing equipment available for the LAN environment and connecting easily into the public networks. A summary of changes for migrating standard ATM to the cable environment is shown in Table 1.

Data transfer in standard ATM is accomplished using a dedicated, point-to-point connection with a symmetrical data transfer rate. In the LAN environment this data rate is 25.6 or 52 Mb/s. It also appears at data rates of T1 , T3, OC-3 and above in the public networks. Security is not a primary concern because the links are usually secure. If security is a concern, bulk encrypters that encrypt all the connections on the link simultaneously are used.

The new paradigm of cable modem places different constraints on the network transport system. The primary link causing this difference is the connection between the headend and the home which is the cable plant. The forward data transfer is well maintained and transfers existing television signals in 6 MHz channels typically in the 50-750 MHz band. Quadrature Amplitude Modulation (QAM) is a prominent digital modulation technique in this situation and it has an effective 27 Mb/s data rate.

The reverse path is usually not activated, or, if it is, presents a challenging RF environment operating in the 5-40 MHz band. Some types of splitters in the home have notches in the 5-l 0 MHz band. CB transmitters, electrical appliances, and amateur radio are sources of in-home contamination in that band. Outside the home, international short-wave, corona discharge from nearby power lines, changes in temperature, corroded connectors and nicks in the line all cause degraded signal quality. To avoid these noise sources, 1 MHz channels scattered around the band are used with the robust modulation technique known as Quadrature Phase-Shift Keying (QPSK) which gives an effective 1.5 Mb/s data rate.

The bandwidths available for data transfer fundamentally cause the asymmetric data transfer that alters standard ATM data rates. The broadband system offers a significant cost advantage over conventional telephone links by spreading of the cost of a single cable among many users. Downstream that means a single point going to multiple users and upstream that means having multiple users share a single cable.

Having multiple users share a single cable implies a loss of privacy and thus the need for encryption of the data to regain security. For a standard telephone line, users feel somewhat more protected because each station has its own dedicated wire. ATM has no specific mechanism for dealing with alternating between security keys and maintaining proper decryption synchronization. This feature will need to be added.

ATM employs a mechanism to keep buffers from overflowing during heavy network usage. OAM (Operations, Administration, Maintenance) information cells are the feedback mechanism whereby the status of the buffers is sent to the originator. If the same mechanism were used in the asymmetrical environment of the cable modem, there would be little upstream bandwidth left for user data. For example, the OAM cells downstream could be 5% of the bandwidth and at 27 Mb/s that would mean 1.35 Mb/s would be consumed by the OAM cells. If this were fed back upstream, the 1.35 Mb/s would consume 87% of the 1.5 Mb/s leaving little bandwidth for data. The pacing of upstream packets using the 'leaky bucket' algorithm and having the switch block

OAM cells from going downstream may be an adequate solution to the problem.

Another interesting change to standard ATM is the assumption that a station address (E. 164 address) remains fixed to a specific switch port. Under a multiple user-per-modulator scenario, there will now be several E. 164 addresses on a single port and it is likely the station addresses will move between RF frequencies (ATM ports). If a user requests a new connection and the existing channel over a QAM modulator does not have sufficient remaining bandwidth to satisfy the request, the station should move to an alternate QAM channel. It is likely that a modem will only have one downstream tuner and thus all of the existing connections must move to the new QAM channel which is connected to a different ATM port.

PERFORMANCE EXPECTATIONS

The broadband network will have data traffic that will take on the characteristics of both the local LAN and ATM trunk links. It will also have unique characteristics to handle interactive video games, services that vary by time of day and billing that will alter the users' traffic patterns. A variety of experimentation will occur over the next few years as providers try to offer the right combinations of services, cost models and feature sets.

The basic applications that may shape the architecture of the network are:

- Internet access
- Access to CD-ROM library including video-clips
- On-line chat services
- Interactive games with a server and peer-to-peer
- Access to community services such as schools and event calendars
- Work at home (LAN emulation)
- Videoconferencing

Application	Packet Size	Traffic Shape	Query Times	Peak Speed
Web Browsing	1500 Bytes	Asymmetric (100:1) / bursty	once / min	Available
On-Line Chat	300 Bytes	Asymmetric (20:1) / bursty	once / 20 min	Available
LAN Emulation	70% - 64 Bytes 30% - 1 KByte	Asymmetric (6: 1 server) / bursty Symmetric (peer) / bursty	once / 30 min	Available
Interactive Game	1500 Byte 500 Byte	Asymmetric (100: 1 server) / bursty Symmetric (peer) / bursty	once / 5 sec once / sec	90 Kb/s
CD-ROM Video	1 KByte	Asymmetric (100: 1) / Continuous	once / 2 min	2 Mb/s
Video Conferencing		Symmetric / Continuous		384 Kb/s

Table 2. Application Characteristics

An interesting aspect of the network usage is that these applications will have peak demands (Table 2) at different times (Table 3). This implies that there will be a need for dynamic bandwidth management and cost recovery by the operators will allow creative solutions. Business applications can support higher cost rates than interactive video games and access to community information may be free to build initial usage.

The operators must recover the cost of building and operating the network consisting of the above applications. It is possible to have billing based on tiered performance, usage or connect time. The proper selection of the

Application	Peak Time of Day	Network Traffic (Peak)	Application Value
Web Browsing	Evening	30 % Users / 30 % data	Moderate
On-Line Chat	9-11 PM	70 % Users / 10 % data	Moderate-High
LAN Emulation	10 AM / 3 PM	70 % Users / 20 % data	Very High
Interactive Game	Afternoon	40 % Users / 5 % data	Low
CD-ROM Video	Afternoon / 7 PM	15 % Users / 30 % data	Moderate
Video Conferencing	10 AM / 3 PM	20 % Users / 70 % Data	High

Table 3. Application Usage

Note: % Users are the % of active users for that application during that time period.
% Data is the % of the network data attributable to that application during that time period.

method is based on the cost models of the network and the customer's perception of the worth of the application.

CONCLUSION

Although the term 'cable modem' implies a device that should behave like today's modem which is attached to a telephone line, the new data paradigm requires a new awareness of the needs of customers. We are moving from a communications paradigm where users have a dedicated link and are assured privacy and dedicated bandwidth. We are moving toward one in which security, performance tiers and shared bandwidth are the norm. Applicable elements of older LAN, ATM and network management protocols are being reshaped to fit the new network paradigm.

Author

Frank R. Koperda
Scientific-Atlanta, Inc.
4386 Park Drive
Norcross, GA 30093

Telephone: (770) 903-5057
Email: Frank.Koperda@sciatl.com

Cascading of Noise and Digital Compression Artifacts

Tony Werner
Vice President
Engineering
TCI Communications, Inc.
Terrace Tower II
5619 DTC Parkway
Englewood, CO 80111-3000
(303) 267-6950
FAX (303) 488-3210

Oleh J. Sniezko
Director
Transmission Engineering
TCI Communications, Inc.
Terrace Tower II
5619 DTC Parkway
Englewood, CO 80111-3000
(303) 267-6959
FAX (303) 488-3210

Abstract

Contrary to popular belief, digital systems are not perfect. This is especially true when like and unlike digital systems are concatenated with other digital and analog transmission systems. These systems add thermal noise, quantization noise and, in the case of compressed systems, compression artifacts. Several questions and mis-beliefs exist as to how these degradations combine on a subjective and objective basis.

This paper presents both objective and subjective results of:

- *cascading quantization noise from linear PCM systems;*
- *the effects of combining quantization noise and thermal noise, and*
- *the effects of combining, thermal noise and compression artifacts.*

The industry is proficient at thermal noise measurements today. Quantization noise measurements are also becoming more consistent through the use of shallow ramp measurements. Compressed signals, however, still offer several challenges. Certain test scenes may look fine on a particular codec, while other scenes may cause significant artifacts.

The analysis of the test results collected during the evaluation indicate that qunatization noise adds similarly to thermal noise both objectively and subjectively. It also indicates that certain compression artifacts contribute subjectively as noise additions. This paper presents these results and analyzes the relationship for cascaded thermal noise, quantization errors, and compression artifacts.

INTRODUCTION

Digital delivery schemes are becoming the backbone for broadband communications networks. Many of the new services are being distributed in a digital format. These services include telephony (voice), data, and video. Through some parts of the cable systems they are distributed in a purely digital form (baseband) using TDM technology, statistical multiplexing, or multiple access schemes. In the coaxial RF part of the network and in some parts of the optical network they are distributed using FDM technology (so-called analog systems). Actually, it is a combination of TDM, statistical multiplexing technologies for efficient channel use, multiple access schemes, some complex modulations schemes (FSK, QPSK, QPR, different levels of QAM or VSB, etc.), and FDM technology.

Any signal, including digitized video, on its way from the originating point to a customer is affected by the transmission media. The industry developed a set of measurements to define analog impairments and correlate these measurements with the subjective picture quality. This correlation is the result of many years of experience and extensive effort by analog video providers. These measurements have also been continuously re-defined as analog video has become more complex, and as the acceptance levels of our customers changed. The challenge for the industry is to evaluate the impact of the transmission medium impairments on digital video in a quantitative way and to find a correlation between the measurements and the subjective picture quality.

VIDEO EVALUATION MEASURES

Analog Measures in Analog World

Over the years, the motion picture and television industries developed a set of measurements for TV picture quality. These measures were correlated to picture distortions and categorized into several categories depending on the required fidelity of the video transmission. Several bodies developed these specifications and measurement techniques, but the most commonly known is one endorsed by EIA [3].

Subjective Measures

Unfortunately, these parameters use test signals that can be processed by most digital video systems without significant degradation. Nevertheless, the same digital system may still provide very distorted signals and inadequate picture quality for certain video content. This inconsistency is well covered in references, [1] and [5]. The only universal and conclusive method of evaluating picture quality is to perform subjective testing. Rigorous subjective test methods were developed by CCIR and are described in [4]. These tests are very expensive and time consuming but cannot be avoided until a set of objective parameters are correlated to the subjective tests results for digital transmission schemes.

In these days of evolving standards for video compression, the subjective tests were unavoidable. Many testing and standard organizations developed a set of "killer" test tapes to test all picture aspects that can be affected by known digital video compression systems.

Many end users cannot afford this type of test every time they want to assess the quality of a digital video system (codecs). To avoid the expense, they usually perform a set of objective tests supported by a subjective evaluation by a limited number of expert viewers for a limited time. This situation must improve, especially with the numerous digital systems available and already in use today.

Objective Measures

The industry has worked hard to develop a set of measures to test digital video quality. The purpose was obvious: to be able to test different digital coding/decoding systems, both compressed and non-compressed, without relying on expensive subjective test methods or being subject to the inaccuracy of simplified subjective tests. The literature on this subject is exhaustive. Two references [1]&[2] present a summary of the current status on objective measures of digital video quality. The following summarizes the measures presented in [2].

Objective parameters using artificial test signals

Similarly to the parameters listed in the EIA standard, these parameters are tested using test signals (artificial video signals — video test waveforms) produced by signal generators. The following parameters are measured:

- average gain and level offset (contrast, brightness, and color intensity distortions),
- amplitude frequency response (affecting resolution),
- active video area,
- active video shift.

Some of these distortions can be easily corrected by adjustments of the units (especially average gain and level offset).

Objective parameters using natural test scenes

The same set of "killer" test tapes used for subjective evaluation can be used for testing objective parameters of the codecs and transmission channels. The set of these parameters is presented in Table 1.

As can be seen from Table 1, many of the parameters quantify the increase in noise level, another estimate resolution degradation, and yet another quantify impairments specific to digital video. The hypothesis is that these subjective impairments that in the analog world cascaded with thermal noise will cascade with thermal noise in the digital world. Similarly, the impairment that in the analog world were masked by the noise (blurring, smearing) will be masked by the thermal noise in the digital world. As much as we would like to test this hypothesis in a formal and rigorous manner, our limited resources do not allow for that and we must leave it to the industry and standardization bodies. This paper presents initial results that support at least part of the hypothesis and were collected over several months during different codecs evaluation tests.

Table 1: Association of Objective Parameters and Impairments

	Objective Parameter	Impairments (as per ANSI T1.801.02)
1	Maximum added motion energy	error blocks, jerkiness, noise
2	Maximum lost motion energy	jerkiness
3	Average motion energy difference	jerkiness, noise, error blocks
4	Average lost motion energy with noise removed	jerkiness
5	Percent repeated frames	jerkiness
6	Maximum added edge energy	spatial edge noise, block distortion, tiling, noise
7	Maximum lost edge energy	blurring, smearing
8	Average edge energy difference	blurring, smearing, spatial edge noise, block distortion, tiling, noise
9	Maximum HV to non-HV edge energy difference	block distortion, tiling
10	Added edge energy frequencies	temporal edge noise, spatial edge noise, edge busyness
11	Maximum added spatial frequencies	spatial edge noise, block distortion, tiling, noise
12	Maximum lost spatial frequencies	blurring, smearing

TRANSMISSION CHANNEL IMPAIRMENTS

The parameters listed above (both for analog and digital video) are only part of the story. They do not include all impairments that can be contributed by the transmission medium, especially by RF coaxial network. The cable TV industry characterizes the network for such parameters as:

- thermal noise,
- phase noise,
- nonlinear products,
- level of ingress,
- level of echoes, and so on.

Through a meticulous effort, the industry set acceptance levels for these impairments. These levels are being updated from time to time when either a new technology evolves (for example, the switch from black & white to color picture required a new, redefined echo curve) or the acceptance level of our customers changes. From time to time, a new impairment is added to the list when industry becomes aware of it or the transmission technology introduces it. The impairments manifest themselves in a distinguished way in a picture and in many cases one type of impairment masks the other.

TEST RESULTS

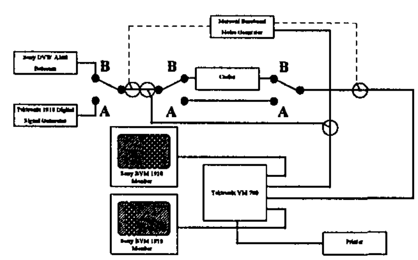
The test set-up diagrams are shown in Figures 1(a) and 1(b). The test set-up pictured in Figure 1(a) was used for most of the tests whereas the test set-up presented in Figure 1(b) was used to verify the results. The test set-up in Figure 1(b) closely resembles the real-life conditions. The test results indicated that both methods yield comparable outcomes. The tests were performed during the last 15 months and the test equipment used may have been different but it was always of equivalent quality and accuracy.

The objective (quantitative) tests were performed with an NTSC signal generator and VM 700, using standard test signals (shallow or full ramp and flat field for SNR result comparison).

The subjective tests were performed with a series of test tapes with the most often used moving and still patterns presented in Figure 2. These patterns were: Moving Zone Plate (Snell&Wilcox), Carousel, Flower Bed, and Still Zone Plate. The pictures were viewed by at least three expert viewers at good viewing conditions.

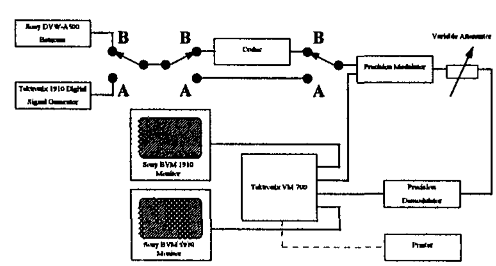
Figure 1: Test Set-Up

1(a) Noise Addition at Baseband



Note: Combiner symbols used to simplify the drawing. A VDA or loop through capabilities of the test equipment was used in actual test.

1(b) Noise Addition at RF



Quantization Noise and Thermal Noise

To separate the quantization noise from other impairments, including compression artifacts, linear PCM systems were used for this test. These systems deliver uncompressed digital video with the quantization noise defined by the resolution of sampling (number of bits per sample) and by additional digital processing of the signal. Systems from two different manufacturers were used. The resulting noise was measured objectively according to the industry standards for quantization noise and subjectively. The tests were performed over a period of four months in 1994 and 1995.

The detailed results are summarized in Tables 3 and 4. The results of the SNR test showed that the cascading of the quantization noise followed the standard power addition rule ($10 \bullet \log$) for a low number of cascaded units. The combined effect for four units was greater than defined by the formula. The thermal noise and quantization noise cascaded according to the rule.

An interesting subjective assessment was that the thermal noise added to the source (codecs input) resulted in higher subjective degradation than the thermal noise contributed by the transmission channel (added to the codecs output).

The test results show that the quantization noise of short cascades of codecs (most common case) adds on the $10 \bullet \log$ basis. The quantization noise and thermal (white) noise also cascade on $10 \bullet \log$ basis. Subjective tests confirm the objective results.

Table 3: Cascading of Quantization Noise

Linear PCM Unit	Measured SNR (Ramp) in dB (unified weighting)	Calculated SNR (Ramp) in dB
#1	59.4	NA
#2	60.2	NA
#3	59.9	NA
#4	60.3	NA
#1&#2	57.2	56.8
#3&#4	55.3	57.1
#1, #2, #3, & #4	52.1&51.4	53.9
#1, #2, #3, & #4 (51.4) plus thermal noise (51.8) on input	48.9	48.6
#1, #2, #3, & #4 (51.4) plus thermal noise (51.8) on output	48.9	48.6
#6	63.2	NA
#6 plus thermal noise (50.3)	50.1	50.1
#7	68.1	NA
#7 plus thermal noise (50.3)	50.1	50.2

Table 4: Subjective Evaluation of Cascading Effect

Linear PCM Unit	Measured SNR (Ramp) in dB (unified weighting)	Subjective SNR (Ramp) in dB
#1	59.2	56.9[1]
#1, #2, #3, & #4	51.7	52.7

Artifacts and Thermal Noise

Similar testing was performed for a series of codecs with compression. The units were from two different manufacturers and employed different intra-frame compression algorithms. The codecs did not employ any inter-frame compression algorithm. Codecs with two levels of compression were used for testing:

- single-channel DS3 codecs, and
- dual-channel DS3 codecs.

The test results are summarized in Tables 5 and 6. The results clearly indicate a cascading effect of thermal noise and the type of the artifacts that were perceived as noise (increased busyness of the picture with noise-like pattern).

NEXT STEPS FOR INDUSTRY

The test equipment industry is working on development of a test set to measure the parameters presented in [2]. We, as the industry that will have to characterize the digital transmission media must challenge the test equipment manufacturers to make these test sets affordable.

Correlation with standards

The standard organizations with our support should correlate the objective measures with the subjective results. Some of the organizations and vendors reported in [5] a correlation factor ranging from 84% to 95%. If the upper correlation factor could be achieved consistently, the industry would be able to perform objective testing with an affordable test equipment without relying on expensive subjective testing. The subjective testing would have to be repeated to achieve new correlation figures when a new technology is introduced or customer expectations change.

Limits

When the correlation is known, we can set impairment limits (parameter values) to define several categories of service at different quality levels, much the same as EIA standard [8] defines it now.

Common Standards

The authors, while reviewing document [2], noticed that many if not all of the parameters listed there (see Table 1) accommodate subjective measures of the picture quality. Moreover, their nature does not seem to be digital-specific. A natural question arises whether these parameters can be used to characterize any video quality whether digital or analog. A positive answer to this question would bring about a uniform standard of evaluating video quality instead having two

[1] The difference due to the difficulty of subjective assessment of high SNRs.

separate standards, one for analog video and one for digital video. This would make the testing process and test equipment independent of the transmission technology.

CONCLUSIONS

Several conclusions become clear. The first is that quantization noise in cascaded linear PCM systems (encode/decode) adds objectively quite close to way the thermal noise does. In short cascades it added exactly on a 10•log basis. For longer cascades noise added on a slightly higher than 10•log basis. The quantization noise that predominated the noise created by these systems presented similar subjective disturbances as thermal noise.

Quantization noise cascaded with thermal noise on a 10•log basis both objectively and subjectively. When the source video had thermal noise added prior to digitization, the addition effects were much worse and unpredictable.

Several compression artifacts, especially those that appear as high frequency busyness, also add subjectively to thermal noise in much the same way as thermal noise does. These artifacts are difficult to measure objectively. Also, these disturbances are often not specified and even when they are, they are typically not calculated into the overall noise budget.

All noise sources must be considered when planning network architecture. Historically these noise sources have been primarily limited to thermal noise, but today they come from a variety of analog and digital sources, all of which accumulate and all of which must be considered in your end to end performance budgets.

ACKNOWLEDGMENT

The authors want to tank Alain Belanger of ABL, Jeff Tokar of Antec, and Jim Farmer of ESP for their contribution in preparing the test processes and for their participation during the tests. Alain and Jim were also active in supporting the authors during the result interpretation process.

References:

[1] Guy W. Beakly (StellaCom), "How to Test Compressed Digital Television", 1994 NCTA Technical Papers.

[2] ANSI T1.801.0x-yyyy, "Digital Transport of One-Way Video Signals - Parameters for Objective Performance Assessment" (Draft).

[3] EIA/TIA Standard 250-C, "Electrical Performance for Television Transmission Systems", 1990.

[4] CCIR Recommendation 500, "Method for the Subjective Assessment of the Quality of Television Pictures", 1990.

[5] Daniel D. Briere, "Quality is in the Eye of the Beholder", Telephony, January 2, 1995.

[6] T1Q1.5 Technical Report No. 16, "A Technical Report on DS3 Transport for Contribution Application of System M-NTSC Television Signals Analog Interface and Performance Objectives", November 1992.

[7] Tony E. Werner, "Regional and Metropolitan Hub Architecture Considerations", SCTE 1995 Conference on Emerging Technologies.

Table 5: Cascading of Compression Artifacts and Thermal Noise

Codecs	Test Pattern	Measured SNR (Ramp) in dB (unified weighting)	Subjective SNR in dB	Thermal Noise Level	Measured Cascaded	Subjective Cascaded (measured)	Subjective Cascaded (calculated)
#1	Several patterns	60.8	57.5, 60.7, & 62.3	NA	NA	NA	
#1	Snell & Wilcox	60.4	52.9	NA	NA	NA	
#2	Snell & Wilcox	58.5	51.4	49.1	48.7	46-48	47.1

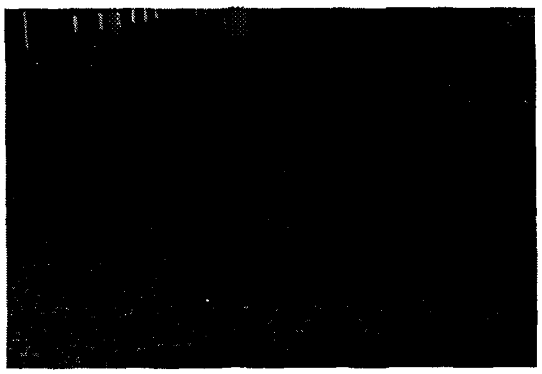
Table 6: Cascading of Compression Artifacts and Thermal Noise

Thermal Noise Level (SNR *unified weighting) in dB*	Codecs #3 (SNR at 59.3 dB, unified weighting)		Codecs #4 (SNR at 62.4 dB, unified weighting)	
	Test Pattern		Test Pattern	
	Zone Plate	Flower Bed	Zone Plate	Flower Bed
Subjective Assessment of SNR in dB	41.7	48.3	61	59.3
62	no degradation		clearly noticeable degradation (≅3 dB)	noticeable degradation
59	no degradation		significant degradation	clearly noticeable degradation (≅3 dB)
56	no degradation	barely perceptible degradation	thermal noise dominates	significant degradation
53	barely perceptible degradation	perceptible degradation		thermal noise dominates
50	perceptible degradation	noticeable degradation		
47	noticeable degradation	clearly noticeable degradation (≅3 dB)		
44	noticeable degradation	significant degradation		
41	clearly noticeable degradation (≅3 dB)	thermal noise dominates		

Figure 2: Test Patterns and Scenes

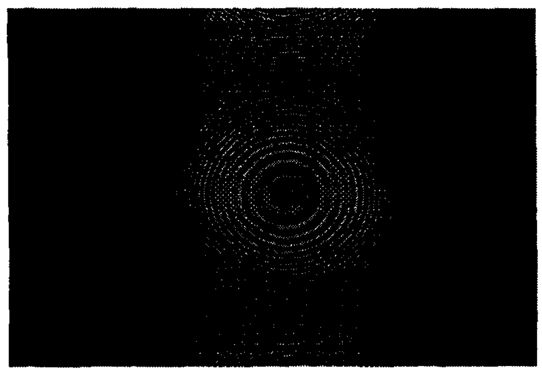
a) **Carousel**

b) **Flower Bed**



c) **Moving Zone Plate (Snell&Wilcox)**

d) **Still Zone Plate**



CHARACTERIZATION OF PEAK FACTOR EFFECTS IN HFC SYSTEMS USING PHASE CONTROLLED CARRIERS

Darryl Schick
RDL
Conshohocken, PA

Abstract

Traditional methods used to characterize multi-carrier distortion products are insufficient for the needs of emerging HFC systems. Noncoherent (free-running) oscillators, which have generally been used to test CTB and CSO products, do not provide the means to characterize performance of HFC system components under various conditions of peak power factor. Since HFC systems are particularly sensitive to distortions and signal clipping resulting from intermittently high peak factors, the ability to control the peak factor of the multi-carrier test signal is crucial.

New test methodology is described that allows the characterization of system components using a multi-carrier signal with precisely controllable peak factor.

BACKGROUND

CATV transmission systems generate thousands of intermodulation distortion products as a result of the multitude of signals that impinge on amplifiers and other nonlinear system components. These distortion products are typically known as composite second order (CSO) and composite triple beat (CTB) products reflecting second and third order distortions respectively.

The level of distortion products generated when a multi-carrier signal passes through a nonlinear component is primarily a function of the power envelope of the multi-carrier signal, rather than a function of the signal's average power level. This is most easily understood by observing the power envelope of a multi-tone signal under different conditions of carrier phasing. A multi-carrier CW signal is represented as:

$$v(t) = \sum_{i=1}^{N} \cos[\,\omega(\mathrm{i}) * t + \theta(i)]$$

The power envelope of a combined 16-carrier signal is shown in figure 1(a) with all phase terms equal to zero. In figure 1(b), the phase terms are randomly distributed. In figure 1(c), the phase terms are adjusted to produce maximum destructive interference. All three waveforms have identical average power levels and would appear identical on a spectrum analyzer, but the peak power factor (peak power to average power ratio) of these waveforms ranges from 15 dB in figure 1(a) to as low as 5.6 dB in figure 1(c). It would be expected that an amplifier will generate higher distortion products when driven from the signal in figure 1(a) than when driven from the signal in figure 1(b) because the amplifier is driven to higher levels of saturation. Similarly, the amplifier will generate even lower distortion product levels when driven from the signal in figure 1(c). As will be shown, CTB products do vary considerably in amplitude in response to varying conditions of carrier phase.

Amplitude vs. time of 16-carrier signal

Fig. 1(a): Carrier phases set to zero

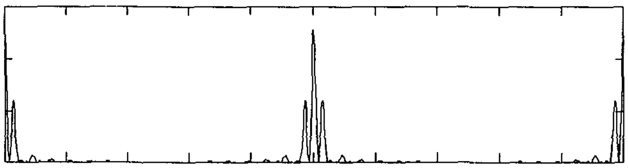
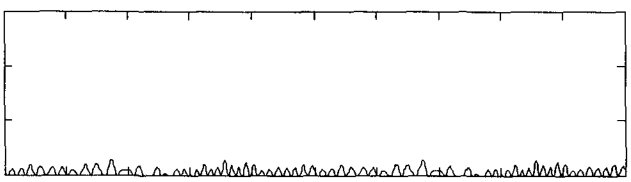


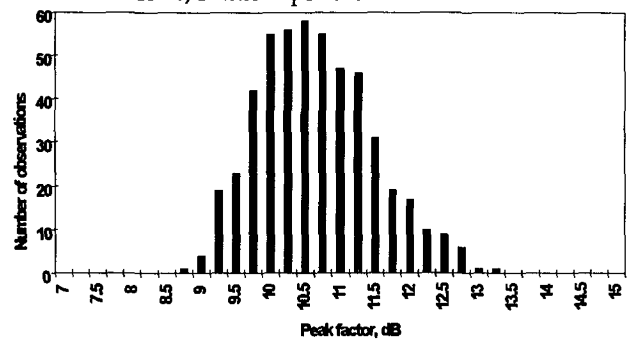
Fig. 1(b): Carrier phases randomly distributed

Fig. 1(c): Carrier phases set for minimal peak factor



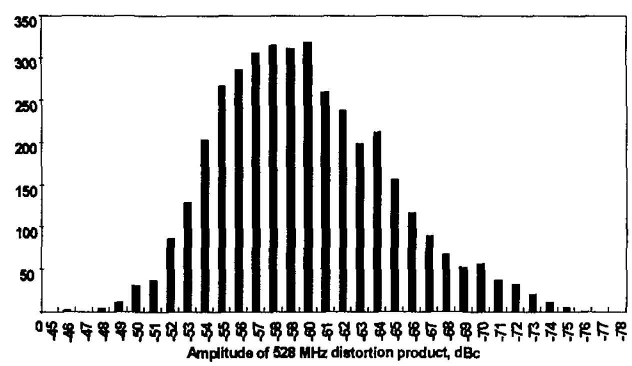
When examined over long time periods, the peak factor of a noncoherent head end varies continuously and is randomly distributed. Coherent head ends also have randomly distributed phases which drift slowly with time. Depending on the phases of the carriers, peak factor of an 83-tone signal can range from a minimum of about 3 dB up to a maximum of 22.2 dB if all carriers have zero phase.[1] The actual peak factor distribution of an 83-tone signal with random phases is shown in figure 2. Within a reasonable certainty, observed peak factors will randomly fall between 8.5 dB and 13.5 dB.

Fig. 2: Histogram of peak factor distribution, 83 carriers, random phases



CTB and CSO product amplitudes vary considerably in response to variations in the peak factor of the multi-carrier drive signal. Figure 3 shows a histogram of 3800 measurements of a 528 MHz distortion product generated when a HRC frequency spectrum was used to drive two standard broadband CATV amplifiers. The measured distortion product varied over nearly a 40 dB range.

Fig. 3: Histogram of measured distortion product at 528 MHz



This result correlates with a common observation: when distortion products resulting from a noncoherent drive signal are observed on a spectrum analyzer, the amplitudes fluctuate randomly over a considerable range. In order to make a measurement, the convention is to use a narrow video bandwidth (typically < 300 Hz) in order to resolve the "average" distortion product level.[2] This method does not therefore resolve the worst peak level that may occur in the multi-carrier drive signal.

While analog CATV systems may be merely degraded by the presence of distortion products, HFC systems employing both analog and digital transmission can be entirely corrupted by peak clipping to the point where BER thresholds are exceeded. HFC systems are more sensitive than analog systems in this respect because a "hard" failure may occur.

Proper characterization of HFC components is essential to insure adequate system performance. When measuring distortion, traditional CSO and CTB measurements do not adequately address the problem of instantaneous peak power spikes. Therefore, the multi-carrier drive signal used for distortion measurements must provide not only the full signal bandwidth, it must provide controlled level of peak power factor.

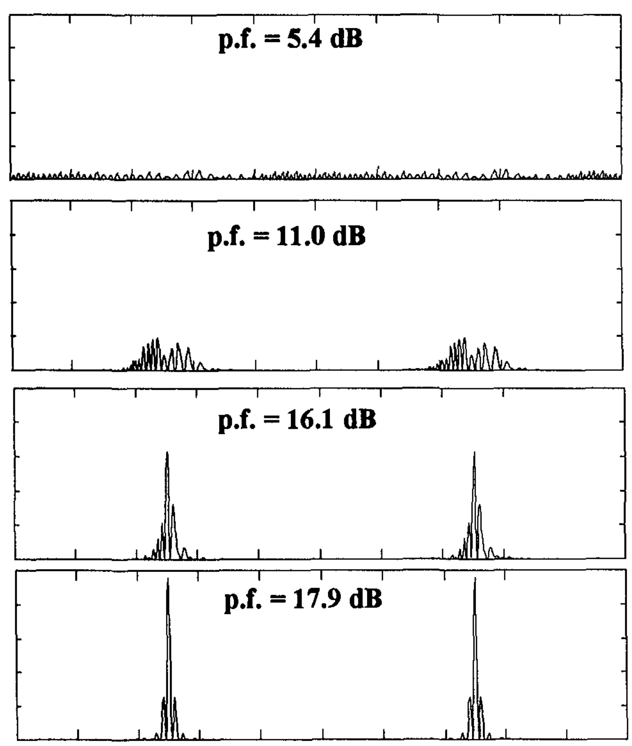
CARRIER PHASE CONTROL

A multi-carrier generator has been developed by RDL which has programmable carrier phase capability. By programming the carrier phases, peak factors can be established over a continuous range from 3 dB minimum up to the maximum theoretical peak factor of 10*log(2*N), where N is the number of carriers. Coherent (phase locked) carriers provide a stable, repeatable drive waveform which is statistically stationary over time.

The selection of carrier phases to provide a particular desired peak factor may be approached in a number of ways. There are infinitely many combinations of carrier phase which will produce a given peak factor. For example, random combinations of phase could be attempted until the desired peak factor is attained. This method is undesirable, however, because it would only produce peak factors over a limited range. Furthermore, the cumulative distribution function (cdf) of peak power level vs. time can take on many forms for a given peak factor, and the cdf of the signal also plays a role in determining the distortion spectrum.

To provide a methodical approach to controlling peak factor, RDL has developed algorithms which can be used to vary the peak factor directly, over its full theoretical range, while maintaining a relatively symmetric signal and consistent cdf. An example of the range of programmable peak factors is shown in figure 4.

Fig. 4: Examples of the range of programmable peak factors for an 83-tone signal.



The proper peak factor to be used for testing HFC components is determined by the system performance requirements. A suggested starting point would be to set the peak factor to a level near the upper tail of the peak factor distribution curve for the given number of carriers. For example, in the 83-carrier spectrum, the likely range of observed peak factors is 8.5 dB to 13.5 dB. Testing with a consistent 13.5 dB peak factor would represent the worst case condition that is likely to be observed. From that point, the peak factor can be adjusted to provide the optimal tradeoff between false rejections, improper acceptances, manufacturing yields, and required levels of system data reliability.

Another benefit of testing with controlled peak factors is measurement repeatability. Once the optimal peak factor and carrier phase combination is established, the same condition can be repeated on numerous multi-carrier signal generators so that results can be better correlated between vendor and customer.

HFC TEST METHODOLOGY

In addition to the traditional CSO and CTB measurements, dynamic operability tests are suggested for HFC components. These tests may take the form of BER testing or observation of waveforms with a high frequency sampling oscilloscope to observe signal clipping.

Dynamic measurements such as these are not practical using noncoherent carriers because there is no repeatable "trigger point" where the peak power spikes can be observed. Using coherent, phase controlled carriers insures that a specific peak power level is repeated at every waveform cycle. The waveform cycle for a CATV spectrum with 6 MHz spacing is $(6 \times 10^6)^{-1}$ or 166 nanoseconds.

Using multi-carrier signals with controlled peak factor, these tests can provide much insight into the effects of peak factor on HFC systems, and can provide more repeatable, meaningful, and accurate characterization of critical system components.

References

1. Schick, Darryl, "Source Tests Next Generation CATV Systems," *Microwaves&RF* , January 1996, Cover story.

2. National Cable Television Association, "NCTA Recommended Practices for Measurements on Cable Television Systems," Second Edition, 1983, I.B. page 4.

Cost Effective Point to Point AM Fiber Links Using High Power Optical Receivers

H. Blauvelt, J. Paslaski, and J. Lipson
Ortel Corporation

Abstract

Fiber optic links for transmitting AM video have typically operated with received optical power levels of about 1 mW. In the past, the most common practice was to split the laser output to serve multiple optical nodes. The higher the output power of the laser, the more optical splitting that could be done, but the received optical power was always approximately 1 mW.

Recently there has been much interest in point to point links without any optical splitting. These links commonly have optical loss budgets in the 2-6 dB range. With conventional 10 mW DFB lasers, the received optical power greatly exceeds 1 mW. To avoid overdriving the receiver, low power DFB lasers are used for these point to point links. Although point to point transmitters have lower output power requirements, the linearity demands on the laser are approximately the same as for high power DFB lasers.

An alternative approach to point to point links is to use medium power (6-10 mW) lasers and higher received power. If the modulation depth is kept the same, this results in significantly higher link C/N. Alternatively, the modulation depth can be reduced to provide the same C/N, with substantially relaxed linearity requirements for the laser. In this paper, the factors influencing receiver linearity and designs capable of operating up to 5 mW received power will be described. The impact of high optical power receivers on laser linearity requirements and on the channel capacity of AM fiber links will also be discussed.

Introduction

Most of the optical transmitters deployed for AM fiber optic links have used 1310 nm DFB lasers. The most common use of 1310 nm DFB technology has been to use 10-15 mW DFB lasers with the output optically split to serve multiple nodes. As the cost of 1310 nm DFB transmitters has dropped, it has become economically viable to use DFB transmitters that are dedicated to a single node. By doing so, the interactive bandwidth available to each node can be greatly increased. The loss budgets for these links are greatly reduced when there is no optical splitting. For this reason, the lasers used for point to point links have had much lower output power levels than lasers used with splitting. In this paper, several of the more important cost and yield issues related to the production of point to point lasers will be discussed. In particular, the impact of higher power optical receivers on point to point link design will be examined. High optical power receivers also can play an important role in increasing the channel capacity of AM fiber optic links.

Point to Point AM Fiber Links

Most point to point AM fiber links have optical loss budgets of 2-6 dB. If optical receivers are constrained to a maximum received optical power of 0 dBm, then the maximum laser power is 2-6 dBm (1.6-4 mW). In most cases, the linearity requirements for point to point lasers are identical to those of lasers used with optical splitting. There has been an expectation within the CATV industry that the cost to produce low power lasers would be significantly less than the 10-15 mW lasers used with optical splitting. This would be true if the yield for a 4 mW laser was much higher than that of a 10 mW laser. However, for standard butterfly modules, almost all 1310 nm DFB lasers are now capable of at least 8 mW. The distribution of optical power for a randomly selected population of 1310 nm DFB modules is shown in Figure 1. Recent advances are pushing this distribution to even higher power. The power shown is the power at which the laser module linearity is optimum. As can be seen, none of the modules had an optimum power less than 8 mW. The standard method for producing a low power laser module is to take one of these higher power modules and add an optical attenuator to bring the power down to the required range. Clearly, this does not result in cost savings. It is possible to achieve some module cost savings by using lower performance optical components, but the savings are generally not large. A much more attractive solution would be to use the power capabilities of the DFB lasers to improve the performance capabilities of the AM links or to relax the linearity requirements of the lasers. This requires high optical power receivers.

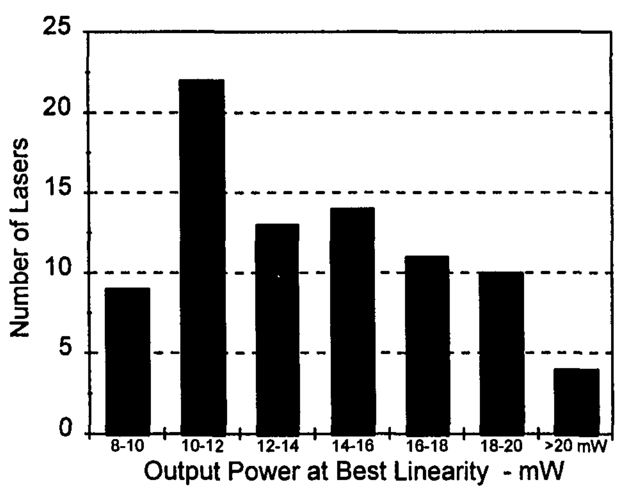


Figure 1. Distribution of optical power for cooled butterfly modules biased for optimum linearity.

If the received optical power of an AM fiber link is increased, then the link C/N increases. The C/N vs received optical power for several link lengths is shown in Figure 2. These calculations are for 80 channels with an optical modulation index (OMI) of 3.5%/ch. The chirp was assumed to be 100 MHZ/mA and the receiver noise was the value appropriate for each received power level (see Figure 5).

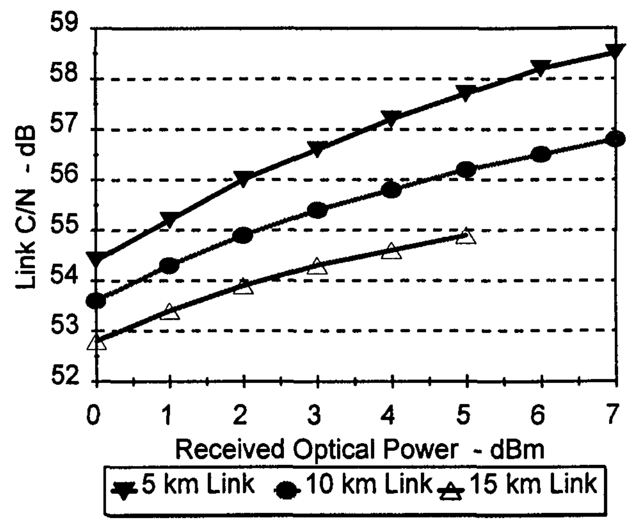


Figure 2. Link C/N vs received optical power for various link lengths.

Figure 2 shows that significant improvements in C/N are obtained when the links are operated at high received power. No results are shown above +5 dBm for the 15 km link because this would require laser output power levels that are sufficiently high that the yield for optical power would become important. The C/N improves as received power increases because the output signal level increases faster than the output noise level. In the range of 0 dBm, the dominant noise source for AM fiber links is photodiode shot noise. The output signal level increases 2 dB for every 1 dB increase in received optical power while the output noise level due to shot noise only increases 1 dB. Most of the benefit of higher C/N is achieved by increasing the received power to the 5 dBm range. At higher received power levels, most of the noise in the link comes from laser RIN or interferometric noise, which increase at the same rate as the signal. Above the power level where these noise sources become dominant there is no significant improvement in link C/N. Figure 2 also shows that although there may not be a direct cost benefit from using lower power lasers, there is a definite benefit from shorter link lengths. The shorter the fiber length, the less degradation there is to the C/N from interferometric noise due to double backscattering within the fiber.

There are many potential benefits from operation at high received power. One benefit is to simply make use of the higher C/N. Another option is to decrease the optical modulation index (OMI) needed to achieve a desired link C/N. The OMI needed to achieve a C/N of 52 dB is shown in Figure 3 for several link lengths. The link assumptions for these calculations are the same as in Figure 2. When the OMI of a laser is reduced, the linearity improves or the channel capacity can be increased. Linearity, and in particular, CSO, is the most important yield and cost factor for 1310 nm DFB lasers. High received power can therefore be used to relax the linearity requirements of the lasers, which substantially reduces the manufacturing costs for AM lasers.

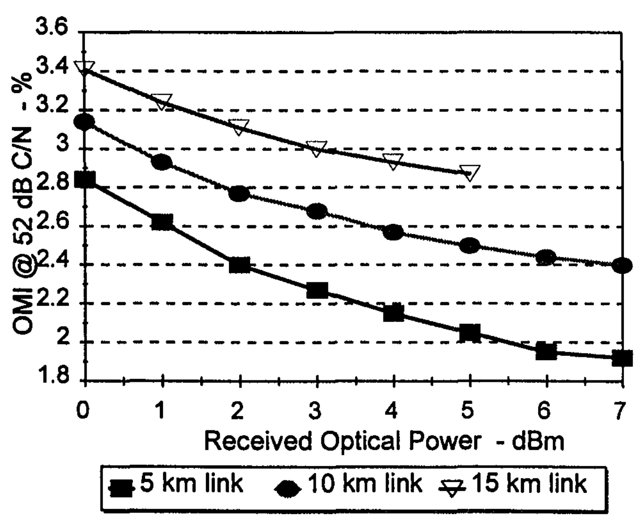


Figure 3. Optical modulation index required to achieve 52 dB C/N vs. received optical power.

For a given OMI per channel, the channel capacity of AM fiber links is limited by laser clipping. If the RF signal modulating the laser is too large then the laser will occasionally be driven down to zero power by large negative peaks in the multichannel signal. Such clipping events cause picture impairments in the AM channels and errors in digital channels being transmitted. By lowering the OMI per channel, the number of channels that can be transmitted before reaching clipping is increased. Figure 4 shows the limit to the channel capacity determined by laser clipping vs received optical power, for a link C/N of 52 dB. In this figure, the RMS modulation depth of the laser, due to the AM channels, is restricted to 23% to allow for the addition of digital channels and still avoid laser clipping. The link assumptions for these calculations are the same as in Figures 2 and 3 except the number of channels is varied. Higher capacity can be achieved using lasers with higher chirp, but the CSO can be degraded due to fiber dispersion[1].

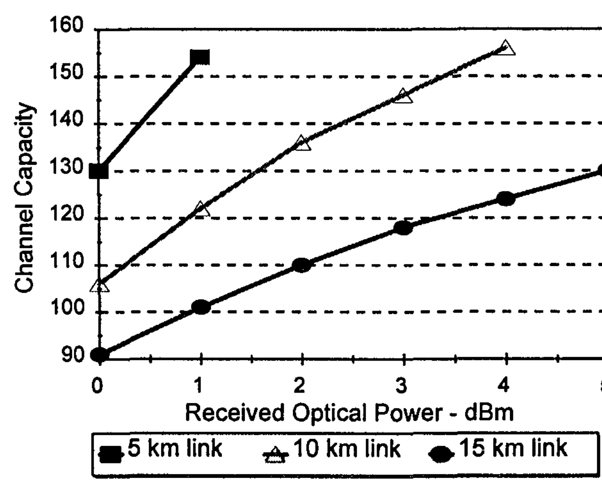


Figure 4. Channel capacity of AM fiber links vs. received optical power.

High Power Receiver Design

The preceding sections clearly demonstrated the benefits of high received optical power to AM fiber links. The only disadvantages are that the linearity, and possibly the reliability, of the receiver are degraded when operated at high power. Until recently, most optical receivers have had rated power levels for linearity of about 0 dBm and conservative absolute maximum ratings of 3 dBm. Recent improvements in the power handling capabilities of linear photodiodes has increased the maximum reliable received power by approximately 5 dB with excellent reliability being demonstrated for power levels in excess of 10 dBm[2].

The second challenge with high power receivers is maintaining linearity. Distortion is produced in optical receivers by both the photodiode and the amplifier following the photodiode. Photodiodes produce primarily second order distortion. With proper design, highly linear photodiodes can operate at received optical power levels up to at least +7 dBm. The second distortion source is the amplifier. In general, the amplifier distortion can be made arbitrarily low by using high power amplifiers or by reducing the gain and therefore the output level of the amplifier. However, both of these degrade receiver noise performance. The fundamental trade-off in high power receivers is therefore between the rated optical power of the receiver and the noise performance of the receiver. We have recently built and tested a family of optical receivers optimized for power levels between 0 and +5 dBm. For this discussion, rated optical power is defined to be the maximum power for which the CSO and CTB of the receiver exceed 75 dB for 80 channel operation at 3.5% OMI. The receivers for 0-2 dBm optical power use low noise integrated amplifiers and have an RF output level of 18 dBmV/ch for 80 channel operation. The receivers for operation at greater than 2 dBm received power use higher power amplifiers and have an output level of 22 dBmV/ch for 80 channel operation. The typical receiver noise versus rated optical power for receiver designs within this family is shown in Figure 5. The dashed portion of the curve represents projected performance for rated power levels up to +7 dBm. In this figure, both the noise of the amplifier integrated with the photodiode and the noise of the following gain stage are included. Although the receiver noise current increases for the high power receivers, the output signal levels and other noise sources, such as interferometric noise, increase much more rapidly at high received power. The fraction of the overall link noise due to the receiver is totally negligible for high power receivers.

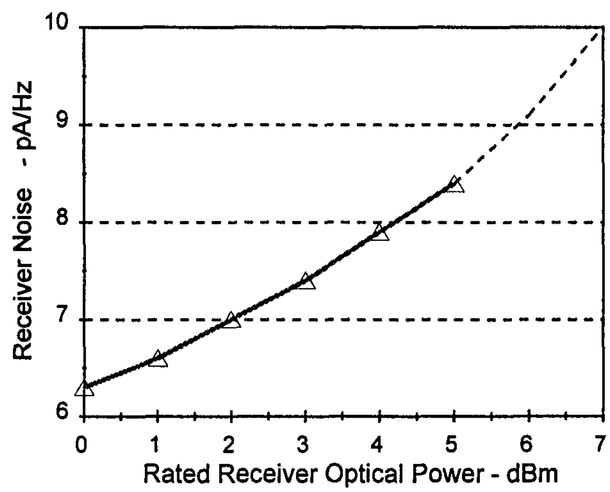


Figure 5. Trade-off between rated optical power and measured receiver noise for optical receivers with integrated amplifiers.

Conclusions

Until recently, most 1310 nm lasers were used with optical splitting to serve multiple nodes. If these lasers are used without splitting using existing optical receivers, the receivers will be overdriven resulting in poor receiver linearity. The current solution to this problem is to use low power lasers. However, low power lasers are almost always simply high power lasers to which an optical attenuator has been added or which are intentionally decoupled. In this paper, an alternate approach that uses the full power capabilities of 1310 nm DFB lasers was discussed. This requires the use of receivers that can operate linearly at high received optical power. Operation at high received power increases the link C/N. If higher C/N is not required, then the use of high power receivers can allow for a relaxation in the linearity requirements for the laser, which can substantially reduce the laser cost. Alternatively, high received power can be used to increase the channel capacity of the AM links. Most of the benefit of high received power is obtained by operating in the range of 5 dBm, but receivers for up to at least 7 dBm appear to be practical.

References

[1] H. Blauvelt, "Laser Chirp in Linear Fiber Optic Links" CED, Sept. 1995, pp. 72-76.

[2] J. Paslaski, et.al. "High Power Microwave Photodiodes for High Performance RF Link Applications" ARPA Symposium on Photonic Systems for Antenna Applications, Jan. 1995, p.628.

Deploying ATM Residential Broadband Networks

Mark Laubach, Chief Architect
Com21, Inc., 1991 Landings Drive, Mtn View, CA 94043
laubach@com21.com http://www.com21.com/

ABSTRACT

ATM residential broadband technology is rapidly entering commonplace discussion. The capabilities provided by ATM network interface devices promise data bandwidth speeds far in excess of those provided by traditional twisted pair public telephone networks. Cable TV operators and Regional Bell Operating Companies (RBOCs), e.g. PacBell, are preparing for this integrated broadband future by installing or rebuilding existing all-coaxial cable plants into two-way Hybrid-Fiber Coaxial plants and by offering a wide range of both data and interactive services which they feel will be most attractive to their subscriber base. Initially these services will only provide Internet access and access to the major information service (e.g., Compuserve, AOL, and Prodigy). These service offerings will quickly advance to support multi-player gaming and collaborative services such as voice and desktop video teleconferencing.

As an introduction to some of the issues surrounding ATM residential access technology, this paper summarizes two of the standardization efforts: the ATM over HFC definition work taking place in the ATM Forum's Residential Broadband Working Group and the standards progress in the IEEE P802.14 Cable TV Media Access Control and Physical Protocol Working Group. Delivering ATM-based integrated services via a Cable TV has its own set of deployment issues and benefits that are briefly overviewed and summarized.

INTRODUCTION

Packet technology has been around since 1964 [8]. Since then, the size of packets has been debated, as well as variable vs. fixed size. Packets are transmitted over any media these days however, the next economic and technical frontier is moving packets over Cable TV (CATV) networks. The driving push is to deliver IP datagrams over cable TV networks. There are several datalink methodologies for delivering IP datagrams via cable modems, This paper overviews the notion of sending small, fixed sized packets over the CATV plant. These small packets are 53-octet Asynchronous Transfer Mode (ATM) cells [1]. In addition to meeting this *Internet rush*, ATM provides an integrated services base on which to offer additional services.

The term *cable modem* has been associated mostly with a description of a CATV Internet access device located in the home. ATM based systems allow the deployment of other services beyond basic Internet access. In this latter use, the ATM-based cable modem can be viewed as a broadband services interface for the home, or an ATM Network Interface Device (NID) or an ATM Premises Interface Device (PID). These terms are synonymous in their use in this paper.

Numerous standards organizations are gearing towards producing cable modem standards. The ultimate goal of each is to drive cable modem availability to commodity status and made available via consumer *off the shelf* purchases at computer boutiques and electronic supermarkets. The minor problem with the commodity process is that these numerous standardization activities are competing and largely uncoordinated and there are about a dozen cable modem manufacturers producing product, some of whom wish to establish defacto standard status by being first to market.

The IEEE P802.14 Cable TV MAC and PHY Protocol Working Group is chartered with providing a single MAC and multiple PHY standard for cable TV networks. The efforts of P802.14 must support IEEE 802 layer services (including Ethernet) and must also be *ATM Compatible.* ATM residential broadband work is currently taking place in the ATM Forum.

The customer network interface *du jour* is Ethernet 10BaseT. There is a mandate for a 10Mbps Ethernet interface in the home. Subscriber access equipment can be a personal computer, X-Terminal, or any such device which support the TCP/IP protocol suite.

Engineering Challenges of ATM over Cable

The standardization and implementation of two-way interactive services on Hybrid Fiber-Cable (HFC) TV networks is fraught with many engineering problems which must be overcome:

- CATV systems are inherently asymmetric in nature, i.e. there is more downstream bandwidth available than upstream and interactive services such as voice or video telephony require symmetric data rates. ATM based services will create an environment where symmetric virtual circuits will be opened and closed frequently.

- High utilization of the upstream bandwidth is necessary to be cost effective and accomplished by sharing bandwidth between stations with the access based on dynamic assignments within a slotted regimentation.

The choice of the allocation protocol and the placement of the bandwidth ownership intelligence is important. A straightforward approach is to place the ownership of the upstream bandwidth under the direction of the head-end controller which is tightly coupled with the ATM traffic and signaling management processes. This also has the effect of reducing complexity in the subscriber unit and by centralizing the allocation intelligence in the network. Communications between the head-end controller and each subscriber unit is important as permission to use the upstream channel is granted by the head-end controller whose allocation algorithm must take into account the needs communicated to it by each subscriber.

ATM in the Residential Broadband Network

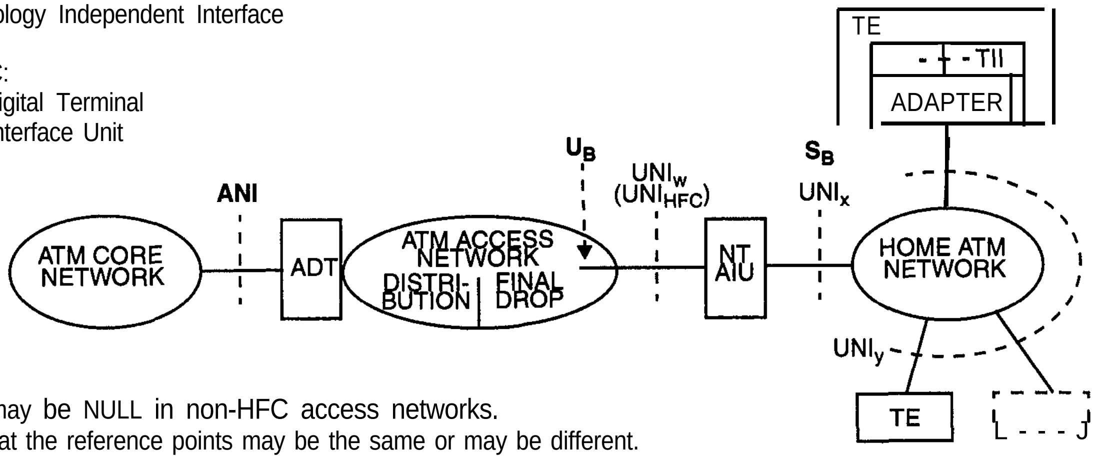
The selection of ATM cells as the data-link layer protocol data unit for Cable TV networks has the advantage in that it provides a suitable integrated multiplexing platform capable of supporting a mix of guaranteed (predictive) traffic flows with best-effort (reactive) traffic flows. In addition, the nature of ATM allows other multimedia applications to be added in the future without requiring iterative changes to the basic ATM protocol. Cable operators can deploy ATM systems as part of an evolutionary path to a fully integrated multimedia bearer service offering.

An ATM data-link protocol can be layered in a straightforward manner for both the downstream and upstream segments of a cable modem network, The challenges are that upstream traffic management and resource management must be creatively controlled to support the guaranteed and best effort Quality of Service (QOS) needs of the cable modem. A residential ATM bearer service easily supports Internet access to the home via the Classical IP over ATM standards of the Internet Engineering Task Force [3] or by providing an IP over Ethernet adaptation overlay service.

While ATM in the home is desired as a future interconnection method by some HFC operators, the cost burden of the ATM interface is not economically feasible today. It is expected that ATM network interface controllers will be decreasing in cost quickly over time so planning a cable modem bear service now to

GENERAL:
ANI = Access Network Interface
TII = Technology Independent Interface

HFC SPECIFIC:
ADT = ATM Digital Terminal
AIU = ATM Interface Unit



NOTES:
1. The NT may be NULL in non-HFC access networks.
2. Interfaces at the reference points may be the same or may be different.
3. The HAN contains physical media and passive devices. It may also contain active devices; e.g. bridges or switches
4. There will be more than one PMD and MAC layer specified for the UNI reference points

Work in Progress - Based on RBB Reference Configuration - August 1995 - Version 3.0

Figure 1. ATM Forum Residential Broadband Reference Model

support both Ethernet and ATM home interfaces can be viewed by some as prudent.

ATM FORUM'S RESIDENTIAL BROADBAND WORKING GROUP

The ATM Forum is focusing attention on delivering ATM over residential broadband distribution systems. This work is being carried out in the Residential Broadband (RBB) Working Group (WC). The material presented in this section represents work in progress in the ATM Forum and is offered as an example of the current thinking on the subject. At some time in the future, the ATM Forum will be producing a published specification which includes the ATM over HFC UNI details. The ATM Forum is a closed industrial consortium requiring membership dues for participation.

The two goals of the RBB WG are to 1) deliver ATM to the home and 2) deliver ATM within the home. These can be euphemistically termed as *the last mile* problem and *the last yard* problem. The current proponents of ATM over HFC systems are concerned with delivering a full function UNI interface to the home via an active Network Interface Unit (NIU) termed an ATM Interface Unit (AIU). Controlling the system is a ATM Digital Terminal (ADT) located at the cable system head-end (see Figure 1) The discussion of ATM within the home is beyond the scope of this paper.

The ATM Network Interface (ANI) defines the connection between the ADT and the ATM WAN network. This interface may either be specified as a Network-Network Interface (NNI) or as a UNI. The AN1 will be based on existing ATM standards and the WG expects complete compliance with existing physical (PHY) interface standards.

The HFC access network is in effect a *black box* to the ATM Forum's design activities. It was decided early in the RBB charter process, that the RBB WG will rely on the efforts from IEEE P802.14 Cable TV MAC and PHY Working Group for the transport of ATM cells over the HFC network. The UNI-HFC will define an RF interface for the ADT and AIU. A possible protocol stack representation of the relationship

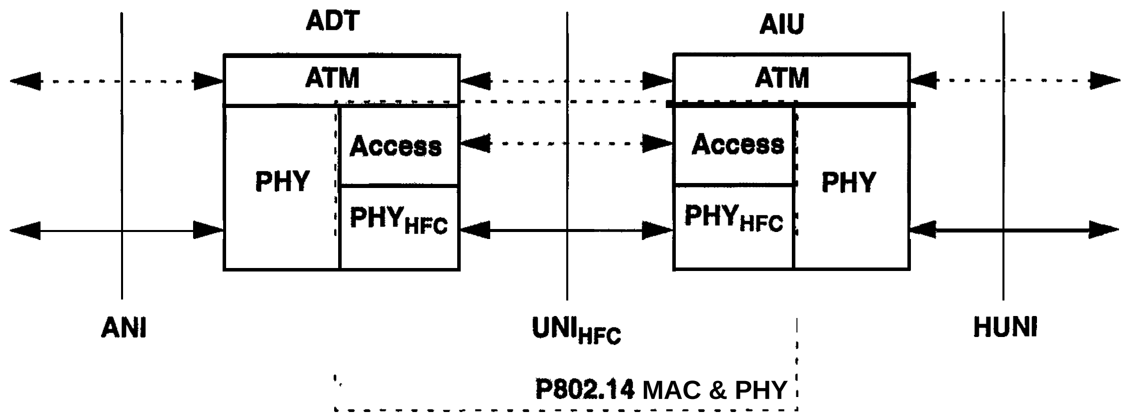


Figure 2. Proposed ATM Transport Protocol Model

between ATM and P802.14 is shown in Figure 2. It has been suggested that communication between the ATM layer and the IEEE access layer be specified via an abstract layer interface definition, which will be referred to as the access interface in this paper.

The AIU provides an ATM UNI to the home. This Home UNI (HUNI) interface is meant to be as standard as possible. It will most likely provide a subset of UNI 3.1 or the upcoming UNI 4.0. It is expected that this interface will deliver the full range of CBR, VBR, ABR, and UBR services. However, the real performance of these will be limited by the available performance offered by the HFC access network and the characteristics of the underlying P802.14 MAC and PHY.

The RBB WG effort is current work in progress. It will produce an implementation reference (i.e., a UNI_{HFC} implementation reference) which is synchronized with the IEEE P802.14 working group. At the time of this writing, the following issues will need to be resolved within the RBB:

- How Virtual Path Identifiers (VPIs) and Virtual Circuit Identifiers (VCIs) will be used within the UNI-HFC and how they are mapped to the IEEE access interface.
- Where and how does ATM UNI Traffic Management (TM) take place in the HFC system, and the nature of the QOS or TM interface to the IEEE access interface.
- What form of UNI signaling will be supported between the ADT and the AIU. The AIU might be passive requiring the ADT to perform proxy signaling on behalf of the home UNI. If all AIU interfaces share a common VCI space, then meta-signaling may be required, etc.
- Will the ATM over HFC system specification include telephony voice over ATM and if so, with what Cell Delay Variation (CDV)?
- What are the required performance goals for ATM peer-to-peer networking when operating over an HFC network.?

The above issues and more will be addressed in the ATM Forum's efforts.

IEEE P802.14 CABLE TV MAC AND PHY PROTOCOL WORKING GROUP

In November, 1994 the IEEE P802.14 CATV MAC and PHY Protocol Working Group met for the first time as an approved project within the 802 standards committee. Previous work had been done in the 802.catv study group in preparation for formal approval. The Project Authorization Request (PAR) charter of the group specifies that it will standardize a single MAC layer protocol and multiple PHY layer protocols for two-way HFC networks. Consistent with the IEEE LAN/MAN 802 Reference Model [5], P802.14 will produce a solution which supports the 802 protocol stack while at the same time supporting ATM in an *ATM Compatible* manner.

The WG has completed a first release revision of a functional requirements document [4] which details the P802.14 cable topology model; defines key assumptions, constraints, and parameters; defines key performance metrics and criteria for the selection of multiple PHY protocols and a MAC protocol; and defines the support of QOS parameters. The working group's work plan called for the close of formal proposals in November 1995, with the recommended protocol defined in July 1996. Seventeen MAC protocol proposals were submitted to the working group. It is anticipated that the WG will select the best features from amongst the proposals and apply appropriate glue to form the standard. The IEEE is a public standards organization and anyone may participate in the standards activities.

The branch and tree topology of a Cable TV single-cable distribution network is divided by RF frequency into a downstream portion (typically 50 Mhz to 550 Mhz or 750 Mhz) and an upstream portion (typically 5 Mhz to 40 Mhz). Both downstream and upstream frequencies are active on the same physical coaxial RF cable, and the use of bandpass filters and diplexors provide the spectral separation necessary for the simultaneous amplification of signals in each direction. A P802.14 subnetwork can be thought of as a single head-end controller communicating with a set of cable modems. via a MAC protocol operating of a collection of downstream and upstream PHY channels (see Figure 3) Home Network Interface Units (NIUs) will communicate with the head-end terminal unit using the agreed upon downstream and upstream PHY. The downstream PHY will support a broadcast, one transmitter many receiver model. The upstream PHY will support a one receiver many transmitter model, requiring that the upstream PHY be shared amongst all participating NIUs in the subnetwork.

The general P802.14 requirements are: support of symmetrical and asymmetrical rates on connections involving the downstream and upstream channels; support of Operation, Administrations, and Maintenance (OAM) functions; support of one way delays on the

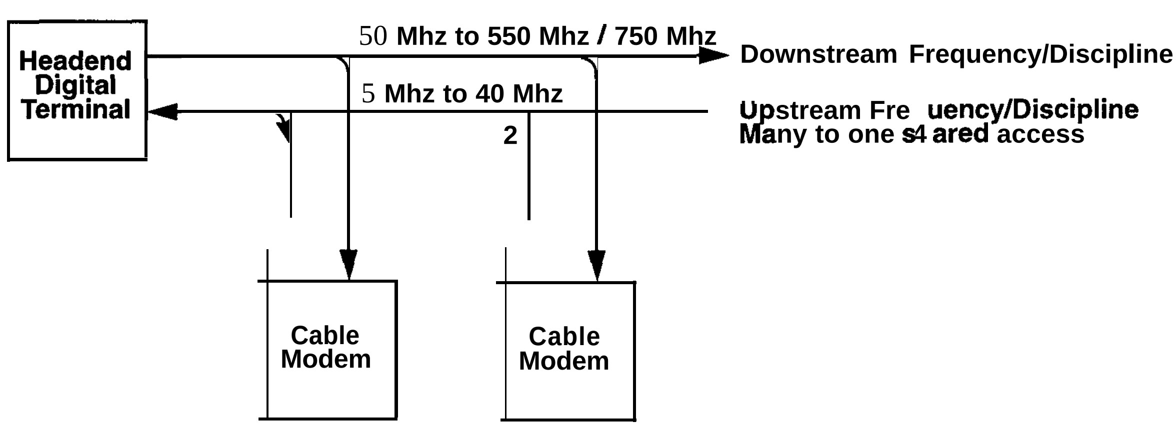


Figure 3. IEEE P802.14 Shared HFC Architecture

order of 400 microseconds (round trip delays to 800 microseconds) support of a large number of users; support for moving data from an originating sub-network to a destination subnetwork which may be the same or a different one; and the option of a customer reference point between in-home and external networks.

The P802.14 MAC layer requirements are: support of both connectionless and connection-oriented services; support of a formal QOS for connections; support for dynamically allocated bandwidth for different types of traffic, including Constant Bit Rate (CBR), Variable Bit Rate (VBR), and Available Bit Rate (ABR); support for unicast, multicast, and broadcast services; interoperability with ATM; predictable low average access delay without sacrificing network throughput; and fair arbitration for shared access to the network within any level of service.

The P802.14 PHY layer requirements are: HFC system size up to 500 households as a reference design point; primary support of sub-split cable plants (5Mhz to 40 Mhz upstream), with optional support of mid-split (5Mhz to -120 Mhz upstream) and high-split (-800 Mhz to ~ 1GHz upstream); frequency reuse in the upstream channel; and co-existence with other home information appliances (e.g., entertainment TV) and other uses sharing the broadband system.

The detailed performance requirements for the MAC and for the PHY have yet to be specified by the P802.14 WG. The majority of the MAC proposals received by the working group will be put to modeling and simulation performance scrutiny with the initial results presented at the March, 1996 meeting.

The P802.14 work plan has been recently finalized, setting the stage for a completely draft work by the end of 1996. This section has attempted to summarize some of the aspects of the challenge faced by the P802.14 WG. At the time of this writing, the following additional issues will need to be resolved by the WG:

- Will Plain Old Telephone (POTS) be a fundamental service; i.e. DSO at 64 Kbps?
- Will ATM be selected as the Protocol Data Unit (PDU)?
- What Forward Error Correction (FEC) algorithm will be used and how much protection?
- If a slotted approach is used, what is the size of the slots?
- Where will complexity be placed in the system? Putting it where is easiest to fix/maintain implies the head-end.
- Many-to-one sharing of a single upstream channel using a slotted approach requires ranging of the home NIUs. How precise will the ranging be and how will it be performed?
- Will the WG specify a set of PHY profiles that may be used or just one downstream and one upstream PHY?
- How will provisioning of authorized stations be performed by the MAC protocol?
- How will the downstream and/or upstream channels be encrypted and how will keys be managed?
- How does the MAC protocol handle errors?
- How will stations will identified in the sub-network?

The above issues are continually being discussed in the P802.14 working group. As of this writing, the P802.14 Working Group has tentatively decided to select Quadrature Amplitude Modulation (QAM) 64 as a mandatory protocol for the downstream channel, The use of

QAM16 and QAM256 are for future study. The modulation technique for the upstream channel is currently under debate in the working group. There is a reasonably likelihood that Quadrature Phase Shift Keying (QPSK) will be one of the selected types due to its ability to perform better in a low signal to noise environment and that there is past industrial experience with this method. The specific choice(es) will be made during the March, 1996 meeting.

Expected Downstream and Upstream Data Channel Rates

For downstream, the QAM64 technique is a 6 bit per Hertz coding scheme which yields a raw data rate of approximately 30 Mbps in a North American 6 Mhz wide standard video channel (data + guard bands). With FEC and the effects of framing, the actual usable information data rate is approximately 27 Mbps. The usable bandwidth is shared amongst all cable modems for both user and management traffic. When ATM is used at the protocol unit, the useful user data rate, the ATM payload rate is 23.9 Mbps.

If the upstream modulation technique is QPSK. The raw channel rate will be anywhere from 1.5 to 3.0 Mbps with the specific rate selected in the near future by the IEEE P802.14 Working Group. For example, a raw rate of 2.56 Mbps can be placed in a 1.8 MHz bandwidth allocation. The upstream channel requires a longer preamble and more FEC as compared to the downstream channel. The requirements for sharing an upstream channel between multiple modems requires the use of a guard band (dead time) is needed between packet bursts. The information data rate a single upstream channel will be approximately 2.0 Mbps. It is expected that several upstream channels can be used with the single downstream channel. The head-end controller will *place* cable modems on the appropriate upstream channels to facilitate load balancing and robustness needs.

An ATM Deplovment Example

An example of using ATM to provide a cable modem based residential Internet service is presented in this section.

Moving IP packets over ATM is straightforward. At the subscriber premises, the user network connection is the Ethernet 10BaseT twisted pair interface. At the home, this means services is delivered via IP over Ethernet. The cable modem must either perform a bridge-like service and transmit Ethernet frames over ATM, or it must terminal Ethernet at the cable modem (like an IP router) and transmit IP datagrams over ATM. Either of these models is support by an ATM based service. Figure 4 illustrates the

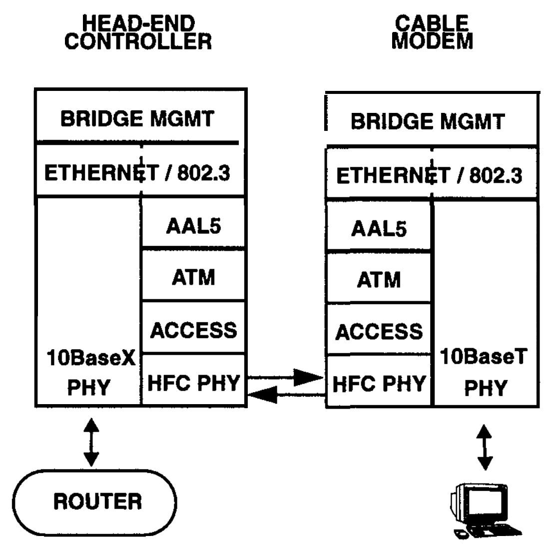


Figure 4. Bridged Ethernet via ATM Example

protocol stacks for a bridged Ethernet over ATM service. For IP directly over ATM, the protocol stacks would appear as illustrated in Figure 5.

The above two examples illustrate different mechanisms for moving Internet IP datagrams over ATM. It should be noted that moving other forms of integrated services, such as telephony is as straightforward. In fact, an ATM cable modem is capable of providing multiple service interfaces from the same box; e.g., an Ethernet

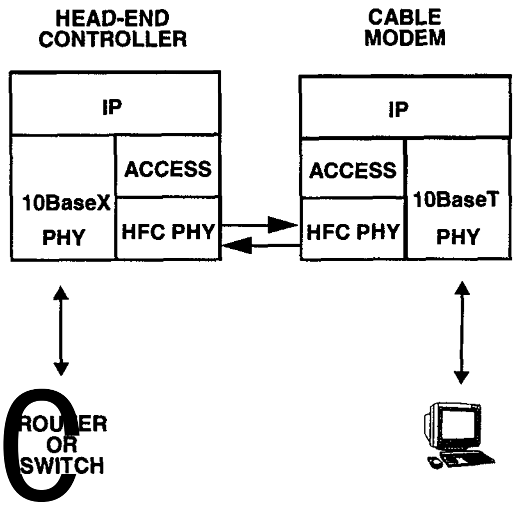


Figure 5. Routed IP Example

port for the home computer, and an RJ11 jack for the telephone. See Figure 6.

The multiple service capability of ATM is a very important consideration for deploying residential broadband services for three primary reasons:

Firstly, deploying ATM based cable modems for Internet services, allows a straightforward addition of future services without having to change out or install completely new additional equipment. That is, the cable modem box the supports Internet data can co-exist on the same downstream and upstream channels as a new cable modem that supports telephony.

Secondly, the ATM cable modems described in this paper operate at the electrical RF interfaces of a CATV network. This means that an all-coax CATV system can begin to deploy revenue bearing integrated services, starting with a small installation size, then growing the system for either data or voice services.

Thirdly, when the service needs of data and voice exceed the capacity of the system, only then must the network be upgraded to a Hybrid Fiber-Coax system. And then, HFC need only be deployed in those geographic areas where the service is required.

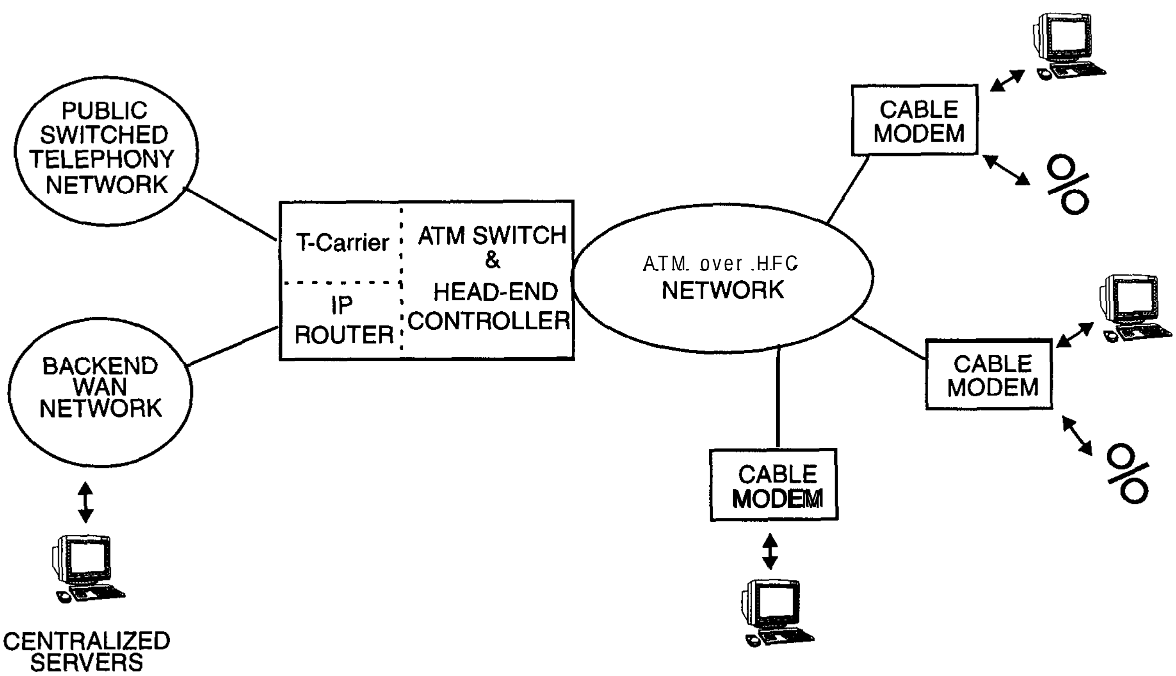


Figure 6. Internet and Voice Deployment Model

SUMMARY

This article has presented an overview of the work in progress of the ATM Forum's Residential Broadband Working Group and that of the IEEE P802.14 Cable TV MAC and PHY Protocol standards Working Group. Initial review of these works is positive and indicate that ATM over HFC systems can be constructed using a MAC layer access approach.

The cable network environment will provide a very usable platform for delivering ATM-based Internet and voice services to and from the home. Actual deployment of ATM-based Internet to the home will occur in many areas of North America in the 1996 and 1997 time frame. Standards for ATM over HFC networks will appear in late 1997 at the earliest.

An ATM based integrated bearer service is well suited to allow a cable operator to install a system that will evolve in time and cost without significant duplication of up front investment,

ACKNOWLEDGEMENTS

The author wishes to thank the participants of both the ATM Forum Residential Broadband Working Group and the IEEE P802.14 Cable TV MAC and PHY Working Group for their efforts to date.

FOR MORE INFORMATION

Information on the IEEE's P802.14 Working Group can be found on the World Wide Web at:

http://www.com21.com/pages/ieee802.14.html

Information the Internet Engineering Task Force's IP over ATM Working Group can be found at:

http://www.com21.com/pages/ipatm.html

The ATM Forum is a closed industrial consortia and non-published work-in-progress documents cannot be distributed publicly to non-members. General information about the ATM Forum may be obtained from the Web at:

http://www.atmforum.com/

REFERENCES

[1] ATM-FORUM, User Network *Interface (UNI) Specification Version 3.0,* ISBN 0-13-225863-3, Prentice Hall, December 1993.

[2] Kim, Albert, *Two- Way Plant Characterization,* Technical Session 23, National Cable Television Association Show and Conference, Dallas Texas, 9 May, 1995.

[3] Laubach, Mark, *Classical IP and ARP over ATM,* RFC1577, January, 1994.

[4] IEEE P802.14 Working Group, *Cable-TV Functional Requirements and Evaluation Criteria,* work in progress, IEEE802.14/94-002R2, IEEE 802 Committee, February, 1995.

l-51 IEEE, *IEEE Standards for Local Area Networks: Logical Link Control,* ANSI/IEEE Std 802.2-1985, Fifth Printing, Feb., 1988.

[6] Abramson, Norman, *Development of the ALOHANET,* IEEE Trans. on Information Theory, vol. IT-31, pp. 119-123, March, 1985.

[7] Roberts, L., *Extensions of Packet Communication Technology to a Hand Held Personal Terminal,* Proc. SJCC, pp. 295-269, 1986.

[8] Baran, Paul, *On Distributed Communication Networks,* IEEE Trans. on Communication Systems, vol. cs-12, pp. l-9, March, 1964.

[9] XEROX, *The Ethernet, A Local* Area *Network: Data Link Layer and Physical Layer Specification*, X3T5 1/80-50, Xerox Corporation, Stamford, CT., October 1980.

[10] IEEE, *Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications,* Standard 802.3-1985 (IS0 DIS 8802/3), IEEE, New York, ISBN 0-471-82749-5, 1985.

[11] Chelehemal, M., Prodan, R., et al., *Field Evaluation of Reverse-Band Channel Impairments,* Society of Cable Telecommunications Engineers, Emerging Technologies Conference, San Francisco, CA, 9-12 January, 1996.

[12] Laubach, Mark, *ATM For Your Internet - But When?,* ConneXions, Vo17, No 9, September, 1993.

[13] Atkinson, Ran, *Towards Real ATM Interoperability,* ConneXions, Vo17, No 8, August, 1993.

[14] Laubach, Mark, *IP over ATM and the Construction of High-Speed Subnet Backbones,* ConneXions, Vol 8, No 7, July, 1994.

[15] Clapp, G. and, Zeug, M. *Components of OSZ: Asynchronous Transfer Mode,* ConneXions, Vo16, No 4, April, 1992.

[16] Rybczynski, Tony, *The ATM Enterprise Network Switch: Enabler of Network Transformation,* ConneXions, Vol 9, No 11, November, 1995.

[17] Heinanen, Juha, *Multiprotocol Encapsulation over ATM,* RFC1483, March, 1993.

[18] Lucien, Rhodes, *The Race for More Bandwidth,* (Interview with Milo Medin of @Home), WIRED Magazine, 4.01, January, 1996, page 140.

BIOGRAPHY

MARK LAUBACH holds a BEE. and M.Sc. from the University of Delaware. He is the Technical Ambassador and Chief Architect at Com21, Inc. in Mountain View California and responsible for the end-to-end systems architecture and ATM over HFC protocol specification of the Com21 product family. Prior to Com2 1, he was with the Hewlett-Packard Company for 14.5 years. Mark is a member of the IETF and is past chair of the IP-over-ATM Working Group. He is the author of RFC1577 *Classical IP and ARP Over ATM.* Additionally he attends the ATM FORUM and is the appointed liaison from the Forum's Residential Broadband Working Group to the IEEE P802.14 CATV MAC and PHY Protocol Working Group. He is a past member of the CSNET Executive Committee and a past member of the board of trustees for the Corporation for Research and Educational Networking (CREN = CSNET + BITNET).

DIGITAL DROP TESTING: What Constitutes Breakage?

Ernest Tsui
Applied Signal Technology Inc.

ABSTRACT

The introduction of digital signals to the cable industry is inevitable—with digital technology you have higher quality transmission, increased channel capacity (more revenue), and ultimately lower costs (via VLSI integration) than analog. These benefits, however, come at a cost: the cost to properly install, maintain, and troubleshoot a complex digital communication system running over a largely non-engineered subscriber premises wiring system and a drop system that can be unintentionally "invaded" by subscriber activities. This paper introduces the key test and maintenance procedural differences between analog and digital signal transmission and the "link breakage" performance differences, as well.

INTRODUCTION

Digital signals perform better than analog signals in noise since the digital "threshold of visibility" (TOV) is obtained at a C/N substantially below that of the NTSC VSB signal. Furthermore, the NTSC signals will have noticeable distortions (albeit, small) in comparison to the digital "near perfect" results whenever the signal is above TOV. Digital signals also provide more capacity with high-quality, efficient compression schemes such as MPEG II. For these two key reasons digital compression/modulation schemes have been widely adopted.

There does exist, however, a "dark side" to the cable network—drop/subscriber premises wiring, governed by a "Mr. Fix-it" mentality, beset by the low-grade, low-performance components. It is in this area, that cable technicians and installers face the significant challenge of properly engineering and maintaining the system for digital signals. Proper training and test equipment will be key to proper qualification and troubleshooting of households for digital.

Headend modulation, fiber transport, and coax/trunk amp links will also impair the signals; however, these segments can be engineered and routinely maintained to yield good performance. In this discussion, we examine potential impairments from headend, trunk, and distribution subsystems to calculate the level of signal impairment accumulated that enters the drop/premises environment. Network operations staff concerned with setting the proper signal power levels and maintaining the network at this level should gain more understanding of the requirements for proper transmission of the digital signals.

Thus, the purpose of this discussion is two-fold: first to present the key differences in the diagnosis of analog versus digital transmission problems and, second, to relate to the cable technician how performance thresholds will differ between analog and digital signals.

DIAGNOSIS

Presently, for analog NTSC transmissions technicians can derive much information from the TV picture such as approximate C/N, hum, ghosting, interference type and level (e.g., Terrestrial TV interference, CTB, etc.). This "pictoral" or visual information is then used to locate/isolate the problem. Digital signals will not afford the technician this effective diagnosis tool. Problem diagnosis for digital signals will need to come—not from the examination of the video—but from the *digital signal itself.* Distortions viewed on a digital video signal are not highly correlated to the impairment type. For example, high enough error rates that are not correctable via the error correction circuits may cause a high number of undetected errors. These undetected errors will cause different distortions depending on their location in the frame structure and picture scene content (dynamic or static), and the burstiness of the errors (impacting the error randomization circuits). NTSC transmissions, with their simple AM (amplitude modulation) transmission scheme and simple framing, retain much of any amplitude distortion in a linear fashion allowing diagnosis via observations of the picture. Table 1 summarizes the issues involved.

Analog	Digital
Impairments and their levels can be inferred from the TV picture.	Impairments and their levels cannot be inferred from the TV picture but must be derived from the digital signal itself.

Table 1. KEY Analog/Digital Transmission Differences

In addition to the diagnosis problem, additional performance threshold differences can occur between analog and digital signals as discussed in the following sections.

SYSTEM TRANSMISSION PERFORMANCE AND DISTORTIONS

Headend

The major transmitted signal distortions appear when the "bits" are "modulated," e.g., placed into 64 QAM modulation format, transmit filtered (for optimal transmission), and further filtered to maintain the signal within the 6 MHz channel allocation so as to not substantially interfere with adjacent channel signals. For digital signals, the modulation process is generally done with digital signal processing circuits. Thus the distortions involve quantization errors and finite digital filter lengths for the transmit filter. In our experience, for 64 QAM good modulator design requires that total distortion products be approximately –40 dbc. At the headend, the digital signal is combined with other signals and upconverted. Figure 1 shows a 64 QAM signal being transmitted at an actual cable headend with an empty adjacent channel slot. As can be seen by the noise in the adjacent channel slot, the average C/N is approximately 36 to 37 dB. Note also that the digital signal was transmitted at an average power approximately 15 dB below the NTSC carrier peak power. Assuming that the transmitted QAM signal can be transmitted at a level of –10 dB (level recommended from [Hamilton and Stonebeck]) below the NTSC carrier yielding 41–42 dB C/N, the combination of modulation distortion (–40 dbc) and transmit C/N (–41 dB) results in total signal-to-distortion-plus-noise-ratio of approximately 37 to 38 dB. In comparison with the NTSC analog system, the equivalent NTSC C/N would be 36 to 37 dB (QAM average C/N) + 15 dB + 7 dB (correction factor[1]) = 58 to 59 dB C/N.

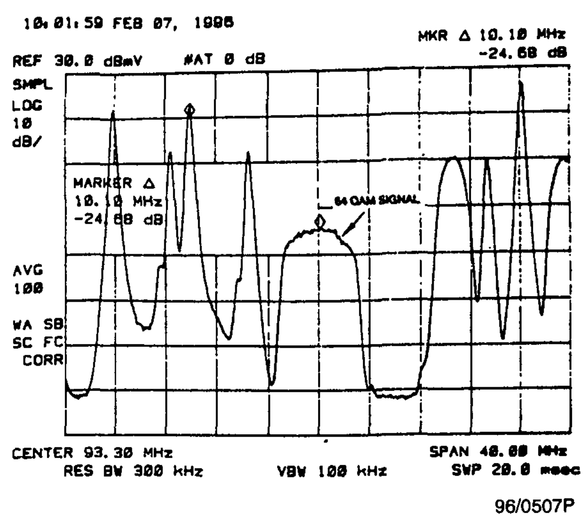


Figure 1. Headend 64 QAM Signal

Fiber "Trunk" Link

After transmission over the fiber optic link (Figure 2), the signal C/N has been seen to degrade by less than 1 dB or so during an actual field test over a 10-mile fiber link.

Distribution System

At the fiber node distribution point, the C/N will degrade proportionally to the number of trunk, bridge, extender, etc., amplifiers that lead to the subscriber drop. Prudent system design should allow the digital 64 QAM C/N to drop to no less than approximately 30 dB or so at the tap for the subscribers going through the maximum number amplifiers in the system.

There is no difference here between the noise affecting the NTSC vs. the digital carriers. However, the CTB, CSO, cross- and intermodulation products will differ from that of the NTSC carriers which are relatively narrowband. The products produced by the digital carriers on the NTSC and Digital Signals will tend to be wide-band or "noiselike." The impact of this characteristic will be two-fold. First, the CTB of the digital signals on the NTSC carriers would appear to be more "snow-like" or noise-like than "liney" or narrowband, with relation to the present NTSC CTB effects which may confuse a technician that was observing the TV picture. Secondly, the "noise-like" CTB would require modified measurement techniques for the NTSC signals and, of course, new procedures for the digital signals to both diagnose and estimate level of impairment would be required.

Drop/Premises

Presently at the subscriber premises, the cable technician observes and diagnoses the impairment(s) on the TV picture and then "repairs" the system so as to achieve the required performance. In most cases, the tolerated impairment levels of the new digital signals will be much higher than the NTSC signals (see Table 2 which presents the approximate differences in performance for typical cable system impairments measured at the drop or at subscriber premises). In some cases, the digital signals will be more sensitive than the NTSC to certain impairments (see Table 3 which shows the impairments that affect digital signals only and not NTSC-transmitted signals).

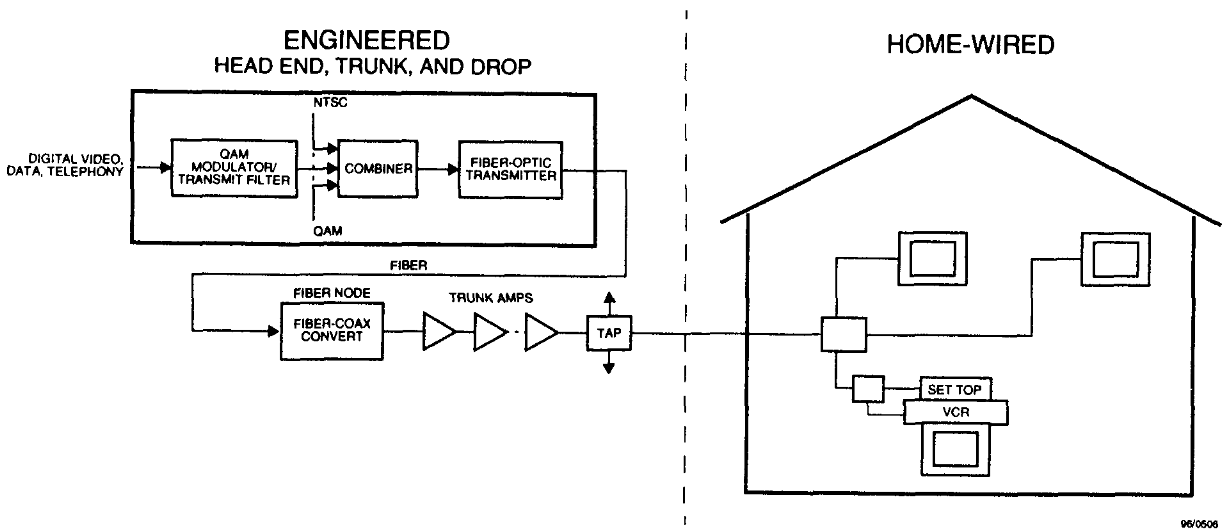


Figure 2. Headend-to-Subscriber Communications Link

Impairment	Analog	Digital
Co-Channel (TV in TV)	30 dB marginal, 25 dB bad ("strong horizontal lines" in picture)	22 to 24 dB at TOV
Composite Second Order Distortion (CSO)	53 dB	28 dB at TOV
Ghosting or Micro-reflections C/G = Carrier to "ghost" power	20 to 40 C/G dB observable; 10 to 15 dB objectionable [Jones]	5 to 15 dB C/G level at TOV
CTB	53 dB, (46 dB is bad: "noise/lines" in picture)	41 dB
C/N	Target values are 48 to 50 dB, 42 to 44 dB just visible, 40 to 41 (objectionable) [Ciciora]	21 to 25 dB
AM hum	3% ("moving bars" in picture)	14% at TOV

Table 2. Key Analog/Digital Performance Threshold Differences

Impairment	Analog	Digital
Residual FM	AM detection makes the NTSC circuits very tolerant of this distortion	Few kHz to 100 kHz dependent on demodulator design
Phase Noise (1/f**2)	Essentially no effect due to AM detection for NTSC	–75 to –80 dbc at 20 kHz away from the carrier at TOV
Minimum Isolation in Splitter (Surfing Problem)	Negligible transient ghost	21 dB required, above which, the adaptive equalizer may lose lock and cause a momentary video outage

Table 3. Impairments Affecting Digital Only

For the digital signals a parameter is used (developed by CableLabs) called TOV or "Threshold of Visibility." This is the level at which an average 3 x 10^{-6} bit error rate is attained. This was experimentally verified by CableLabs to be a threshold at which digitally compressed pictures become unacceptable. This is the parameter we will use to compare the digital performance results to the NTSC performance parameters mandated by good cable operator practices and/or the FCC.

The "NTSC Picture" performance criteria for analog signals was derived from an SCTE video tape [SCTE], and 64 QAM results from a paper presenting Applied Signal Technology, Inc. results on a QAM demodulator test at CableLabs [Laudel].

All the digital performance results are presented in terms of the average QAM signal power to distortion level in dB, and analog results in peak NTSC power to distortion power, again in dB.

The TV interference seems to be comparable between analog and digital since both are marginal at about 30 dB after the digital is corrected from average (shown in Table 2) to peak power (+6 dB) to afford a fair comparison with the analog NTSC performance.

However, the digital signals are significantly more tolerant of CSO (tested with NTSC signals generating the CSO only) since the adaptive equalizer in the demodulator circuit will actually cancel this narrowband interference problem. For "noise-like," digital signal-generated CSO, the results may be more comparable between analog and digital signals since in this case the adaptive equalizer will not be nearly as effective in the cancellation of this wideband interference.

For CTB, the digital adaptive cancellation appears to be less effective in providing comparable rejection to the analog signal (after adding the 6 dB correction for peak vs. average power). This may be due to the fact that the CTB interference is a wider bandwidth signal than CSO and therefore is more difficult to cancel.

AM hum is caused by poorly regulated power supplies, and NTSC signals can be significantly degraded with as little as 3%. The digital signal demodulators have circuits that mitigate this effect and allow a significant 14% of hum before reaching TOV.

Micro-reflections are caused by impedance mismatches, deformations in the cable, etc., and can be a significant problem in the subscriber premises. However, the adaptive equalizers in the digital demodulator, again, are called upon to cancel this type of interference and can attain performance superior to that tolerated in the NTSC signals.

Finally, C/N is a very key performance parameter, and digital signals significantly outperform (by 10 dB) analog signals for this impairment This improvement is primarily due to the powerful error correction techniques employed and the efficient coherent demodulation (need precise phase knowledge of the received signal) technique over the AM detection for the analog signals.

The first two effects listed in Table 3 are caused mainly by set-top box tuner electronics and are included here for completeness. The only exception to this is if an AML link is used for trunking of the digital signals the up- and downconversion process can generate significant phase noise that needs to be addressed. The third effect, isolation, is purely a subscriber problem and may occur when a TV set is connected by a low isolation splitter along with a set-top. This problem is transient since it would only occur as the TV tuner moves through the same channel as the set-top.

As can be observed from the performance data, digital signal transmission techniques have many advantages over analog transmission and may, in fact, be more robust in practice. However, system designers, trying to avoid amplifier overload, dictate that the digital signals be transmitted 10 dB below the analog signals reduce significantly the "robustness margin" of the digital transmissions over analog. Furthermore, when impairments do exceed the ability of the digital demodulator and equalizer to correct them properly, the cable technician will be challenged to infer the degradation without the TV picture.

Ingress Field Tests/Example

A comprehensive study of what to expect of the drop and subscriber environment was presented by [Prodan, et. al]. The most significant finding relevant to performance testing was that spurious power is actually higher than noise power. It was found that 5% of the subscriber premises have less than 36 dB of shielding in their wiring systems. The shielding problem is due, in many cases, to the

F-connectors which will cause degraded shield performance when they are loose, corroded, improperly installed, or damaged. It was shown that loose connectors can cause an additional 20 to 30 dB loss over tight connectors [Bauer]. Since digital signals will be transmitted at a higher frequency than analog, connector shielding will decrease additionally by 30 to 60 dB at 750 MHz vs. 100 to 400 MHz [Bauer].

As an example of ingress that actually occurred on home wiring, Figure 3 shows the spectrum of a 64 QAM signal with an FM radio signal in an adjacent channel (empty) slot. Figures 3a and 3b show the ingress at the tap and set-top, respectively. Observe that at the tap, the FM ingress level is appreciably lower than the level at the set-top. The estimated signal-to-ingress ratio at the tap is approximately 27 dB versus the in-home of approximately 9 dB. This increased ingress was due to poor shielding by the connectors and/or the cable in the home.

SUMMARY

Digital signals afford outstanding benefits in the areas of additional capacity and picture clarity. In order to provide these benefits, operators will need to understand why digital links can break and at what impairment levels

These impairments and levels will need to be diagnosed and estimated from the modulated digital signals themselves before they are decoded into bits. This will be even more important for future cable modem and telephone signals where there is no "picture" to begin with.

Cable technicians/engineers must also understand that although digital signals are very robust, the subscriber premises wiring "hurdle" may require more comprehensive characterization to insure successful installation/maintenance and system power allocation trades.

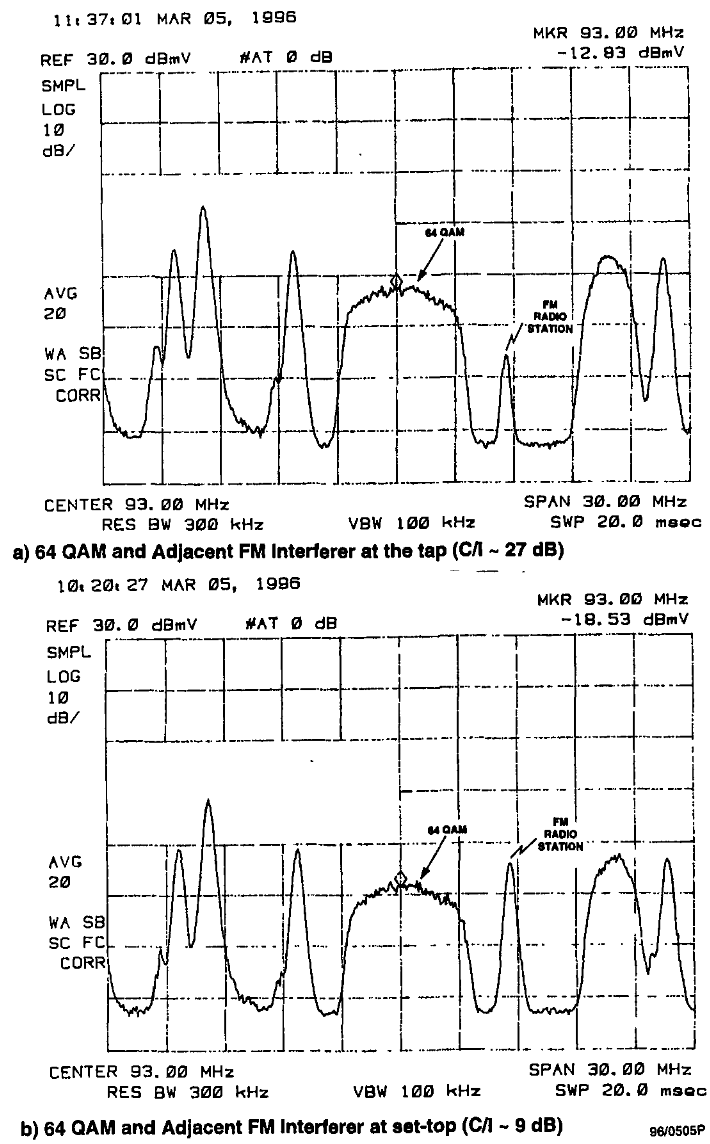


Figure 3. Field Example of FM Ingress Enhancement Due to Subscriber Wiring

REFERENCES

Bauer, B., "Approaching the Seamless Drop," *Communications Technology*, June,1994, p 138.

Hamilton, J. and Stonebeck, D., "How Digital Carriers Affect the Analog Plant," *CED*, June 1994, p. 80.

Jones, Bronwen, "Subjective Assessment of Cable Impairments on Television Picture Quality", *1992 NCTA Technical Papers*, p. 1.

Prodan, R., Chelemal, M., Williams, T., and Chamberlain, C., "Analysis of Cable System Digital Transmission Characteristics," *1994 NCTA Technical Papers*, p. 254.

Laudel, Ken, "Performance Results of a Low-Cost Alternative Equalizer Architecture for 64/256-QAM Demodulation in a CATV Receiver," *1995 NCTA Technical Papers*.

SCTE, "T-1001-A, TV Your Primary Diagnostic Instrument," video available from the SCTE.

FOOTNOTE

[1] To generate the average 64 QAM power from which an SNR can be computed: the average QAM power is approximately 6 dB reduced from the QAM peak power. Finally, for a typical 64 QAM signal, the bandwidth is measured in 5 MHz vs. 4.2 MHz bandwidth for the NTSC signal, which results in an additional 0.75 dB conversion for SNR. Thus, the average QAM SNR or C/N compared to the equivalent NTSC C/N will be lower by approximately 7 dB.

Don't Get Clipped on the Information Highway

D. Raskin, D. Stoneback, J. Chrostowski and R. Menna
Transmission Network Systems
General Instrument Communications Division
Hatboro, PA

Abstract

Signal degradation due to clipping is a threat to the performance of both the forward and the return path that has been given too little attention by our industry. In the forward path, the performance of fiberoptic transmission systems for broadband services is often limited by statistical clipping of the RF signal at the laser source. This means that systems meeting NCTA distortion and noise standards may nonetheless exhibit unacceptable video performance. Furthermore, the nature of these impairments is very different for directly-modulated, as opposed to externally-modulated optical sources. We report first on video tests of DFB lasers, indicating the nature and cause of clipping related impairments. Based on subjective picture quality tests, we have defined a straightforward quantitative test for determining acceptable clipping performance of a transmitter. The results of a similar test program for externally-modulated sources used in 1550 nm systems are described and compared with DFB's. Criteria for successful downstream data transmission jointly on lasers with analog signals are reported, as are preliminary investigations of 12VDC vs 24VDC RF amplifiers. Upstream digital transmission tests that investigate the effects of clipping both in return amplifiers and lasers are discussed, as well. Quantitative tests will be proposed as supplementary recommended practices for the industry.

INTRODUCTION

What causes clipping in CATV systems?

When an applied electrical signal amplitude goes beyond the level capability of a device, we say that the signal is "clipped" by that device. A simple, but extreme example would be an AC signal applied to a circuit through a diode. In that case the diode would clip 100% of one polarity of the AC current signal. In cable television systems, laser transmitters cause more subtle clipping of the RF signal when the amplitude of the composite signal exceeds either the bias current, in the case of a directly modulated semiconductor laser, or the modulation limits for an external modulator (ex-mod) in a 1550 nm or YAG solid state transmitter. This is shown graphically for a DFB laser and an ex-mod in Figures 1 and 2, respectively. Keep in mind that these diagrams are somewhat misleading because of their over-simplification of the highly complicated signal that is applied to a CATV analog laser. In a real CATV transmitter, clipping is supposed to occur only when all of the many individual signal amplitudes "line up." Fortunately with properly designed equipment in modern high-channel-count systems, statistical averaging works to make this a relatively rare event.

RF amplifiers can cause clipping, as well, when the input signal exceeds the bias of transistor elements. In nearly all cases this is not a problem in CATV, since the RF levels that cause noticeable clipping would already be significantly beyond the distortion limits of the amplifier. Recently, however, two changes in CATV operations and technology have raised questions about the importance of amplifier clipping: (1) the much greater signal loading of return systems has driven the RF levels at return node amplifiers close to clipping limits and (2) the desire to use GaAs amplifiers has encouraged the use of lower bias voltages (e.g., +12 VDC, rather than the traditional +24 volts). We have found situations in which clipping can be a problem in the return amplifiers. On the other hand, our initial investigation of 12V amplifiers does not indicate reasons for concern.

What are the effects of clipping?

From Figures 1 and 2 one can see that DFB lasers cause clipping only in the downward direction whereas the ex-mod clips bidirectionally. We can understand the effect of this difference by imagining the clipped output signal as the sum of an unaffected ("pure") signal and an error signal (Figures 3 and 4). For a sinewave input, the error signal in the DFB case is an even function, while that of the ex-mod is odd. This means that the clipping events will show up on a spectrum analyzer at the CSO beat frequencies for a DFB and at the CTB beats for an ex-mod.

Figure 1a: DFB Laser in Linear Operation

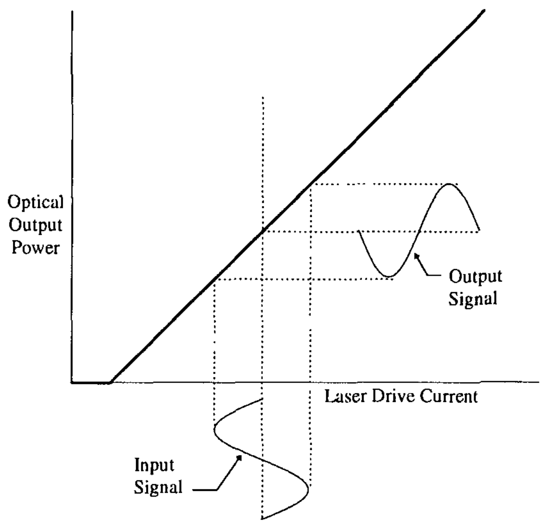


Figure 2a: Externally-Modulated Signal in Linear Operation

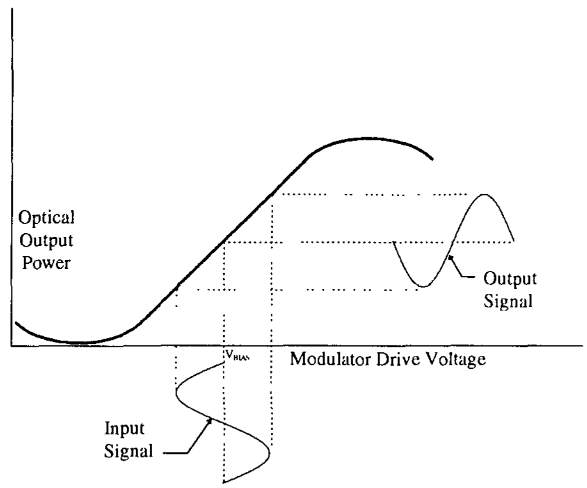


Figure 1b: DFB Laser in Clipping

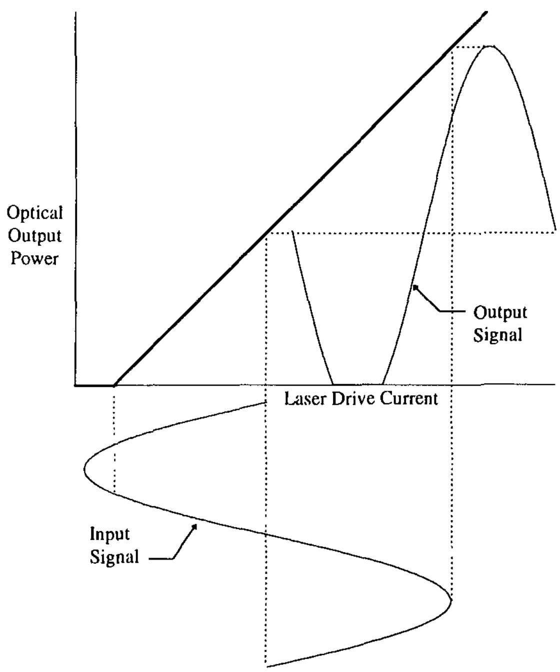


Figure 2b: Externally-Modulated Signal in Clipping

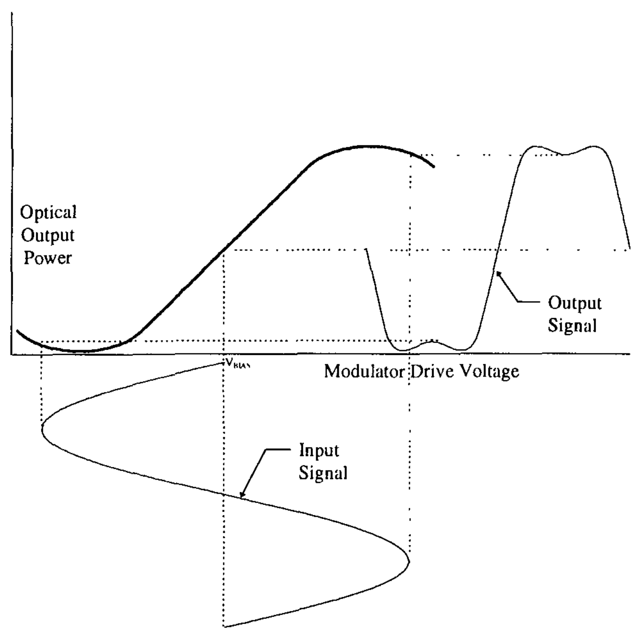


Figure 3: Decomposition of DFB Clipping

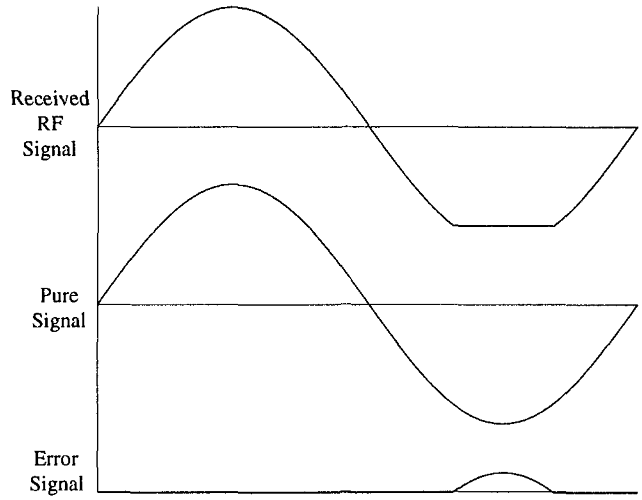
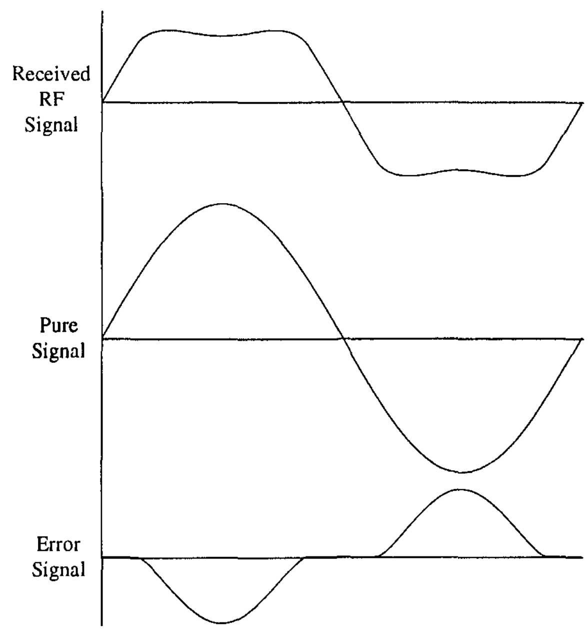


Figure 4: Decomposition of External-Modulator Clipping



A more critical difference between the clipping effects in these two types of transmitters results from the fact that in DFB's the laser is actually turned off (biased below threshold) when the RF signal gets too large in the negative direction. Unfortunately, the RF optical signal out of the laser does not appear to recover instantaneously. Typically this time delay is several microseconds, which is a significant fraction of the television horizontal scan time (62.5 μsec). Thus a laser clip looks like a white horizontal flash on a TV screen. This is easily seen in a display of a medium brightness flat field (≈ 50 IRE), which is how we made subjective evaluations of analog clipping. On the other hand, the light source in the ex-mod case is never turned off. When the modulator is over-driven, the optical power is still modulated in a continuous fashion. Although the signal during the clipping event may be rich in harmonics, the duration is very short. Thus analog clipping in an ex-mod shows up as random "busy-ness" on a flat field TV display. RF amplifier clipping would manifest itself in a similar way, since it is bi-directional and there is no latency in the active device.

Clearly these clipping events also have the capability of interrupting the flow of digital data, hence to cause errors. We have used uncorrected bit error rate (BER) measurements to determine the severity of clipping on these signals.

FORWARD PATH TESTING

DFB lasers

Over the past three years we have found that the subjective picture quality of signals transmitted over fiber by high quality DFB analog transmitters is limited by clipping in the lasers, rather than by the traditional second and third order distortions. As noted above this clipping shows up as annoying horizontal flashes in a picture. Because of the obvious need for a quantitative test to determine when a laser is operating properly, we examined a number of approaches. Ultimately a time domain measurement was found that correlated well with the flat field tests: a spectrum analyzer is set at zero span on a mid-frequency CSO beat frequency at resolution and video bandwidths of 30 kHz and is allowed to make a single sweep of 30 sec duration. The display will show a rough baseline with several distinct peaks due to clipping (Figure 5). We have found that the amplitude of the highest peak is not a good indicator of subjective picture quality. The height of the tenth highest peak, however, turns out to be a very good indicator of clipping quality: if the tenth highest peak measures greater than -45 dBc we observe noticeable clipping in the picture.

Figure 5a: Clipping Test Passing

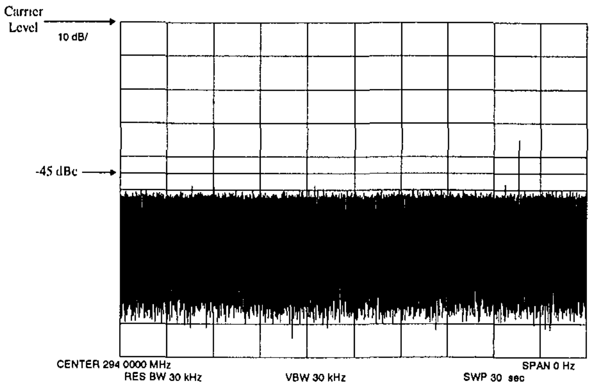


Figure 5b: Clipping Test Failing

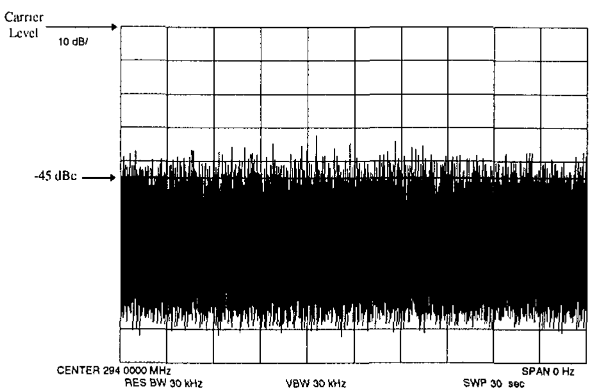


Table 1 gives the results of a number of different tests of forward lasers carrying 77 analog channels from 50 to 550 MHz and 200 MHz of simulated digital loading (including one channel of 64-QAM) from 550 to 750 MHz. These show the effects of different operating levels and of relative levels of digital to analog. The column "CSO Clip" gives the result of the time domain test described above. As stated above for DFB's: (a) there are no measured clips at the CTB points and (b) CSO clip numbers larger than -45 dBc are reliable indicators of observable video clipping. Interestingly, for the digital signals, if we define (somewhat arbitrarily) the acceptable level of uncorrected BER's as being $<10^{-9}$, then clipping appears to become unacceptable for the digital at essentially the same laser operating point as for the analog. (That statement applies only if the digital levels are set within 0 to 10 dB below the analog.) Since the 10^{-9} criterion is likely to be very conservative for error-corrected forward signals, it appears that digital signals in the forward path will generally not be limited by clipping.

TABLE 1: Directly Modulated DFB

Drive Level	Subjective Clipping: Ch 36 (Mid Freq.)	Subjective Clipping: Ch 77 (High Freq.)	BER	Mid Frequency: CSO Clip	Mid Frequency: CSO	Mid Frequency: CTB	Mid Frequency: C/N	High Frequency: CSO Clip	High Frequency: CSO	High Frequency: CTB	High Frequency: C/N
CW Loading; Noise & Digital 10 dB down:											
Nom	Very Visible	Very Visible	1.60E-06	-36	78	65.3	52.0	-34.5	62.5	64.9	53.2
-2	None	None	3.00E-10	-49.8	79	72.4	50.0	-46.2	65.5	71.1	51.1
CW Loading; Noise & Digital 5 dB down:											
Nom	Visible	Visible	1.30E-08	-39.6	78	66.5	51.3	-34.5	62.9	66.2	51.8
-1	Barely Visible	Barely Visible	>4E-10	-45.6				-42.7			
-2	None	None	>4E-10	-50.4	78.5	73.7	49.6	-51.5	66.1	72.1	50.2

Note: Drive Level is referenced to the "nominal" setting. Nominal refers to the level which is automatically chosen by the transmitter's AGC circuit to match the factory calibrated operation point.

TABLE 2: External Modulation

Drive Level	Subjective Clipping: Ch 36 (Mid Freq.)	Subjective Clipping: Ch 77 (High Freq.)	BER	Mid Frequency	Mid Frequency: CSO	Mid Frequency: CTB	Mid Frequency: C/N	High Frequency: CTB Clip	High Frequency: CSO	High Frequency: CTB	High Frequency: C/N
Nom	None	None	1 E-09		-67.7	-64.8	49.9	-45.5	-66	-64.8	48.4
+1	None	Some CTB	3 E-09		-68.2	-62.7	50.2	-40.5	-66.4	-63.1	48.8
+2								-37		-62	

Notice that the laser input level that gives acceptable analog clipping performance is approximately 2 dB less than the level that gives a CTB of 65 and a CSO of 62. We have found that this is generally the case.

As a further test, the fiber link was connected to the peak-to-RMS ratio measurement system described in Appendix 1. This system allows a measurement of the peak-to-RMS ratio of the device being tested. The maximum ratio observable is about 15 dB. At nominal levels, the laser's ratio was about 14 dB. When the drive level was increased by 1 and 2 dB, the ratio decreased to 13 and 12 dB, respectively. The histogram clearly showed that only one side of the waveform was clipping. This is expected with a directly modulated DFB. The ratio of the other side of the waveform remained at 14 dB.

Externally-modulated transmitters

Limited tests of a 1550 nm ex-mod transmitter (Table 2) confirm our expectation that clipping is less of a concern in these units. Note that the clips occur at the triple-beat frequency, as expected. Essentially all three drive conditions give tenth-clip-event heights greater than -45 dBc, but the pictures did not appear to have objectionable clipping. As the modulation depth is increased, CTB becomes noticeable in the picture before any clipping impairment is observed. Thus some other test or criterion would have to be established to quantitatively judge clipping performance in ex-mod transmitters, if this were deemed worthwhile.

RF amplifiers

We have confirmed the expectation that clipping degradation is highly unlikely for conventional push-pull and power-doubled hybrid amplifiers in a wide range of normal operating points. The test setup is described in Appendix 1. The multiple carrier generator was inserted into a 750 MHz power doubled amplifier. The output voltages can be predicted as follows:

For an output level of +44 dBmV per channel for 110 channels into 75 ohms, the RMS value will be:

$$RMS_{110} = 10^{\left(\frac{44}{20}\right)} * \sqrt{110} / 1000 = 1.66 \ volts$$

and assuming a peak-to-RMS ratio of 14.8 dB, the peak value will be:

$$PEAK_{110} = 10^{\left(\frac{14.8}{20}\right)} * 1.66 = 9.1 \ volts .$$

The peak-to-RMS ratio remained constant at 14 to 15 dB until levels were increased to more than +52 dBmV per channel at the amplifier's output. The CTB was an unacceptable -55 dBc at +51 dBmV output. At +55 dBmV the peak-to-RMS ratio decreased to about 12.5 dB. The calculated peak voltage at this level exceeds the 24 volt supply, so it is no surprise that the peaks are clipping. The conclusion is that the amplifier does not clip until the CTB is intolerably bad.

As a further confirmation, a 12 volt amplifier was tested at +44 dBmV output per channel for 110 channels. The peak-to-RMS ratio was 14.9 dB, which is taken to signify that there was no clipping of the output.

RETURN PATH TESTING

QPSK digital

The effect of clipping on an HFC communications network is not limited to the forward path. Both the amplifiers and lasers in the return path can adversely affect the performance of digital signals if they are being operated too far into clipping. To investigate this phenomena, various components were loaded with a signal consisting of a real QPSK signal and filtered noise, all at a constant-power-per-Hz loading, to yield a total bandwidth of 35 MHz. In order to measure the amount of Composite Intermodulation Noise (CIN), there was a 50 dB notch in the middle of the noise. The depth of this notch directly correlates to the CNR of the link for any system that is loaded on a constant-power-per-Hz basis[1]. The CNR was compared to the average Bit Error Rate (BER) for various drive conditions. A typical result is shown in Table 3. Note that throughout this paper BER's are measured with no error correction.

Table 3: QPSK at Various OMI

dB Relative to 35 %	RMS OMI (%)	Peak OMI (%)	CNR (dB)	BER
---	35	124	44	$< 1*10^{-9}$
3	49	174	36	$< 1*10^{-9}$
6	70	248	27	$< 1*10^{-9}$
9	99	351	19	$< 1*10^{-9}$
10	111	394	---	$2*10^{-2}$

The RMS-to-peak ratios for various types of modulation and for noise are given in Appendix 2. The Peak OMI in Table 3 is calculated with an assumed average ratio of 11 dB. Note that the laser is always very far into clipping. However, the

BER does not get bad until the CNR degrades. Essentially, the limiting factor here is distortion, not clipping. The table demonstrates that a measurement of CNR should be sufficient for predicting BER, with no special consideration needed for clipping.

A common problem with the return path is the existence of large interferers (ingress). To determine how much interference can be tolerated, a CW interferer was injected into the laser along with the QPSK and noise payload discussed previously. The result is that, once again, the CNR is a good representation of the achievable BER. As the CW level was increased, the CNR would go down (due to an increase in CIN), and eventually, so would the BER. However, if the interferer was injected at a much higher frequency (such as 100 MHz) where the spectrum of the beats with this carrier did not fall on top of the QPSK signal, the CNR did not degrade and neither did the BER. This was true even with the CW at an OMI of 350%. This clearly indicates that clipping is not directly a threat.

n-QAM digital

The QPSK tests described above were repeated for 16-QAM and 64-QAM in an effort to determine the effect of clipping on amplitude-dependent signals. The data in Table 4 shows only slight differences between the drive-level dependence of the QPSK and 16-QAM data streams, which may be attributable to measurement error. The 64-QAM is somewhat more sensitive, due to its higher carrier to noise requirement. However, for a given BER, the higher order QAM modulation formats require a higher CNR as the laser is driven into clipping. This is most likely due to the amplitude component in n-QAM which does not exist in QPSK.

Table 4: Required CNR to obtain 10^{-6} BER

	QPSK	16-QAM	64-QAM
Direct	23	24	28
Laser Overdrive	19	23-28	35-39
Amplifier Overdrive	21	25	40

The data given in Table 4 has a several dB margin of error because the actual modulation signal was at 44 MHz, whereas the CNR was measured at 22 MHz. Nevertheless, some trends are evident. The term "overdrive" refers to a signal which is higher in amplitude than the recommended operating level of the device. The drive levels were increased until the BER degraded to 10^{-6}. Both QPSK and 16-QAM are relatively unaffected by either the laser or the amplifier overdrive. The 64-QAM, however, requires an increased CNR for equivalent performance when passed through a device that is being overdriven. This can be explained by the peak-to-RMS data given in Appendix 2. At a BER of 10^{-6}, the laser is 3 dB above nominal and the hybrid is at +67 dBmV total power out. Table A2 demonstrates than at these levels, the peak-to-RMS ratio is being compressed. This compression is likely to cause errors in the amplitude component of the 64-QAM signal. Similarly, 64-QAM is not as immune as QPSK to large out-of-band CW interference (such as the 100 MHz CW carrier at 350% OMI mentioned in the previous paragraph).

Amplifiers

The ability of return path amplifiers to handle large amounts of high level data was tested. The tests conducted were very similar to those described above for return path lasers. Table 4 indicates that 64-QAM signals are sensitive to overdrive conditions. As previously mentioned, this can be explained by referring to Table A2. Note that as the amplifier's output level is increased above 65 dBmV total power, the peak-to-RMS ratio decreases. Regardless of whether this is due to compression or clipping, the result is that some of the amplitude information gets compressed.

In order to determine how high the output levels should be in return path amplifiers, the output level was compared to CNR and BER. Table 5 shows performance for a 25 dB gain hybrid return path amplifier when subjected to overload.

Table 5: BER and CNR vs Output Level for Return Path Hybrid

Total Output Power (dBmV)	CNR	QPSK BER	16-QAM BER	64-QAM BER
45	>50	$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
65	48	$< 10^{-9}$	$< 10^{-9}$	$< 10^{-9}$
68	39	$< 10^{-9}$	$< 10^{-9}$	$2*10^{-6}$
73	24	$< 10^{-9}$	$1*10^{-6}$	
75	21	$1*10^{-2}$		

A fully loaded return path will have a total energy of approximately 60 dBmV at the output of an amplifier. Table 5 demonstrates that at levels of approximately +65 dBmV, the amplifier is getting close to affecting BER's of higher order modulations. Since CIN is dominated by third order distortion which cascades on a 20*log factor, one must be careful not to use amplifiers with inadequate distortion performance

CONCLUSIONS

We have determined that the upper drive level limit for forward analog DFB transmitters is determined by clipping, rather than by the conventional distortions. We have defined a simple quantitative test that correlates well with subjective evaluations and we are proposing that it be considered for general use by the industry. Within practical limits, it is likely that digital signals carried along with analog will not be limited by clipping (that is, not before the analog signals themselves will be impaired by clipping). We have found that clipping does not appear to be a performance-limiting problem in externally-modulated transmitters or in forward RF amplifiers.

In the return path, signal levels into Fabry-Perot lasers are limited mainly by CIN and not by clipping. The upper limit for DFB lasers is no higher than that of FP's. The limits for 16-QAM digital signals do not appear to be any tighter than those for QPSK. 64-QAM signals, however, appear to be somewhat clipping sensitive.

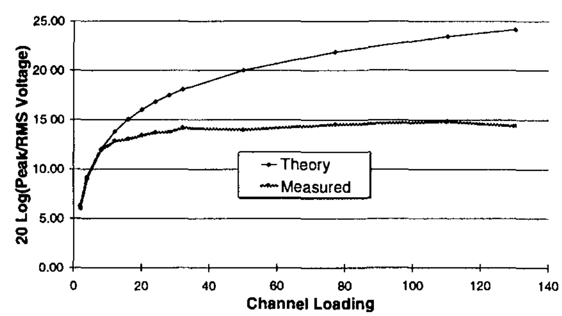
APPENDICES

Appendix 1: RMS vs peak

A test method was devised which would allow a measurement of the peak-to-RMS ratio of muticarrier signals. Our evaluations consisted of a multiple carrier generator which produced up to 130 independent CW carriers and a high speed (35 GHz) sampling oscilloscope. First, a baseline was established by measuring the output of the carrier generator directly. The output of the generator was connected to the scope and the level was adjusted to maximize the dynamic range of the scope. The scope was set to free run at a rate of 500 kHz and the timebase was 100 ps/div. Several million samples were allowed to accumulate in the infinite persistence mode (about 20 minutes). Since the scope was free running at a very high acquisition rate, the distribution of dots which accumulated on the display represented the distribution of voltages present in the composite signal. A histogram was then obtained from the scope and relevant points were recorded. This was done for many different channel loadings from 2 to 130 channels. In each case, the value of the highest and lowest points were recorded as well as the values for the mean, and one, two, and three standard deviations. The ratio of the peak value to the RMS value was calculated in each case. The results are compared to the theoretical values in Figure 6.

Figure 6

dB Ratio of Peak/RMS Voltages vs Channel Loading



Note that the theory predicts a continuous increase in peak-to-RMS ratio as the number of channels grows. However, the test results indicate that the ratio levels out between 14 and 15 dB. We believe that the reason the peak-to-RMS ratio does not continue to increase as predicted has to do with noise in the system. In particular, the phase noise and frequency stability of the CW generator are likely fluctuating enough to prevent any more precise coherent adding of the carrier phases.

Appendix 2: Peak-to-RMS ratios for various signals

The following table summarizes the peak-to-RMS ratio for some common signals. The measurement method was described in the previous appendix.

Table A2: Measured Peak-to-RMS Ratios

Type of Signal	Peak-to-RMS Ratio (dB)
CW signal	3.2
Unfiltered Noise (5-1000 MHz)	7.8
Filtered Noise (5-40 MHz)	13.5
Modem Brand "A"	
QPSK @ 10 MB/sec	9.2
16-QAM @ 20 MB/sec	11.3
64-QAM @ 30 MB/sec	10.2
Modem Brand "B"	
QPSK @ 256 kB/sec	6.7
QPSK @ 2 MB/sec	6.6
Laser with Filtered Noise Loading	
35% RMS OMI	13.8
62% RMS OMI	9.2
78% RMS OMI	7.8
Hybrid with Filtered Noise Loading	
+45 dBmV total RMS Power	13.5
+65 dBmV total RMS Power	12.4
+70 dBmV total RMS Power	9.6
Hybrid with CW Loading	
+45 dBmV RMS Power	3.2
+65 dBmV RMS Power	3.2
+70 dBmV RMS Power	3.2
+75 dBmV RMS Power	3.0

ACKNOWLEDGMENTS

The authors wish to thank Dr. Henry Blauvelt of Ortel Corp. and Dr. James Farina of Photonic Applications, Inc. for insights into the clipping of DFB and ex-mod transmitters, respectively. Dr. Blauvelt was a key contributor in the development of the quantitative clipping test for DFB's. The skillful assistance of Vipul Rathod, Mark McCusker, Albert Bachman and David Mascaro of General Instrument in obtaining the data was invaluable.

---

[1] D. Stoneback and W. Beck "Designing the Return System for Full Digital Services". SCTE Conference on Emerging Technologies, San Francisco, CA, January 1996.

Fiber induced distortion and phase noise to intensity noise conversion in externally modulated CATV systems

Dogan A. Atlas
Scientific-Atlanta, Inc.

Abstract

Fiber nonlinearity and chromatic dispersion can severely degrade the signal quality in externally modulated long-haul AM-VSB CATV transport systems. Standard single-mode fiber chromatic dispersion will convert inherent laser phase noise fluctuations to intensity noise. This dispersion-induced intensity noise increases with frequency, and thereby, imposes a fundamental limit on CNR. Even in the absence of stimulated Brillouin scattering, self-phase-modulation and chromatic dispersion induces CSO distortion. For transmission over 70 km fiber, CNR and CSO degradation associated with fiber nonlinearity and chromatic dispersion limits laser linewidth and fiber launch power to less than 1 MHz and less than 18 dBm, respectively.

I. INTRODUCTION

Efforts to enlarge the dynamic range of multichannel CATV systems has received attention with the advancement of linear lasers, linearized external modulators and low noise erbium-doped fiber amplifiers (EDFAs). To improve network reliability, upgradability and video signal quality, fiber is being deployed deeper into the CATV architectures, serving a smaller number of home pockets. Operating at the 1550-nm window is attractive where EDFAs can be employed to increase the power budget of future CATV transport systems.

For a directly modulated 1550-nm laser system, significant composite-second-order distortion (CSO) is generated when the laser chirp interacts with the standard single-mode fiber chromatic dispersion. For an externally modulated 1550-nm system, CSO distortion due to chromatic dispersion is negligible. However, fiber nonlinearites such as stimulated Brillouin scattering (SBS), self-phase modulation (SPM) and modulation instability (MI) can degrade the CSO distortion in high power cascaded amplifier CATV transport systems. SPM causes frequency chirp which interacts with fiber chromatic dispersion [1]-[3] which, in turn, degrades CSO in proportion to fiber input power and fiber length.

When the externally modulated light is transmitted through a long fiber length, laser phase noise is converted to intensity noise (PM-AM) due to chromatic dispersion [4]. This dispersion-induced intensity noise depends on the laser linewidth, fiber length and increases with frequency.

In this paper, the potential CNR and CSO degradation associated with fiber nonlinearity and chromatic dispersion is experimentally observed and analytically verified.

II. ANALYSIS

A. SPM-Dispersion Induced Distortion

The SPM is generated when the intensity modulated optical carrier transmitted through a single-mode fiber undergoes a nonlinear phase shift through a third-order fiber nonlinear process. The incremental nonlinear phase shift per unit length is given by

$$\Delta\phi_{NL}(t,z) = e^{-\alpha \cdot z}\gamma\,|a(t)|^2 \qquad (1)$$

where α is the fiber loss, $\gamma = 2\pi n_2/\lambda A_{eff}$ is the fiber nonlinearity coefficient at the wavelength λ, n_2 is the nonlinear index, A_{eff} is the fiber effective core area and $a(t)$ is the optical field envelope which carries the amplitude information. The intensity of the optical field envelope is expressed as

$$|a(t)|^2 = P_O\left(1 + m\sum_{i=1}^{N}\cos(\omega_i t + \theta_i)\right) \qquad (2)$$

where P_o is the average optical power of the optical signal, m is the modulation depth, N is the total number of CATV channels. The *total* nonlinear phase increases as the field propagates through the single-mode fiber distance. The optical field of the signal at a distance z without any interaction with fiber chromatic dispersion is given by

$$e(t,z) = e^{-\alpha \cdot z/2}a(t)\exp j\{\omega_O t + \phi(t,z)\} \qquad (3)$$

where ω_o is the angular frequency of the optical carrier,

$$\phi(t,z) = \phi_{NL}(t,z) + \phi_{PN}(t) + \phi_{PM}(t) \qquad (4)$$

denotes the phase information of the optical field, $\phi_{NL}(t,z)=\int_0^z \Delta\phi_{NL}(t,z')\,dz'$ is the nonlinear phase variation due to fiber nonlinearity, $\phi_{PN}(t)$ is the laser phase noise and $\phi_{PM}(t)$ is the phase modulation applied to increase the SBS power threshold. The propagation of the optical field $e(t,z)$ through a single-mode fiber can be described by the propagation term $exp\{-j\beta(\omega)z\}$ where

$$\beta(\omega)=\beta_0+\beta_1(\omega-\omega_o)+\frac{1}{2}\beta_2(\omega-\omega_o)^2 + \frac{1}{6}\beta_3(\omega-\omega_o)^3+\cdots+ \quad (5)$$

is the propagation constant expanded around ω_o. The terms β_0, β_1, β_2,..., denotes the derivatives of the propagation constant with respect to the angular frequency. For a standard single-mode fiber, the dominant dispersive term is $\beta_2 = -D\lambda^2/2\pi c$, whereas for operation near zero dispersion or for a dispersion shifted single-mode fiber, β_3 and higher terms dominate.

It is generally difficult to solve for the combined effect of SPM and fiber chromatic dispersion. Several numerical and analytical approaches have been used in the past [1],[2]. We will approximate the interaction of SPM with chromatic dispersion in a similar way described in [5].

The optical field after interacting with fiber chromatic dispersion, to first-order, can be approximately described as [6]

$$e(t,z)=e^{-\alpha\cdot z/2}e^{-\frac{1}{2}\frac{d\tau(t,z)}{dt}}a(t-\tau(t,z)) \cdot \exp j\{\omega_o t+\phi(t,z)\} \quad (6)$$

where

$$\tau(t,z)=\beta_2 z\left(\dot{\phi}_{NL}(t,z)+\dot{\phi}_{PN}(t)+\dot{\phi}_{PM}(t)\right) \quad (7)$$

is the total group delay generated by the interaction of the nonlinear chirp, phase noise, and phase modulation with fiber chromatic dispersion. From expression (6) the intensity of the optical carrier detected at a distance z=L can be expressed as

$$i(t)\cong e^{-\alpha L}|a(t-\tau(t,L))|^2\left(1-\frac{d}{dt}\tau(t,L)\right). \quad (8)$$

In this analysis, we have assumed that the transmitted optical field undergoes an incremental phase shift $\Delta\phi_{NL}(t,z)$ at distance z. The field is then assumed to interact with the dispersive fiber from z to L. Therefore, the nonlinear group delay can be found as

$$\tau_{NL}(t,L)=\beta_2\int_0^L (L-z')\Delta\dot{\phi}_{NL}(t,z')\cdot dz'. \quad (9)$$

The CSO due to fiber nonlinearity, (i.e., SPM), at channel $\omega_k = \omega_i \pm \omega_j$ can be found by using Bessel function of expansion:

$$CSO_{NL}(\omega_k)=\frac{K_2}{4}\left(D\frac{\lambda^2}{2\pi c}\gamma\, P_O m\omega_k^2 \cdot\frac{(\alpha L+e^{-\alpha L}-1)}{\alpha^2}\right)^2. \quad (10)$$

B. AM-Dispersion Induced Distortion

To increase the optical power threshold of SBS the optical field from an externally modulated transmitter is phase modulated. Thus, the envelope of the optical field which carries the amplitude modulated (AM) information signal is spectrally translated at multiple octaves of the phase modulating frequency. When the spectrally translated optical envelope interacts with chromatic dispersion at long fiber lengths, CSO distortion is generated. Unlike the work in [7], the AM-dispersion induced CSO distortion, when the optical spectrum is spectrally translated at multiple octaves, is found to be low. However, using a similar approach to [8] we have derived the AM-dispersion induced distortion as follows:

$$CSO_{AM}(\omega_k)=K_2\left(\sum_{n=-\infty}^{\infty}J_n(m/2)J_{n+2}(m/2) \cdot m^{-1}\cos\left(\frac{\omega_k^2}{2}\frac{\lambda^2}{2\pi c}DL(n+1)\right)\right)^2. \quad (11)$$

The amplitude of the second-order AM distortion is lower in magnitude with an opposite sign compared to the amplitude of the SPM distortion. Thus, the system CSO distortion which is composed of contributions from SPM- and AM-dispersion induced distortion. However, CSO_{AM} is much lower in magnitude than CSO_{NL}.

C. PM-Dispersion Induced Intensity Noise

The group delay generated by the interaction of the laser phase noise and fiber chromatic dispersion is determined as follows:

$$\tau_{PN}(t,L) = \beta_2 L \dot{\phi}_{PN}(t). \quad (12)$$

The variance of the laser frequency noise is

$$\sigma^2_{\dot{\phi}_{PN}} = 4\pi\Delta\nu \quad (13)$$

where $\Delta\nu$ is the laser linewidth. The signal current at the photodetector output, after traveling through a fiber of length z=L, is given by

$$i(t) \cong e^{-\alpha L} |a(t-\tau(t,L))|^2 \cdot \left(1 + D\frac{\lambda^2}{2\pi c} L \ddot{\phi}_{PN}(t)\right). \quad (14)$$

The RIN due to laser phase noise to intensity noise conversion can be easily found as follows:

$$RIN_{PM-AM}(\omega) = 4\pi\,\Delta\nu \left(D\frac{\lambda^2}{2\pi c}L\right)^2 \omega^2. \quad (15)$$

Inspection of expression (15) shows that the RIN is proportional to $\Delta\nu(\omega L)^2$.

III. RESULTS

Figure 1 shows the schematic of our experimental setup. The 1550-nm distributed-feedback laser (DFB) source has an optical power of 20-mW and a laser linewidth of 1 MHz. The optical signal is externally modulated by 77 AM-VSB NTSC channels, amplified and transmitted through L_1= 25, 50 and 70 km of standard single-mode fiber. The input power launched into the fiber was about 18 dBm. In the second part of the experiment, we employed a 16 dBm output power in-line amplifier. In this experiment, the externally modulated optical signal is transmitted through L_1=50 km of fiber, amplified and then passed through another L_2=45 km of fiber. The total fiber length employed in the second part of the experiment was about 95 km.

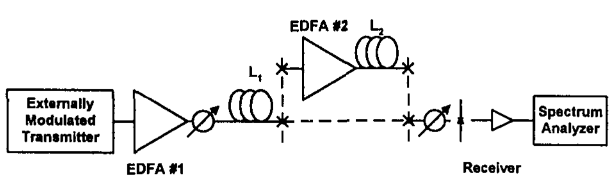


Figure 1. Experimental setup employed to investigate SPM-dispersion induced CSO distortion.

To increase the optical power threshold limited by SBS, we have both phase and frequency modulated the optical field by applying out-of-band RF tones. The optical modulation depth measured over the CATV band has a down tilt of about 1.8 dB due to the transmitter response. The modulation depth was measured to be 2.7%, 2.5% and 2.2% at channels 2, 41, 78, respectively. In our calculations, we have used the modulation depth measured at the mid-channel (i.e., m=2.5%).

Table 1. List of Parameters used in our Model

Parameters	
λ	1553 nm
n_2	$2\cdot10^{-20}$ m^2/W
A_{eff}	80 μm^2
α	0.0576 Np/km
D	17 ps/nm·km
K_2	29

Figure 2 plots the CSO of the system as a function of fiber input power. Inspection of Figure 2 indicates that as the input power to the fiber is increased the SPM-dispersion induced CSO degrades. At 18 dBm fiber input power the measured CSO is -66.8 and -62 dBc when m=2.5% and 3.4%, respectively.

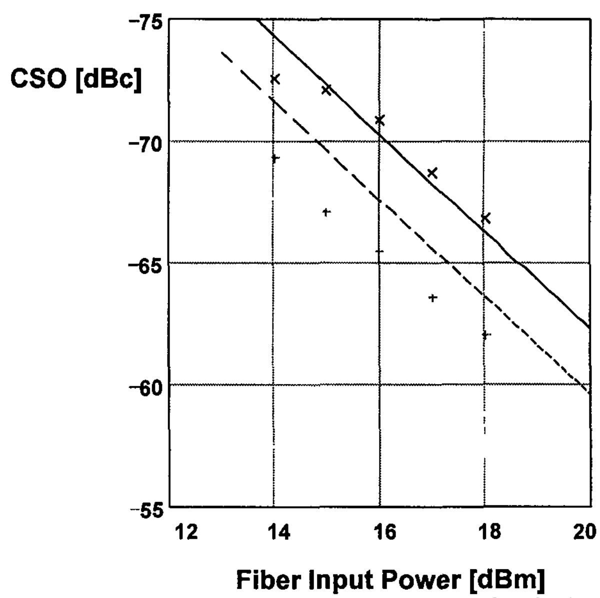


Figure 2. Plots of measured (x and +) and calculated (solid and dashed lines) CSO versus fiber input power at channel frequency 547.25 MHz and 50 km fiber. Solid line and (x) indicate CSO for m=2.5%; and dashed and (+) indicate CSO for m=3.4%.

Figure 3 plots the CSO of the system as a function of fiber length. Inspection of Fig. 3 indicates that at fiber lengths greater than about 50 km the measured and calculated CSO behave somewhat differently as the fiber length is increased. The discrepancy observed between the measured and calculated CSO is currently under investigation.

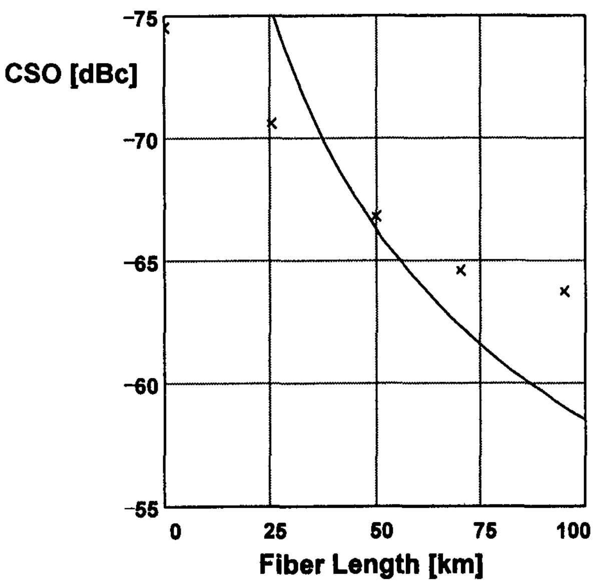


Figure 3. Plots of measured (x) and calculated (solid line) CSO versus fiber length for 18 dBm fiber input power and at channel frequency 547.25 MHz. Solid line and (x) indicates CSO for m=2.5%.

Figure 4 shows the system noise spectrum within the CATV band for various fiber lengths. The system noise spectrum is composed of a noise floor due to shot, thermal, laser intensity and amplifier noise; and frequency dependent noise due to fiber chromatic dispersion induced PM-AM noise. At lower frequencies, the noise spectra converges to the noise floor, while at higher frequencies the noise increases proportionally to the square of the frequency and fiber length. The system CNR penalty due to PM-AM noise is reduced by using a source laser with very narrow linewidth. Figure 5 shows the CNR penalty as a function of $\Delta\nu L^2$. The CNR penalty is calculated using the following parameters: received power 0 dBm, thermal noise 7 pA/$\sqrt{\text{Hz}}$, laser RIN -164 dB/Hz and amplifier RIN -156 dB/Hz. To achieve a CNR penalty of less than 0.5 dB for transmission over 50 km or 70 km of fiber length, the required laser linewidth is 440 kHz or 225 kHz at 550 MHz, respectively.

IV. CONCLUSIONS

We have investigated the interaction of SPM and laser phase noise with chromatic dispersion in an externally modulated AM-VSB lightwave CATV system. The CNR degradation associated with PM-AM noise is minimized by employing narrow linewidth laser sources. The CSO performance due to SPM-dispersion interaction can be a limiting factor in high power long-haul CATV trunk systems. The CATV system must be designed according to these limitations.

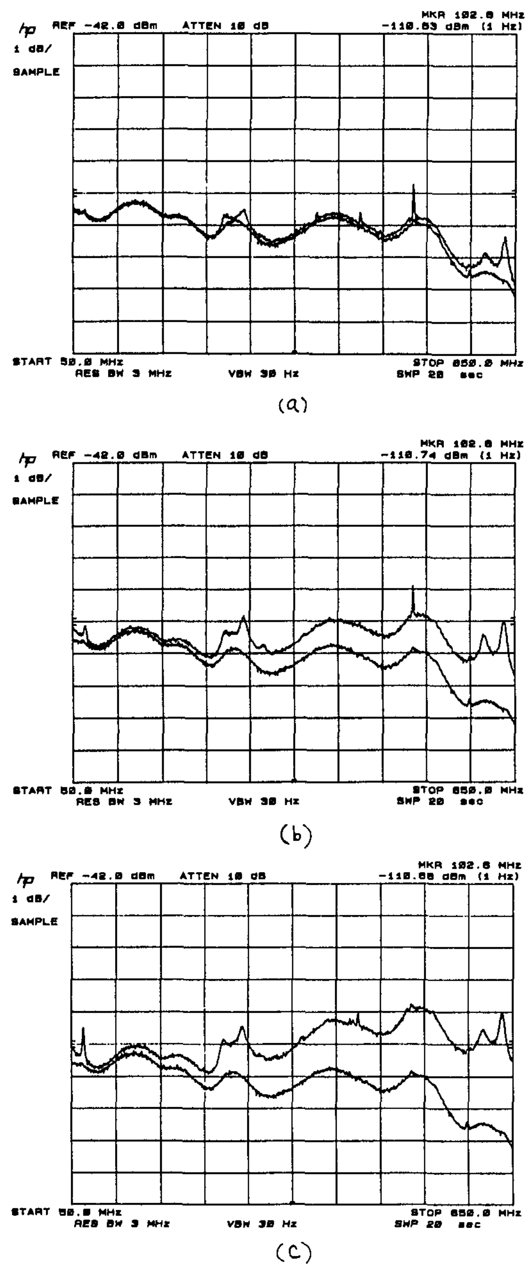


Figure 4. Plots of system noise spectrum with (a) 25 km, (b) 50 km and (c) 70 km of standard fiber (upper traces) and short fiber (lower traces).

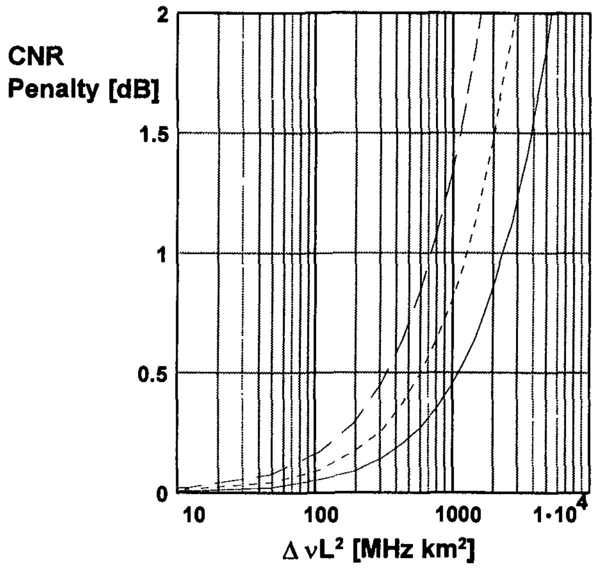


Figure 5. Plots of calculated CNR penalty versus ΔvL^2. Solid, dotted and dashed lines indicate the penalty at 550 MHz, 750 MHz and 1000 MHz, respectively.

REFERENCES

1. M. R. Phillips et al, "Nonlinear distortion generated by dispersive transmission of chirped intensity-modulated signals," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 481-483, 1991.
2. C. Desem, "Composite second-order distortion due to self-phase modulation in externally modulated optical AM-SCM systems operating at 1550-nm," *Electron. Lett.*, vol. 30, pp. 2055-2056, Nov. 1994.
3. F. W. Willems et al, "Experimental verification of self-phase modulation induced distortion in externally modulated AM-VSB lightwave systems," in *Proc. OFC'96*, San Jose, CA, Feb. 1994, paper ThR4, pp. 281-282.
4. J. Ming, K. Petermann, "Small signal analysis for dispersive optical fiber communication systems," *IEEE Lightwave Technol.*, vol. 10, no. 1, pp. 96-100, January 1992.
5. N. Kikuchi, S. Sasaki, "Analytical evaluation technique of self-phase modulation effect on the performance of cascaded optical amplifier systems, *IEEE J. Lightwave Technol.*, vol. 13, no. 5, pp. 868-878, May 1995.
6. T. G. Cross, "Intermodulation noise in FM systems due to transmission deviations and AM/PM conversion, " *Bell Sys. Tech. J.*, vol. 45, pp. 1749-1773, Dec. 1966.
7. M. R. Phillips et al, "Chromatic dispersion effects in CATV analog lightwave systems using externally modulated transmitters," in *Proc. OFC'96*, San Jose, CA, Feb. 1996, postdeadline paper PD17-1.
8. A. R. Chraplyvy et al, "Phase modulation to amplitude modulation conversion of CW laser light in optical fibers," *Electron Lett.*, vol. 22, no. 8, pp. 409-411, April 1996.

HEADEND MANAGEMENT OF MPEG TRANSPORT STREAMS FROM MULTIPLE SOURCES

Caitlin Bestler, Manager Architecture & Advance Development
Zenith Electronics, Network Systems

Abstract

Traditionally the CATV Operator has had the option of grooming the signals carried over their plants. Applications have included arranging the RF channel line-up, the logical channel line-up, Conditional Access tiering and local ad insertion.

Digital material is carried within MPEG Transport Streams, each with multiple services. Without the ability to separate those services the CATV Operator is faced with a choice between two undesirable options: surrendering control of many of these options to a remote satellite content provider, or paying for their own real-time MPEG encoding.

However there is a third option. MPEG Transport Streams can be managed in the headend, without decompression or recompression. With the techniques described the CATV Operator will once again have the tools to exercise per-service control over packaging the material carried on their network.

Techniques covered include selecting material from incoming MPEG Transport Streams, creating new MPEG Transport Streams from a mixture of remote and/or local sources, adding data and other services using the same downstream RF channels, and handling of local ad insertion in the digital domain.

BASIC TRANSPORT STREAM MIXING

The ability to obtain combine services from multiple MPEG sources is vital if CATV Operators are to maintain control over the services provided over their network.

This section will review the need to obtain material from multiple sources, the possibility granularities when mixing them, and the business implications of less than per-service mixing.

The Need For Multiple Sources

MPEG Material will be available from several sources to each CATV headend:

Satellite feeds will be available. Because of their wide delivery base satellites can be the most cost effective means of obtaining most digital programming. However for those very reasons it cannot be the source of all material. Satellites cannot carry local material, and certainly not local ads. They cannot even adapt schedules of nationally distributed material to local tastes.

Public SONET and/or ATM networks may be carrying MPEG Transport material primarily intended for Switched Digital Video (SDV) deployments.

Local Video Servers can be cost effective if their capacity is kept under control. When dealing with local ad insertion or the top titles for NVOD (Near Video On Demand) services they could prove to be very effective. However, the cost of storing hundreds of movies at **each** headend will simply not compare with the cost advantages of a satellite feed that is distributing its signal to a much wider market. Further, almost by definition, Video Servers cannot provide MPEG material for live events.

Real-time MPEG encoding is required to provide digital service when the material is received in analog format, for example with current off-air channels and existing local production equipment.

The need for multiple MPEG sources can be limited In a fully hybrid analog/digital network, supporting hybrid analog/digital set tops, along with analog only set tops and cable ready TVs and VCRs. Such a network has numerous advantages in preserving current equipment investments and efficient spectrum utilization.

However, there are several reasons why the CATV Operator may wish to provide services received in Analog format via MPEG to those customers with digital set tops:

- DBS and DVD (Digital Video Disc) marketing may succeed in creating consumer demand for "digital video" quality *per se*, not merely as a means of accessing more material.
- A digital-only set top box costs less than a hybrid analog/digital set top box.
- Digital conditional access and security is inherently superior to analog security.

Once they are being broadcast, off-air HDTV and/or SDTV signals will be MPEG Transport streams. Whether or not any portion of these channels are "must carry" services is not known. However they will be the easiest way to provide local broadcast services to customers with digital-only set tops.

Pure Data Sources will be available to supply data feeds independent of MPEG A/V programming on serial, ethernet or other data networking interfaces. Applications would include utility data for the set top Navigation and EPG programs as well as full Cable modems.

Selection Granularity

Given that the CATV Operator will be mixing and matching material from multiple sources the question is at what granularity: full 6 MHz channels or on a per service basis?

Material arriving from a satellite feed or other network will typically be pre-bundled into what the Satellite Distributor believes should be 6 MHz Transport Streams. This will typically be 28 Mbits worth of material, and could easily represent seven movie channels.

This is a fairly large chunk of material. Without the ability to separate these externally supplied feeds the CATV Operator is faced with a take-it or leave-it choice on the material.

Should CATV Operators' technology choices leave themselves in such a "all or nothing" position the Satellite Providers would be foolish if they did not take advantage of the situation and bundle weaker material with strong material.

The CATV Operator should retain the option of picking which services will be carried on their network. This can apply to whole services, portions of a multiplex, or even NVOD theaters. The supplied satellite feed may be showing movies every 15 minutes. If that is overkill in this market wouldn't it make sense to have the technological option of dropping every other showing to provide 30 minute NVOD?

Requiring local material and data services to be an even multiple of 6 MHz, carrying either 28 or 38 Mbits, would be especially wasteful.

Consider local community access programming. Will you **ever** need a full 28 Mbits of local community access programming per community?

Data services may be a valuable mix, but will they always consume an even 28 or 38 Mbits capacity? This is especially true for HFC plants where data services can be carried on a per-node basis.

LOCAL CONDITIONAL ACCESS

The CATV Operator should not only retain control over which of their subscribers is authorized for what material, but also on how that material is packaged. What services are part of what tiers?

Without these capabilities the CATV Operator is reducing their role to that of a pipe provider, or at most an order taker. Traditional leading roles in marketing material and building a customer base will be lost.

Obviously these issues must be negotiated with the Copyright Holders of the source material, but without the capability of controlling tiering locally there isn't anything to negotiate about.

Local Ad Insertion

Local Ad Insertion should remain a major part of the business plan even as material shifts from analog to digital.

CRITICAL ISSUES

Given that the ability to control and manipulate MPEG Transport streams is desirable, the question then becomes what types of equipment should be considered for these jobs.

This section will introduce some of the technical issues in handling MPEG Transport streams as a lead-in to discussing possible solutions.

Split, Merge or NxM Multiplexing

The next question is how much mixing of Services is required. The basic options are:

- **Split Only**: (1 to N) each incoming Source Transport Stream is split into N outgoing Transport Streams on the basis of services. Each outgoing Transport Stream contains material solely from a single incoming stream, but not all of it. Additional data services can be fill the RF channels capacity.
- **Merge Only** (N to 1) each outgoing Transport Stream combines material from N incoming Transport Streams, dropping specific services as configured to fit within the channel's capacity.
- **Full Mix** (NxM) any service from any incoming Transport Stream can be placed within any outgoing Transport Stream, subject solely to the capacity of the outgoing Transport Stream.

The Merge and Full mixing options require the ability to relabel MPEG Program Numbers and PIDs. Further, since these numbers are themselves referenced within System Information tables, those will have to edited and/or replaced to carry the modified references.

Any downloaded applications that reference MPEG Program Numbers or PIDs directly, without going through the proper directory access procedures may pose a problem. This can be solved by restoring all MPEG Program numbers and PIDs for the selected MPEG service either while demultiplexing or through software illusions, or preferably just by simply refusing to support the practice.

When splitting an MPEG Transport Stream based upon service it may be necessary to duplicate certain elementary streams. Of course this should be avoided whenever possible, and never done if it requires duplicating a video elementary stream.

The likely conditions when it would occur are for data streams associated with multiple services. A data stream associated with an NVOD service would be an example. Each showing might want access to the data stream, but in the bandwidth allocation chosen the NVOD service may have been split between two RF channels.

Far more common will be sharing of the PCR tracking PID. Depending on what PID was selected, the selected set of services may or may not have use of the data on the designated PCR PID. If not, the ideal split would copy only the adaptation headers containing the PCR values and reduce the total bandwidth required.

The choice between Split, Merge and NxM mixing has a direct impact on redundancy planning. Split and Merge modes can support N+1 redundancy schemes only at the level of the entire set of equipment supporting an RF channel.

If one real-time MPEG encoder fails either the entire set of equipment supporting an RF channel must be switched to the hot standby equipment set, or there must be full (N+N) redundancy on the MPEG encoding equipment.

With NxM mixing each component of the system can have N+1 redundancy. Whenever work is switched to the hot standby it's inputs/outputs can be routed as required.

Mixing Is More Than Packet Switching

As you have probably already read many times each MPEG Transport Stream is a multiplex of several Elementary Streams. At first glance that makes the problem of manipulating MPEG Transport Streams appear to be very similar to other packet-switching applications such as ATM.

However MPEG Transport Streams cannot be "switched" in the way that ATM can. That is because each Transport Stream is not only a collection of Elementary Streams, it is a **self described** and **self-timed** collection of Elementary Streams.

Transport Stream Self-Description: System Information

Not only are there Elementary Streams carrying Video, Audio and Private Data for applications, there are other Elementary Streams that describe what Elementary Streams are present. Further that self description can take place at two different levels: a here-and-now level to support basic user navigation and another to support Electronic Program Guides.

All of this self-description is carried in a set of tables on various elementary streams. The MPEG Transport standard defines a standard format for download of tables that allows for updating of tables in sections, identification of new versions of tables and even pre-loading of 'next' versions of tables before they are required.

This self-description of an MPEG Transport Stream is contained in tables such as these:

- The Program Association Table (PAT), which is always carried in PID 0, identifies the Program Map Table (PMT) stream for each MPEG Program (or service) carried in this Transport Stream. It also identifies the PID where to Network tables will be found.
- The Conditional Access Table (CAT), always carried in PID 1, identifies where EMM messages are to be obtained for each service in this Transport Stream.
- Program Map Tables (PMTs) are used to specify what elementary streams are used for what portions of an identified service. For example what PID is the video in, etc.
- Program Information and Program Name tables are placed in the same stream as the PMT and supply further display and navigation information about the current program.
- The Network table stream is used to provide information about services available on other Transport Streams in the system and maintenance functions such as the System Time. Virtual Channel lineups would be included here.
- Electronic Program Guide data will usually be supplied as a Private Data feed as part of a pre-defined service.

Dynamic vs. Canned System Information

When creating a new Transport Stream from multiple sources, new System Information must be created for the new Transport Stream.

The old System Information cannot just be simply merged. If the multiple sources are truly independent they may have chosen the same PIDs or MPEG Program Ids. These numbers must be re-assigned to avoid conflicts.

Most current equipment solves this problem by the following steps:

- Each incoming packet is either thrown away or relabeled to the correct PID. All incoming System Information is thrown away.
- New System Information is inserted with the correct PIDs.

This approach is dependent on having correct information delivered off-line **in advance**. It requires a highly co-operative relationship with the source of the MPEG Transport stream.

Should the DBS system adjust its schedule dynamically in anyway, even for something as trivial as the PID selection, then the inserted System Information will be incorrect.

This approach also requires a cumbersome distribution of Program Guide information. It can be argued that as long as viewers rely upon printed schedules that the data cannot be updated that rapidly. However, even while maintaining compatibility with printed schedules the frequency of NVOD showings can be adjusted transparently to the user. An unexpectedly popular title could be bumped up to every 15 minutes, while another that is proving to be more of a cult favorite could be dropped to every hour.

There are ultimately only a few choices:

- Contractually lock your source to adhere to a precise schedule that must be provided off-line well in advance. Since this may restrict their flexibility in supporting other customers, such as DBS, expect negotiations to be difficult and/or expensive.
- Provide a separate real-time communication path from the Satellite headend to your headend and contractually obligate the supplier to provide prompt notification of all changes via this link. Not only will this link be expensive, but expect the contract negotiations to be even worse. The DBS Operator is now agreeing to maintain and

keep operational extra equipment that is not relevant to their DBS customers.

- Parse the System Information already being provided to the DBS boxes within the incoming Transport Streams.

Requiring CATV headend equipment to be able to parse the incoming System Information is clearly the only viable long-term solution.

Dynamic Source System Information

Real-time parsing of incoming System Information can be used in two different ways.

Canned System Information can still be used, only all material is identified relative to the System Information. The source of the output Transport Stream is no longer Port X PID Y, but Port X Virtual Channel Y's video PID. Should the mapping change, the new relabellling will adjust automatically, and hence no new system information is required.

However the ultimate solution is to merge the System Information base on CATV Operator specific rules. For example the CATV Operator would specify which MPEG services, or Virtual Channels were desired from each source on each output Transport Stream. The MPEG handling equipment would have to create the merged System Information on the fly from the source material.

Divergent System Information Formats

The MPEG Transport standard itself only deals with a portion of the System Information tables.

Other standards, such as DVB or the proposed ATV standards are required to define the exact set of tables in use. The earlier examples of System Information tables was based upon General Instrument's proposed System Information for Digital Television.

Dealing with multiple sources will mean dealing with multiple formats for System Information. With luck, only the ATV and DVB formats will be required, but there is always the chance of encountering proprietary formats, particularly for Electronic Program Guide data.

No matter how many formats System Information arrives in, the system set tops should only deal with a single format. The logic to deal with multiple formats would be error prone and drive up set top Flash and/or DRAM requirements. It is far preferable to require the information to translated in the headend CATV equipment to one of the standards based formats.

Variable Bit Rates

Putting together a Transport Stream from multiple sources requires knowledge that the proposed streams will fit within the capacity of the 6 MHz channel, whether it is 28 or 38 Mbits.

When the Source Transport Stream is statistically multiplexed there are some special considerations.

For each set of programs that are being switched as a group the following must be known:

- The **Sustained Bit Rate** for these service: what bit-rate must be available on a sustainable long-term basis for these services.
- The **Peak Bit Rate**: what is the maximum bit rate that these services will ever reach. Also associated with any peak rate is a defined period over which the actual rate must fall back down to the Sustained Rate.

For variable delay services, such as data, the two rates are a **Guaranteed Rate** and a **Sustainable Rate**. The Guaranteed Rate is lower than the Sustainable rate, meaning that for any short period the data service may be restricted to this lower rate, but it can count on catching up to the Sustainable rate within the defined time window. The headend equipment must offer sufficient buffering to hold the extra material until it can catch up.

In order to plan an output Transport Stream the component services must be picked such that the available capacity is not exceed by either.

- the sum of all Guaranteed and Peak Bit Rates.
- the sum of all Sustained Rates.

This solution complements the traffic requirements of Data and MPEG services naturally. MPEG Transport benefits from variable bit rate encoding to deal with more difficult scenes. The Peak Bit Rate is always greater than or equal to the Sustained Bit Rate. Data services, by contrast can survive with occasionally delayed delivery as long as the aggregate Sustained rate can be guaranteed.

Conditional Access

MPEG Transport Streams received from other sources may contain their own Conditional Access logic. According to the MPEG Transport standards Conditional Access is enforced in two types of elementary streams: ECM and EMM.

EMM (Entitlement Management Messages) streams are used to entitle individual set tops to receive (or not receive) specific services and elementary streams. This typically involves editing authorization bitmaps and/or lists and distributing decryption keys. EMMs typically individually addressed.

ECM (Entitlement Control Message) streams are used to control the decryption of other elementary streams. They enable set tops that have been previously authorized by EMMs to correctly decrypt the controlled material.

How these two stream types are handled will determine how much control the CATV Operator will have over the Conditional Access packaging of material on their network.

The ability to substitute EMM streams is required to have control over which set tops are authorized. However the EMM stream can be delivered out-of-band, even though doing so makes things easier for any cryptographic attackers.

ECM streams, however, must be delivered in-band within the MPEG Transport stream. Unless the ECM streams are **locally generated** the tiering of that Transport Streams material will be dictated by the source.

The CATV Operator would loose control over decisions such as what was part of an Extended Basic tier.

Transport Stream Self-Timing

MPEG Transport Streams are self-timed. Material that is jitter-intolerant, such as the Video and Audio, carry PCR (Program Clock Reference) values. PCRs allow the receiving MPEG decoder to lock its 27 MHz clock to the source clock.

An MPEG decoding subsystem can only lock to a single source clock at a time (of course a set top could theoretically contain multiple MPEG decoding subsystems, but each would still only be capable of locking to a single clock). All PCR sensitive streams for the selected service must be based upon the same clock.

When a single Transport Stream contains services from multiple sources there will be more than a single clock within the Transport Stream. This is allowed in MPEG Transport syntax, although it is seldom implemented that way when a single back of MPEG encoding equipment generates the entire Transport Stream.

Any form of multiplexing inherently causes at least a minor amount of jitter. Ideally the headend equipment should minimize the jitter it causes, if possible dejitter any jitter caused by ATM or other networked delivery of material, and properly not any jitter it caused by adjust PCR values in the packets it handles.

Audio Format Mixing

Sources carrying fully legitimate standards-compliant MPEG-2 Transport Stream can still have incompatible audio formats: some may be in MPEG or Musicam format, others will be in Dolby AC-3 format.

There are only two sure solutions to this problem:

- Restrict material selection to sources that includes the audio format you have selected. Some material may be available with dual audio streams to address this problem.
- Specify set tops that are capable of either. The cost premium of this option will continue to decline, and will eventually be the unbeatable solution.

It is **theoretically** possible to convert between the two formats without altering the matching video stream. Such a product would be highly desirable, if anyone were working on it. However I have yet to encounter even one engineer who claimed that it would be easy. Given the general optimism of engineers it would be prudent to rely on one of the first two options.

POSSIBLE SOLUTIONS

Full mixing of MPEG Transport Stream requires a vast amount of processing and communications horsepower.

Each incoming Transport Stream must:

- be captured according to the framing and error detection/correction of the delivery protocols. This could be a simple CRC as from a Satellite, or it could be ATM AAL-5 from an ATM network.
- packets must be relabelled and routed to the correct output Transport Streams.
- be tracked as to maintain lock with its 27 MHz clock and ensure that the rates supplied are within the configured range.
- have all incoming System Information tables parsed.
- dejittered as necessary and possible. This function is dependent on the type of network port material is being received from.

Each outgoing Transport Stream must:

- merge portions of the incoming Transport Streams routed to it.
- make PCR value adjustments to reflect any unavoidable jitter it creates.

Additional processing is required to update relabelling and routing instructions based on the parsed incoming System Information and/or to rebuild the new System Information to be output.

The solution should also be scaleable. The number of Transport Streams requiring mixing may start low and be expanded as time goes on. The solution should not require buying equipment capable of remixing 70 RF channels when only two will be required.

Add/Drop Remultiplexers

The first generation of solutions are variants on the equipment used to generate the original Transport Streams after real-time MPEG encoding.

This equipment typically is capable of relabelling PIDs and re-timestamping PCRs. Management plane interfaces must specify how each PID is to be handled, and supply new System Information for insertion.

To support dynamic tracking of sources these devices would have to learn how to parse and kick-out updated tables for other headend equipment. That equipment would then interpret the new tables and issue modified instructions to the Mux.

Each remultiplexer generates a single output Transport Stream. Implementation of Merge mixing is relatively easy. Each Remultiplexer is simply overfed, and drops PIDs as configured.

Splitting can also be performed on a limited basis. Each remultiplexer simply drops a different streams.

To accomplish full NxM mixing extensive splitters and wiring is required. Each remultiplexer would typically be overfed several times the material it required and drop virtually all of it. Such solutions are scaleable, but only to a certain degree before full ATM switching equipment is required in addition to the remultiplexers for each RF channel.

Backplane base NxM Mixers

An alternate approach is to perform the Transport Stream mixing function in a cardrack with processor cards connected on a backplane.

Modular processing cards communicate with each other over the backplane, eliminating the need for splitters and special wiring between each source and each multiplexer.

Each processing card deals with the processing requirements of a single input or output transport stream (or if very powerful possible two). Processing power need only be added as required.

Use of an industry standard backplane, such as PCI or VME-64, would allows easy addition of more processing power later. Network connection cards can be obtained that are not developed solely for this project. The processor cards can be replaced with later more powerful cards as new applications are found and/or costs per MIP drop.

Two types of backplane architectures are worth consideration: switched and shared. Switched backplanes allow multiple processor cards to feeding material to other cards concurrently. This is preferable, but currently results in far higher costs on a per port basis.

The shared backplane approach requires blocking backplane transfers, resulting in a fixed delay on the order of 100 msecs on all switchable services. With careful design portions of the multiplex which are not switched can be passed through without this fixed delay.

ADDITIONAL POSSIBILITIES

Local Ad Insertion

A processor based Transport Stream handler would naturally be capable of handling functions such as Ad Insertion in the digital domain.

However, it can also do so in a way that could radically reduce the cost of interfacing with an Ad Server.

Storing local Ads in digital format is already desirable for random access to them. More ads can be available for insertion, with no restrictions on which can be played in what order or when.

A Transport Stream Mixer as described, would be capable of handling a more flexible and cost effective interface. After loading a very short portion of the Ad into its memory, the Transport Stream Mixer could **pull** the remaining material at the clock rate of the overlaid material. This material could now be pulled over 100 Base T or switched 10 Base T network lines from a plain File Server. No special Video Server or OC-3 ports would be required.

Data Routing

Data received from network ports can be encapsulated within MPEG Transport packets and sent downstream as opportunistic data.

Additionally a Transport Stream Mixer could perform several routing functions:

- Tracking acknowledgments.
- Routing packets to an in-band Transport Stream when it knows the set top is tuned there and there is available capacity, rather than using up the limited capacity on any out-of-and Transport Stream.

Optimized Data Circulation

Many tables downloaded via MPEG Transport are circulated as a data carousel allowing the set tops to collect the data as required and/or able.

This type of data repetition is required to deal with plant noise, the set top being busy while changing channel, loss of power to set tops and to minimize the amount of DRAM required in each set top.

However it is hardly desirable to provision network inputs to accept these data feeds at their full speed. Accepting only the content of each data carousel table, its priority and duration would greatly simplify the interfacing between external servers, management systems and the MPEG Transport handling sub-systems.

Feedback to Video Server

The output of a Transport Stream Mixer could be routed back to a Local Video Server, rather than just being delivered.

This could allow for delayed rebroadcast of local material, capture of satellite distributed ads, and capture of new material for locally controlled NVOD re-broadcast.

CONCLUSION

The ability to handle MPEG Transport Streams within a CATV headend can provide valuable options to the CATV Operator. A single system based upon standard interfaces can provide an upgradable cost and space effective solution.

Incremental Deployment of Network Powering Infrastructure

Paul D. Brooks – Staff Engineer
Engineering and Technology – Plant Engineering Department
TIMEWARNER CABLE
Englewood, Colorado

ABSTRACT

Delivery of network powered lifeline telephony service over existing HFC plant requires deployment of an infrastructure capable of providing reliable power to both system actives and telephony equipment.

This power must be available on a continuous basis during both momentary and extended utility outages. The need for continuous power requires deployment of alternate power sources. The economics of continuous power point to changes in the way power is distributed, The increased loads imposed by telephony equipment require commiserate increases in installed capacity.

These facts cause powering system enhancement to be the largest new plant cost item required in connection with telephony service. Because a certain amount of this new powering infrastructure must be in place before the first telephone customer is turned up, much of this cost is fixed.

The challenge before us is to deploy the lowest cost powering infrastructure which will meet the immediate needs of service launch, and to develop a plan to incrementally increase power capacity as the business grows. This paper will suggest network design techniques which address this incremental deployment.

INTRODUCTION

Preface

The analysis presented here is focused on the delivery of network powered lifeline telephony service. Conceptual descriptions of network powering design techniques are provided for universal application to HFC architectures. The powering infrastructure costs presented include provision of power to system actives, telephony network interface units, and the cost of long duration standby power sources. Cost models are based on the generic pricing supplied to Cable Labs and used by participating designers at the recent design conference.

Business Model

When telephony service is launched, the penetration rate will be zero percent. Power will be consumed only by system actives. As the service becomes more successful, the power consumption will increase. This is contrary to the predicament faced by our telco competitors, who must provide for 100% take rates up front with the expectation that market share will be lost as new players enter the marketplace. This business model offers an opportunity for significant reduction in the initial deployed cost of powering infrastructure.

OBJECTIVES

The following objectives have been identified as key to the control of capital expenditures and the provision of sufficiently robust powering infrastructure.

Minimizing Supply Locations

Long duration standby power systems are typically large, relatively expensive units that require easy maintenance access, and often need connections to natural gas and 240V secondary lines. Status monitoring is also required in these complex units.

Use of natural gas fired generators creates a situation where the materials cost of a supply is no longer a linear function of output capability. Basically, small generators cost nearly as much as large ones.

Time Warner's experience in Rochester NY has demonstrated the difficulty and expense involved in finding suitable locations and purchasing required easements.

For these reasons it is critical to minimize the total number of supply locations needed.

Localized Capacity Expansion

Because of the high cost of powering infrastructure, capacity increases must be separately deployable on an as-needed basis as the business dictates. This requires the discipline to monitor business success on a node by node basis, and a responsive design and construction operation capable of rapid network re-configuration.

A topology aware Operation and Support System capable of localized penetration reporting is an essential part of incremental power deployment. At a minimum, the OSS needs the capability to report the number of telephony devices deployed on a node by node basis. This provides for periodic comparison of existing power consumption (based on penetration level) to installed capacity. When existing penetration reaches a predetermined "alarm" point, power provisioning activities are initiated. These provisioning activities may alternately be triggered by remote monitoring and reporting of actual field power consumption, with the trigger based on a comparison of this reported consumption with the calculated consumption at the deployed capacity level.

Avoidance of Stranded Investment

The act of localized infrastructure capacity expansion will be repeated multiple times in the most successful service areas. This expected behavior makes it useful to examine in advance the network changes required beyond a given immediate business need. Such an examination will prevent stranded investment by revealing the nature of facilities which may be needed in the future, thereby minimizing the duplication of construction labor and the installation of undersized cables and other inadequate equipment.

CHALLENGES

Power Consumution Variables

Significant variables exist in expected power consumption in connection with telephony service offering:

- It is certain that some HFC service area nodes will be more successful than others.
- Telephony customer distribution will not be uniform along all streets in a given node, but is expected to be erratic in nature, with clusters of high penetration intermixed with lower penetration areas and areas without any subscribers.
- Telephony traffic levels will affect power consumption. Traffic effects will vary based on demographics and on the use and degree of RF concentration techniques.

The infrastructure design must be sufficiently conservative to accommodate these variables.

Stress testing techniques are available to examine the effects of traffic and subscriber distribution, but are beyond the scope of this paper.

60 Volt System Actives

In plant where upgrades have recently been completed, system actives will typically not be capable of operation above 60 Volts. The capital cost advantages afforded by the use of 90 Volt powering are unlikely to offset the cost of the needed equipment changeouts, particularly at low initial telephony penetrations.

The biggest liability of 60 Volt powering is the lack of "reach" compared to 90 Volts. This limitation creates two problems when compared to powering at higher voltages:

- A greater number of supply locations are needed, with the associated cost, reliability, and maintainability penalties.
- For a given service area, longer and/or larger coaxial cables must be added to attain a given design penetration level. This is both costly and time consuming.

90 Volt System Actives

Use of 90 Volt powering offers substantial deployment cost efficiencies compared to 60 Volts. However, caution must be used in design such that network instabilities such as multi-state oscillations ("motorboating") or network voltage collapse and lockup are avoided.

Effects believed to contribute to these undesired behaviors are:

- Low end of line design voltages
- Excessive total power source loading
- Excessive power dissipation in coaxial cables
- Switch mode power pack designs which continue to draw significant currents at very low ac input voltages before cut-off occurs and restart is initiated.

Variable Power Source Output Voltages

It has been suggested that the voltage of the power source could be adjusted such that the first amplifier would see only 60 Volts when fed by a long coaxial cable run. This technique appears to be attractive in cases where 60 and 90 Volt equipment is intermingled, or in upgrade scenarios where 60 Volt equipment is being progressively replaced with 90 Volt gear.

However, a closer examination of this technique reveals a serious flaw. If a portion of the load were to be lost, the voltage seen by all remaining devices could increase significantly, with potential damage to 60 Volt equipment. Safe application of the variable voltage technique will require a control system capable of protecting against equipment overvoltage. The added complexity and cost of this equipment is likely to negate the marginal advantages of variable voltage powering.

Power Provisioning Outage Reduction

Incremental infrastructure deployment must take into account potential interruptions in service during provisioning. Techniques such as temporary load segmentation and backup, and pre-placement of critical power inserters will minimize provisioning related outages.

TECHNIQUES

The following discussion illustrates how network powering infrastructure can be incrementally deployed in a typical service area node.

Samnle Node Characteristics

A sample node was selected to be representative of the "Fiber Rich" HFC architecture currently being deployed at Time Warner. The selected node has the following characteristics:

- 553 Homes Passed
- 6.59 Strand Miles
- 84 Homes per mile

- 37 Active Devices
- 5.6 Actives per mile
- A mixture of P-l and P-3 cables
- 750 Mhz upgrade with analog loading of 77 channels and 200 MHz digital.
- Minimum tap port outputs of +1 1 dBmV (virtual 750 MHz) over +8 dBmV (55 MHz).
- Cable sizes ranging from .4 12" to .875"
- Maximum RF cascade of one optical receiver/launch amp, two trunk amps, one terminating bridger, and three line extenders.
- A mixture of aerial and underground plant, mostly aerial.

Telenhonv Power Consumntion and Traffic Assumntions

While penetration is variable and is incremented in the design process, rules must be established to define how power consumption is allocated to each home passed. The following assumptions were selected to represent near term deployment of network interface units "NIUs" typically attached to the sides of single family residences.

Traffic assumptions:

- 3.6 ABSBHCCS. This is perhaps better explained as a design capable of a maximum of 10% of the deployed lines being off hook at any one instant.
- 2% Ringing.

Penetration assumptions:

- 10% of the NIU population has activated second line capability.
- Penetration is evenly distributed throughout the service area. This can be explained as follows: If a given pole location passes five homes, and the penetration is at 10%, then one half of one NIU power budget described previously is consumed at that pole location.

Power consumption:

- Idle State: 3 Watts
- Off-Hook: add 1 Watt per line to idle
- Ringing: add 10 Watts per line to Off-hook/idle combination.
- NIU design power allocated on a per telephony subscribing home basis: 3.63 Watts

Design Strategv

Powering designs were performed on the sample node using 60 Volt supplies, and again at 90 Volts. Real world active devices were selected with wide input voltage ranges, such that the same actives could be used in each case. The only specification difference was in the operating voltage range of the active devices. The voltage ranges used were as follows:

- 60 Volt design operated the actives (including NIUs) over an input voltage range of 40 to 60 Vac rms.
- 90 Volt design operated the actives (including NIUs) over an input voltage range of 45 to 90 Vac rms. The 45 Volt minimum was selected as a safety precaution against network stability concerns.

As stated earlier, an important objective is to avoid stranded investment. To this end, the design task is made considerably easier when the network is optimized at 100% telephony penetration initially, and lower penetration deployment strategies are developed based on the nature of the infrastructure needed at 100%. This conclusion was arrived at after considerable experimentation with designs starting at 0% and working upwards. When the next penetration level is already known, guesswork is eliminated as to cable sizes and service area segment boundaries.

The Design Process

The first step in powering design is driven by another of the objectives stated earlier, the need to minimize supply locations, The ideal node service area is fed from one location only.

A design at 0% telephony penetration is attempted, using a location central to the power load. The best site is not necessarily co-located with the optical receiver. If the entire node cannot be served from one location with minimal addition of coaxial power feeder cables, the node is then divided in half and the exercise repeated. This was the case for the sample node at 60 Volts, where two locations were needed.

Once the minimum number of supply locations is established, the most efficient layout capable of 100% telephony is determined for each supply location, using conventional design techniques. This design is then saved for future use.

Portions of the 100% telephony penetration design infrastructure are successively removed, and each power supply location service area is tested to determine its new, reduced penetration capacity rating. It becomes quickly evident to the designer which changes are associated with given penetration capabilities.

Each penetration level is saved with a unique name, and used as the basis to produce the next lower level. This process is repeated until all infrastructure other than that required at 0% penetration has been removed.

The 60 Volt Examples

Figures 1.) through 15.) illustrate the progressive penetration design technique applied to an actual HFC node service area design, using a supply voltage of 60 Volts. Multiple output ports are implemented at each power supply location to provide power into each power feeder cable, and to segment the load in order to avoid device current rating violations. In each figure, the thick black lines indicate power feeder cables that were required to attain the particular stated telephony penetration level. Tags attached to these lines indicate the type and length of cable needed. The squares containing a sine wave symbol are ports or modules at a common power supply location. The 60 Volt design, as stated earlier, required two such locations.

In Figure 2.), the replacement of a "weak link" of high resistance .4 12 cable, and the addition of one more output port were all that was needed to achieve 11% penetration capability. In Figure 3.), 15% was achieved by replacement of two other such weak links.

In Figures 4.) through 15.) power feeder cables and ports were added, and power stops moved to attain progressively greater capacities. Note that in Figure 15.), two power feeder cables were required in both the 1337 foot span and the 383 foot span. While this does represent labor cost incurred twice, It is possible that this particular node will never attain the penetration levels necessary to mandate the addition of these second cables.

Also note the long series of power feeder cable sections successively added in the trunk run at the bottom of Figure 15.). These power feeder extensions were required due to a voltage starvation problem in the longest cascade in this node. The underground sections of this run increased the cost substantially.

The 90 Volt Examples

Figures 16.) through 24.) illustrate the progressive penetration design technique applied with a 90 Volt supply. Note that the penetration jumps are larger, and the 100% configuration shown in Figure 24.) is no more complex that the 60 Volt version in Figure 15.), even though only one location was used.

Figure 1.60 Volts, 0% Penetration

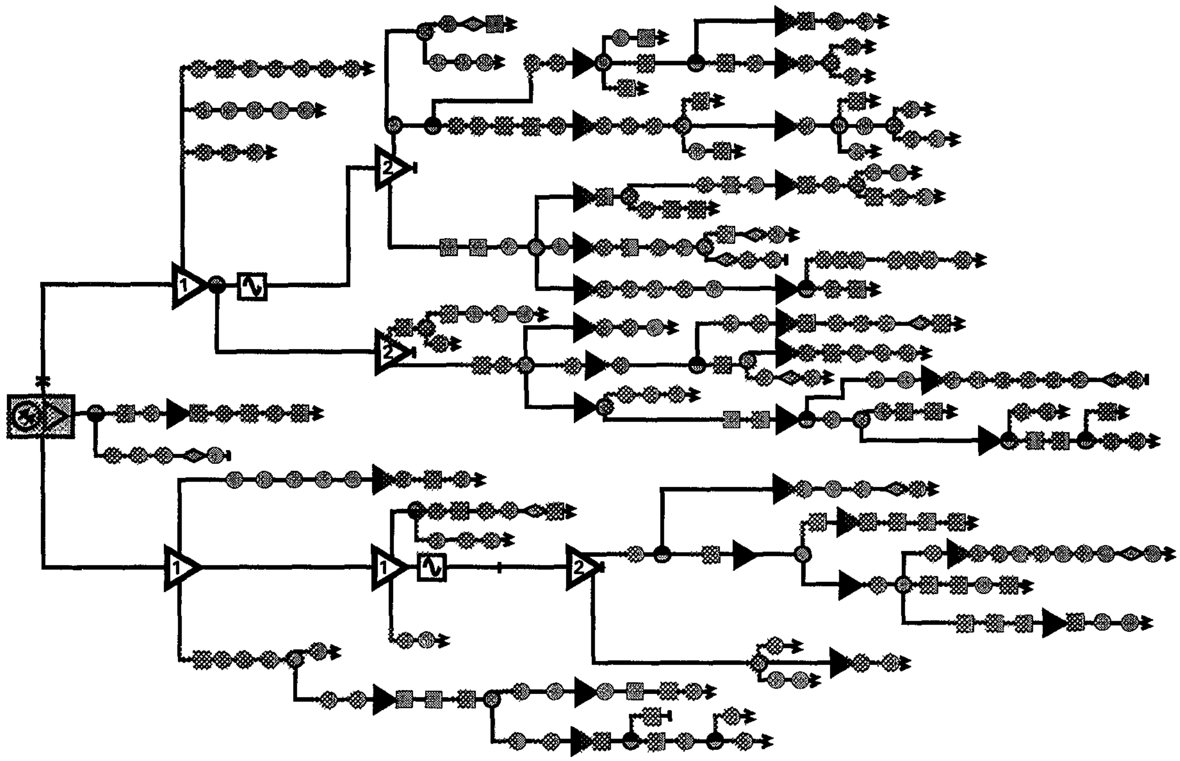


Figure 2.60 Volts, 11% Penetration

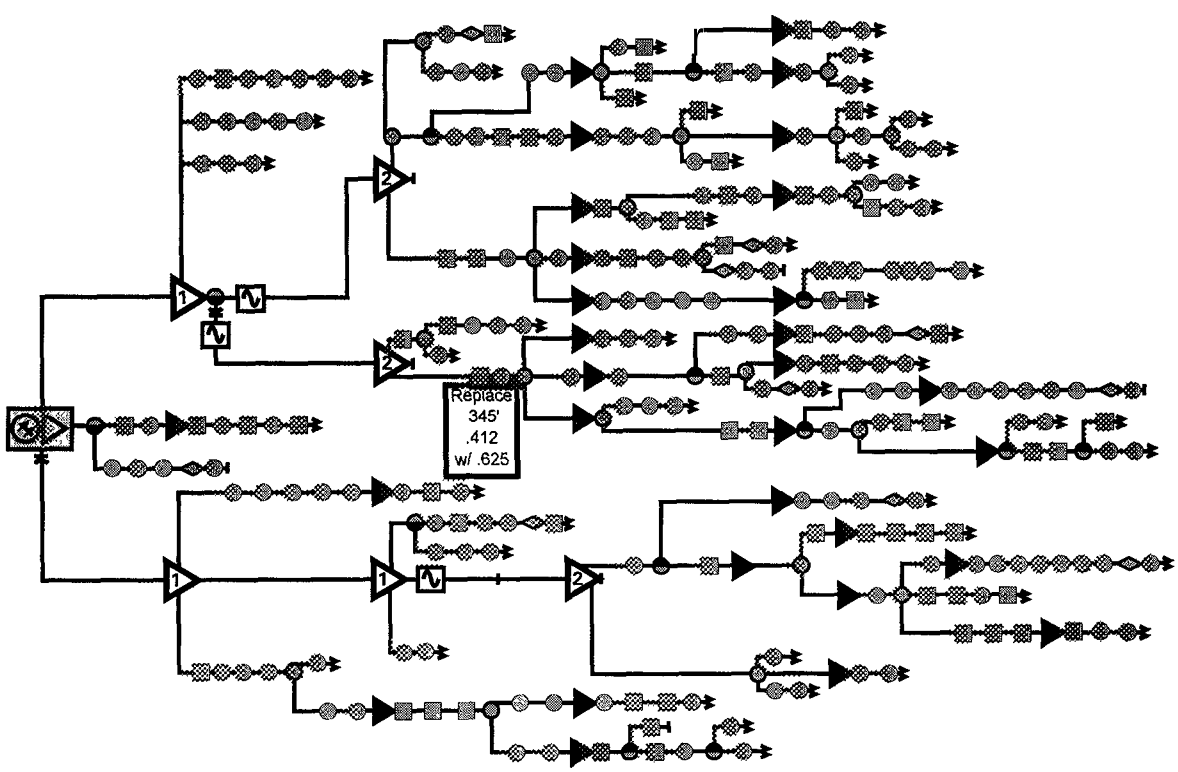


Figure 3.60 Volts, 15% Penetration

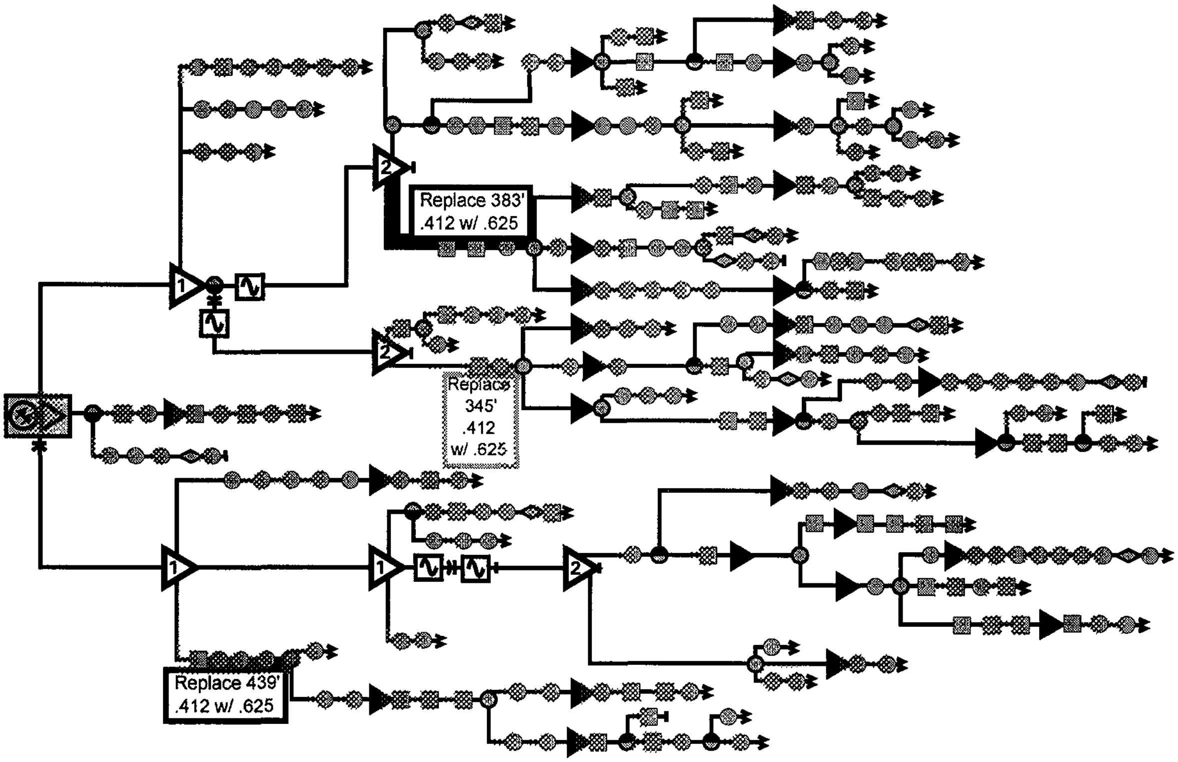


Figure 4.60 Volts, 21% Penetration

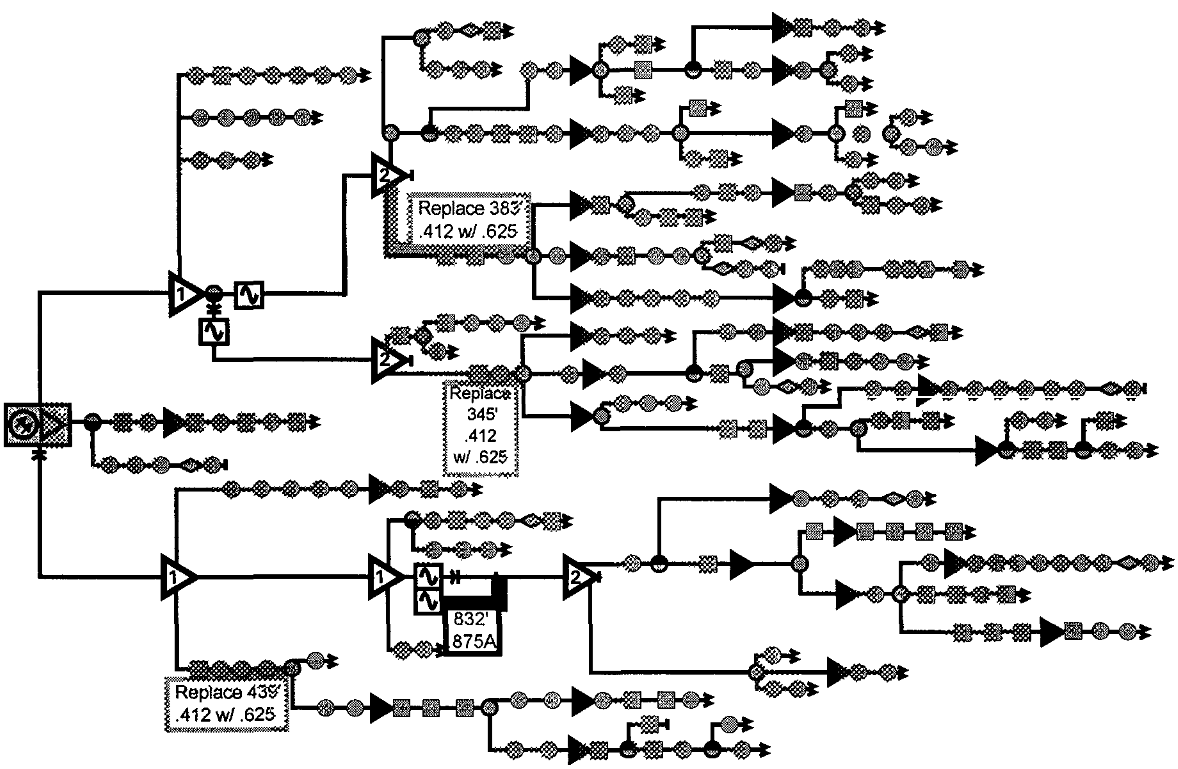


Figure 5.60 Volts, 25% Penetration

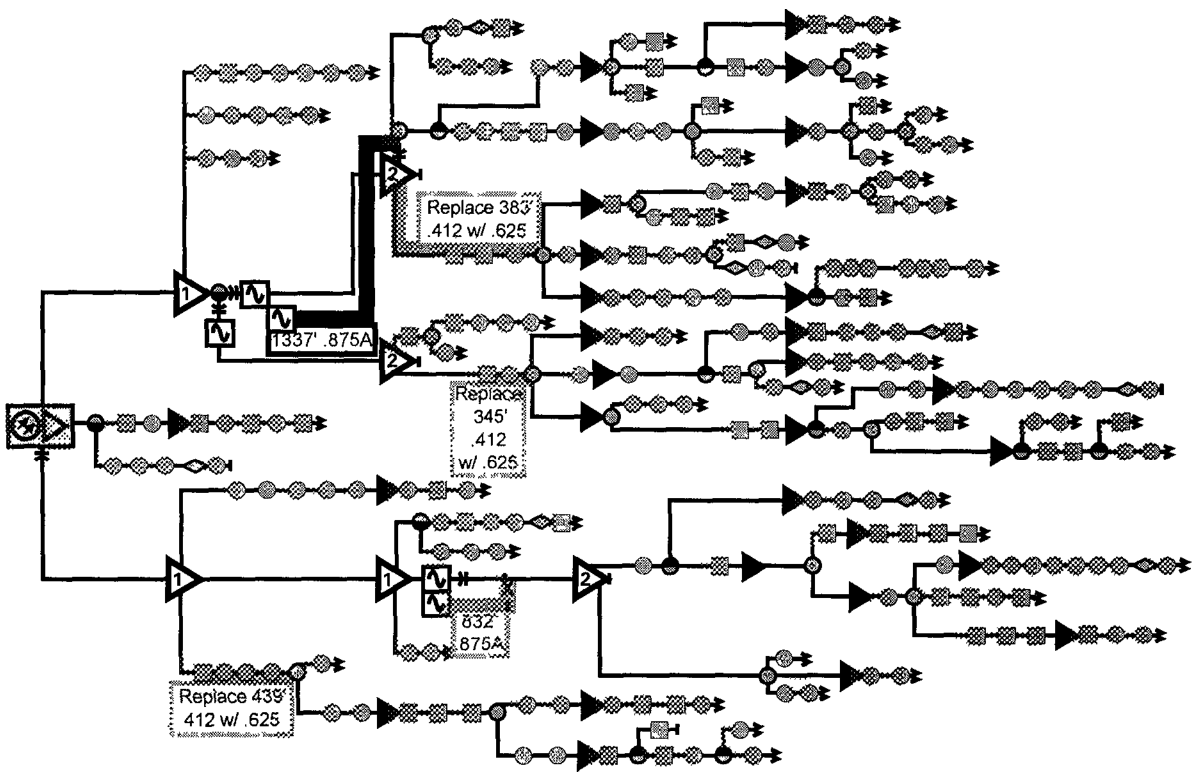


Figure 6. 60 Volts, 35% Penetration

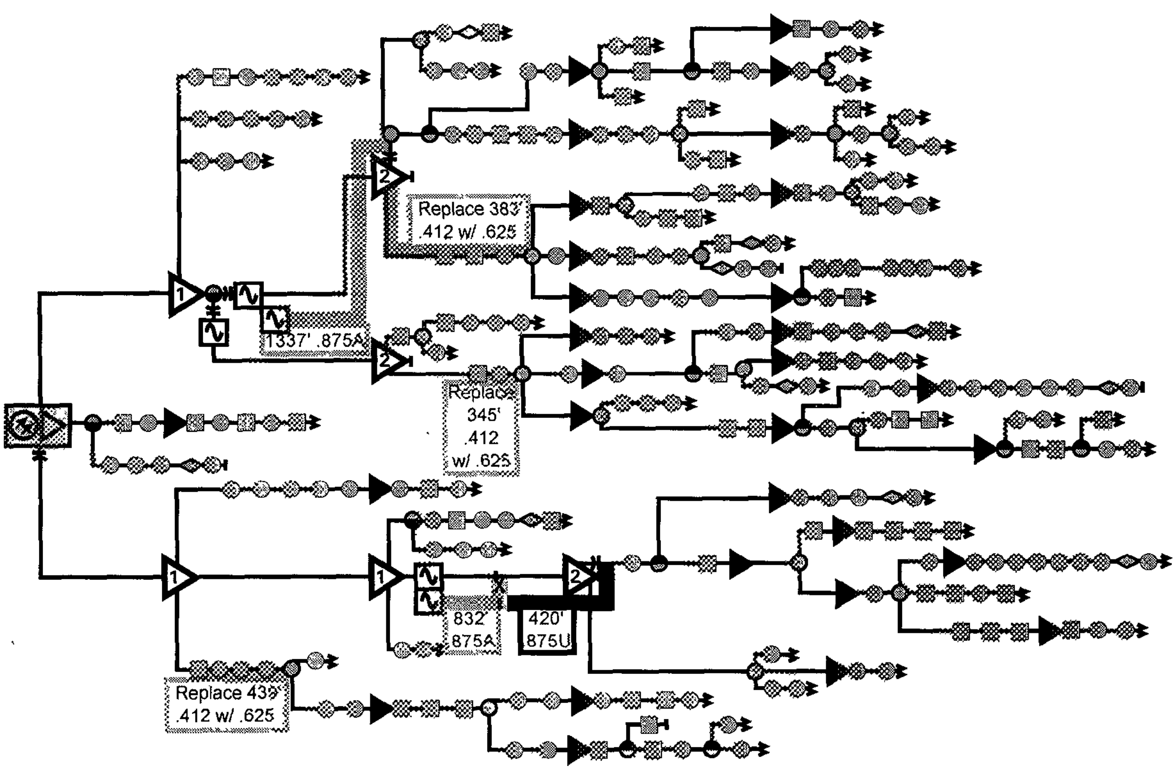


Figure 7.60 Volts, 46% Penetration

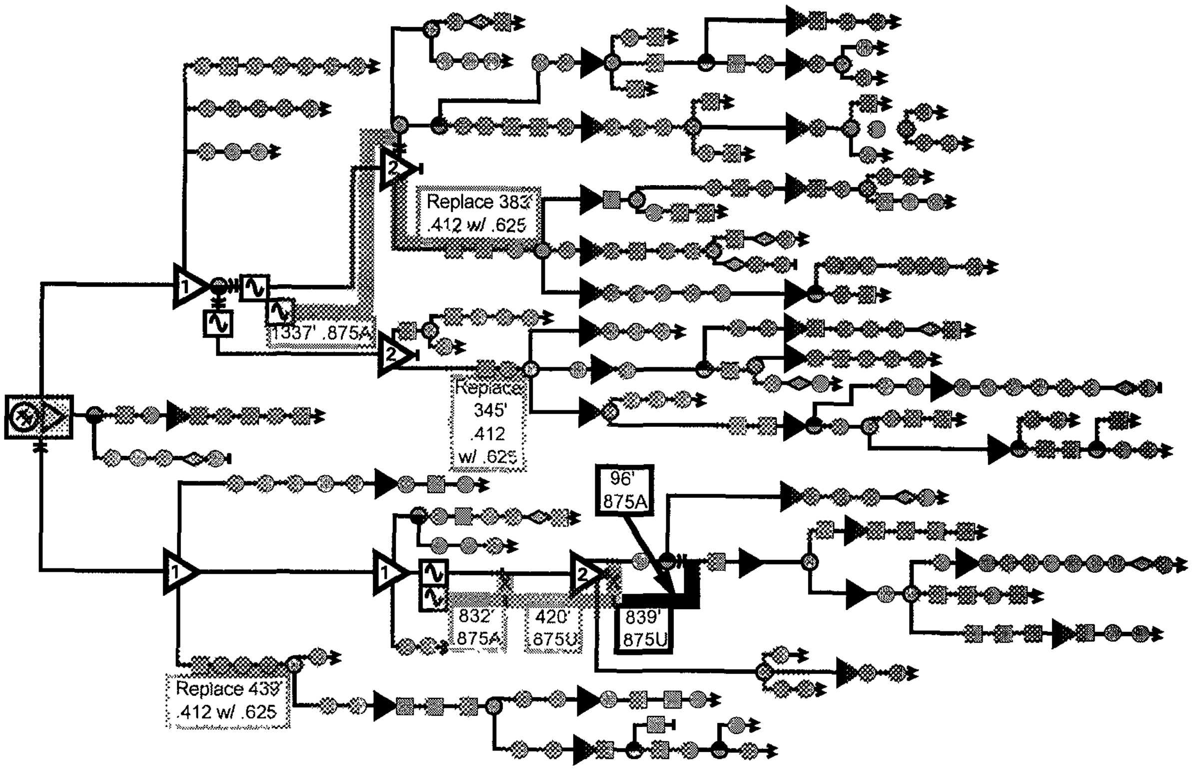


Figure 8.60 Volts, 48% Penetration

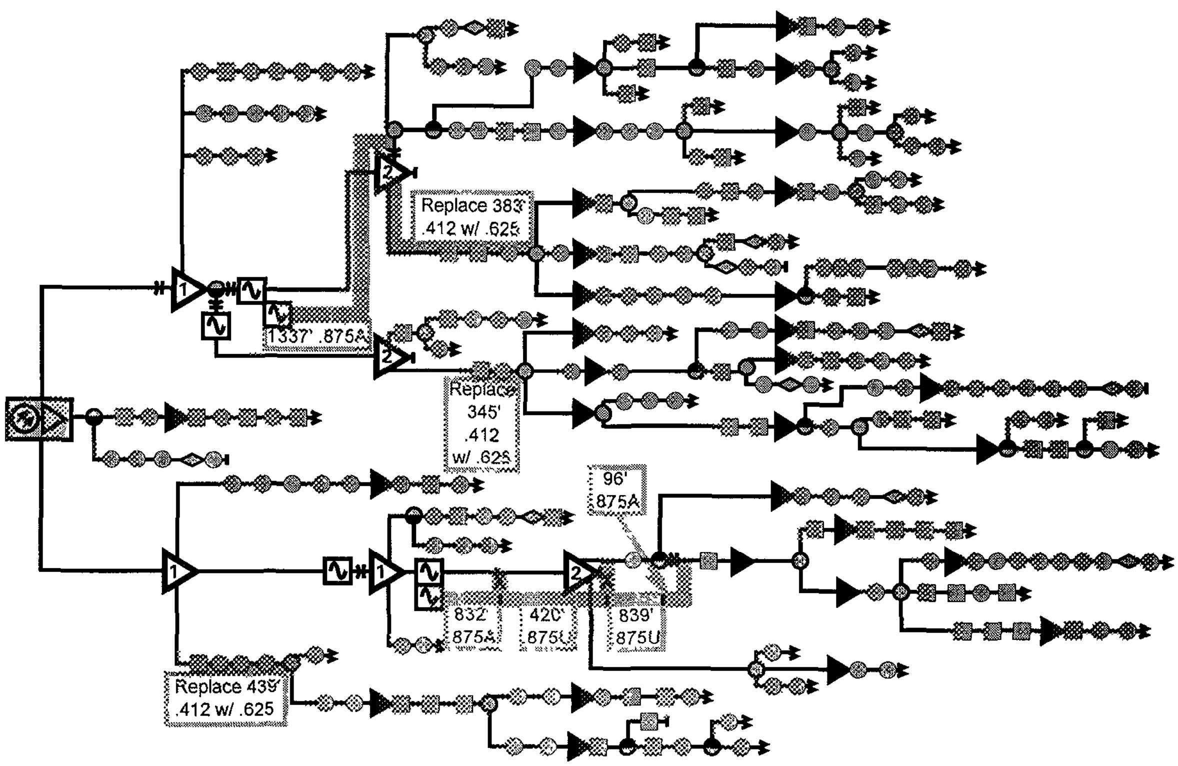


Figure 9.60 Volts, 52% Penetration

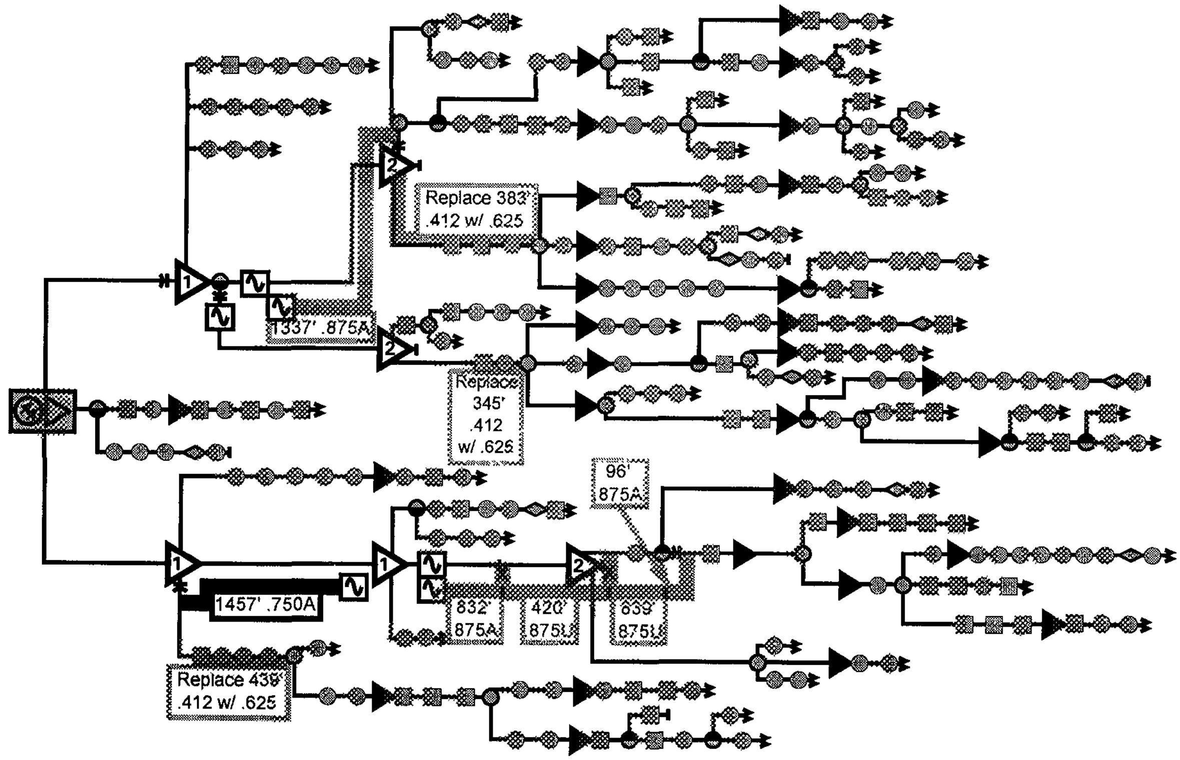


Figure 10.60 Volts, 57% Penetration

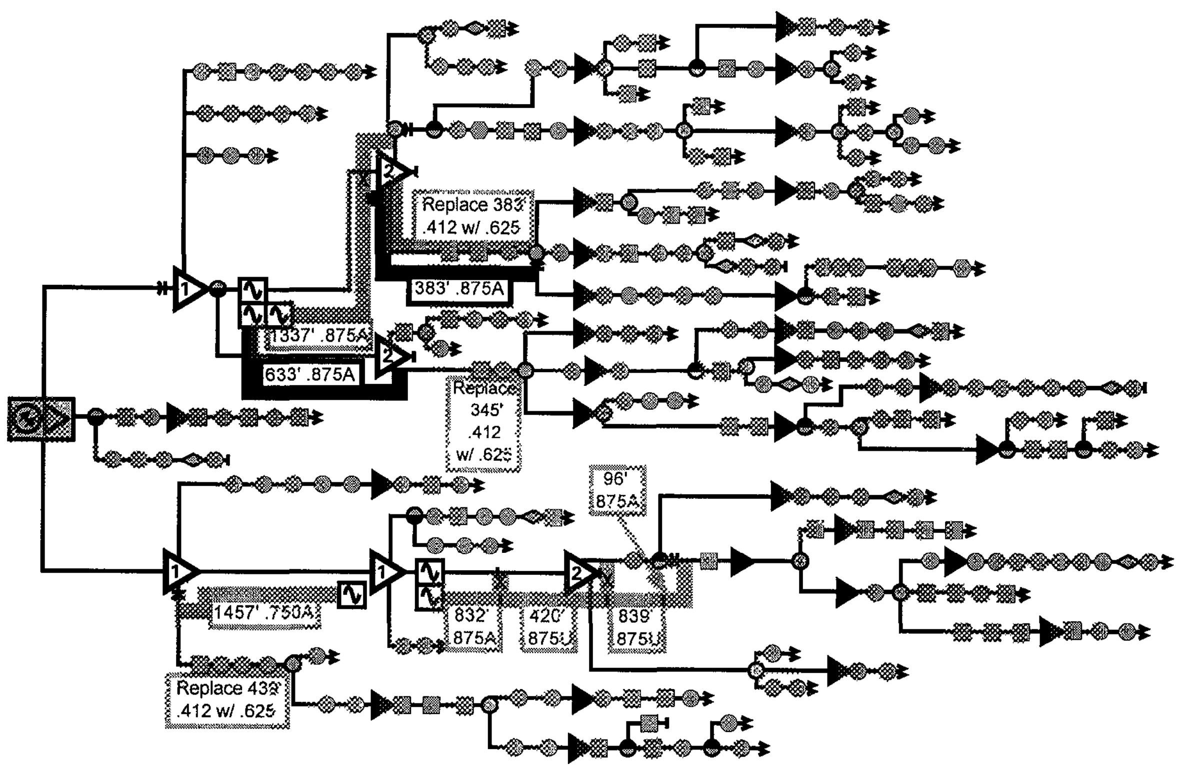


Figure 11. 60 Volts, 69% Penetration

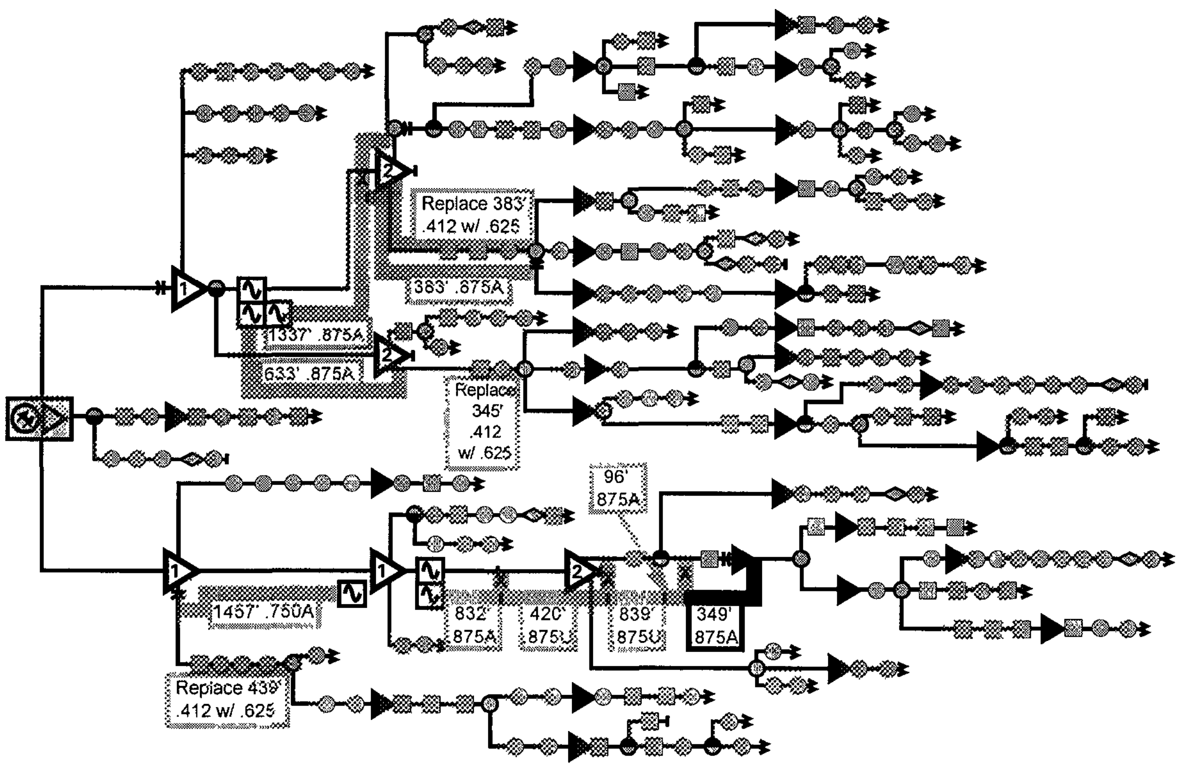


Figure 12.60 Volts, 71% Penetration

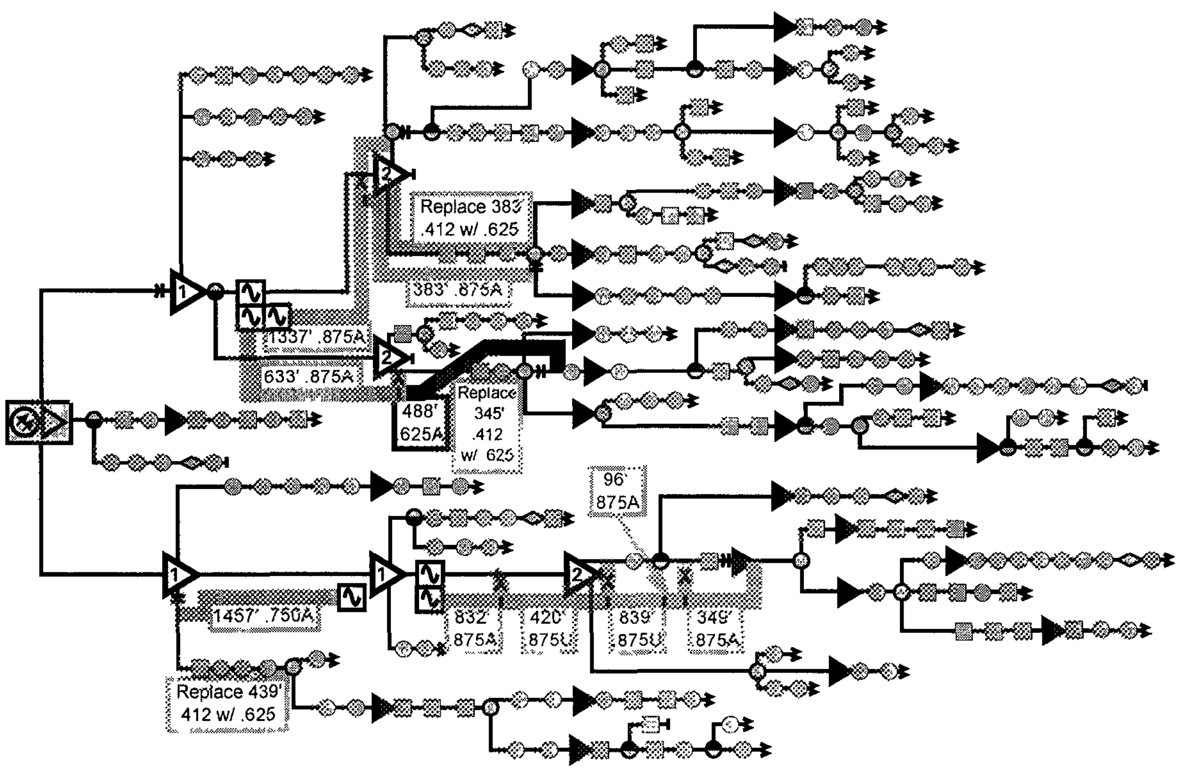


Figure 13. 60 Volts, 76% Penetration

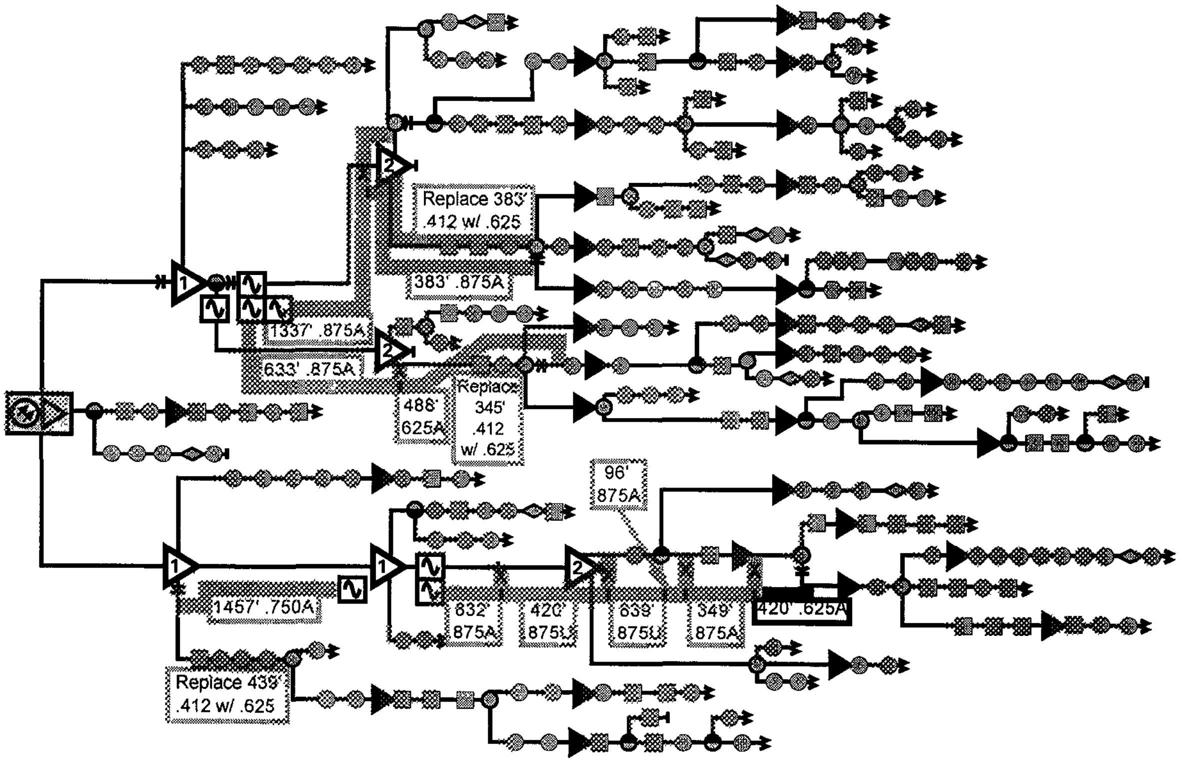


Figure 14. 60 Volts, 84% Penetration

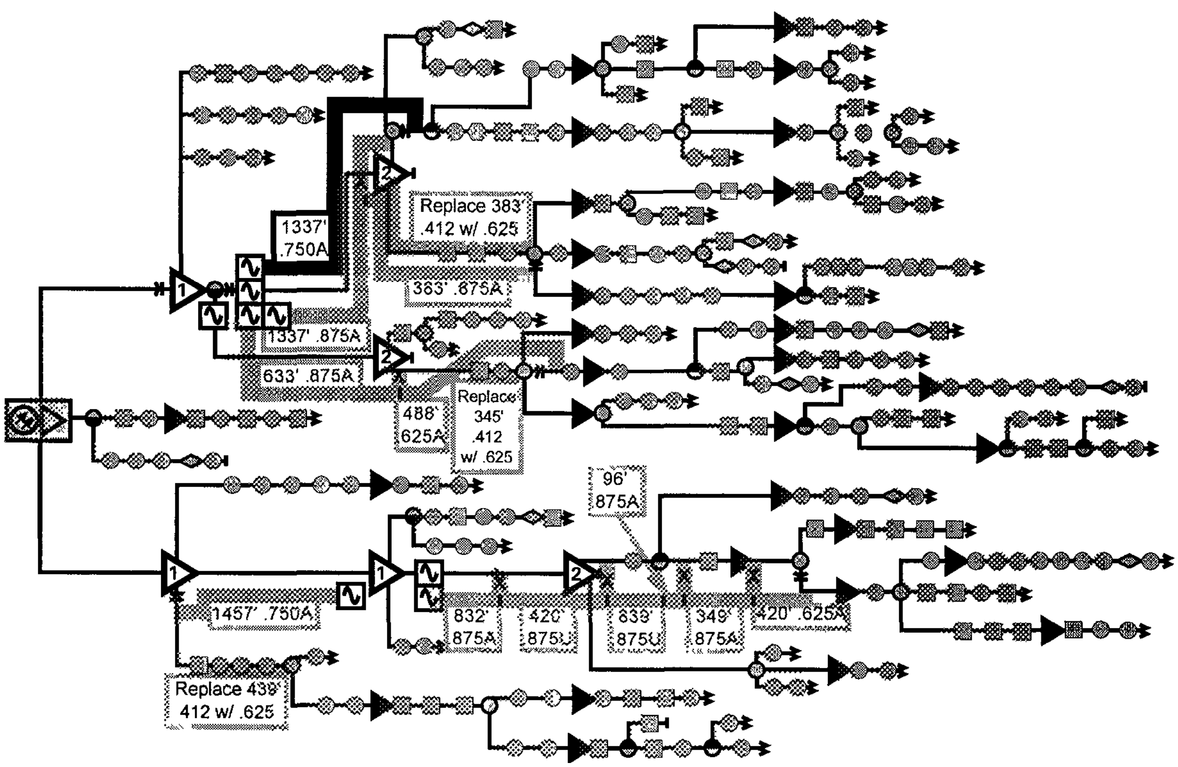


Figure 15. 60 Volts, 100% Penetration

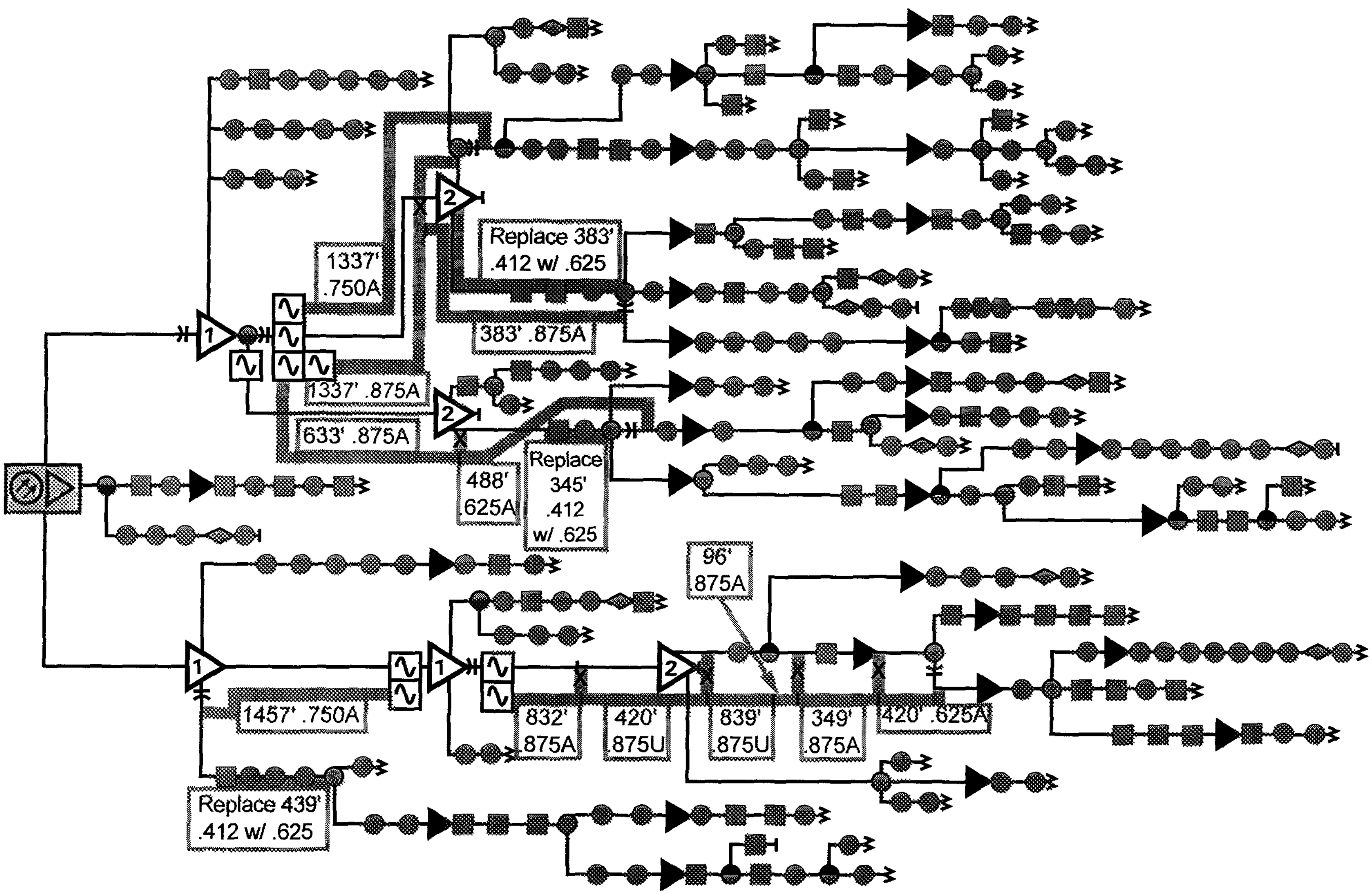


Figure 16. 90 Volts, 0% to 5% Penetration

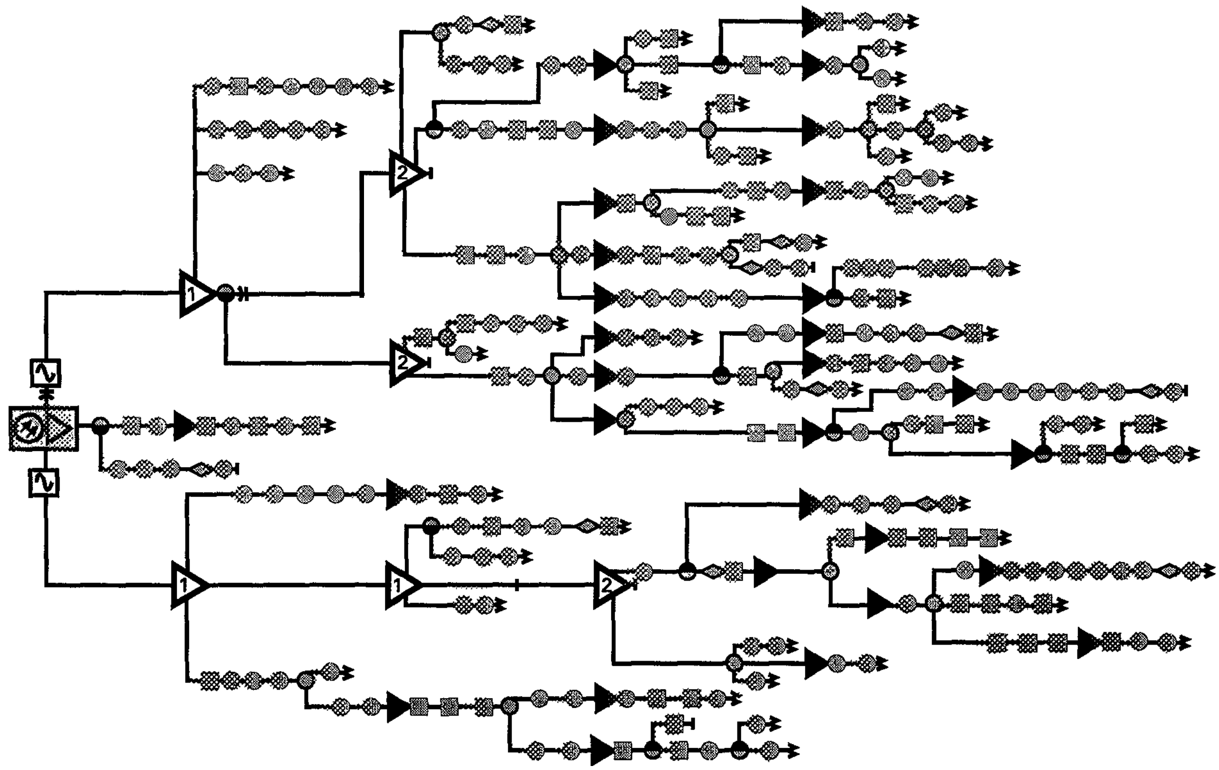


Figure 17. 90 Volts, 6% Penetration

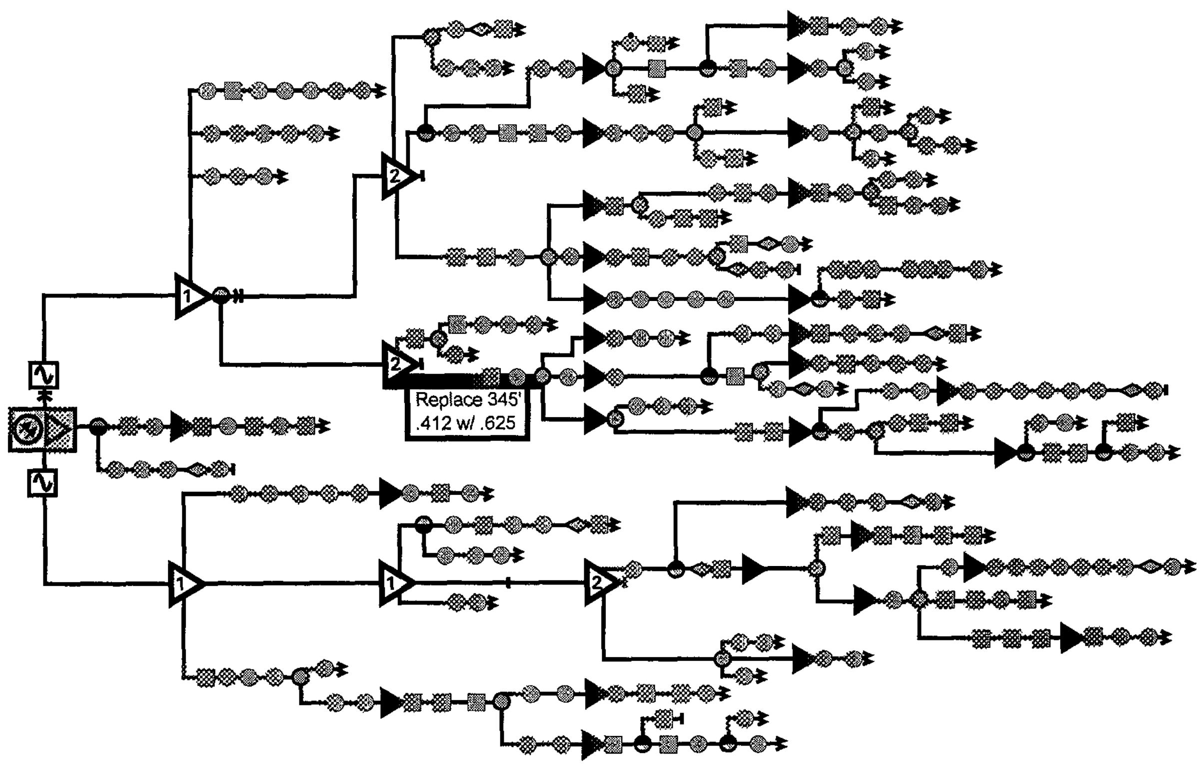


Figure 18. 90 Volts, 22% Penetration

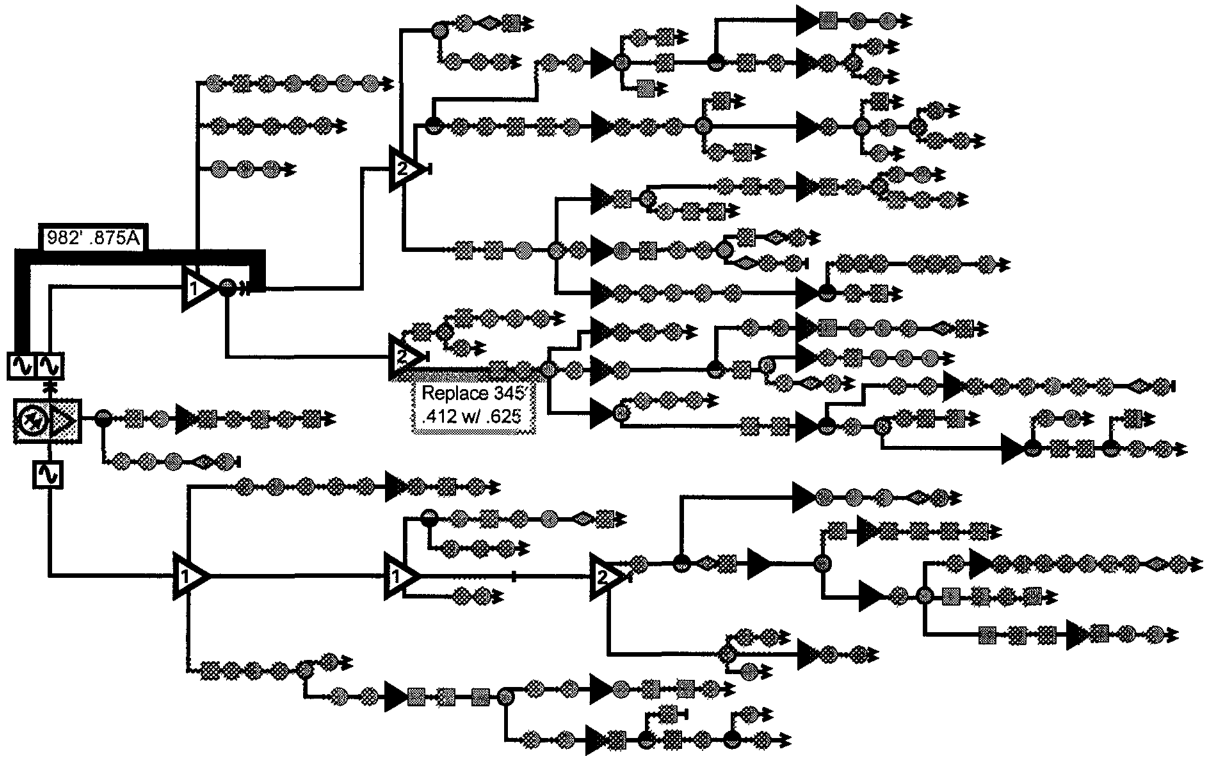


Figure 19. 90 Volts, 46% Penetration

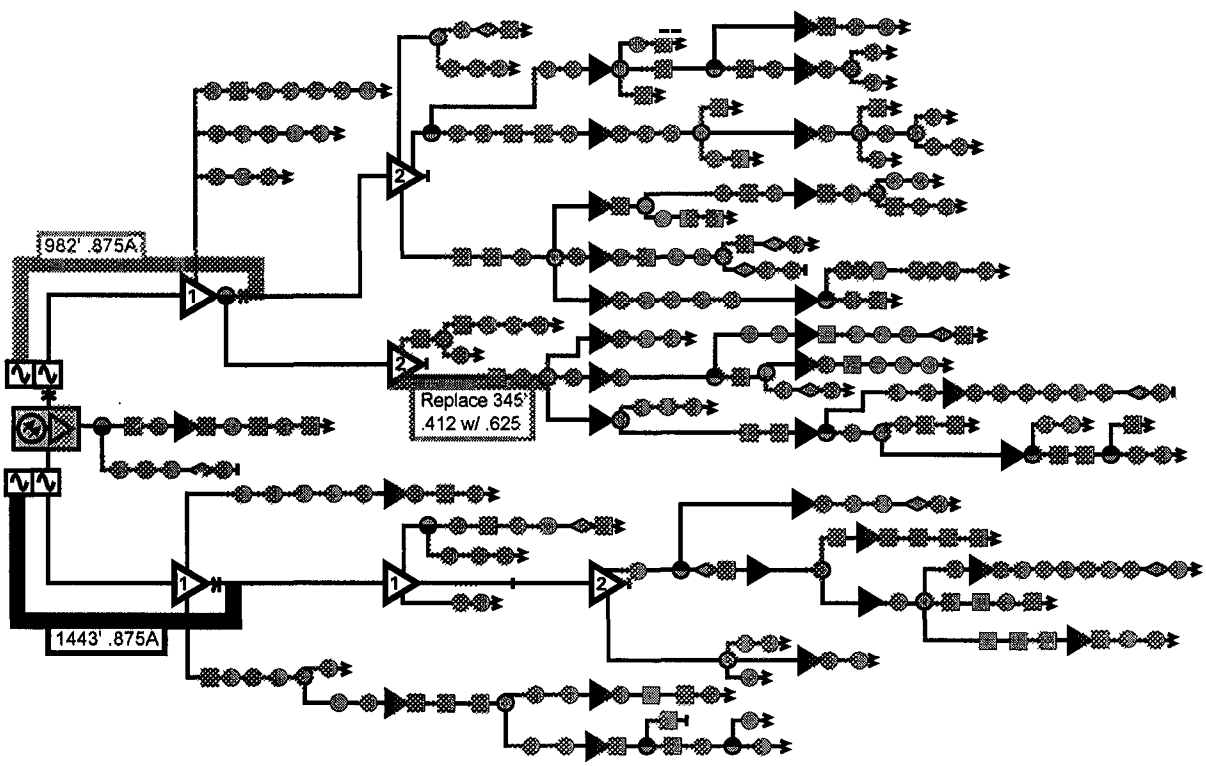


Figure 20. 90 Volts, 53% Penetration

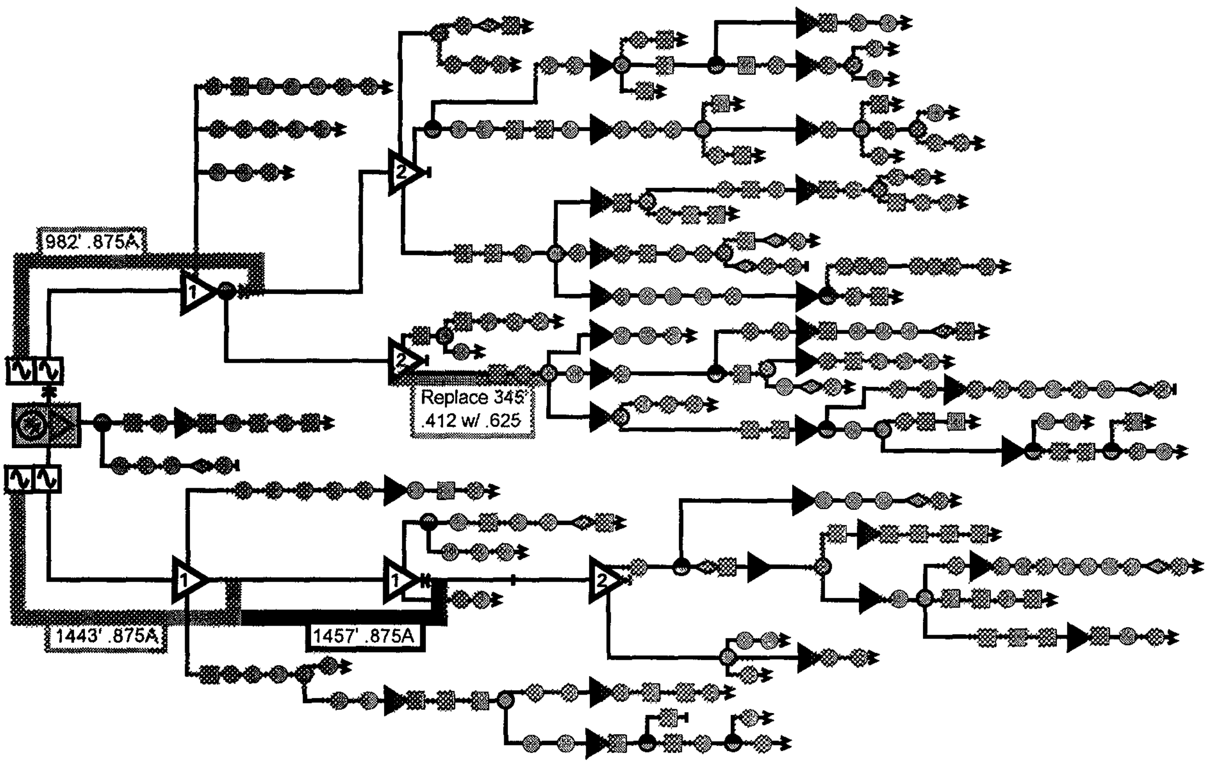


Figure 21. 90 Volts, 69% Penetration

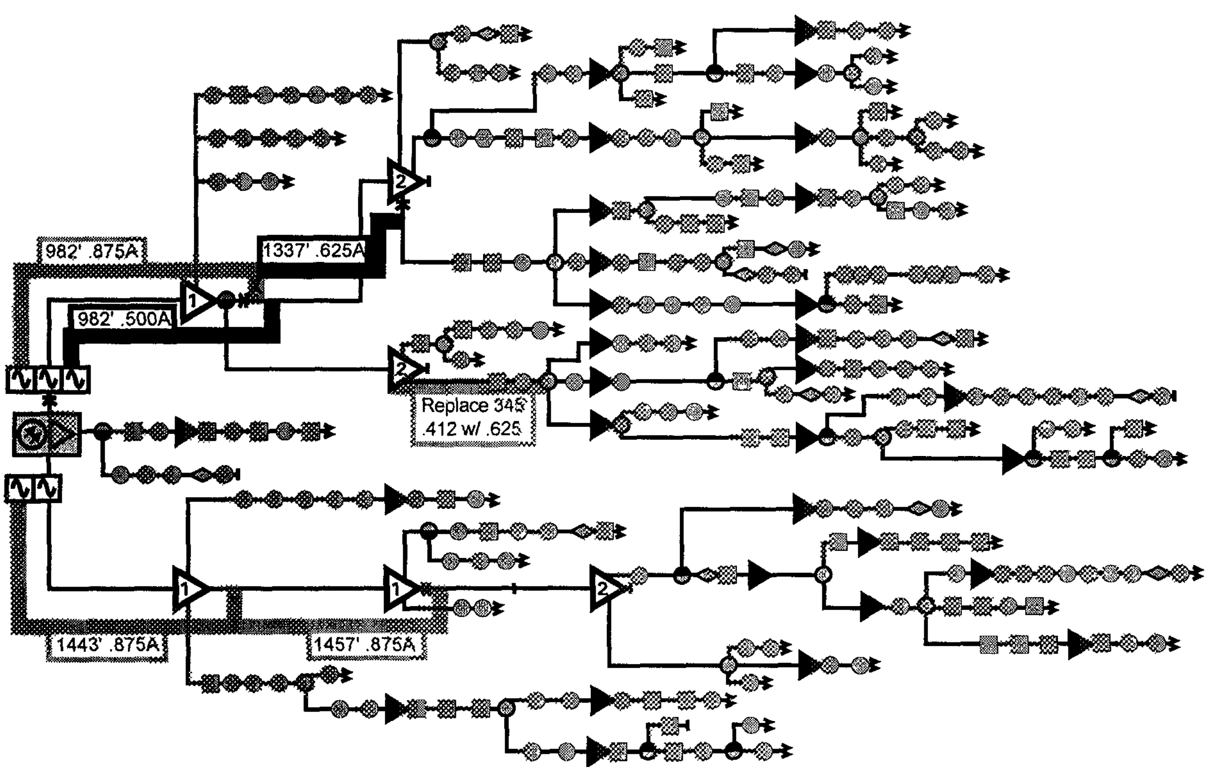


Figure 22. 90 Volts, 78% Penetration

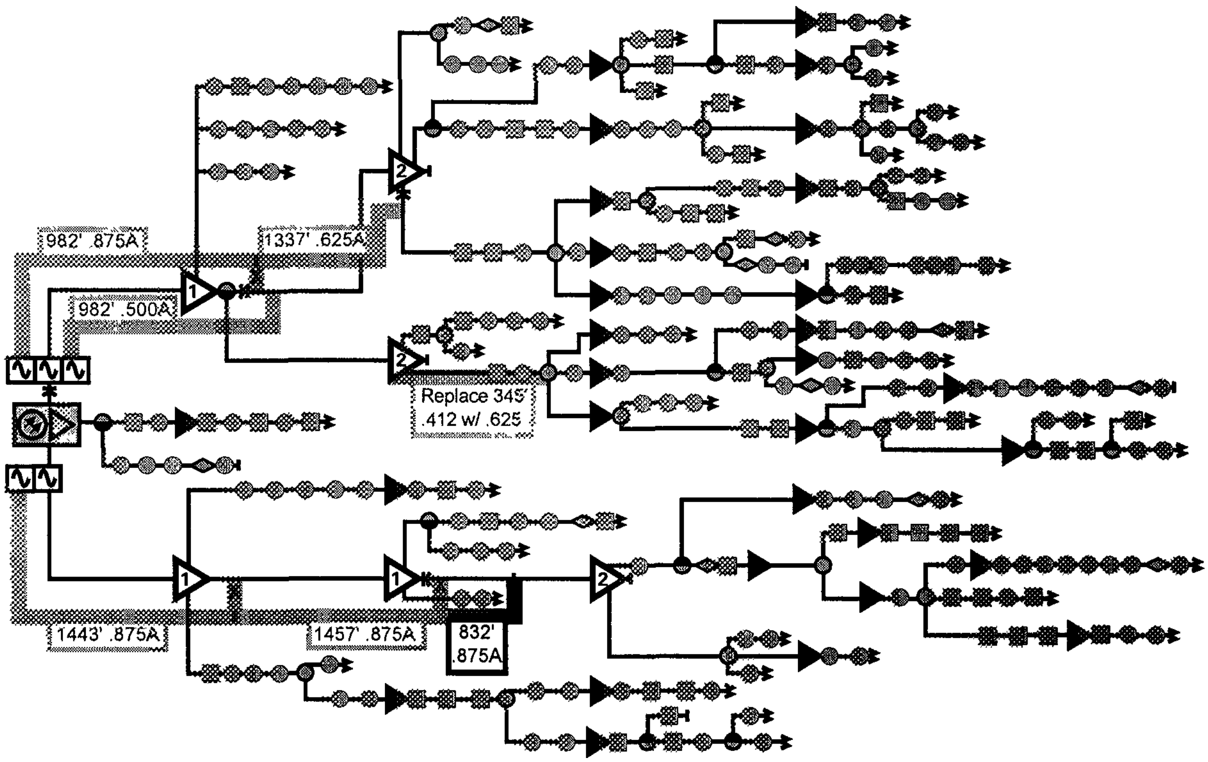


Figure 23. 90 Volts, 91% Penetration

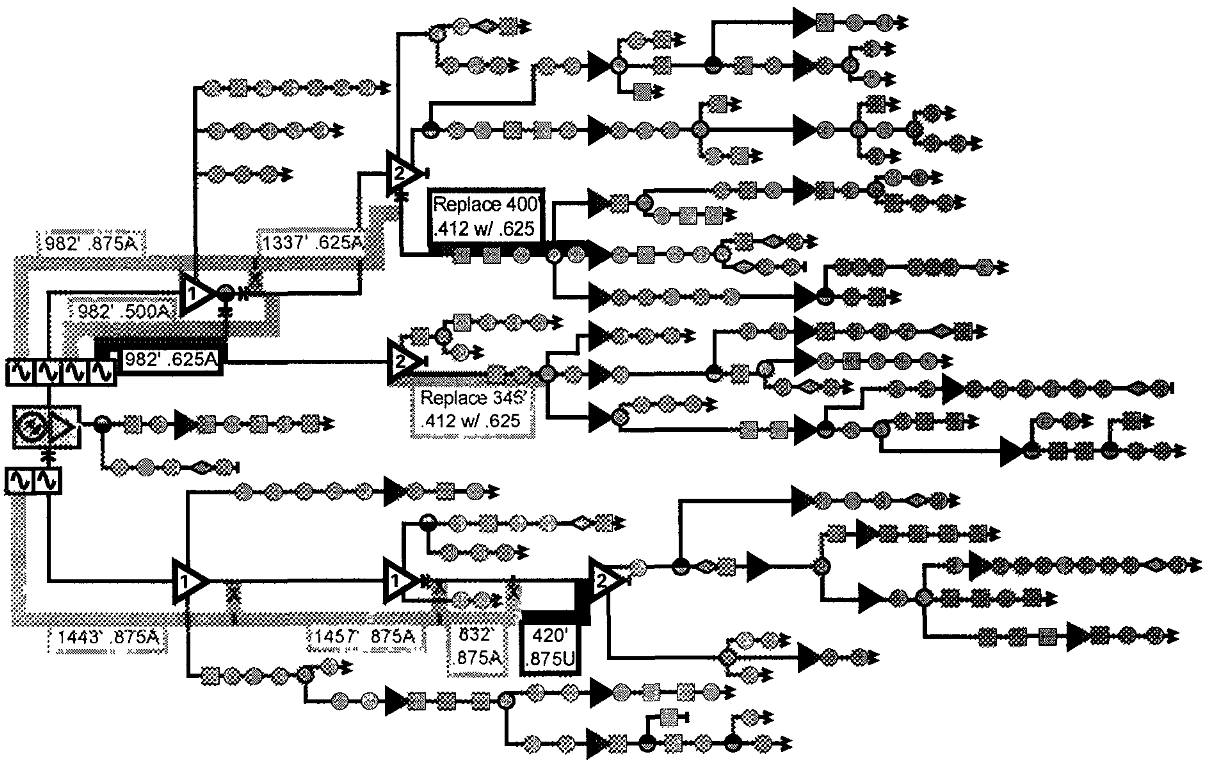


Figure 24. 90 Volts, 100% Penetration

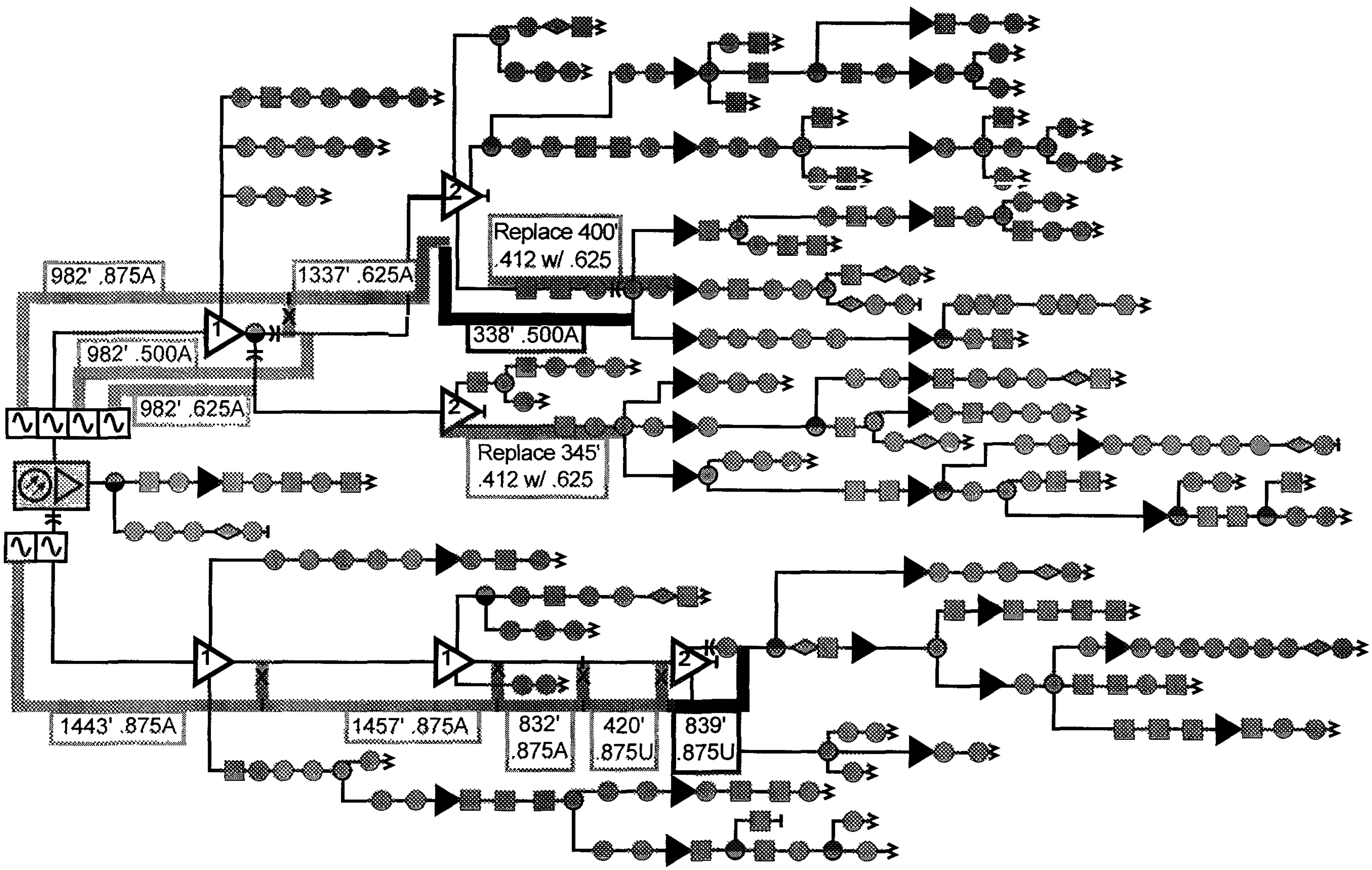


Figure 25. Infrastructure Cost Per Passing

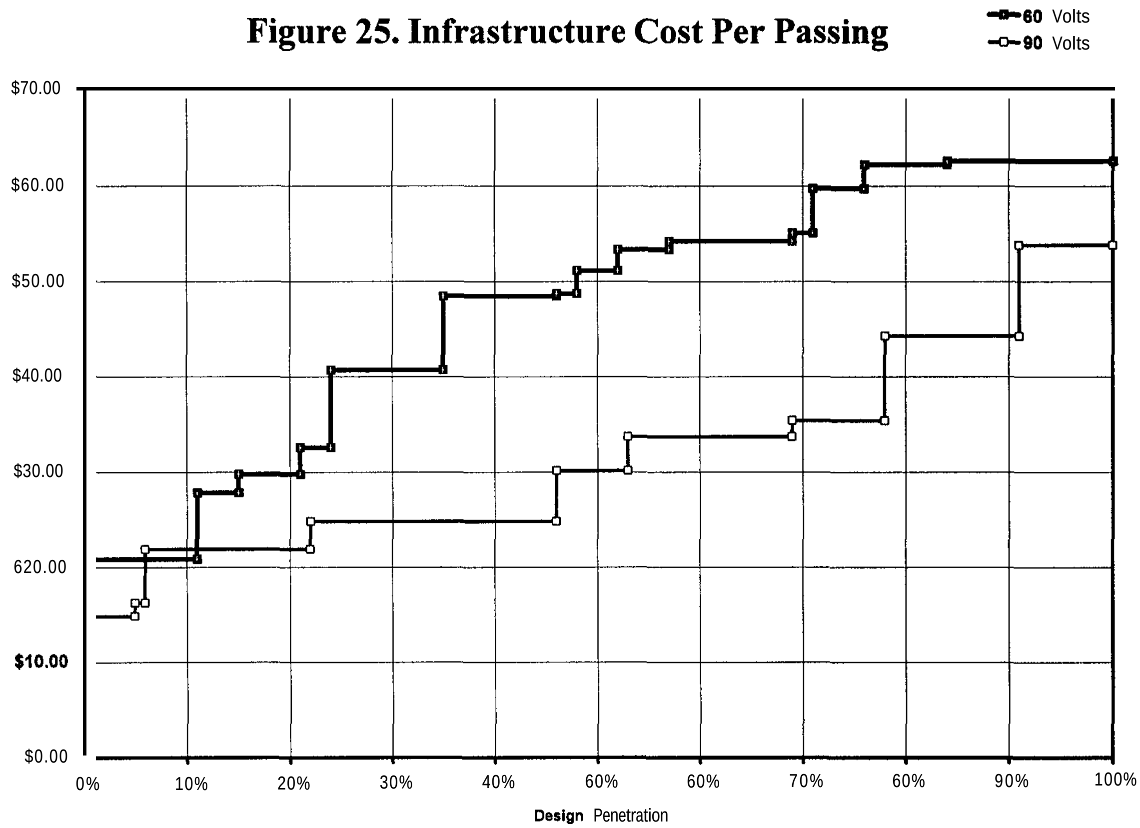


Figure 26. Power Consumption Per Passing

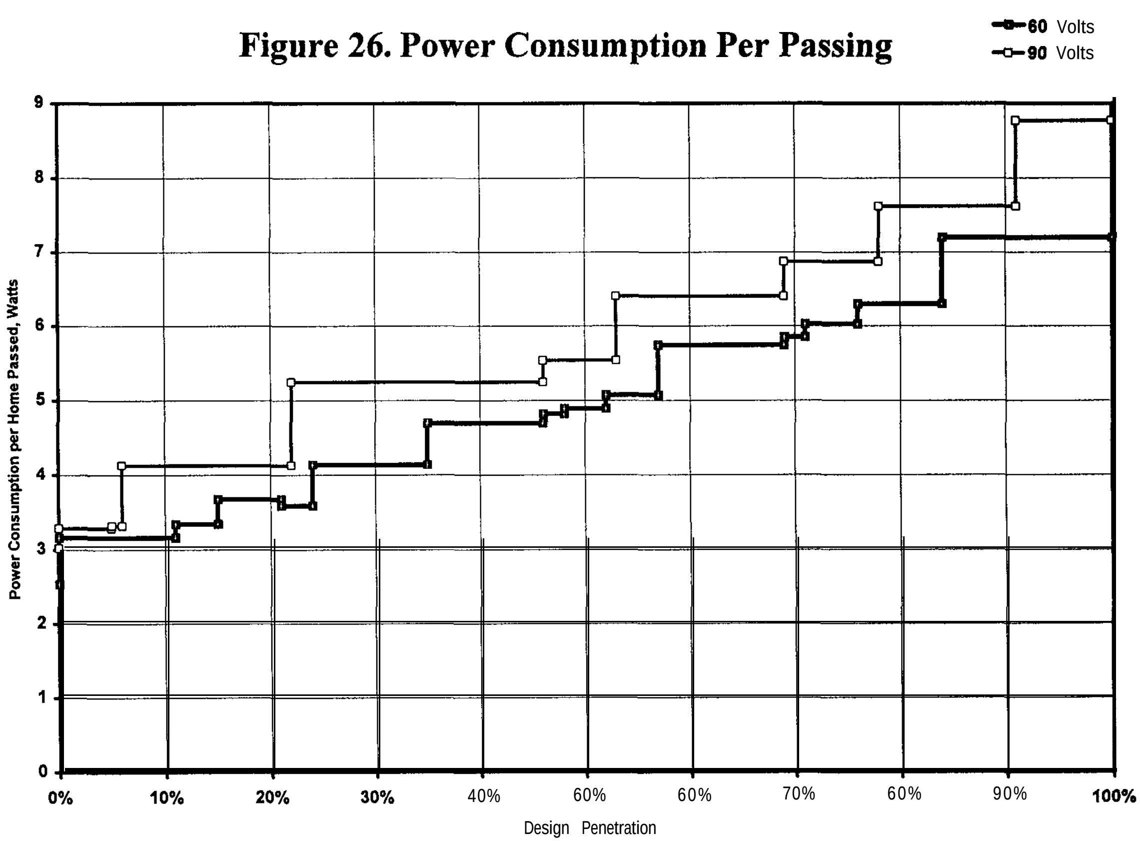


Figure 27. Cost Per Watt of Power Consumed

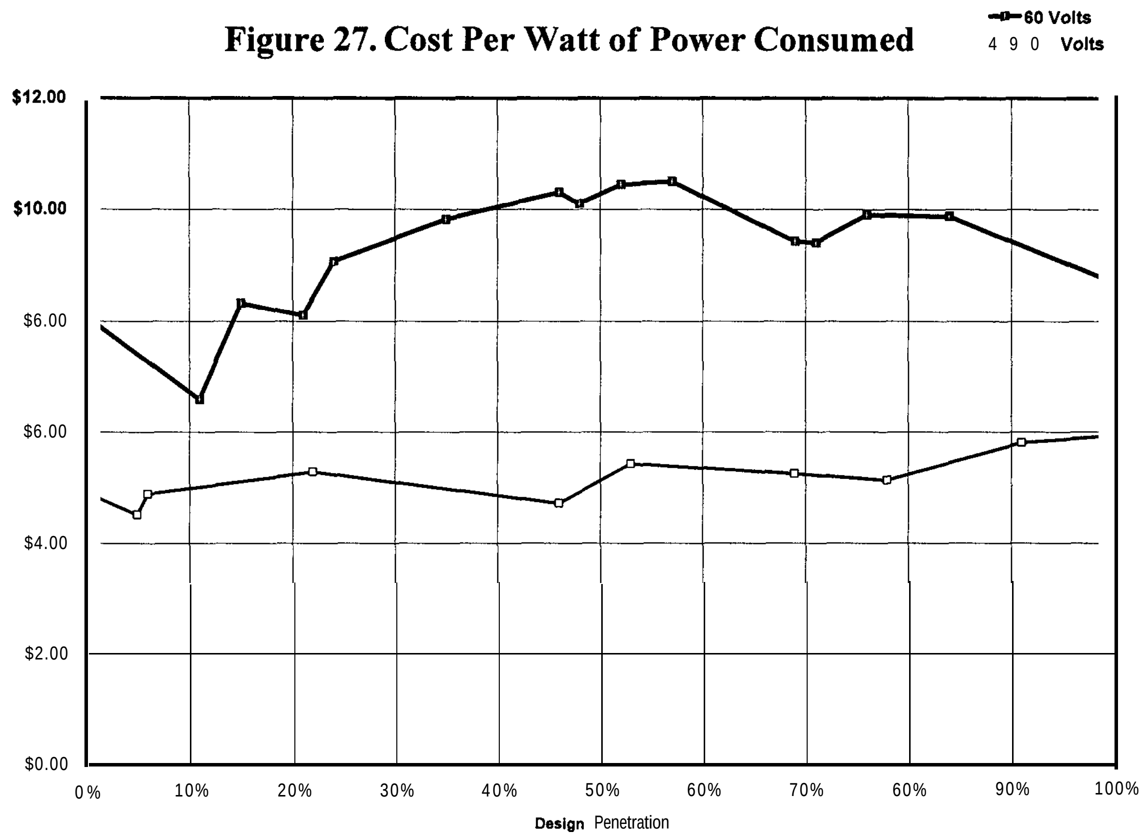
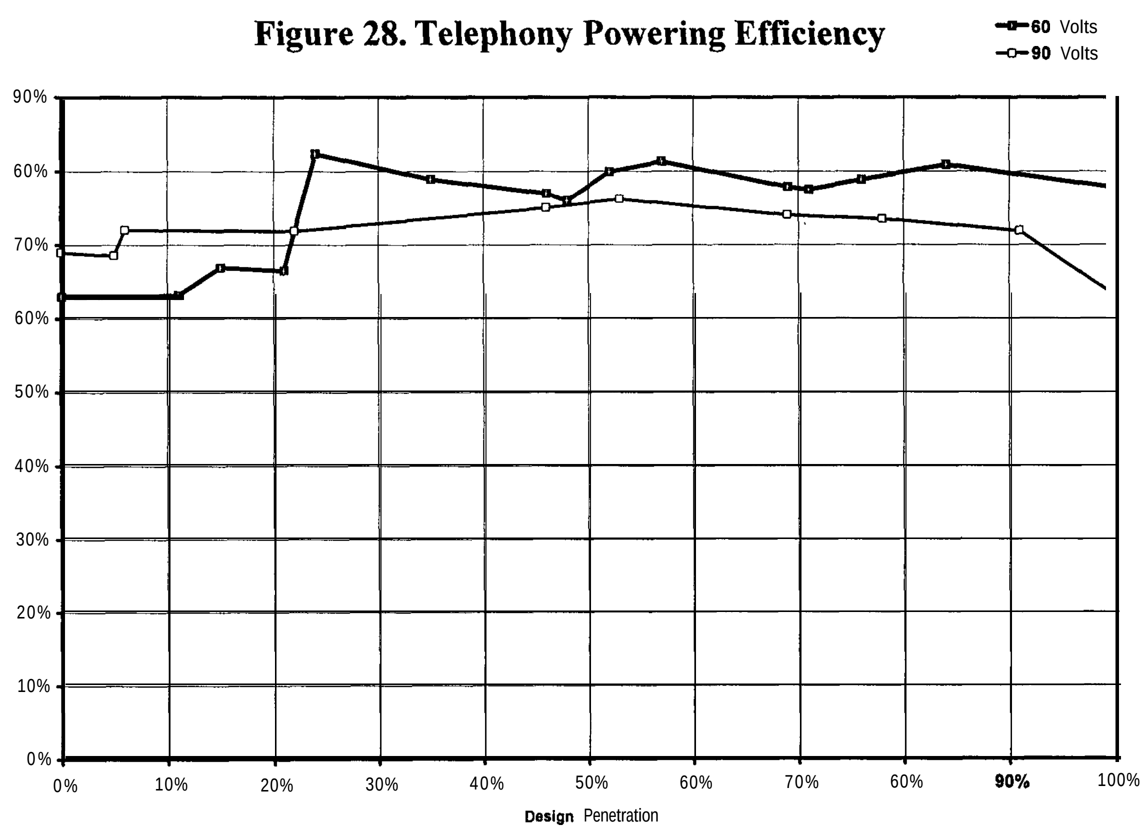


Figure 28. Telephony Powering Efficiency



Cost, Power! and Efficiency

Cost Estimation

Cost estimates were performed for the infrastructure added in Figures 1.) through 24.). These cost estimates were then combined and plotted against the associated penetration levels.

Deployment Cost Graph

Figure 25.) depicts the cost per passing of this incremental deployment. Note the stepwise function. The initially deployed capacity is determined based on the highest penetration number attained on a horizontal line that passes the minimum design requirement. For instance, If 30% penetration was selected as the minimum criteria, at 90 Volts a 30% design would actually operate to 46%.

The high starting costs represent the initial installation of long duration standby supplies which are needed even at low penetration levels in order to power system actives.

Power Consumption Per Passing

Figure 26.) depicts the wattage consumed per home passed at various penetration levels. Note also the high starting levels due to amplifier powering.

Infrastructure Cost Per Watt

Figure 27.) shows the capital cost of deployed infrastructure needed per watt of installed design capacity. Note the cost advantage afforded by use of 90 Volt powering.

Network Power Losses

Figure 28.) displays network powering efficiency curves. Efficiency is calculated as follows:

1. The design power allocated to an NIU (3.63 Watts) is multiplied by the number of deployed NIUs at each penetration level.
2. The network power consumed by actives only (0% telephony) is subtracted from the total network power consumed at each penetration level.
3. Item 1. is divided by item 2., converted to a percentage, and plotted against penetration.

Note that use of 90 Volts does not offer a significant efficiency improvement over 60 Volt powering. This is due to the wider coverage area, and therefore greater resistive cable losses incurred as a result of selection of a single powering location.

CONCLUSION

Network powering design techniques are available which provide for incremental deployment of network powering infrastructure. Proper application of these techniques can result in substantial reduction in the investment needed to deploy network powered lifeline telephony services.

Use of these techniques requires the discipline to monitor service area penetrations and expand network powering infrastructure as needed.

Increased emphasis must be placed on the production of quality powering design. This emphasis will require a commitment to appropriate design staffing levels in order to benefit from the techniques described here.

ACKNOWLEDGMENT

I wish to again acknowledge the contributions of Trygve Lode, president of Lode Data Corporation. The current release of the Design Assistant@ network design software was specially modified to allow completion of research work necessary for this paper. The features used are now available in the commercial release version 1.5.

Ingress Margin Improvement in Sub-Low Return HFC Plant

Paul D. Brooks – Staff Engineer
Engineering and Technology – Plant Engineering Department
TIMEWARNER CABLE
Englewood, Colorado

ABSTRACT

Many engineers regard the task of overcoming ingress as the most difficult in reliable operation of return plant. Because ingress levels are dependent only on source characteristics and physical coupling "receiving antenna" mechanisms, every dB of operating level is one additional dB of ingress immunity. Since the physical coupling mechanisms have historically been difficult to control, it is essential to operate signal sources at the highest possible levels in order to "buy margin" over ingress sources.

This paper will introduce three techniques of ingress margin improvement:

1. *Leveling of tap port sensitivity through purposeful introduction of loss in the sub-low band.*

2. *Leveling of tap port sensitivity by a new approach to selection of return amplifier input levels.*

3. *A brief examination of the upper limits of existing terminal device output levels and power densities, and the implication of these limits on the selection of the operating points of return amplifiers as these operating points pertain to ingress margin.*

INTRODUCTION

Ingress is only harmful when it becomes high enough in power to affect the transmission of information. This can occur in two ways:

1. High levels of ingress can cause clipping or intermod products in an active device.

2. When located within its passband, the ingress can overcome the wanted signal such that there is no longer sufficient margin to reliably transmit the desired information.

Since ingress power and coupling efficiency typically cannot be controlled, the margin can only be increased by raising the power of the desired signal, while decreasing the sensitivity of the network such that the power received at the return laser transmitter remains constant.

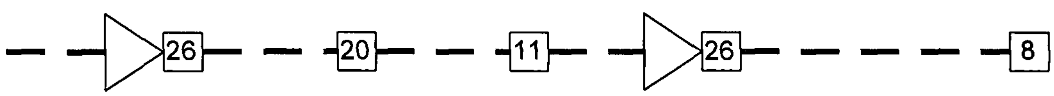
The sensitivity of the sub-low return plant is not equal at all potential ingress input sites. Instead, this sensitivity varies based on the location-dependent loss between a given tap spigot and its respective amplifier return input. Sensitivity also varies based on the operating point selected for each amplifier return input and its relationship to the forward output level upon which tap value selection rules are based.

It is well established that in a properly maintained plant, most ingress enters the sub-low return spectrum in the house drop, and specifically in the subscriber premises portion of that drop.

For the purpose of simplification, this paper will assume that all ingress fields are uniform throughout the service area, and that each tap has a total ingress sensitivity independent of the number of tap spigots and the nature of the drop plant. In practice, certain ingress fields have a strong local component, and tap sensitivity is dependent on the number of connected drops at a given tap location, and the summed coupling effectiveness of these drops. Ingress sensitivity is examined at 40 MHz in the return spectrum. While it can be shown that a greater disparity in sensitivity exists at 5 MHz, the value of these techniques is evident even at the higher cable loss frequency of 40 MHz.

Location Dependent Tap Spigot Sensitivity

In light of the above assertions, it is entirely possible that a large portion of the return plant ingress "gets in" at low value taps which are deployed downstream of longer cable segments. This assertion leads to one potential means of significantly reducing total ingress power. Additional loss could be purposely introduced in the sub-low band to force remotely located return signal sources to operate at levels comparable to devices proximate to amplifier outputs. This technique would also reduce the need for wide output attenuation ranges, which require complex circuitry and can consume power, particularly in NIUs with relatively wide, spectrally dense outputs.



Losses in dB	Tap	Cable	Tap	Cable	Tap	Gain	Tap	Cable
750 MHz	.8	10.85	1.1	8.68	4.8	29.48	.8	26.04
50 MHz	.3	2.70	.7	2.16	3.4	10.07	.3	6.48
40 Mhz	.3	2.45	.7	1.96	3.4	9.55	.3	5.88
5 MHz	.3	.80	.8	.64	3.5	6.28	.3	1.92

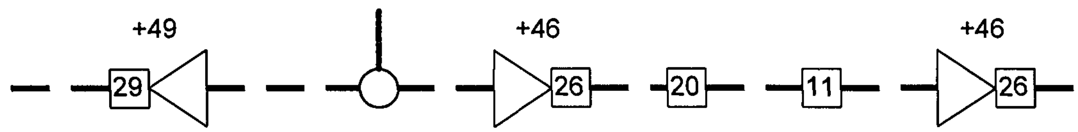
Levels in dBmV	Amp Port	26-4 Tap Spigot	20-4 Tap Spigot	11-4 Tap Spigot	Amp Port	Amp Port	26-4 Tap Spigot	8-4 Tap Spigot
750 MHz	46	20	14.35	13.57	16.52	46	20	13.33
50 MHz	35	9	12.0	18.14	24.93	35	9	20.76
40 Mhz	17	**43**	**39.75**	**33.41**	26.55	17	**43**	**30.69**
5 MHz	17	**43**	**38.10**	**30.54**	23.28	17	**43**	**27.06**

Figure 1.

Figure 1. shows the losses and levels in an example coax feeder. The required tap spigot return inputs are shown in bold. This example is included to illustrate the reason for the variation in tap port sensitivity. In this case, the 8dB 4 port end of line tap is over 12 dB more sensitive to ingress than the 26 dB taps at 40 MHz.

The Practice of Forward Output Derating

Amplifier return inputs in today's HFC networks are normally set to a uniform level at all actives in the feeder portion of the plant, with no regard to the typically non-uniform forward output levels. This practice causes taps downstream of lower output feeder actives to have greater ingress sensitivity than taps fed by higher output feeder actives. The return input level of any active which directly feeds taps should be scaled in inverse proportion to its forward high channel output level. For instance, if a given line extender has a forward high channel output of +49 dBmV and a return input of a flat +17 dBmV, then a "derated" line extender with a forward high channel output of +46 dBmV (3 dB lower) should be set to a return input of +20 dBmV (3 dB higher). This technique will reduce sensitivity to ingress on taps fed by derated line extenders.



Levels in dBmV	29-4 Tap Spigot	Amp Port	Amp Port	26-4 Tap Spigot
750 MHz	20	49	46	20
50 MHz	9	38	35	9
40 Mhz	**46**	17	17	**43**
5 MHz	**46**	17	17	**43**

Figure 2.

In Figure 2. the coax feeder of figure 1. is joined by a non-derated feeder fed from the same bridger output. The required tap spigot return inputs are shown in bold. Note that taps in the derated feeder are 3 dB more sensitive to ingress. This example is included to illustrate how tap port sensitivity is affected when forward levels are derated and return levels are left alone. If return amplifier inputs are increased in proportion to the decrease in forward outputs, the disparity in tap sensitivity is eliminated.

Plant Design Analysis

Example Plant Characteristics

An existing 350 mile cable system was analyzed to examine the value of these techniques. This typical, midwestern, mostly aerial plant was comprised of 38 service area nodes passing about 23,000 homes. Density ranged from 10 to 190 homes per mile. The RF cascade limit was two trunks, one terminating bridger, and three line extenders. Minimum tap port levels were 11 dBmV to 13 dBmV at 750 MHz depending on drop length class. Tilts were a uniform 11 dB, with bridger ports at +48, single line extenders at +49, and both two and three extender cascades derated to +46 dBmV. In should be noted that if a third deration step of 5 dB down were to be used for the three extender cascades, a greater improvement to ingress margin would be attainable than in this design, where LE levels were chosen at a uniform +46dBmV.

Software and Methods

The design software tool employed here provides a "test window" function. This optional function reports instances where the difference between the specified maximum tap spigot return input and the calculated spigot input requirement exceeds a user defined threshold.

If this threshold is set to .001 dB, then the network location and magnitude of each calculated spigot return input level variance is reported for every tap. These differences, reported in dB, indicate differences in tap port sensitivity.

The software provides automation capability which allowed rapid evaluation of these large plant sections. While separate "test window" thresholds are available at 5 MHz and 40 MHz, only the 40 MHz test was performed. As stated earlier, the 5 MHz test produces more dramatic results, but the 40 MHz test results represent the minimums which hold throughout the entire sub-low spectrum.

Results of Plant Analysis

A baseline was established to represent the sensitivity of this plant to a power summation of all signals present at the tap spigots, and therefore to global ingress sources. With this baseline normalized to zero dB, the table (Figure 3.) represents the results of various ingress margin improvement techniques.

While all numbers in the table are based on analysis of actual design, the "theoretical" equalizer step size improvements are extrapolated based on the assumption that with a lossless device, the "not to exceed" step size will produce a sensitivity reduction such that the summed power of all tap ports will average one half of the step size. Also, the "theoretical" numbers assume equalization of tap ports independent of the through path.

Ingress Margin Improvement Results

Item	Description	dB
1.	Theoretical port eqs, continuous	7.19
2.	Theoretical port eqs in 1 dB steps	6.69
3.	Theoretical port eqs in 3 dB steps	5.69
4.	Theoretical port eqs in 6 dB steps	4.19
5.	Theoretical port eqs in 9 dB steps	2.69
6.	Increased levels at derated LEs	1.53
7.	6 dB feederline eqs to 5 MHz	3.22
8.	Items 6. & 7. combined	5.10

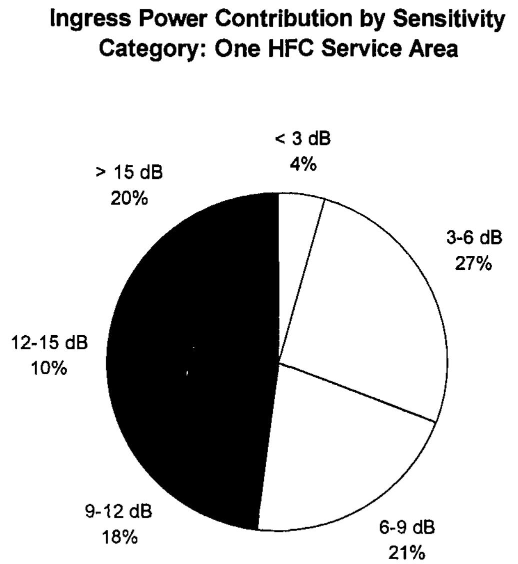
Figure 3.

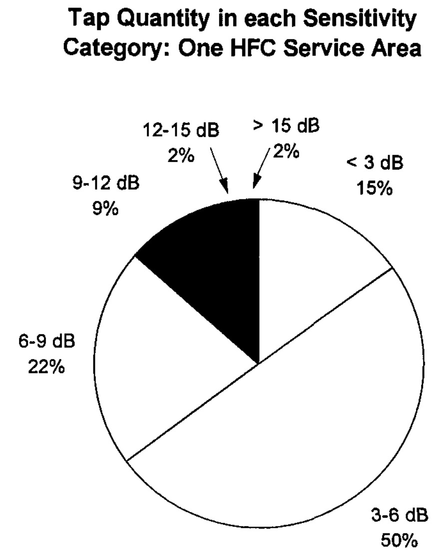
Design Runs

Items 6, 7, and 8 represent individual design runs where all 350 plant miles were redesigned each time. Item 8 produced results better than 6 and 7 summed; this is due to the fact that with higher return outputs, feederline equalizers could more often be placed between the bridger port and the first extender. The margin improvement of 3.22 dB shown in item 7 shows generally good agreement between implementation and the 4.19 dB theoretical limit of item 4.

Feederline Equalizers to 5 MHz

In item 7, a fictitious but reasonable feederline eq was specified to have 1 dB of insertion loss, and a characteristic equal to the inverse of 6 dB of cable at 750 MHz. Where this device differs from today's line eq is that no diplex filters are used; the response is well behaved down to 5 MHz. The elimination of diplex filters should allow for the low insertion loss (1 dB) and should also provide lower cost of manufacture. In item 7, only tap plates were allowed to be changed. All cables and amplifier locations were retained, and no new amplifiers or cables were added. Line equalizers were only placed where the existing plant had sufficient gain to allow re-balancing to correct forward and return levels. A total of 274 equalizers were used, about 0.78 per mile.





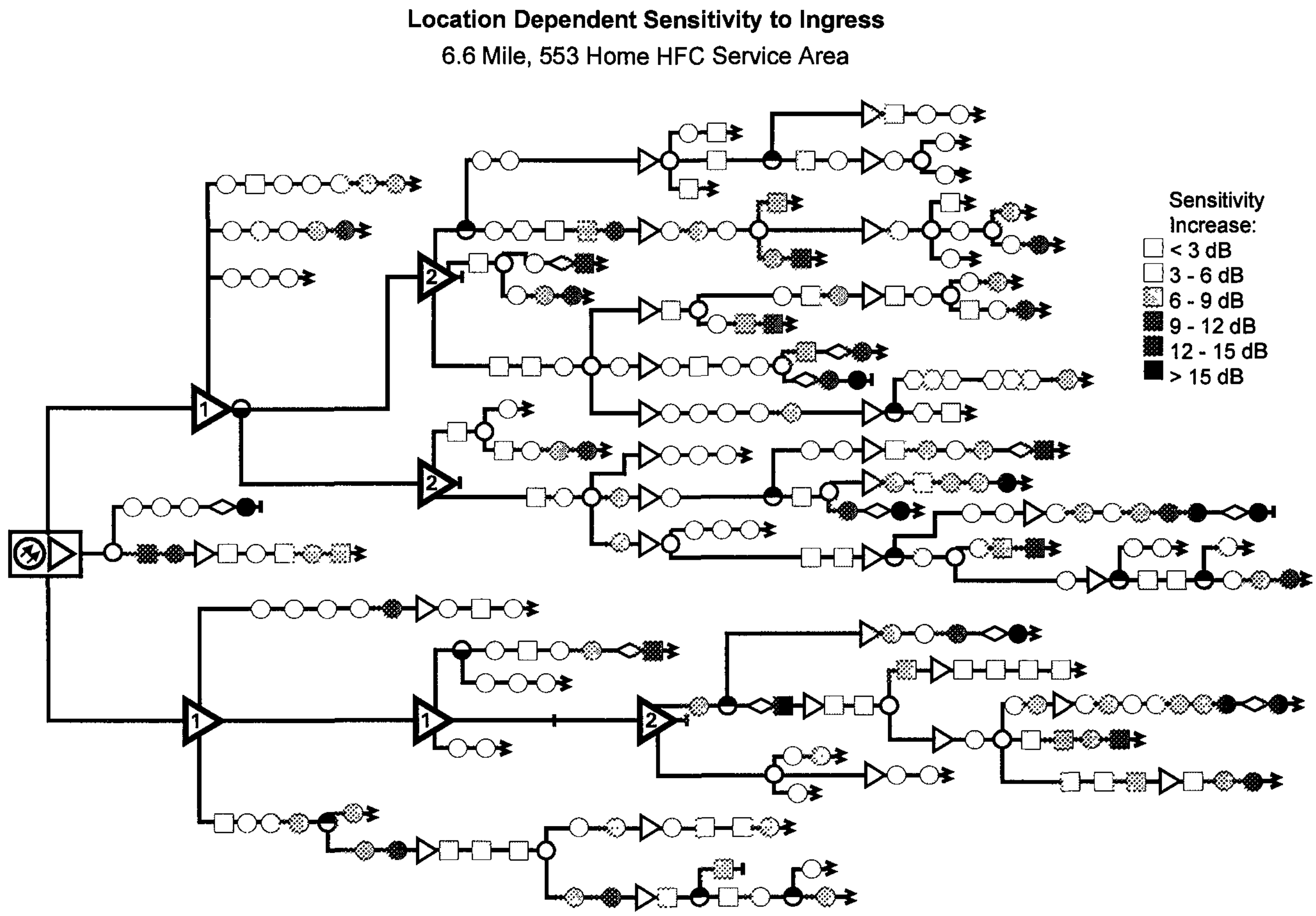


Figure 4.
A Typical Service Area Node

A Typical Service Area Node

Figure 4 provides insight into how location dependent ingress sensitivity operates in a typical node. The pie charts and node diagram were developed from one existing node designed without use of margin improvement techniques. The sensitivity variance is a combined result of tap location and the derating effect. The pie charts, taken together, contrast the ingress power contribution vs. the quantity of taps in each sensitivity range. The diagram shows where these taps are located in a real design.

Optimizing Return Operating Points

Return Plant Alignment Theory

Return operating levels are shown in figures 1 and 2 of this paper for the purpose of clarity. In practice, levels in the return plant are arbitrary and meaningless without a reference to the bandwidth of the signals being carried.

For the purpose of the following discussion, the terms "upstream" and "downstream" refer to the cable and equipment between the amplifier under test and the headend, and further from the headend, respectively.

When return plant is aligned, the gain of each amplifier is adjusted to exactly compensate for the cable and passive losses of the upstream "span" or "spacing." As long as the alignment signal is locally injected at the amplifier, it makes no difference what levels are used, provided no interference is caused, and the levels are within the useful dynamic range of the amplifier under test. These test signals are measured at the local amplifier return output, at the input to the upstream amplifier, or at the headend with certain automated test equipment. Return pads and equalizers are then selected based on these measurements, thus establishing a "unity gain" system.

Modern terminal devices provide remote output level adjustment capabilities which are typically controlled from the headend. As such, final operating level adjustment is normally performed "after the fact" of plant alignment.

The Alignment Reference

The Fabry-Perot return laser transmitter, collocated in the field with the optical receiver, typically determines both the upper and lower limits of the dynamic range of HFC sub-low return plant. Therefore, the input power levels supplied to this transmitter establish the system operating point, based on a tradeoff between noise floor margin and overload margin. Once this operating point is determined, the amplifier gains can be specified to best utilize all of the available terminal device return output power. The methods of establishing this operating point are beyond the scope of this paper.

If the optical receiver/transmitter supplies signals directly to tapped feeders, it will likely be necessary to select appropriate pads to balance these return inputs against other ports which feed only amplifiers. These pads provide for the use of high terminal device outputs by preventing the laser transmitter overload point from imposing a false limit on the maximum useable terminal device output levels.

Ingress Margin Improvement

Every dB of usable terminal device output level improves ingress margin by one dB. The degree of improvement possible is limited only by the compression point performance of the amplifier return modules.

Amplifiers used in HFC plant should have their return operating points selected based on knowledge of compression points of the return gain blocks. The chosen operating points are located below compression by reasonable safety factors. Terminal devices such as converters and telephony NIUs need adequate return output capabilities to attain the desired return amplifier output power.

5 to 40 MHz Operating Levels: Power outputs in dBmV, gains and losses in dB						
Power bandwidth, MHz	**35.0**	**6.0**	**2.0**	**1.0**	**0.1**	**0.05**
Amplifier return output compression point	71.4	63.7	59.0	56.0	46.0	43.0
Compression safety factor	-6.0	-6.0	-6.0	-6.0	-6.0	-6.0
Amplifier return output operating point	65.4	57.7	53.0	50.0	40.0	37.0
Amplifier return gain	18.0	18.0	18.0	18.0	18.0	18.0
Amplifier return input operating point	47.4	39.7	35.0	32.0	22.0	19.0
Spigot loss, tap on amplifier forward output	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0
Spigot return input, tap on amplifier forward output	76.4	68.7	64.0	61.0	51.0	48.0
Exterior drop loss, 150' of series 6, 1.0 dB per 100'	-1.5	-1.5	-1.5	-1.5	-1.5	-1.5
Return output of NIU at house bonding block	77.9	70.2	65.5	62.5	52.5	49.5
Four way drop splitter loss	-7.0	-7.0	-7.0	-7.0	-7.0	-7.0
Interior drop loss, 50' of series 6, 1.0 dB per 100'	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
Return output of set top or computer modem	85.4	77.7	73.0	70.0	60.0	57.0

Figure 5.

Figure 5 illustrates the effect of bandwidth on operating level, with the spectral power density held constant. A typical amplifier compression point of 71.4 dBmV in 35 MHz was selected as a convenient reference.

Spectral Power Density

The lack of sufficient power density (output level per Hertz) in terminal devices can limit the chosen return amplifier outputs for these specific services to levels significantly below the compression points of the gain blocks in place. While there is generally more than adequate noise margin available in the coax portion of the plant, ingress margin is sacrificed due to this power deficit.

Today's RF Converter

For the first example, examine the 0.1 MHz column in Figure 5. This bandwidth represents a typical RF return converter. The bottom row in the table shows a level of 60 dBmV, a typical maximum converter output. Thus, today's converters can be operated to effectively utilize an amplifier compression point performance of about 72 dB over 35 MHz.

If RF converter return levels were used in the design software as a reference, then the maximum tap port return input levels would be set to +51 dBmV for this service (assuming that the drop described in Figure 5 is the worst case), and the return amplifier input level would be established as +22 dBmV (assuming that the 29 dB tap is the highest loss path between any tap spigot and its upstream amp input). Basically, an output level of +60 dBmV works pretty well for a 100 kHz wide service.

The Telephony NIU

For the second example, examine the 2.0 MHz column, which is representative of a typical time division multiple access telephony network interface unit. The NIU has an advantage in that it is located at the bonding block, and thus does not need to drive through the lossy drop splitters found in the house. However, use of the wide bandwidth required by TDMA imposes a strict requirement on output level. In order to operate at the optimum ingress margin point, the NIU must have an output power capability of +65.5 dBmV in the 2.0 MHz of occupied spectrum.

Today's NIUs can commonly attain power levels of only +50 dBmV. The result is a failure to realize 15 dB of potential ingress margin improvement that is incremental to the techniques discussed earlier in this paper.

The Computer Modem

For the third example, examine the 6.0 MHz column in Figure 5. To make a long story short, the computer modem must produce an output of +77.7 dBmV to best utilize the ingress margin available to it. A +50 dBmV capable modem experiences a penalty of 27 dB of wasted potential ingress margin improvement.

Margin Against Wideband Interference

As a service requires simultaneous use (TDMA, for instance) of a large chunk of contiguous bandwidth, the sensitivity to wideband interference increases in direct proportion to the bandwidth occupied. Localized electrical interference generated by small appliances is an example of such wideband interference. The power of such an interference source is often relatively flat with respect to frequency, and therefore total power increases with increasing bandwidth.

An Additional Safety Factor

Because few amplifiers will ever see anywhere near full spectrum loading, a built-in power loading safety factor exists. This safety factor is due to the fact that only one path can use a given portion of the spectrum at a given instant in time. The number of unique paths through the network is nearly as great as the number of inputs, by virtue of the tree and branch topology of HFC networks. Therefore, power loading is statistical in nature, and as a device is located further out in the forward network, the probability of high return spectrum power loading is vastly reduced.

ACKNOWLEDGMENT

I wish to acknowledge the contributions of Trygve Lode, president of Lode Data Corporation. The current release of the Design Assistant® network design software was enhanced to allow faster completion of research work necessary for this paper.

Integrated Wireless PCS - Hybrid Fiber Coax Network Architecture

Douglas E. Hohulin, Robert W. Hammond

Motorola, Pan American Wireless Infrastructure Division

ABSTRACT

The following paper explores the trends, issues and opportunities for telephony service providers to leverage broadband Hybrid Fiber Coax (HFC) networks to deliver integrated wired and wireless services to consumers. In particular, specific attention is focused on the industry trend toward the use of HFC networks; the complementary architecture; and the benefits as well as issues associated with the use of HFC networks for the integration of personal communications services. This paper will include an overview of Motorola, STV, TCI and Cox's PCS trial over the cable plant in San Diego and Dallas using GSM and CDMA technologies.

INTRODUCTION: TELEPHONY IS GOING BROADBAND

At present, there are three service providers with direct connections to the home: telephone, power and cable. As each of these respective industries make architecture decisions to compete for a share of the $120 billion telephony market, HFC networks will continue to expand, improve and proliferate. Additionally, with the introduction of new services and spectrum hungry applications, such as: virtual shopping, video telephony and megabit network access, the requirement for telephony bandwidth will continue to grow. Recognizing the prevalence of and bandwidth available from HFC networks, many companies are now exploring the use of the HFC plant for delivering not only end user services but also distributing radio frequency (RF) signals for wireless telephony.

THE MOTOROLA EXPERIENCE

Motorola has performed trials with both CDMA and GSM access technologies in live cable systems using two different approaches for wireless signal distribution. Specifically, Motorola has trialed GSM RF transport with Cox Cable Communications in San Diego, CA and more recently CDMA RF transport and wireless integrated baseband transport with Tele-Communications, Inc. (TCI) in Dallas, TX and Arlington Heights, IL.

The remaining sections of this paper describe the similarities between HFC and wireless networks as well as the advantages and issues associated with HFC wireless network integration.

HFC ARCHITECTURE

Cable passes greater than 95 percent of the homes in the United States with approximately 60 million homes subscribing to cable services. The main operators of these HFC networks are cable operators.

Cable operators use the cable plant, a broadband bi-directional pipe, to simultaneously deliver multiple services to consumers. Currently, cable operators are upgrading their existing infrastructure, as illustrated in Figure 1, to provide up to 700 MHz of bandwidth in the downstream direction (to the home) and 37 MHz of upstream bandwidth (from the subscriber to the cable operator.

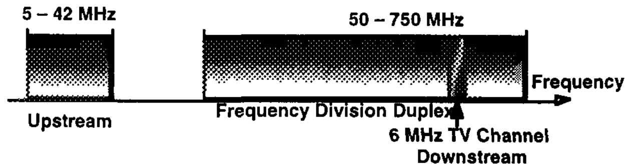


Figure 1.

As shown in Figure 2, the HFC cable plant is divided into three fundamental components:

- Headend
- Fiber Node
- Coax Feeder-Drop Distribution

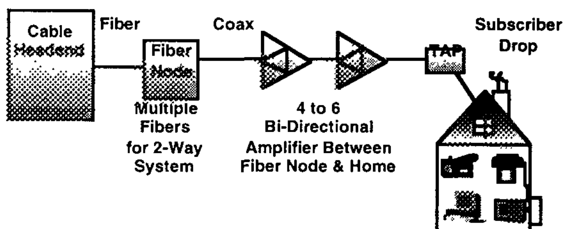


Figure 2.

The Headend

At the headend, all video signals are combined from satellite receivers, local television receivers, fiber optics and other headends for distribution to the subscriber. The headend is the concentration point for the HFC network. Local commercials may also be added at the headend.

Fiber Nodes

Banks of lasers located at the headend distribute the video signals to fiber nodes. Each fiber node typically supports between 500 and 2,000 homes and is spaced between one and three miles apart. At the fiber node, the video signals are converted from light back to RF for use in the coax plant.

Coax Feeder-Drop Distribution

The coax plant is distributed using bi-directional amplifiers to the home subscriber. Prior to the use of HFC cable plants, 20 amplifiers may have been used from the headend to a subscriber's home. This distribution caused the system to be less reliable since any one amplifier outage would affect service on the entire coax run. With the implementation of HFC into the cable plant, the number of amplifiers between the headend and a subscriber's premise was significantly reduced to contain commonly less than five amplifiers.

To date, much discussion in the industry has focused on how far toward the home fiber should be run. The contention revolves around the added cost of fiber versus coax; consumer demand for bandwidth; and network reliability. In some cases the application and issues of economics reliability suggest minimal, if any, coax in the network. In the case where no coax is utilized the network can be broadly classed as Fiber To The Home or a FTTH network. Popular discussion introduces the less expensive alternative Fiber To The Curb or FTTC architecture.

Under scrutiny, both FTTH and FTTC reveal themselves as variants of the same broader classification — hybrid fiber coax network. The only differentiation between these two classifications is the amount of bandwidth and associated cost. In the traditional cable industry, there is a great deal of discussion related to the optimum amount of bandwidth required for meeting the customer's ever increasing thirst for applications. In cable language this is described as the number of homes served per fiber node.

At present, it appears that the cable industry is settling on 600 home fiber nodes. At 600 home fiber nodes, cable could be described as Fiber To The Neighborhood or FTTN. Stated alternately, fiber to the home, fiber to the curb and today's traditional cable networks are all variations of a similar architecture intended to deliver sufficient bandwidth to the home to meet both today's and tomorrow's service demands. The key differentiation being the smaller the number of homes served from any one fiber means the greater the available bandwidth and cost for the served area.

Beyond today's cable Multiple System Operators (MSOs) some of the Regional Bell Operating companies have begun deploying HFC networks as well. Most recently, Pacific Bell, Bell Atlantic, Ameritech and US WEST have all announced, or are in the process of building, HFC networks. This when coupled with Sprint's recent partnering with Comcast, Cox Cable Communications and TCI for Personal Communication Services (PCS), demonstrates significant industry momentum toward HFC as the chosen architecture for delivering the next generation of telephony services.

WIRELESS ARCHITECTURE

There is a great deal of similarity between today's wireless systems and HFC networks. Figure 3 depicts a current generic wireless network. As shown, the architecture is composed of three main subsystems: network, switching and radio.

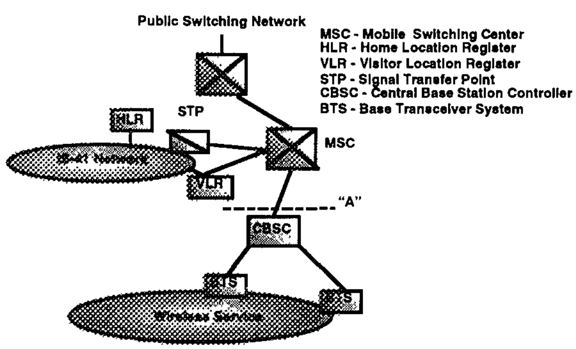


Figure 3. A-Interface and Architecture

In the radio subsystem, the cell sites or Base Transceiver Station(s) (BTS(s)) are connected to Centralized Base Station Controller(s) (CBSC(s)) via T1 interfaces. There are two types of cell sites: macro and micro.

A macrocell typically uses high gain sectored antennas mounted on 100 to 300 foot towers to provide between one and 10 miles of coverage. The difference in range is primarily a result of tower height, power output, terrain and morphologic conditions (urban, suburban and rural environment). A typical macrocell may consist of between 50 and 100 traffic channels capable of supporting between 2,000 and 5,000 subscribers. The traditional macrocell uses a shelter or in-building space to contain the radio equipment. In contrast, a microcell typically uses omni antennas placed on a smaller tower or the top of a building (under 100 feet). The range of the microcell is under one mile with the microcell radio equipment in a small outdoor enclosure about the size of a small refrigerator that can be placed easily in many outdoor environments. The microcell may have from eight to 20 traffic channels and most likely supports less than 1,000 subscribers.

System deployments over 100 macro and microcells are often needed to cover a large metropolitan city. This requires an extensive distribution system to interconnect the BTS(s) to the CBSC(s). The CBSC concentrates the T1s and controls the mobility aspects of the radio subsystem. The Mobile Switching Center (MSC) is the interface between the CBSC, the Public Switched Telephone Network (PSTN) and other wireless networks. The MSC further concentrates the T1 interconnect from the CBSC(s). The MSCs in today's wireless systems are unique to handle the special mobility aspects of the wireless switching systems.

Like most system architecture the wireless system continually evolves to best serve end customer needs. The next generation of architecture looks to more fully integrate the wired and wireless networks through the reuse of wired network switching and intelligent peripherals. Although not fully explored in this paper this architecture is similar in its similarity with HFC networks as well as the benefits associated with HFC-wireless network integration.

INTEGRATION OF HFC AND WIRELESS NETWORKS

There is an interesting similarity between HFC networks and wireless networks. Typically in today's wireless network there is a single MSC per city. The wireless MSC is analogous to the main headend located in each franchise or city. The HFC network also connects to individual headends which are intended to support between 20 and 100 fiber nodes. Each fiber node can support up to 2,000 homes. The headend and fiber node are very similar in function and size with wireless systems base station controllers and cells.

APPROACHES FOR USING HFC NETWORKS FOR PCS

Over the past few years, wireless operators and vendors have investigated the use of the cable plant as a distribution system for wireless telephony. In general, there are two techniques to leverage the cable plant infrastructure for PCS:

- RF transport
- Baseband Transport

RF Transport

RF transport uses the cable plant to distribute RF energy (that would normally go to an antenna of a macrocell) through the cable plant to specific points on the cable plant and radiates this RF energy by simulcasting it to a larger coverage area than could be covered by a traditional macrocell. In an alternate approach, baseband transport is integrated with microcells to provide cable based backhaul instead of the more traditional leased line or microwave.

Cox Cable Communications has been a leader in the development of wireless - HFC network integration and was granted a PCS pioneers preference license for its work in the integration of PCS with cable. A significant portion of Cox's activities were RF transport demonstrations with multiple vendors. In particular the Motorola/Cox trial in San Diego used RF transport to integrate GSM RF signals with the cable plant.

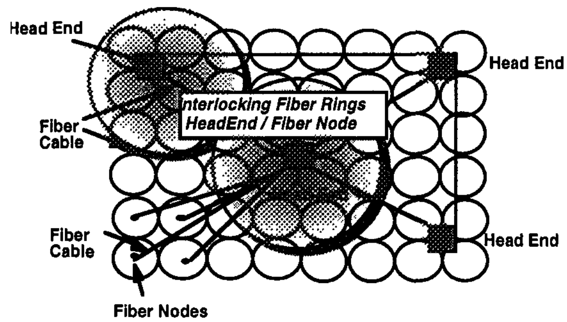


Figure 4.

As shown in Figure 4, the trial used what was earlier called RAD/RASP (Remote Antenna Driver / Remote Antenna Signal Processor) or is now referred to as CMI/HIC (Cable Microcell Integrator/Headend Interface Converters) equipment to transport RF energy from a base station through the cable plant to an antenna system. Additionally, Motorola has performed a CDMA CMI/HIC trial as well as baseband transport with TCI in Dallas, TX and Arlington Heights, IL respectively. The experience and results of these trials have been combined with the Cox Cable Communications trial to formulate many of the conclusions in this paper. Further information regarding the Cox Cable Communications GSM CMI/ HIC trial is available in the FCC Quarterly Report entitled, Cox Enterprises, Inc. Sixteenth Quarterly Progress Report to the FCC for Experimental Licenses, March 1995.

CMI/HIC

With CMI/HIC a HIC is co-located at the base and converts RF energy (CDMA, GSM or other air interfaces), that would normally go to an antenna, to an upstream and downstream frequency allocated for wireless communication on the cable plant. The RF energy is injected into the cable plant at a very low signal level (10 dB below a normal video channel). In the cable plant, CMIs convert the RF energy at cable plant frequencies back to the original frequency of the wireless system. The CMI is a 40 pound box that is located on the cable coax strand at an average of 23 feet above the ground level.

To facilitate understanding, Table 1 delineates up- and down-frequency conversions for a CMI/HIC system.

Table 1. Representative Frequency Conversions for CMI/HIC System

CMI/HIC System Component	Down Stream	Up-stream
Base Station Frequency	1862.2 MHz	1942.2 MHz
Cable Plant Frequency	541.9 MHz	20.5, 23.5 MHz
CMI Frequency	1862.2 MHz	1942.2 MHz
Portable Frequency	1862.2 MHz	1942.2 MHz

In this example, a base station modulates a PCS RF carrier at 1942.2 MHz. The RF energy is then sent via transmission line to the HIC which converts it to 541.9 MHz. The signal then travels through the cable plant to a CMI where it is converted back to the original 1942.2 MHz. This is the path for the downstream frequency translation.

In an upstream path (CMI to the base station) the conversion works exactly the same, only in the reverse. A CMI receives a portable's RF energy at 1862.2 MHz from two (space diversity) separate receive antennas. Antenna A's RF energy is translated to 20.5

MHz and antenna B's energy is translated to 23.5 MHz. The RF energy is delivered back to the head end using the cable plant and is converted to 1862.2 MHz by the HIC.

The CMI is a low power device with a one watt power amplifier. Since it is only located 23 feet above the ground, it has a very small coverage area (less than one-half mile range for GSM, larger for CDMA).

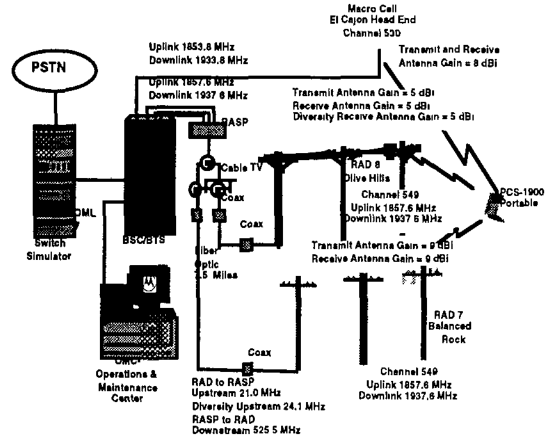


Figure 5.

For the CMI system to cover an area as large as a macrocell, as shown in Figure 5, multiple CMIs are simulcast from a single HIC and base station. From the base station's point of view, the RF energy is sent to only one antenna and its receiver is gathering RF energy from only one antenna (two if spatial diversity is employed). The RF transport system uses multiple CMIs (from four to 10) to simulcast and sum RF energy at each CMI to simulate a macrocell sector. In total, the CMI/HIC system is transparent to the radio system and acts as a distributed antenna array with cable friendly frequency conversions.

Baseband Transport

The second technique to use the cable plant for wireless distribution is baseband transport. Baseband transport uses the cable plant to distribute base band (DS0 level) information through the cable plant.

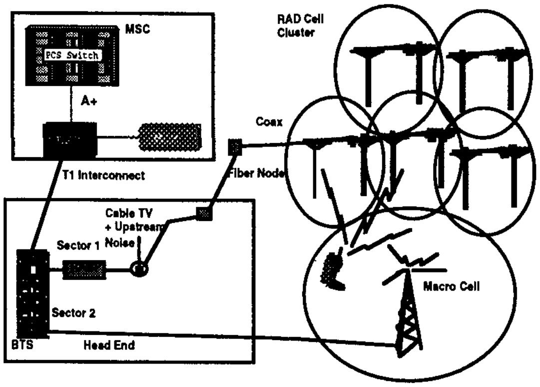


Figure 6.

As shown in Figure 6, it connects the CBSC to the BTS. Instead of using microwave or leased T1s to interconnect all the cells to the CBSC, the cable plant is used. The key benefit of this technique is that the cable plant is already in place and ready to provide this distribution. The CBSC would be located at the headend and the T1 is converted by a cable modem to a channel on the cable plant as shown in Figure 7.

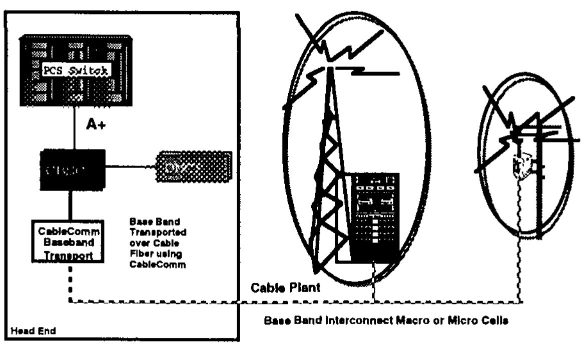


Figure 7.

The cable plant distributes this baseband information to the cell site location. The cable modem retranslates this data to a T1 format that can interconnect to the base station. Motorola has developed its CableComm™ product to meet this interface. CableComm™ can support up to 28 traffic channels in 600 kHz of cable bandwidth.

The trend in the wireless industry continues toward small self-contained microcells. These microcells typically weigh between 40 and 150 pounds and can support from 200 to 500 subscribers. The key advantage to microcells is that they can be located in various environments: top of bill-boards, top of buildings, in buildings and on telephone poles. The cable strand may support up to 60 or 70 pounds. At this weights and coverage area, a microcell coupled with the extensive existing cable plant creates an eloquent service solution for coverage, capacity and backhaul.

The difference between RF transport using CMIs and baseband transport using microcells is illustrated in Table 2.

Table 2. Comparison Summary of RF and Baseband Transport

Issue	RF transport	Baseband Transport
Number of Radiators	Multiple CMIs simulcasting a single carrier	Single Microcell
Cable Plant noise	Reduces the coverage of the CMI	No effect on the air interface
Location of the base station	Headend	Cable Plant
Cable Spectrum Required	2 to 4 MHz	600 kHz
Number of CMIs or Microcell per Macrocell coverage	10 to 30	5 to 10

BENEFITS OF INTEGRATED HFC AND WIRELESS NETWORKS

Recognizing that approximately 80 percent of a wireless network is the radio system, integrating HFC and wireless networks can provide many competitive advantages for the operator. First economic models have shown that CMI/HIC deployment is potentially one-half the cost of a macrocell deployment on a cost per square mile basis. Cable based microcells may be even less expensive, approaching one-quarter cost of a macrocell deployment on a per subscriber basis. The cost savings are the result of removing the imbedded costs associated with towers, transmission line, power systems, shelters and site improvements.

Beyond one time deployment cost savings, HFC - wireless integration provides continuous operational cost savings by using the HFC network leased lines, land lease improvements are avoided. This issue becomes even more critical with the PCS or 1900 MHz wireless operators. At 1900 MHz, a system will require up to two or three times the number of cells for coverage versus a similar 800 MHz deployment. Additionally, an operator has the opportunity to more fully utilize an existing cable maintenance staff for both HFC and wireless tasks.

Yet another benefit to HFC based wireless system deployment is the potential for a radically lower time required for site deployment. This belief is derived from HFC network ready availability for backhaul along with the absence of zoning/site approval, coupled with HFC operator's right of way access for in-building systems. Compounding these advantages is the relatively short time required to commission a CMI or cable based microcell, on the order of one hour.

Last, and potentially most important of all, is the opportunity to unlock the hidden asset value in the HFC network by demonstrating its versatility and value for wireless. Under analysis, one would expect that by owning a more valuable asset, the stock value of the owner would be positively impacted.

ISSUES ASSOCIATED WITH HFC DEPLOYMENT

As commented previously, Motorola has conducted field trials with both GSM and CDMA CMI/HIC as well as baseband transport. This coupled with Motorola's extensive history as a wireless infrastructure and subscriber unit manufacturer provides Motorola with the broadest view of the issues associated with HFC integration with wireless telephony.

These extensive trials have validated both Motorola's theoretical analysis and system simulations regarding the limits and viability of specific technologies for use in HFC networks. Depending on the particular technology chosen for wireless service, there are significant issues that must be addressed for HFC system deployment. In summary, our trials have demonstrated that these issues are substantially more difficult for GSM CMI/HIC deployments versus CDMA CMI/HIC or cable based microcells with baseband backhaul.

RF Transport Issues (CMI/HIC)

The issues associated with GSM CMI/HIC are fundamental to the cable plant and the GSM air interface. These issues can be summarized under the following four broad categories.

- Dynamic Range
- Noise Summing
- Fiber Node Differential
- Powering

Although other TDMA based air interfaces have not been examined in field trials by Motorola, one would expect that these technologies (IS-54 and IS-136) would deliver similar results as the trialed GSM air interface.

Dynamic Range

Depending on the network usage and configuration the level of ingress or noise becomes problematic. The wireless signal must overcome the noise level if the signal is to be detected at the base station. Effectively, the noise level or ingress in the HFC network reduces the sensitivity of the system. One option to address this issue is to insert gain in the CMI to increase the sensitivity of the unit. Unfortunately, the use of this option is limited on the upper level by the requirement that the upstream lasers not be overdriven to saturation. Upstream laser saturation would interrupt not only the wireless signal but also all video service. Obviously this is not an acceptable condition.

Exacerbating the limited dynamic range is what TDMA based GSM originates at full power (1 watt). This creates the situation were a subscriber located directly next to a CMI would saturate an upstream laser. Conversely, CDMA originates at minimum power and allows for a larger gain in the CMI. This minimizes the concern of a call origination occurring too close to a CMI and interrupting upstream services.

Noise Summing

The issue of noise summing is similar to the law of diminishing returns. To elaborate, with the introduction of each CMI, the overall noise level of the cable plant is increased and therefore requires additional CMIs to overcome the increased noise floor. To minimize this effect, it is important to decrease the number of CMIs as much as possible. Effectively, the coverage advantage delivered by CDMA is magnified by the noise summing condition thus producing substantially more CMIs for GSM versus CDMA.

Fiber Node Differential

During our trials, it has been observed and measured that due to the distance traveled in the HFC network the vast majority of fiber nodes are not within the equalization tolerated in a TDMA based GSM network (16 µsec). The impact of this problem is that some form of equalization circuitry must be calculated and installed in the CMIs prior to a GSM system deployment. Beyond the added complexity in the CMI this issue raises tremendous operational and system planning issues for the operator. Such operational and planning issues include: time, location and skill required for CMI placement; viability of disaster recovery and self healing fiber rings due to the dynamic delay changes; and system growth planning.

Powering

As more centralized power systems are introduced for reliability, it becomes increasingly important to limit the number of CMIs and the collective current draw per fiber node. Beyond the simple economics of requiring more power supplies for more CMIs, it is also important to monitor the network design to insure that maximum current is not reached on any particular power feed. This issue is compounded with the larger number of CMIs required for GSM based system deployments.

Baseband Transport Issues

The issues noted above are less of an impact for baseband transport due to the fact that baseband transport is performed in a known digital format that allows lower power levels, error correction, auto noise tuning and auto-level setting. The primary issues for baseband transport are microcell pricing, security of cell and weight/size versus power output tradeoffs.

HFC NETWORK INTEGRATION PLANS

Motorola is continuing its program of developing cable based microcells with integrated baseband transport as well as additional improvements and system level trials with CDMA CMIs.

SUMMARY

Hybrid Fiber Coax networks continue to proliferate as the chosen architecture for meeting the broadband service requirements of consumers. These HFC networks provide significant cost and time savings for operators wishing to more fully leverage the value of the network to distribute wireless signals. Nonetheless, the technology and approach used to distribute the wireless service play dramatic roles in the operator's competitive cost and time advantage in future operation plans.

Of particular note there remains significant uncertainty and work with regard to TDMA based CMI/HIC deployment. In contrast due largely to the noise immunity and coverage enhancements provided by CDMA's spread spectrum, either CDMA RF or baseband transport appear to be viable and economically attractive solutions for an integrated HFC-wireless system deployment.

ABOUT THE AUTHORS

Robert W. Hammond

Mr. Hammond is a member of Motorola's PCS Core Marketing group within the Pan American Markets Division. During his five year tenure with Motorola Mr. Hammond has held various positions in the engineering, sales and marketing groups. He holds a Bachelor of Science Degree in Electrical Engineering from Purdue University, a Master of Science Degree in Engineering from West Coast University and a Master of Business Administration from The Lake Forest Graduate School of Business. Prior to joining Motorola, Mr. Hammond held positions in missile flight test and engineering with General Dynamics in California.

Mr. Hammond has left Motorola and is now working for Sprint Telecommunications Ventures.

Douglas E. Hohulin

Mr. Hohulin has been with Motorola for over seven years and his present role is Business Development Manager. While at Motorola, Mr. Hohulin has been an instrumental technical member of multiple account teams throughout the world. Through these diverse and expansive experiences, Mr. Hohulin has gained extensive real world experience in radio frequency system design and network architecture analysis. Included in Mr. Hohulin's experiences are assignments in Europe, South America and the United States as both a technical expert and sales engineer. Prior to joining Motorola Mr. Hohulin was employed by Magnavox in Ft. Wayne, IN. He holds a Bachelor of Science Degree in Electrical Engineering from Purdue University.

Acronyms

HFC	Hybrid Fiber Coax
GSM	Global Systems for Mobile Communications
CDMA	Code Division Multiple Access
RF	Radio Frequency
TCI	Telecommunications, Inc.
FTTH	Fiber To The Home
FTTC	Fiber To The Curb
FTTN	Fiber To The Neighborhood
CBSC	Centralized Base Station Controller
MSC	Mobile Switching Center
MSO	Multiple System Operators
PSTN	Public Switched Telephone Network
PCS	Personal Communication Services
TDMA	Time Division Multiple Access
RAD/RASP	Remote Antenna Driver/Remote Antenna Signal Processor
CMI/HIC	Cable Microcell Integrator/Headend Interface Converters

INTELLIGENT CABLE DATA SERVICE ARCHITECTURE: ENSURING COMPETITIVE PARITY WITH OPTIMIZED CAPACITY

Terry L. Wright
Convergence Systems Incorporated (CSI)

Abstract

*Cable has a **"once-in-an-era opportunity"** to leap to the head of the class in Internet (and other information) services by optimizing and leveraging its bandwidth advantage in the delivery of data services. Cable's challenge is to deliver competitively-priced yet **"superior informational experiences',** to increasingly diverse customer communities, before Local Exchange Carriers (LECs) can realize returns on investments in more efficient local loop technologies (e.g., ADSL) and HFC overbuilds. The key to this opportunity is the development of an integrated cable data services architecture. Such an architecture, when implemented, would optimally leverage existing network assets while maximizing return on HFC capital expenditures, creating a data service environment with the broadest of market segment appeal.*

This paper explores key aspects of cables' architectural challenge. It first explores the potential rewards associated with achieving an optimal data services architecture. It then examines a data services architecture designed to optimally leverage cables' existing broadband assets, while maximizing return HFC investments. This architecture attempts to show how some latent (and generally un-used) HFC component features, might be combined with improved packaging (i.e., reduced form factor) of certain legacy LAN functions (expected in the very near term), to create a flexible data services infrastructure capable of supporting the broadest array of applications and related market segments

PERSPECTIVE

The deployment of Internet access and other data services capabilities over cable television introduces a whole new competitive dimension to the cable industry (and, we might add, just in time). These services represent the most promising near-term cable offerings for creating profitable new revenue streams, enhanced system valuations, and competitive parity against other local telecommunications companies. The importance of these services is further amplified by the market uncertainties associated with new entrants and emerging threats to traditional core cable entertainment offerings. After all, the future success of traditional cable and recent competitive entertainment services such as video dialtone, MMDS, LMDS, and DBS, depends entirely on the shifting consumer perspective re: how to spend leisure time. To surf, learn, and grow, or sit and risk mental atrophy.

Perhaps we shouldn't be so surprised at the Internet phenomena; after three decades of distributing information and information technologies, society is beginning to get hooked on the knowledge and potential opportunities that can be won via ready access to information, Access to the Internet represents a valuable "portal" into the largest and most dynamic global repository of humanly consumable information in existence. The cable industry has the unique opportunity to define a new and enhanced class of access portal using cable television networks, a portal that enables the Internet experience to take on whole new dimensions and entertaining attributes. The enabling of these new dimensions will, in turn, cause the Internet to respond with new experiences that take advantage of them. The players providing the most *value-enabling* portals onto the Internet will help shape the Internet's evolution by continually enhancing the potential of the Internet experience.

Cables' data-oriented opportunities are, however, accompanied by non-trivial architectural challenges associated with the deployment of robust, reliable, scaleable, and manageable data services, in multiple classes and quality levels, in an ever-changing technical and marketing climate. If cable is to effectively compete in the exploding data services space, it must embrace these architectural issues in such a way as to maximize the value of its present bandwidth advantage, across diverse customer communities, while maintaining the integrity of its traditional entertainment delivery capability.

With the steady advance of communications technology, and the continuing integration of computing into the telecommunications and entertainment infrastructures, cable can draw on an increasing array of advancing capabilities to embrace its architectural challenges. However, even with today's truly incredible arsenal of technologies at its disposal, the absence of key functions in critical areas still prevent the realization of the optimal high performance infrastructure. It is important to note that most (if not all) of the technologies needed to implement such a data service infrastructure already exist. Indeed, some of them simply need to be replicated and relocated within the infrastructure, while others are repackaged and deployed in ways possibly beyond their creators' visions.

Although some of the functional integration concepts envisioned may seem radical at first, they essentially represent traditional data networking concepts common to the traditional computing industry market, applied to cables' broadband HFC environment. When viewed in the broader context of today's convergence economy, an economy that is in no small way influenced by how and in what quantity we choose to package sand and glass (silicon chips), these concepts are not that unfamiliar.

Finally, this paper should be viewed as posing "what if' architectural questions to an industry already wrestling with a variety of changing forces, from new realms of opportunities, to new and very real threats beyond its control. These and other dynamics not withstanding, this paper advocates a proactive approach to the architectural challenges that stand in cables' way of delivering "superior informational experiences" to growing market communities anxious to participate in them. It is estimated that the Internet grows by some 25,000 users each day, with almost 22,000 of these users adding their own web page (one page every 4 seconds). With this as an incentive, and with the knowledge that we are for the most part simply exploring concepts that combine re-packaged legacy LAN functions with existing (but little-used) HFC features (versus having to invent new technologies), the context and thrust of this paper is established. Physical (electrical) issues such as ingress, common path distortion, laser clipping, power line hum, and so on, have been thoroughly analyzed by many others in a variety of quality works, These issues are the center of much debate throughout the industry and in several formal standards groups, and are not discussed in this paper. However, as important as these physical challenges are in the overall cable data services agenda, the value of their imminent solutions may be lost without a clear understanding of how best to deploy them.

INTRODUCTION

Most will agree that the cable industry's most definitive (current) competitive advantage over other data service and network access providers is its broadband spectrum, and the superior locally-deployed bandwidth it represents. A well-designed cable network *architecture* is the key to leveraging cable's bandwidth advantage.

A good architecture will leverage cable's bandwidth advantage for many years to come; a poor (or lack of an) architecture will likely sacrifice this initial advantage in the wake of poor reliability, contention-oriented service access denials, congestion-oriented poor service performance characteristics and related quality of service (QOS) issues. Problems of this nature would undoubtedly result in serious growth impediments and high customer attrition rates for cable-based data services.

The primary "data services" cable agenda in 1996 must therefore be to successfully overcome these architectural challenges in such a way *as* to achieve a *future-ready* technical posture consistent with long term prosperity in the data services space.

Many cable companies are already deploying there best estimate of what a hybrid fiber-coax (HFC) network should be. The absence of key formal technical standards, and other industry dynamics, are creating reluctance on the part of many cable operators to invest heavily in such an uncertain market climate. However, an expanding markets' growing appetite for higher performance information access, and a confidence that competing interests will certainly respond to that appetite with any performance improvements they can muster, creates a serious paradox for cable.

Should cable wait for a more stable technical climate while competitors' capabilities (and market shares) grow, or should it simply set course and weigh anchor with the knowledge that it may have to build a new ship while at sea? Would a cable architecture that ensured realization of maximum bandwidth efficiency and related market value be compelling enough to move cable forward with serious energy focused on its data services agenda?

It is probably fitting that, consistent with cables' traditional entrepreneurial spirit, it must pursue its data services architectural challenges in the most entrepreneurial of environments where new industries can spring up overnight, and traditional architectural concepts are constantly being tested and sometimes redefined. The very definition of information architecture itself is evolving. Before examining these challenges, a brief excursion into the nature of the reward such an architecture might bring seems in order.

PART I: THE POTENTIAL REWARD

The Nature of Association: A Look Back

In the early days of telecommunications, telegraph operators provided a service where the delivered value was a decoded message on a piece of paper. In this case, the service provider was the one with the telecommunications device. Then came the telephone, where the telephone company provided everyone buying voice service with a device called a telephone. Radio, and then television, marked the beginning of scenarios where user-owned devices played an active role in the delivery/consumption of telecommunicated information or other value. When we were all allowed to own our own telephones, then basic telephony joined the scenario where customer-owned devices in the home played an active role in the delivery of a telecommunications service. Next came the cable set-top converter, yet another device in the home, that greatly expanded the programming choices available to the viewer.

These devices became associated with the nature of the service being delivered and consumed, e.g., telephones became associated with voice services, the television with visual entertainment and news services, and the radio with audio entertainment, news, and other programming. The cable set-top converter became associated with greater *variety* in television content, and importantly, gave viewers far greater options with respect to control over what they watched on television.

The important point here is the *association:* people associate different devices with the consumption of different services (e.g., radio with the reception of audio entertainment, television with the reception of televised content, and so on). Both bounded (e.g., POTS) and less bounded (e.g., cable television) services have associated with them a device involved in the use of those services. The computing revolution, however, added a new twist to this "service-device" association.

Information *Manipulation* Association

People bought personal computers to do useful and/or interesting things with information, whether using a spreadsheet program, a word processor, and/or playing a game. Computers basically manipulate, transform, package, and present information at very high speeds. Storage and other limitations of affordable personal computers, and the need for people to exchange information electronically (i.e., Email, document exchange, and so on), fueled the concept of *networking,* making it possible to *access* remote information and even collaborate and exchange information with others in real time

The phenomenal popularity of the Internet's World Wide Web has even spawned a challenge to the association of information manipulation with the PC; the new low-cost *Internet appliances* grabbing headlines recently are contenders for this association, although the PC's have a formidable headstart (i.e., 70 million plus PC households) to overcome.

Information *Access* Association

Since the telephone line represented the incumbent two-way media in and out of buildings and homes, the telephone modem naturally became the device associated with traditional data networking.

However, the steady evolution of computing, software, and networking around the markets' growing need for more and more sophisticated methods of packaging and exchanging information, **has led to an unprecedented opportunity for cable to *re-define* the device associated with *access to information*.** Visual information (e.g., graphics, video, imaging) has always been superior to mere words, with respect to the exchange of ideas and information (i.e., a picture is worth a thousand words). As information becomes more visual in nature, and is reliably delivered over cable networks, its consumers will very likely find it natural to associate this kind of information with the television industry.

Since the telephone system was designed around the 4 Khz wide voice-band, it contained its own built-in limitations re: useable bandwidth which, to the detriment of telephone companies, falls far short of the bandwidth needed to reasonably deliver today's popular multimedia information. The telephone companies, however, are not sitting still. Intense work on expanding the information-carrying capacity of the local loop, both in compression algorithms and line encoding efficiencies (e.g., ADSL), are beginning to pay dividends as the telephone industry responds to cables' emerging high-performance data services threat. In addition, the LECs have announced plans to build out broadband networks to cover more than 30 million homes nationwide on planned expenditures of over $30 billion.'

The cable industry, on the other hand, has its own set of built-in limitations associated with the traditional one-way nature of cable networks, and the less sensitive (versus data) nature of the signals involved in the delivery of traditional cable entertainment programming. The development of bi-directional HFC networks and related 2-way active coaxial components go a long way towards eliminating cables' primary obstacles. But will HFC networks and 2-way components enable cable to win the battle for the markets' association of cable-based data solutions with the preferred method of accessing information? Perhaps, but more probably if cable-based information access represents a superior information *experience* regardless of whether that information resides on the Internet, Online service provider system, or on the office, hospital, or school LAN just a few miles down the road.

The stakes are high in the competition to earn the markets' industry association for data and information services, and the telephone industry has a running head-start.

Cables' Opportunity

In order to effectively compete with the telephone industry's' certain continued innovation and capital investments, cable companies must deploy data services capabilities architected to optimally leverage their local bandwidth advantage in the areas of

- maximum service capacity (user concurrency and system load),
- quality of service (minimized service disruption impact),
- ease of management (fault isolation), and
- maximized functional value/utility for both the data services user community as well as the cable operator.

In the author's view, cable's near-term success in delivering Internet access and other data services is a prerequisite to the market's acceptance of cable as a viable alternative for the many existing (e.g., POTS) and future complex services the Information Age is yet to produce.

PART II: CABLE DATA SERVICES ARCHITECTURE

As with any complex technical system, many instantiations of a variety of architectural approaches can be made to work in the delivery of data services over cable networks. In fact, existing data-oriented cable products, such as existing cable modems and related headend equipment, impose a good deal of system. architecture with their deployment. This is due largely to the current absence of formal standards for this type of equipment. These existing architectures are by definition product-oriented, and are limited in their context to a particular vendor's offerings, and its view of the problems to be solved.

While these early product-dictated architectures are important aspects of cable's thrust into the data services markets (i.e., they are available and useable today), they are only one element in the context of architectural thinking cable must consider in order to sustain any early data services success these products enable.

Background

When most people think of architectures, they think of an orchestrated suite of functional and interface specifications that, when realized through technology deployment, support the interoperability and interoperation of these elements towards accomplishing some meaningful goal that markets will find of value. Traditionally, architectural thinking in information systems design has concentrated on the physical and logical dimensions of the information system objectives. Interoperable hardware and software have long dominated architectural concepts.

As mentioned earlier, however, the all-transforming wave of information and information technologies sweeping over nearly every aspect of society demands that we expand our thinking with respect to the *scope* of information systems architecture. It is no longer sufficient to think about information architectures in terms of hardware and software interoperability among the components of a specific information system solution. The growing role across the board of information and its technologies is establishing new priorities for those that provide information access infrastructure, priorities that can transform existing assets into future mainstream wealth creation systems. This requires information access providers to create flexible service environments where the changing needs of increasingly diverse market segments can be accommodated with new applications, not just the Internet or other Online service. These applications must not be constrained by underlying architectural limitations that restrict access capacity to information repositories due to their location within the architecture.

Current Thinking in Cable Data Architecture

The prevalent thinking today in cable data architecture revolves around the HFC network with asymmetric bandwidth allocation favoring outbound over inbound spectrum (e.g., 30 Mb/s outbound versus, say, 1.5 Mb/s more or less inbound). In addition to all the known benefits of HFC networks (e.g., reduced amplifier cascades, improved signal quality, etc.) this thinking is rationalized as

- stretching the utility of scarce useable return spectrum,
- taking advantage of the "small request/ large response" nature of the way the Internet and other Online services are accessed,
- accommodating the subsplit design of the majority of deployed cable systems, and
- accommodating the performance limitations associated with troublesome return spectrum..

The well-known benefits of HFC networks, as well as the ongoing massive capital investments in HFC equipment by MSOs and LECs alike, suggest that this technology will continue to enjoy wide-spread acceptance and deployment. However, the market opportunities that may be sacrificed to the limitations imposed by asymmetrical data solutions should be thoroughly explored and understood before large scale deployment of technology, especially if the intent is to utilize this same technology to support services opportunities across all market segments.

Additional Thinking on Data Services

Were it not for the limited amount of return spectrum available in the majority of deployed cable systems today, the argument for the asymmetric allocation of this bandwidth would weaken considerably. Since there is no magic wand that can be waved to suddenly create more useable return spectrum or improve the quality of existing spectrum in predominantly subsplit systems, architectural methods that reduce system-wide contention for this bandwidth (and/or technical breakthroughs in return spectral efficiency) offer the only realistic alternative to asymmetric bandwidth allocation. Before addressing an architectural approach with the potential to overcome existing return spectrum constraints, it is appropriate to understand the potential importance of such an approach in terms of the market ramifications that may occur as a result of the use of asymmetrical technology for all data service applications and markets.

"If cable creates a [data] services access environment that supports only a hand-full of asymmetrically-oriented information services prevalent today, it is betting the proverbial farm that asymmetrical information applications will continue to be all a community will ever want or need."[2] Investment resources tracking the evolution of the Internet forecast that "the number of individuals with full Internet access will grow from approximately 20 million today to more than 400 million in ten years, with the greatest growth coming from the corporate sector as businesses move to put themselves and their employees Online" [3]. The important point here is that, **unless these businesses are planning to relocate their servers and storage farms to cable headend facilities. or cable operators are planning to pull fiber to most of these institutions as thev deploy HFC networks, the asymmetric bandwidth model will limit access capacity to this corporate information as a function of the amount of return spectrum asymmetrical connectivity devices can support.** This could be a serious impediment to the use of cable-based data services by the business communities as they look to cut costs and boost productivity through work-at-home scenarios and Internet-based marketing. In addition, when one considers the voluminous nature of the information generated by many of these commercial institutions (e.g., X-rays and EM1 by hospitals and medical labs, graphical product descriptions and promotionals from retailers and grocery stores, and so on), restricting access capacity to these potential data services customers may be an expensive proposition in terms of lost market share. Commercial use of the Internet (especially the Web) to

- enable better employee collaboration,
- market products and services,
- provide access to corporate information such as employee handbooks, standards, product literature, and
- manage and administer software updates and other Workgroup tasks

offer compelling rationale to consider architectural approaches and technologies that accommodate the limited return spectrum situation with high-spectral efficiency on return channels versus compromising the value and related appeal of cable-based data services to commercial segments.

It is also clear from the estimates of numerous industry sources that applications such as video-conferencing, and the rising popularity of multimedia content on the Web, will continue to drive up demand for bandwidth in both directions. A realistic view recognizes that virtually all current sources of Internet content (including web pages and other multimedia) are distributed server systems located somewhere other than cable television headend facilities. An architecture that offers the highest capacity access to only those server systems located within cable headend facilities, or directly off an Internet backbone, significantly limits the appeal of cable-based data services as an Internet access and local data transport mechanism.

Leveraging Available Technology

Virtually all the technologies necessary to deliver data over cable, *without* imposing the limitations of an asymmetrical bandwidth model, exist today. As previously noted, to accomplish this will require the creative packaging of some of this technology in combination with latent yet currently un-used capabilities available in some HFC technologies (see Figure 4).

Although many new offerings, especially in the cable modem space, are expected to be available in the late 1996/early 1997 timeframe, at 25,000 new Internet users a day, a lot of market share will likely go elsewhere in the interim if cable opts to wait for these offerings to become available and stabilize. In addition, the large majority of these anticipated new cable modems offer asymmetrical service only. While these products will certainly play a large role in the providing robust access to the Internet and other Online services, cable operators are encouraged to consider the current crop of modem solutions to establish an immediate market presence (and gain meaningful intelligence) in both consumer and commercial markets.

The large majority of currently available (and deployed) cable modem solutions provide surprisingly reliable symmetric services over existing coaxial and HFC cable networks. Although these currently available modems may not offer the super high-performance in the outbound direction as those expected later this year, they nevertheless provide as much as 10 Mb/s symmetrical performance *now*, with further improvements expected as these products continue to mature. Not only does currently available technology enable cable's immediate entry into the data services space, its symmetrical capabilities should be especially appealing to commercial customers (e.g., ISDN and dedicated circuit users) hoping to utilize cable's superior capacity to facilitate work-at-home programs, enable employee collaboration, and market products and services on the Internet *now*.

The Challenge

Before exploring a viable architecture for cable-based data services, it is important to review primary architectural objectives, and the challenges that must be overcome to accomplish them. Our primary architectural objectives are to

- leverage cables' superior bandwidth capacity (relative to traffic load) while maximizing *access probability* and sustained session *performance through-put,*

- enable delivery of robust data services, in multiple classes and at different quality of service (QOS) levels, that support the applications and needs of as many market segments as possible,

- deploy *manageable* technology with maximum cohesion (functional isolation) and minimum coupling (inter-function dependency) to accommodate anticipated changes in available technology, and importantly

- deploy *profitable* solutions that can be efficiently and inexpensively transformed by the still uncertain macro and micro trends and forces the "infoconomy" is yet to exhibit and/or create.

Rationalizing a Model

Just as the arrival of the Industrial Era brought with it new thinking about the nature of commerce and social agendas, giving rise to such concepts as assembly lines, mass production for mass markets, and rapid urbanization, the creative challenge accompanying the infoconomy is about to sweep over us like the tidal wave it is. If we thought the move from the farm to the factory was significant, the rise of intellect over muscle that is rapidly taking root around us will make the farm-to-factory shift seem like a momentary preoccupation with human materialism.

Information technology has, in its relatively short existence, grown from a simple tool in the game of business, to the game itself. The plow has become both the field and the crop, where a persistent rain of creative thinking falls on mankind's' experiments with sand and glass, yielding crops of new perspectives on not only *how* we do things, but on *what* we do. And with the disparity in the distribution and use of information technology around the world, the infoconomy wave will produce different effects in different places at different times, giving it many meanings to many different communities.

As different as the fast-paced assembly line was from the season-long cycle from seed to crop, the Intellectual Era upon us will be far more profound. New entrants and competitive offerings call for adjustments in strategy. When whole new industries, methods of commerce, sovereign relations, and human priorities are created and transformed by

- the changing commercial and social significance of Information Age phenomena like the Internet and the World Wide Web, and

- the changing climates of various horizontal and vertical market segments.

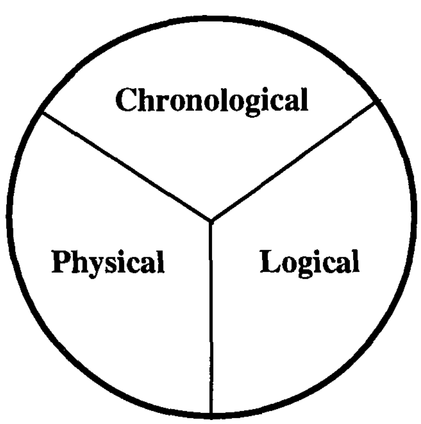


Figure 1. Architectural Thinking in an ***Infoconomy***

the combination of sand, glass, and innovative thinking, it's time to invent new tools and adopt new perspectives consistent with that kind of world. The concept of an architecture is a good place to start.

Figure 1 depicts a perspective on an architectural model that is both simple and compelling. It suggests an expanded definition for information systems architecture that incorporates a temporal dimension in recognition of

- the increasing amount of information available about information technology (for infrastructure planning),

- the growth of ED1 and its implications (on value chain and other strategic analysis),

- the central role information and its technologies now play (at the macroeconomic and geopolitical levels),

The elegance of this architectural model lies in its comprehensive simplicity. While traditional architectures have always embraced the logical and physical dimensions shown here, there wasn't captured information about temporal events to attribute any significance to it as an architectural consideration. The addition of the chronological dimension acknowledges the increased role of information technology in nearly everything any of us do, the increased predictive and planning capabilities it endows us with, as well as it self-perpetuating nature. Figure 2 provides a more detailed view of the kinds of functions that lie within each dimension of this architectural model.

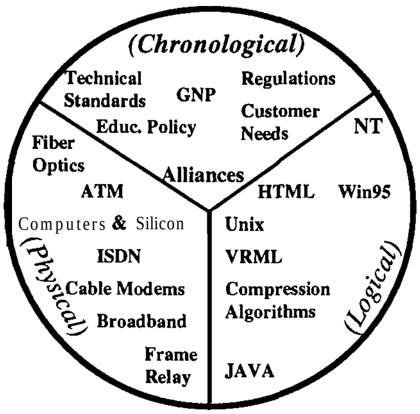


Figure 2. A Few Architectural Components

The key differentiator between the above architecture, and traditional architectures, is the presence of a chronological dimension. This is in recognition of, and respect for, the transforming central role information and information technology is increasingly playing in every facet of society from global commerce to entertainment, sovereign relationships to politics and social priorities. In other words, the *permeation* of networked information and information technology throughout the fabric of all these commercial, social, and institutional segments is creating increasing opportunity to include temporally-oriented forces, processes, and events in our architectural thinking. As this permeation continues across the board, can anyone afford to ignore the potential for linking these chronological issues into our automated information? Considering the pace at which this "info-permeation" is evolving, we are compelled to examine the possibility of making room for the implications of value chain analysis in our core business and technical architectures. With this architectural context in place, we can now explore an architectural approach for cable-based data services that embraces all three dimensions.

A Cable Data Architecture

"A network must be both flexible and reliable enough to allow for future services".[4] The goal of the architecture pictured below is to profitably leverage all the capacity and performance capabilities of a 2-way HFC cable network, supporting as many market segments and applications as possible. This must be accomplished without compromising delivery of entertainment, while accommodating known cable characteristics such as the scarcity of quality return spectrum in most plants, the shared nature of the cable network medium.

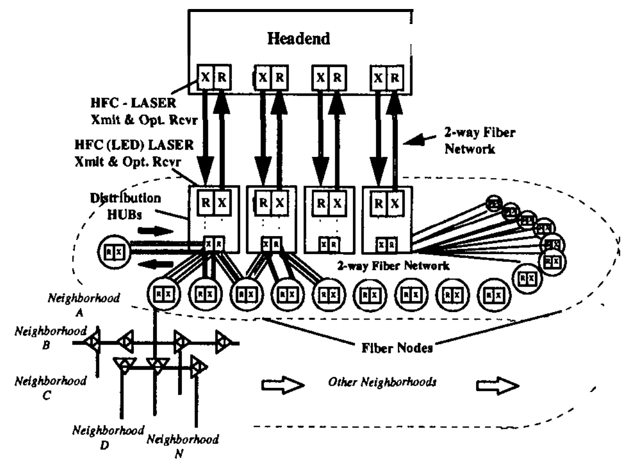


Figure 3. A Typical Cable HFC Architecture

Figure 3. shows a typical HFC implementation utilizing a popular hierarchical approach (e.g, master headend, distribution hubs, fiber nodes, and feeder coax). (Hub-to-fiber-node ratios shown are simply for illustration; real implementations have been found as high as 1:40. Amplifier cascades counts are typically 5 or less.)

In order to maximize available capacity across an entire cable system, while minimizing contention for that capacity, routed segmentation of the cable network is required. Routed segmentation architecture requires special packaging of some existing router, cable modem/bridge MAC/PHY layers, and frequency translation technologies, as well as the utilization of frequency stacking capabilities available in some existing HFC fiber optic technology. A fiber node capable of frequency stacking, combined with resident routing, cable bridge MAC/PHY, and frequency translation functions introduces a new class of broadband data communications device called a "SPRouter™" (for Spectrum Parallel Router). Figure 4 below a possible SPRouter implementation.

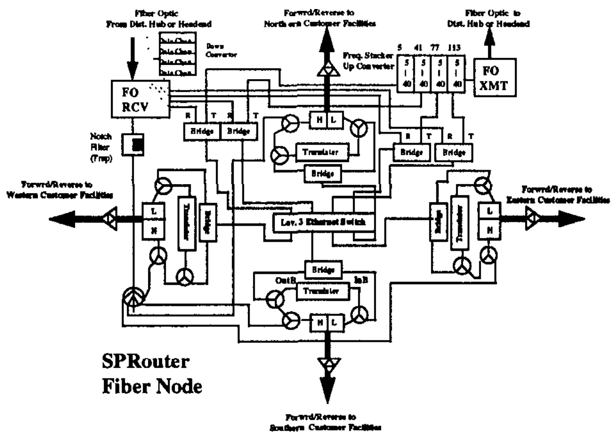


Figure 4: A APRouter Implementation

A *sprouter-bused* (routed segmentation) cable data network takes advantage of recent developments in fiber node technology (frequency stacking) and router function siliconization (i.e., router chips). The frequency stacking capabilities inherent in some recent fiber node implementations would be used to simultaneously transport (i.e., in parallel) the 5 to 40 MHz return spectrum, from each physical neighborhood feeder network, back to distribution hubs (DH) or headends (HE) for further routing. Emerging router chips would be useful in constructing DH or HE resident routers to accommodate the parallel networks created with the SPRouter approach.

The challenge in implementing a sprouter-based network is the need to locate router and other "LAN-defining" equipment (e.g., frequency translator, MAC/PHY layer modem/bridge device) functionality as close as possible to each physical neighborhood feeder junction along the cable distribution network. In order for the optimal implementation of this approach to be feasible, a low-cost router chip, combined with an intelligent MAC/PHY cable modem device, would be required that could be integrated into two-way trunk amplifiers, line extenders, and fiber nodes. However, until market demand (characterized in terms of data traffic analysis, specific applications usage, customer expectations and requests, and so on) proves that an investment in sophisticated trunk amps and line extenders would be warranted (and rewarded), "neighborhood size" cable-based LANs would be premature. (If cable acquires significant data services market share in high densities, it would not be surprising to see developments in this area occur.)

In the interim, however, many of the pieces exist to create routed LANs out of individual fiber node coverage areas (i.e., neighborhood groups). In addition, this approach would take advantage of spectrum stacking capabilities already available within existing fiber node technology, and the use of an intelligent switching capability at distribution hubs and/or headends. Fortunately, silicon-based router functionality is nearing availability. This will create opportunities to at least explore other related issues (e.g., weather-proofed enclosure, powering, space, etc.) associated with packaging router functionality within key active components.

The SPRouter approach requires complimentary functionality at the DH or HF. Figure 5. shows the DH or HE-based SPRouter that provides the functional counterpart to the SPRouter Fiber Node shown in figure 4.

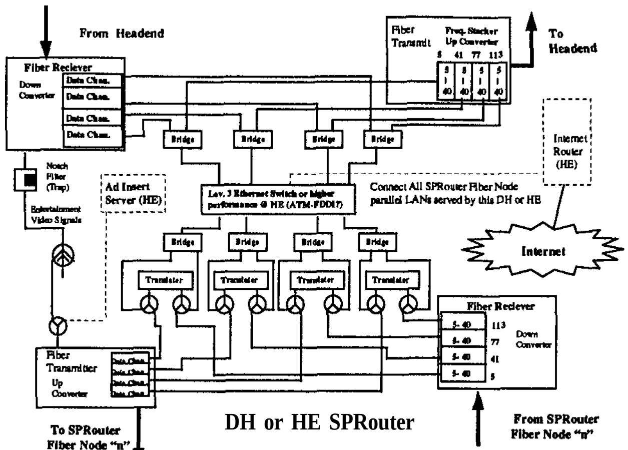


Figure 5: A HE or DH APRouter Implementation

This effect of a SPRouter network is the creation of a number of distributed LANs interconnected by the cable network where locally-destined traffic (i.e., within a LAN) is contained, and only remotely-destined traffic traverses beyond a LANs' boundary router.

The benefits of this approach are numerous:

1. More customers served by a cable system can engage in more locally-oriented bandwidth-intensive applications (e.g., a home to neighborhood school conference).

2. System-wide traffic is reduced to only that traffic needing (based on router decisions) to traverse the entire system in order to reach its destination (e.g, a home office to a corporate LAN, or an Internet user to the Internet access portal).

3. Each distributed LAN can utilize the same highest-quality return spectrum.

4. Each neighborhood LANs receive only the data traffic (from DH or HE level) destined for it by router decisions.

5. By integrating intelligent (level 3 switches) and modem/bridges within the DH and Fiber Node architectural roles, cable's data services delivery capability can easily take advantage of other technologies (e.g., Fast Ethernet, FDDI, SONET) as they become available and cost effective.

CONCLUSIONS

An intelligent cable network architecture, such as the SPRouter concepts described above, will help maximize cable's broadband spectrum bandwidth advantage by leveraging the functionality of typical HFC system-wide active devices and data equipment (modems, translators, switches/routers) relative to their role in the delivery of data services. This architecture identifies and maps minor (and reasonable cost) enhancements to mainstream cable network active trunk devices (e.g., fiber nodes) that will, when integrated into the functionality of an overall data service delivery system, transform the typical HFC cable network into a competitive weapon for the industry re: the optimized delivery of concurrent data and entertainment services. Importantly, the proposed device enhancements and system roles are technically feasible today, with many of the features that will be leveraged already latent but under-utilized in existing devices.

As stated at the beginning, one of the primary goals of this paper is to stimulate "outside the box" thinking among cable operators and technology providers. BY deploying networks that are capacity-optimized for data services, the cable industry will create a formidable challenge to competitors who would take advantage of any HFC capabilities not leveraged by the cable industry. As new technologies (e.g., ATM, others) become available, and as HFC systems evolve, cable must continue to strive for architectural approaches that leave little room for improved efficiency. With the stakes being what they are, and likely to get higher, why leave competitive advantage un-exploited?

Cable modem solutions are, indeed, a necessary part of cables' data services agenda, and their increasing availability is an important aspect of the challenges cable faces. However, cable's thrust into data services involves much more than the right cable modem. Cable must consider infrastructure enhancements to address the complete context of necessary functions like help desk support, data archival and storage management, billing services, software version control and licensing management, routing table management and domain administration, server performance tuning and system optimization, to name but a few. These and a variety of other issues must be considered when seriously contemplating provisioning Internet access and other data services. It is through the efficient use of information in all facets of delivering Internet and data (or any other) services that competitive advantage will be determined.

---

[1] J.P. Morgan Securities Inc.; T. Savageaux, Equity Research, November 28, 1994, p.14

[2] A Case for Symmetrical Bandwidth; T. Wright, Convergence Systems, Inc.; Proceedings of the 1995 2nd International Workshop on Community Networking

[3] Institutional Research; Robertson, Stephens, & Company; The Internet Age, December 20, 1995, p.2.

[4] Considerations for Development of Existing CATV Networks for Future Telecommunications Services; D. Gall & P. Brooks, Time Warner Communications; 1995 NCTA Technical Papers

LESSONS FOR THE INTERACTIVE RETURN SYSTEM

Dr. Kerry D. LaViolette
Philips Broadband Networks, Incorporated

Mr. Tom Staniec
The Excalibur Group, a Time Warner Company

Abstract

Activating and using the CATV return system as part of a profitable communications network has inspired tremendous work, discussion, and demonstration. This activity level tends to concentrate on known problems with subsystem components but often loses direction toward a complete system architecture. This paper presents laboratory results for a full system implementation as well as for subsystem components. Then data on impairments from real implementations is presented. Together these allow the larger system view while focusing on key areas for success.

INTRODUCTION

The interactive return system is the emerging technical challenge in CATV architecture. The conceptualization and implementation of a functional system requires a thorough understanding of the desired goals, the subsystem components, and the interaction between these components. In addition, the ability to overcome external impairments (like ingress) and the alignment/maintenance of the interactive network pose significant challenges.

This paper concentrates on the most prevalent modulation implementation in the return system. It is no surprise that Quadrature Phase Shift Keying (QPSK) data transmission is the modulation scheme of choice for most current implementations. This modulation scheme is extremely tolerant to impairments based on its low modulation order and use of FM techniques for data transmission [1]. Indeed, 20 T1-rate (1.544 Mbits/sec) carriers can easily fit within a 10-40 MHz bandsplit providing essentially a is a telephone line for every home on a node without any blocking! Therefore QPSK is not only robust, but also provides enough bandwidth for today's applications.

This paper begins by showing test results of return amplifiers as independent units while also introducing to the testing methods, equipment, and graphs that are used. The next section looks at return lasers, both Fabry-Perot (FP) and Distributed Feedback (DFB) type. Following this is a look at a cascade performance built with currently available equipment. The cascade not only shows the complete system performance, it also shows results over temperature.

The next two sections are devoted to practical impairments and draw on experience from actual field trials. Both sections focus on ingress with the first as a discussion of actual field experience. The next looks at potential mitigation strategies, followed with a conclusion on the lessons learned.

RETURN AMPLIFIER PERFORMANCE

Before showing the performance of a return amplifier, the testing methods need to be explored. Figure 1 shows the standard test set along with the second-order and third-order products that fall within the 19 MHz test signal. Twenty-one T1 QPSK modulators are used as inputs to the device under test. These modulators are spaced from 5 to 40 MHz at 1.75 MHz spacing, while the tested frequency is at 19 MHz. The frequency plan is graphically shown in figure 2. The level into the device under test (DUT) is adjusted by attenuators and (if needed) gain blocks, while the output is filtered and adjusted to keep the input to the QPSK demodulator within its AGC range. The system is fed by a Bit Error Rate Test set (BERT) at 1.544 Mbit/s with a 2^{23}-1 PsuedoNoise (PN) sequence. A PN sequence of bits simulates a random pattern, but eventually repeats. In this case the pattern repeats every 2^{23}-1 bits. The QPSK modulator/demodulator pair, at the 19 MHz test frequency, was specifically selected because it operates within 1 dB of theoretical performance. *Therefore, the device or system is being characterized rather than modulator/demodulator design.* In addition, any gain blocks used were also tested to make sure they were not impacting the performance of the DUT.

When performing this characterization, two factors are considered. The first is noise performance. At low input levels the DUT noise will produce errors, referred to as the noise side or left side of the curve. The second consideration is high input levels. At high input levels the distortions generated will fall under the tested carrier eventually rising enough to cause errors. For QPSK and items tested in this paper this behavior happens only after the DUT goes into compression or clips. This is referred to as compression or the right side of the curve. These two curves create a washtub graph that shows the input dynamic range that can be used for error-free performance. For this paper Bit Error Rates (BER) below 1×10^{-8} are considered error-free. This error-free dynamic range is critical because it sets the levels of operation that overcomes noise and ingress while maintaining margin for environmental and system related impairments.

This technique was applied to a typical return amplifier, and the results are shown in figure 3. Keep in mind that this graph is similar to following graphs so the descriptions, provided here, need only to be stated once. The X-axis of the graph is RF power (dBmV) per channel as presented to the device input. The device input is a valued reference when aligning the return path [2] and will be used for all the following graphs.

The washtub curve is clearly evident. The noise side has two different lines plotted. The first is with just one QPSK (the 19 MHz test channel) present, while the other has all 21 QPSK channels on. For this device there is no discernible difference between the two conditions and since it is a thermal noise limited device, this similarity is expected. At $1x10^{-8}$ the dynamic input range is nearly 68 dB. This is an extremely wide operating window; however, it will be narrowed by noise funneling in a cascade situation. This narrowing will be shown in the system test section.

Also, it should be noted again that errors are not observed on the right side curve until the device is in compression. The compression sets the absolute upper limit that the input of the amplifier can tolerate and along with any input padding determines the highest level available in the coax. This in turn sets the carrier to ingress level. Input padding is a recent approach to improve tolerance to ingress. Aligning the system for unity gain by padding at the input of the following amplifier allows the highest RF levels possible in the coaxial cable, thus maximizing the signal to ingress level.

RETURN LASER PERFORMANCE

This section shows the performance of typical lasers and/or transmitters that can be acquired for systems today. Discussion continues over which type of device, either an FP or DFB, should be used. The comparison that was made in [3] still applies. A DFB can inherently give better performance in terms of link length and overall operating widow. However, this performance is not free. Coolerless DFBs have been sampled to the market place and may narrow the cost differential. The choice, as always, is performance versus cost and can only be decided when one knows the system needs, the upgrade path, and the upgrade timetable. This section will show the relative performance of these devices and point out some practical considerations that have not been previously mentioned.

Figure 4 shows the performance of a typical FP transmitter with the same testing methodology mentioned in the previous section. The device was biased at 0.5 mW and transmitted over 6 dB of singlemode fiber. Although a return receiver was used, it plays no part in the performance and, therefore, will be regarded as transparent for QPSK data. However, the return receiver does have a role: its gain adjustment is used to maintain the proper level for operating in the demodulator's AGC range.

As we look at the figure we see a well defined operating window of 30 dB. The right side of the curve was determined by laser clipping while the left side of the curve is determined by laser noise. Also, note that again there is no discernible difference between single-channel and multiple-channel operation. This noise is sporadic in nature. Note that the slope of the curve is not steep as would be expected for white noise. This sporadic noise is caused by laser chaos when Rayleigh backscatter from the fiber plant is presented at the FP device[4].

Clearly, this transmitter performance is poorer than the single return amplifier. However, 30 dB is an extremely wide operating window and should be sufficient for practical systems, as will be shown in a following section. It should be noted for future reference that this transmitter design has only a minor shift over temperature for QPSK data.

Since this performance is lower than that for a return amplifier, methods to improve it were considered. One method was to ask a supplier to change the design parameters of the FP. It must be recalled that these devices were designed and optimized for ON/OFF digital performance and although this allows "technology piggybacking," it may be a suboptimal design for QPSK carriers. Figure 5 shows a redesigned device with two sets of curves. The first has an output of 0.5 mW. This device' 10 dB improvement over the device in figure 4 is significant and is on the noise side. The slopes of the curves are now extremely steep, so white noise performance has replaced sporadic noise performance.

A practical consideration is shown in the second set of curves. Another "improved" FP laser was biased at 2 mW and then tested over 6 dB of fiber. This device was rated to 4 mW, and the thought was to increase the right side of the curve by increasing the clipping point. Although this step was accomplished, the noise side of the curve was extremely degraded. The dynamic input window on this device is 30 dB, 10 dB less than the 0.5 mW improved FP! The additional current used to bias the device 6 dB higher has allowed additional laser modes. These additional modes increase the laser mode partition noise, as they compete for gain, causing the degradation. This outcome demonstrates the importance of defining and characterizing both sides of the washtub curve.

A return DFB is shown in figure 6. This device was biased at 2 mW and is both cooled and isolated. The transmission distance was 6 dB. The device performs as expected with an operating window of 45 dB, which is 15 dB better than the typical FP in figure 4 but only 5 dB better than the improved FP, biased at 0.5 mW, in figure 5. The noise curve is white noise limited, and there is no difference between a single carrier and a multiple carrier.

A practical matter is illustrated by the DFB in figure 7. This device was operated over 11 dB of fiber and performs as expected with regard to clipping. However, the noise side shows an interesting dichotomy. There is a large difference between single-carrier performance and multiple-carrier performance. In fact multiple-carrier performance is 15 dB better! The supposition is that the device has a quieter noise level when it is operated under a RF load; multiple carrier operation is, therefore, better. However, in practical systems the dynamic input range would be limited to 30 dB which is the difference between the single-carrier noise performance and clipping. This reasoning assumes that the system allows dynamic carrier allocation and, therefore, there will be times when only a single carrier is operating. This can be mitigated by forcing some carriers to always be operational; however, these carriers take away useful information capacity. Again, we see the importance of defining both sides of the washtub curve and not assuming what is a worst-case scenario.

CASCADE RESULTS

Using the return amplifiers and standard FP transmitter from the previous sections, a cascade was built. The cascade, shown in figure 8, has 5 amplifiers in cascade with additional noise added to simulate 50 amplifiers in the node. The return RF cascade was aligned for unity gain and a flat response. Its characteristic curve, at room temperature, is shown in figure 9. This figure shows that the cascade's dynamic input window is 50 dB, which is 18 dB less than a single return amplifier, with the change due almost entirely to noise funneling.

Since the input window is still larger than that of the FP transmitter and when compared to figure 4 has different RF levels for compression, alignment for unity gain from the last return amplifier to the return transmitter did not make sense. Therefore, alignment was made so that the compression levels were the same. This was facilitated by padding the input to the return laser transmitter. With this alignment the entire transmitter dynamic range is available while using the upper portion of the cascade. **Therefore, the signal-to-ingress level is maximized in the RF section where ingress occurs.** Due to the very nature of fiber transmission, no additional ingress occurs in the glass, so the carrier-to-ingress level is set before the optical transmitter.

The performance of the cascade at room temperature is shown in figure 10 and is referenced to the input of the first return amplifier. Due to the alignment described in the previous paragraph, the full system performs with a 30 dB window. The dynamic input range's width, as expected, is the same as the FP transmitter and is still extremely wide, providing room for system operation. The system was cycled over temperatures from -35 °C to +55 °C.

The resultant worst-case performance is shown in figure 11. The dynamic input range is reduced to 20 dB. An investigation into this reduction showed that the laser transmitter performance was nearly stationary over the entire temperature range and contributed very little. However, the changes in gain and cable length in the RF section were significant. The right side of the curve is set at low temperature, where gain is the highest and therefore the laser clipping point is reached more quickly when referenced to the cascade input. The right side of the curve is set at high temperature, where the additional cable lowers the signal level at the transmitter, bringing it closer to the noise.

In summary, the cascade has a dynamic input window of 20 dB over temperature. This window is still very large, and system operation can be robust. Work in thermal gain control, raising the RF amplifiers compression point, and widening the return transmitter's input range will make the system even more robust.

DISCUSSION OF INGRESS

Ingress, the unwanted by-product of poor shielding integrity in coaxial networks, poses a problem that has to be reconciled before a network can function as a reliable communications platform. This is really a very simplistic statement about a fairly complex problem. Fifteen years of history return networks proves that you can have an operational network that will support telecommunications. However, what is required are a sound view of the problem's principally origin and a steadfast attitude on how to eliminate, mitigate or diminish its impact on the network.

First, how does ingress get into a network? By empirical observation, 70% of all ingress comes directly out of subscribers' homes. That number could be higher, or possibly lower (doubtful but possible), as indicated in [5]. Further, the drop cable from the tap port to the ground block at the house contributes another 25% to the problem. An interesting study, done by CableLabs originally for digital network quantification of high-order modulation schemes, dealt with RF drop shielding as measured in the 88 to 108 MHz FM band. A number of conclusions were drawn from study, but the most significant is up to 60% of the drops hanging in the air today have shielding effectiveness of less than 50 dB! This is a statistic with significant implications. It points to signal leakage in the return portion of the plant where previously problems have not been seen. The remaining 5% of ingress problems are in the physical hard coaxial plant and come from critters, craft and catastrophes.

Examining the home and the subscriber link, one issue is immediately apparent. By Federal Communication Commission rule, the MSO has no right to deal with the internal house wiring once it has crossed the wall boundary of home unless it causes a leakage problem. That mandate puts the operator in a difficult position: how to deal with a problem coming out of the subscribers home but which does not affect forward network operation, cause a Cumulative Leakage Index (CLI) problem, or impact anything the customer maybe doing internally to the home. If the customer has a paid up and current bill, the operator can be and *has been* denied access and be left without recourse.

More recent work by CableLabs shows that RF shielding is a major source of high-level impulsive strikes in a network. Figure 12 is an extreme example of this problem. The spectrum analyzer print shows the effect that a thermostat, used in an apartment complex's electric baseboard heating, has on a return network. Obviously the picture is not a pretty one to start with and is compounded by the intermittent nature of the problem.

The above deals with problems internal to the home, but the greatest problem in the network is the subscriber. Contemplate how many times your system personnel have had to troubleshoot direct pick-up problems, poor video performance and numerous other subscriber complaints, which subscribers caused themselves on there own. A sobering thought considering this is an obvious chink in your network armor.

The drop cable from the tap to the home suffers from many of the same problems, but the subscriber is not the culprit. The crafts person, cable age and weather are the greatest threats in this area. These should be manageable because a bad drop can be repaired; unfortunately, it is less often managed than it is found by accident. CLI and leakage ride outs are touted as the equalizer. However, a large percentage of systems never get ridden because they are off easement. Thus, problems show up as a "blip" on leakage test equipment screen. Roughly speaking, the drop wire from the pole throughout the house covers as many miles as your hard coaxial plant. In a very large number of cases this "plant" never gets measured.

Additionally, fly-overs are promoted as a way to feel secure. Contemplate the following: a recent "blip" was closely investigated and a leak of 1000 μV/m was found about 400 feet off the road easement. It was caused by a maintenance man "making" a five-way splitter to serve new television sets. During the same time, a fly-over took place and no leak was registered. Why? The building had a steel roof and structure. The ingress caused by the "splitter" was causing a problem in the return network. This situation took place in a system which operators believe is a tight network because of CLI and a their fly-over results.

A closer look at ingress in the forward system plant shows that signal leakage, in the air navigational portion of the band, is in large part created by short wavelength slot aperture antennas. Return ingress is made up of a long wire antenna that, particularly in the low end of the return band, acts like AM radio. It is well known that attaching a piece of wire from an AM receiver to a cold wire pipe increases your reception capability. The same thing happens with the shield on coaxial cable and house electrical wiring. This implies that you are dealing with over- the-air radiated signals and ground-wave radiated signals. A look at the construction of an AM radio station or a short-wave broadcaster such as Voice of America further illustrates the situation. The result is an opportunistic signal that gets into the system by poor shielding characteristics. By interpolation, 95% of all problems come from the coaxial cable from the tap down through the home where the shielding integrity is the worst and where CLI is not performed with any consistency. Thus, it is no mystery that a CLI that can meet or exceed specification yet not give you a clean return system. This is a simplified view of the nature of ingress but one that is accurate.

PRACTICAL SYSTEM OPERATION

Fixing the ingress problem and setting up the network are directly related to each other. If the network is set up improperly, it may not be operating as far above the noise floor and ingress as it could. What options are available in the network to overcome noise and ingress? There are three possible methods. First, talk louder than the ingress giving yourself headroom above the problem. Second, repair the ingress to knock it down, thereby increasing the headroom. The third choice is to use a combination of the previous two. All of the above assume a modulation scheme that has a good chance of surviving in the return environment.

As discussed in the previous section, the nature of ingress largely centers on poor shielding. Objectively, there are two ways to deal with the problem: filter it or fix it. Filtering is the single most effective method to clean up and control the problem. Fixing the problem does not ensure it is gone but only momentarily mitigates it. As can be seen, using this approach never takes control of the network. No control leaves the network continuously vulnerable to whatever happens behind the closed doors of the subscribers' homes.

Utilizing the filter approach in number two or three above has the effect of either killing ingress in total or controlling (the lossy-filter approach) the total level allowed into the return network. The best possible position for either type of filtering is at the tap port. Positioning the filter there removes the need to repair all ingress problems in a node caused by the drop system through to the house. Interestingly, while it is often stated that the return network there is 35 MHz of bandwidth, without some type of ingress mitigation strategy, a system may have only 20 MHz or less that is functional. That significantly impacts the total number of services that the network can provide.

What can be done to increase the bandwidth and provide a quality of service that will provide long-term success? The best method is a combination of approaches. Fundamentally, it must be recognized that network must be owned by the provider and protected from all interlopers who are either ill-informed, accidental or malicious. Second, control of ingress is paramount and works in concert with the above strategy. Once the above are realized, then operational methods can be devised. Primarily, the return network should be set up to reference flat inputs to the return amplifiers and at the highest level possible without operating in compression. By input-referencing the return amplifiers, as opposed to the forward amplifier output-referencing, maximum carrier-to-noise and distortion can be achieved. By properly selecting a modulation scheme, a network can be established that is capable of providing high-performance telecommunications.

Logic indicates that increasing the output-level capability of the communications device in or on the home will provide a way to build more headroom and margin into the network. While desirable, that solution opens the door to the problem in drops with less than 50 dB RF shielding. In short, a signal leakage problem will arise in a frequency band where it was never experienced in the past. A recent letter[6] points out the shielding problem, only in reverse. With amateur radio transceivers having a 0.5 μV/m receiver sensitivity and high transmitter output power, a significant problem can materialize both from ingress and egress in the return network. The proper operation of the network requires control, maximized input reference levels, minimized ingress levels and resilient RF modulation schemes.

CONCLUSIONS

What lessons can be gleaned from this discussion? First, today's subsystem components, when configured in a typical system, allow a wide operating window for return system performance, this includes FP transmitters as well as current return amplifiers. Current work in thermal control, amplifier-compression point, transmitter design (both house modulator and laser) will only strengthen this system. Second, in order to operate a network, the provider must own the network. This ownership responsibility must then include must be taken by a combination of "protective" strategies, such as filtering, maintenance and the provision of gateway access. Indeed, a fundamental shift in system design is required for network operation. The ground-up strategy, that has been so successful, for broadcast video must be changed to a network-down strategy, thus concentrating on the most critical issues without allowing subsystem "solutions" to adversely affect network operation.

REFERENCES

[1] LaViolette, et. al., "System and Design Considerations for the Cable Return System," Proceedings of the NCTA, May 1994.

[2] Staniec, "Return Path: Perspective from Field Trial Experience," Cablelabs Seminar, HFC Return Path: Paving the Way, Dallas Feb. 1996.

[3] LaViolette, "Fiber by Design," Proceedings of the NCTA, May 1995.

[4] LaViolette, "The Impact of Rayleigh Backscatter Induced Noise on QPSK Transmission with Fabry-Perot Lasers," IEEE Photonic Technology Letters, May 1996.

[5] Cable Television Laboratories (CableLabs), "Return Path Characterization."

[6] QST Magazine, March 1996

Test Diagram

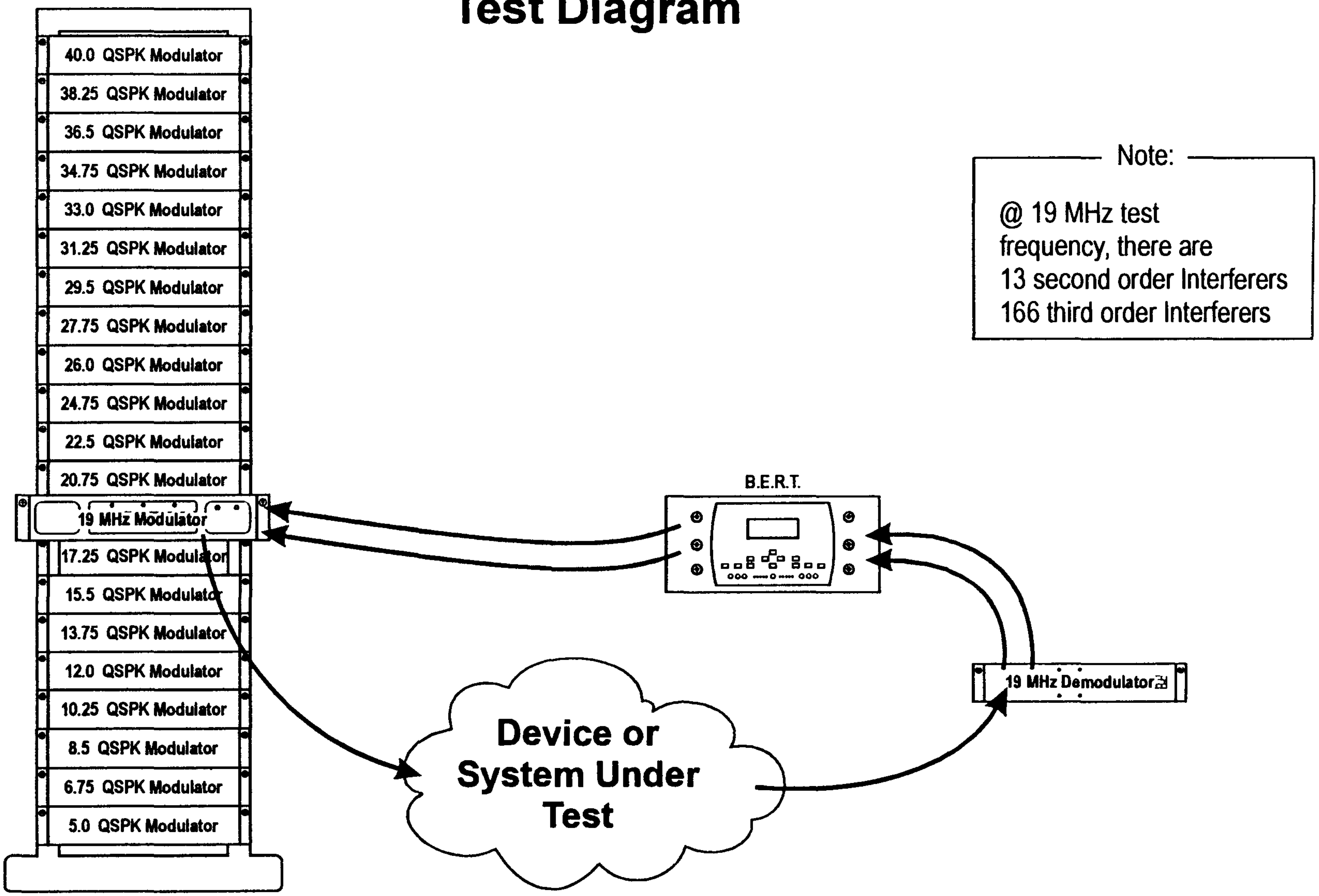


Figure 1

Frequency Plan

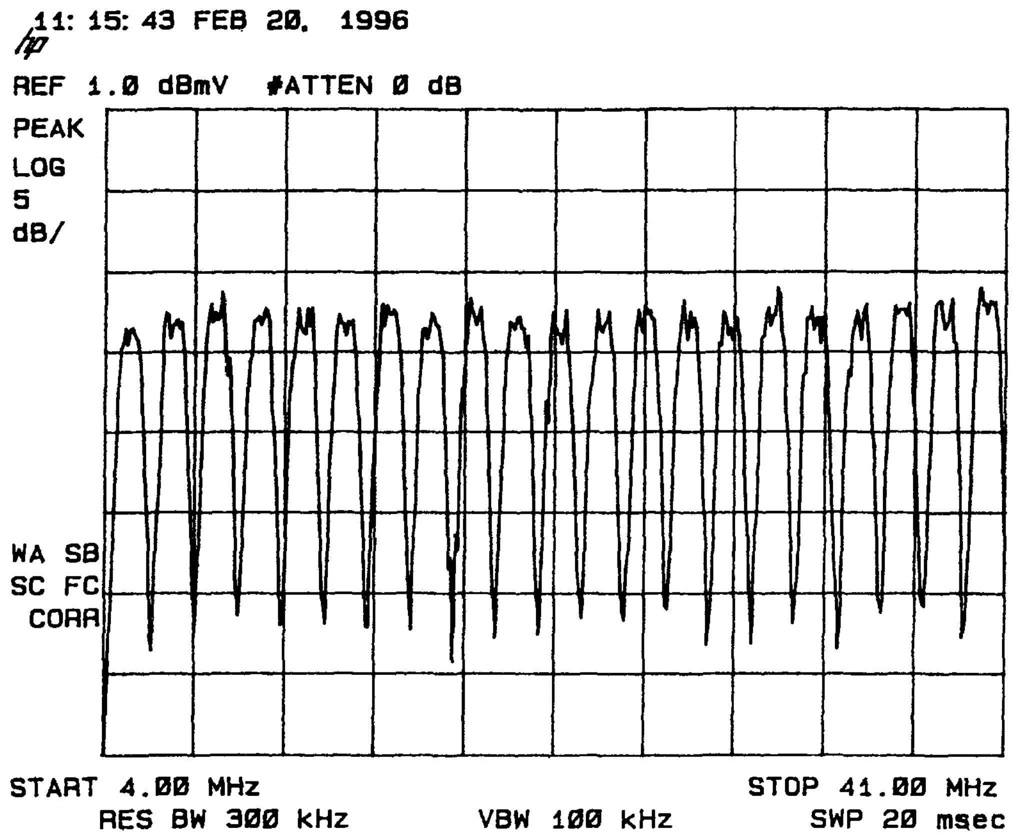


Figure 2

RETURN AMPLIFER TYPICAL PERFORMANCE

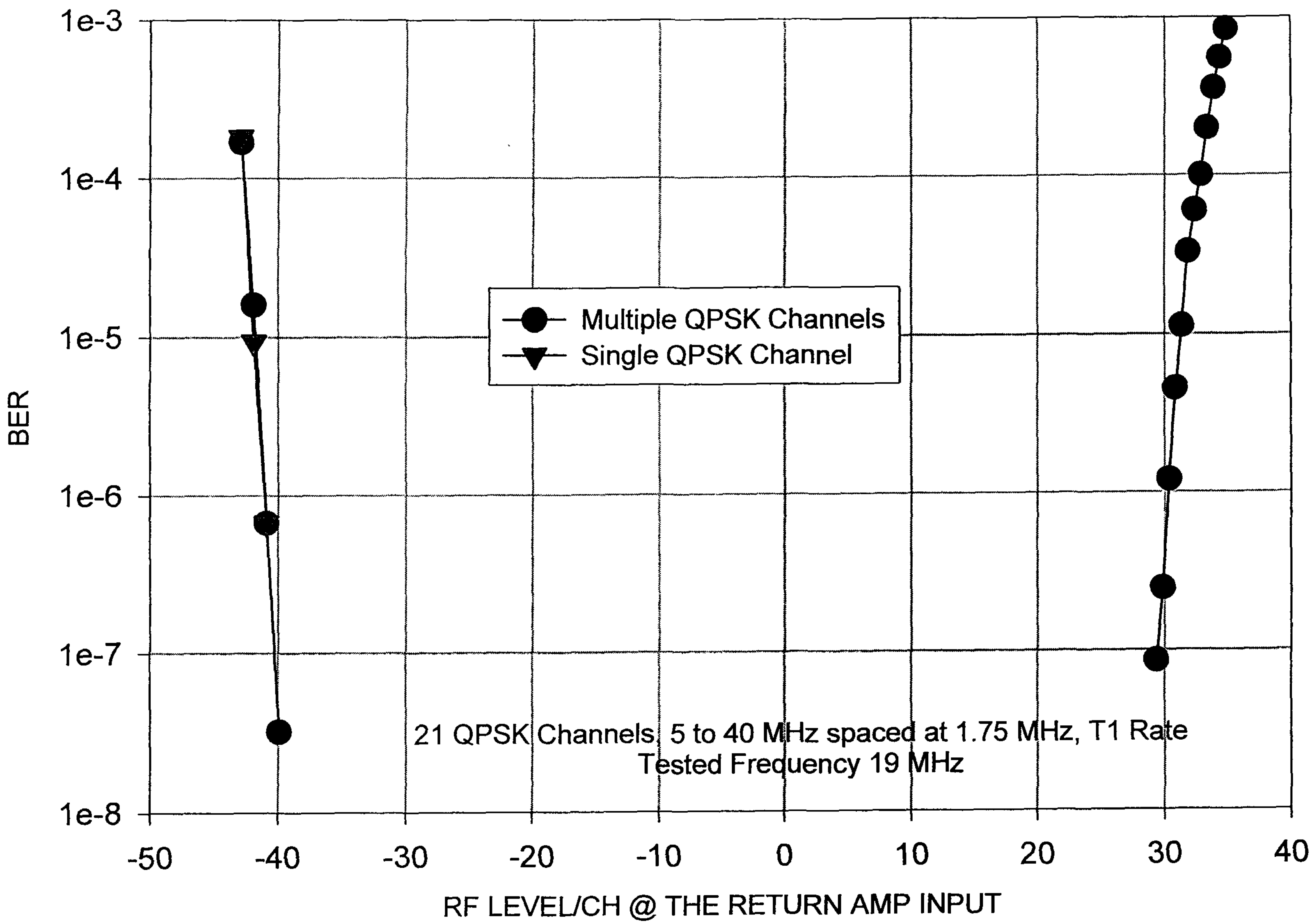


FIGURE 3

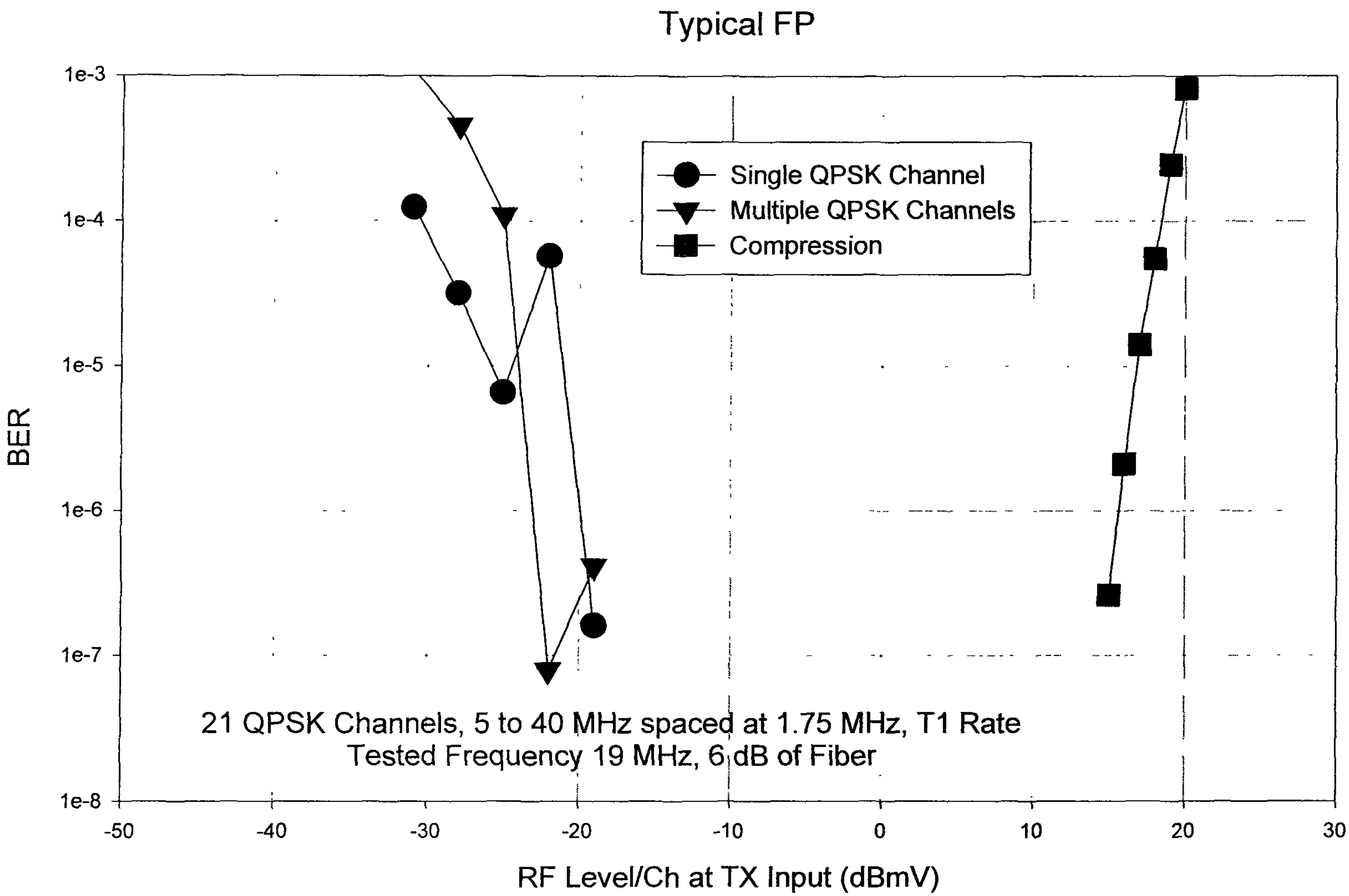

Typical FP
1e-3
1e-4
1e-5
1e-6
1e-7
1e-8
BER
Single QPSK Channel
Multiple QPSK Channels
Compression
21 QPSK Channels, 5 to 40 MHz spaced at 1.75 MHz, T1 Rate
Tested Frequency 19 MHz, 6 dB of Fiber
-50
-40
-30
-20
-10
0
10
20
30
RF Level/Ch at TX Input (dBmV)

FIGURE 4

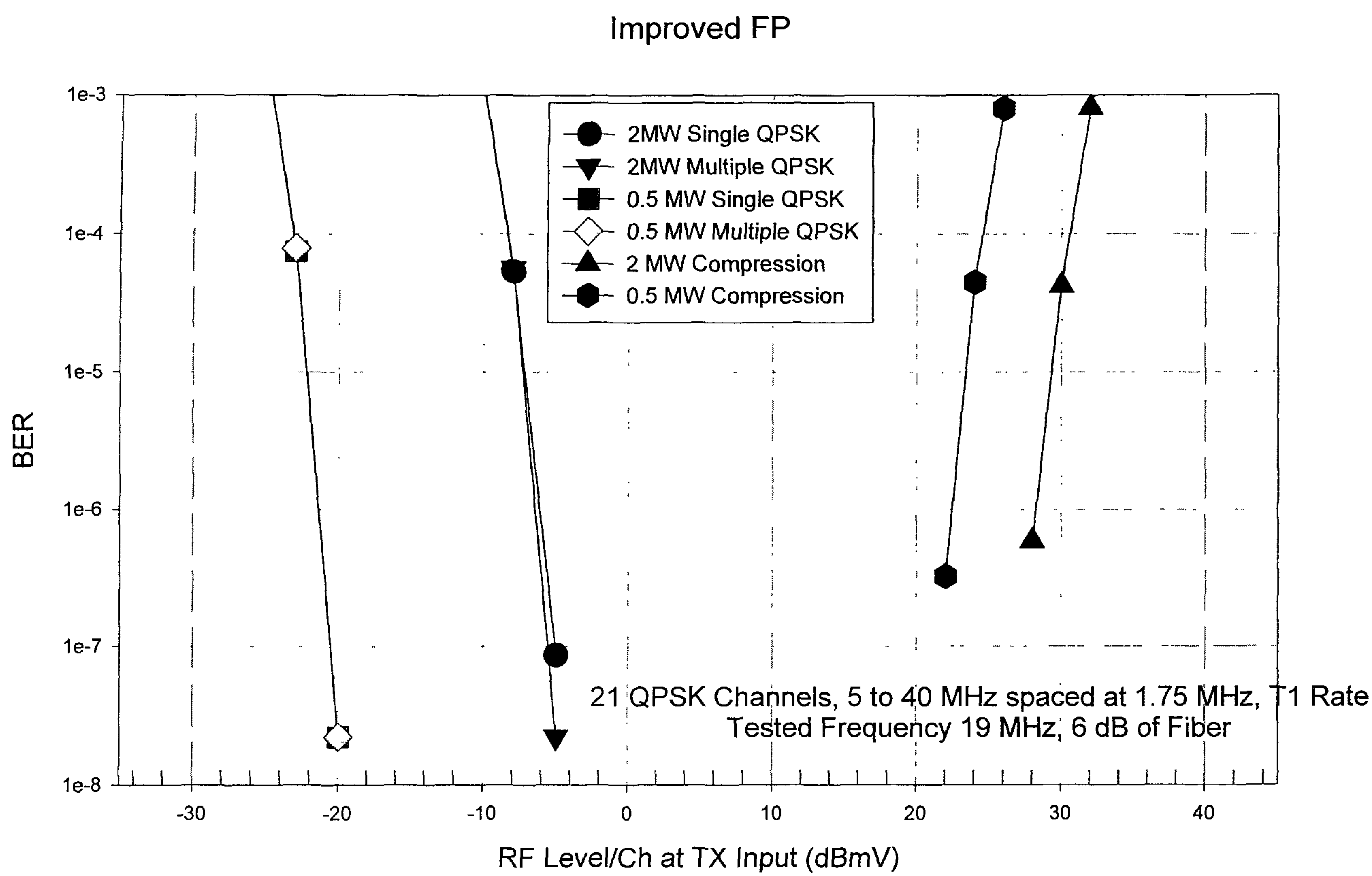

Improved FP
1e-3
1e-4
1e-5
1e-6
1e-7
1e-8
BER
-30
-20
-10
0
10
20
30
40
RF Level/Ch at TX Input (dBmV)
2MW Single QPSK
2MW Multiple QPSK
0.5 MW Single QPSK
0.5 MW Multiple QPSK
2 MW Compression
0.5 MW Compression
21 QPSK Channels, 5 to 40 MHz spaced at 1.75 MHz, T1 Rate
Tested Frequency 19 MHz, 6 dB of Fiber

FIGURE 5

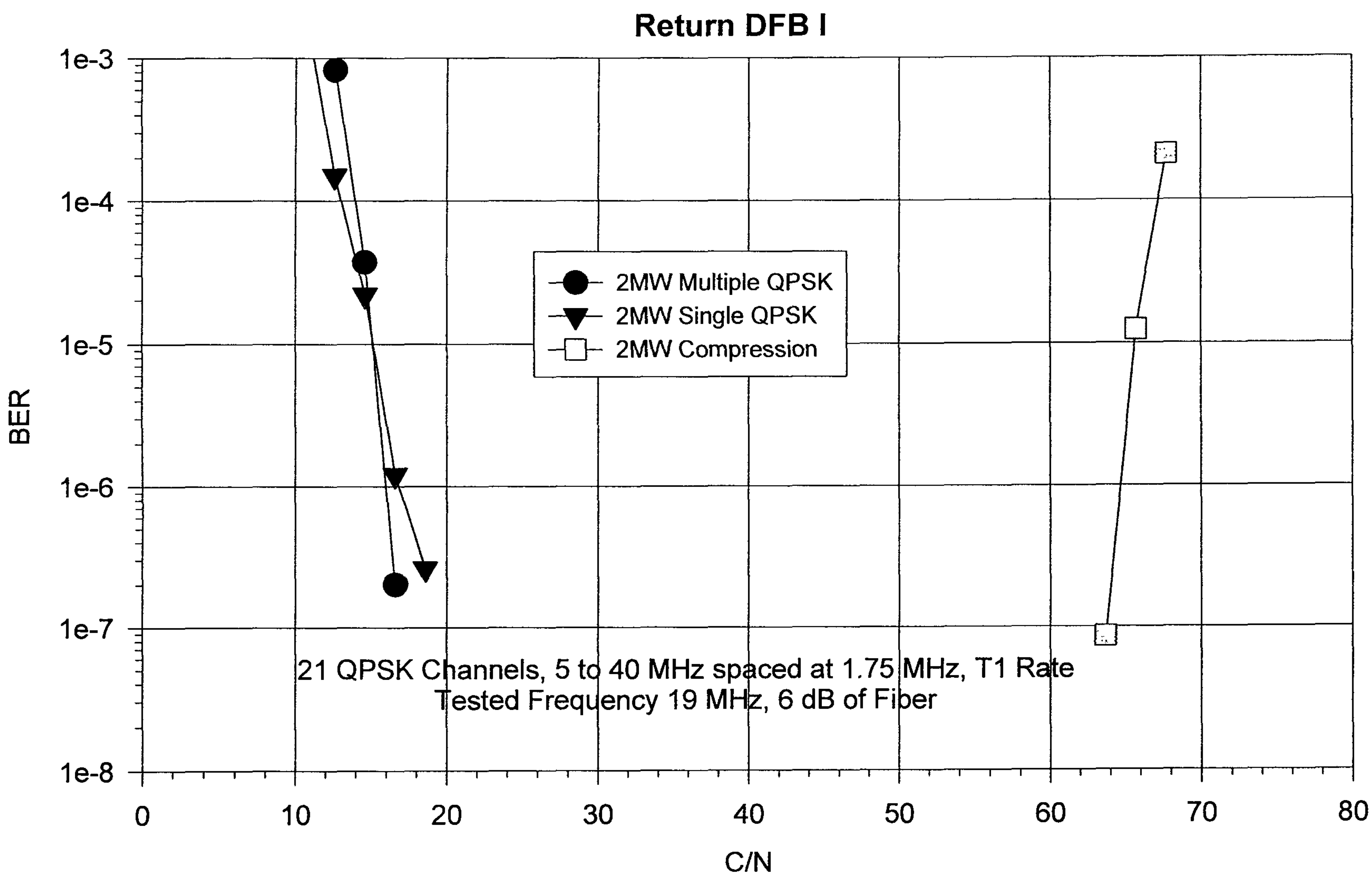

Return DFB I
1e-3
1e-4
1e-5
1e-6
1e-7
1e-8
BER
2MW Multiple QPSK
2MW Single QPSK
2MW Compression
21 QPSK Channels, 5 to 40 MHz spaced at 1.75 MHz, T1 Rate
Tested Frequency 19 MHz, 6 dB of Fiber
0
10
20
30
40
50
60
70
80
C/N

FIGURE 6

RETURN DFB II

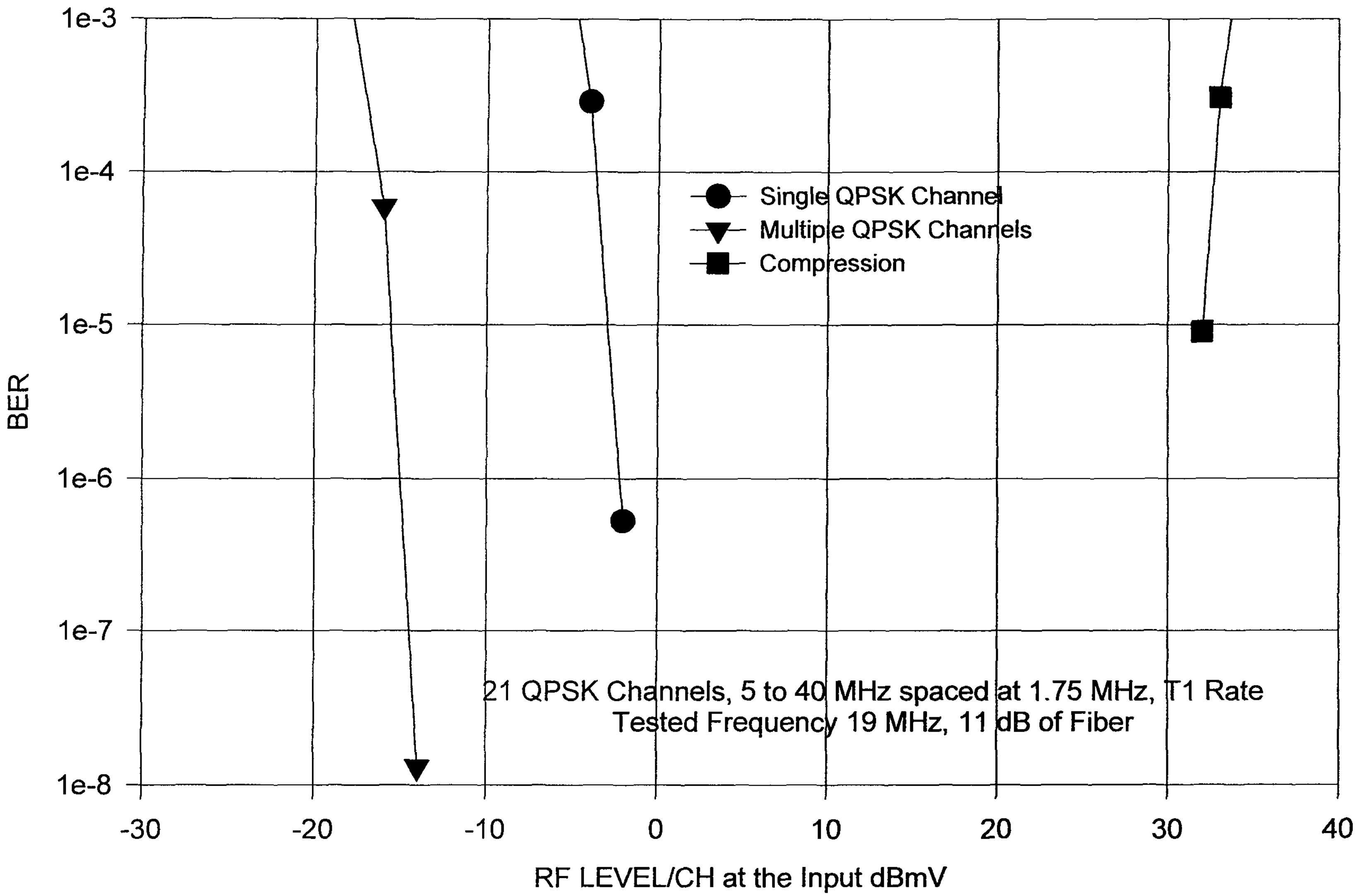


FIGURE 7

System Diagram

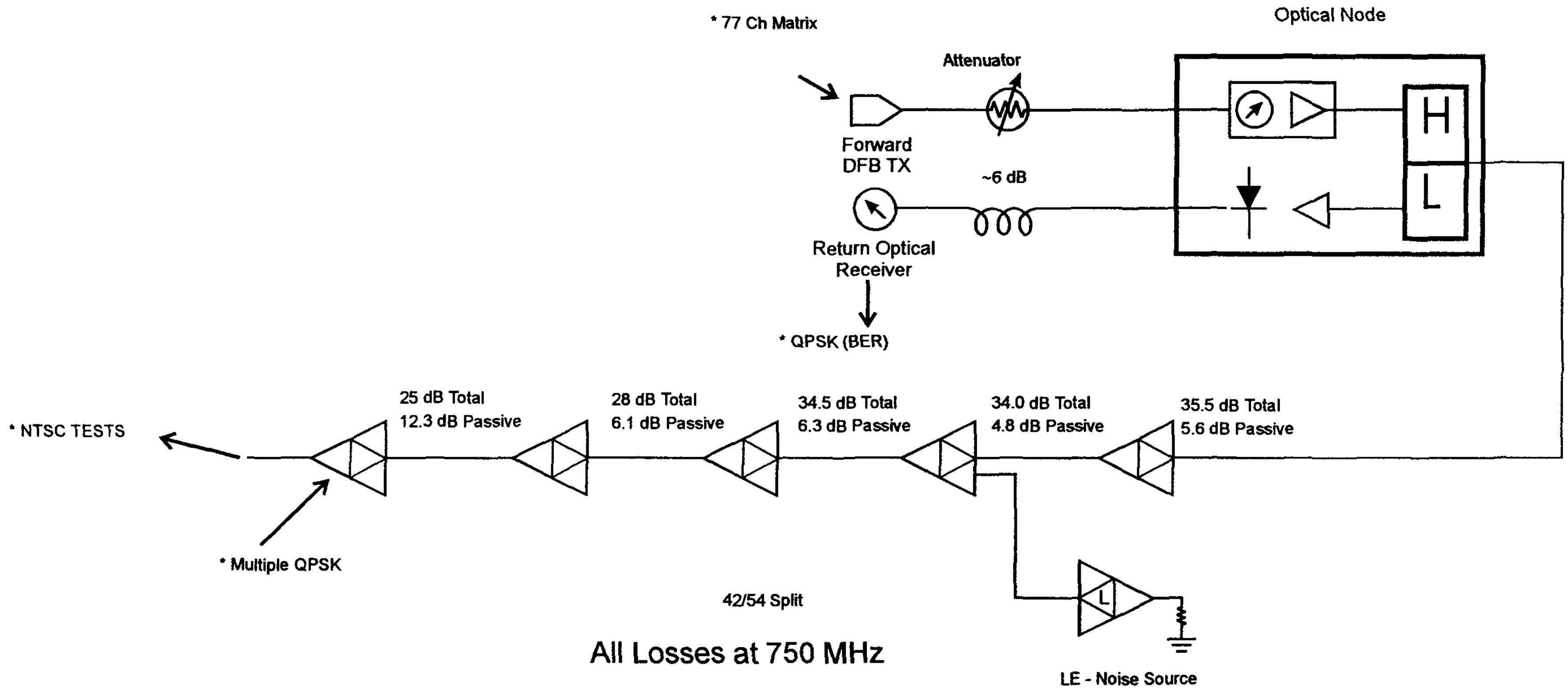


Figure 8

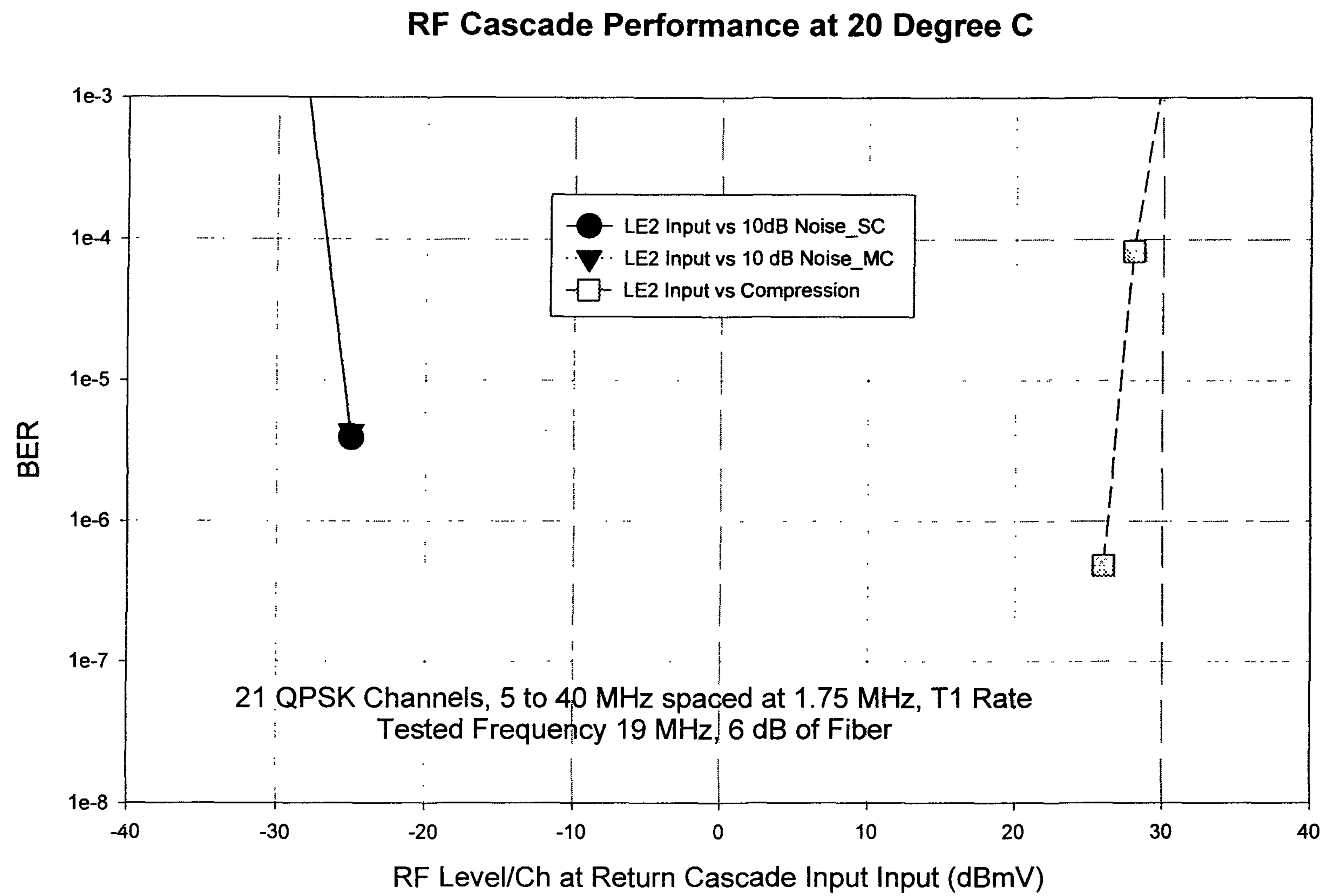
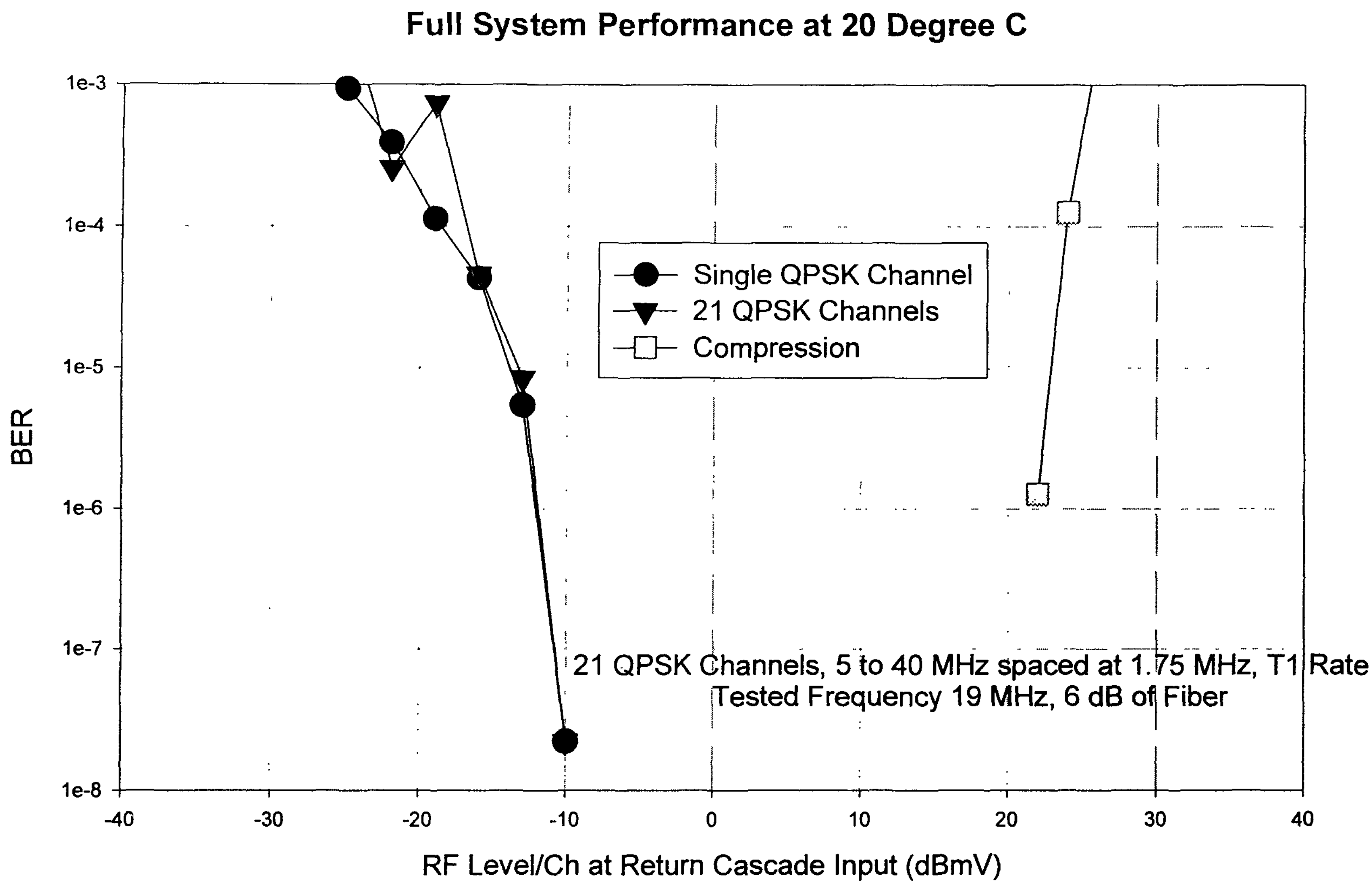

RF Cascade Performance at 20 Degree C
1e-3
1e-4
1e-5
1e-6
1e-7
1e-8
BER
LE2 Input vs 10dB Noise_SC
LE2 Input vs 10 dB Noise_MC
LE2 Input vs Compression
21 QPSK Channels, 5 to 40 MHz spaced at 1.75 MHz, T1 Rate
Tested Frequency 19 MHz, 6 dB of Fiber
-40
-30
-20
-10
0
10
20
30
40
RF Level/Ch at Return Cascade Input Input (dBmV)

FIGURE 9



Full System Performance at 20 Degree C
BER
1e-3
1e-4
1e-5
1e-6
1e-7
1e-8
Single QPSK Channel
21 QPSK Channels
Compression
21 QPSK Channels, 5 to 40 MHz spaced at 1.75 MHz, T1 Rate
Tested Frequency 19 MHz, 6 dB of Fiber
-40
-30
-20
-10
0
10
20
30
40
RF Level/Ch at Return Cascade Input (dBmV)

FIGURE 10

Full System Performance Over Temperature

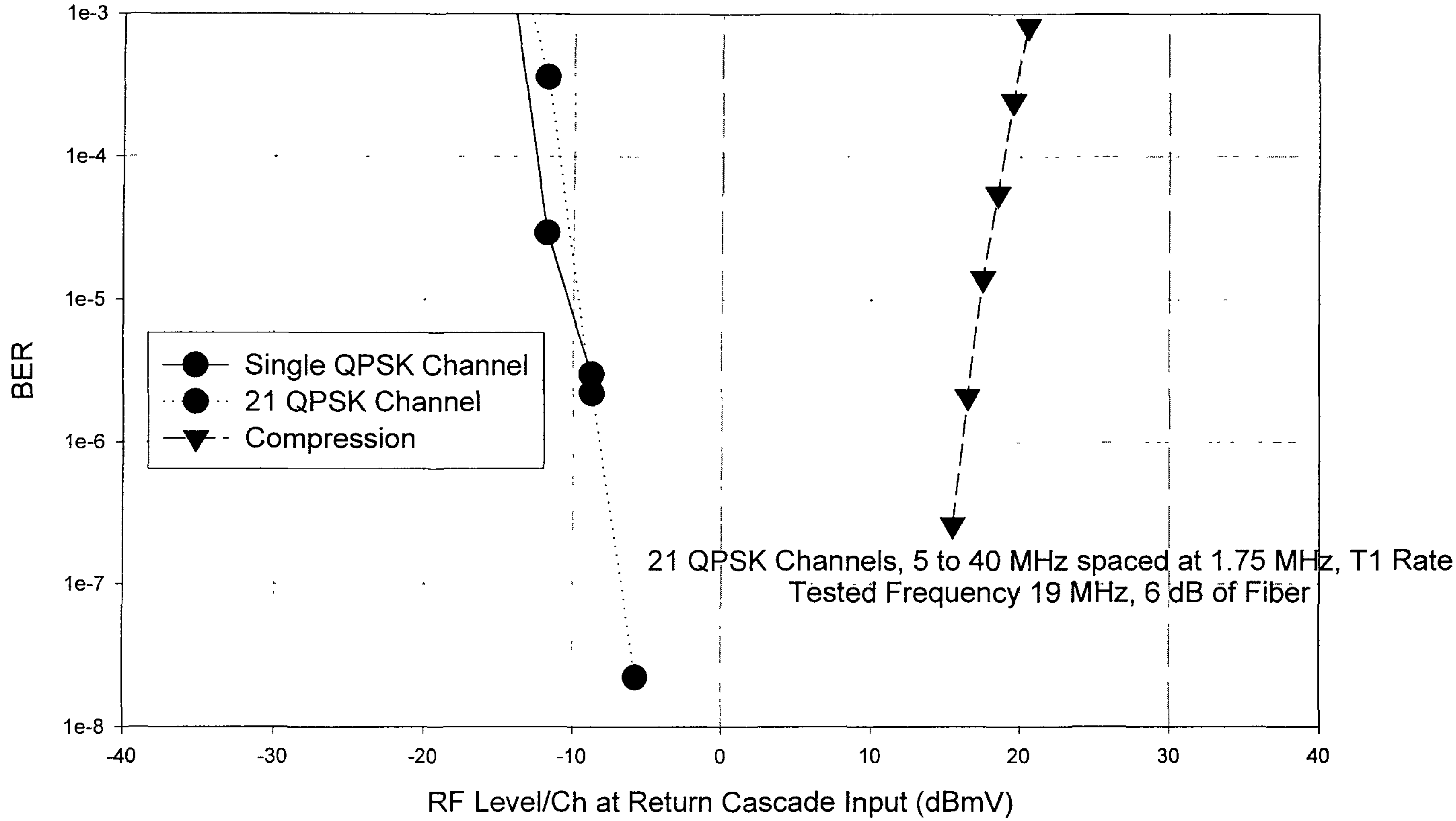


FIGURE 11

Ingressing Event

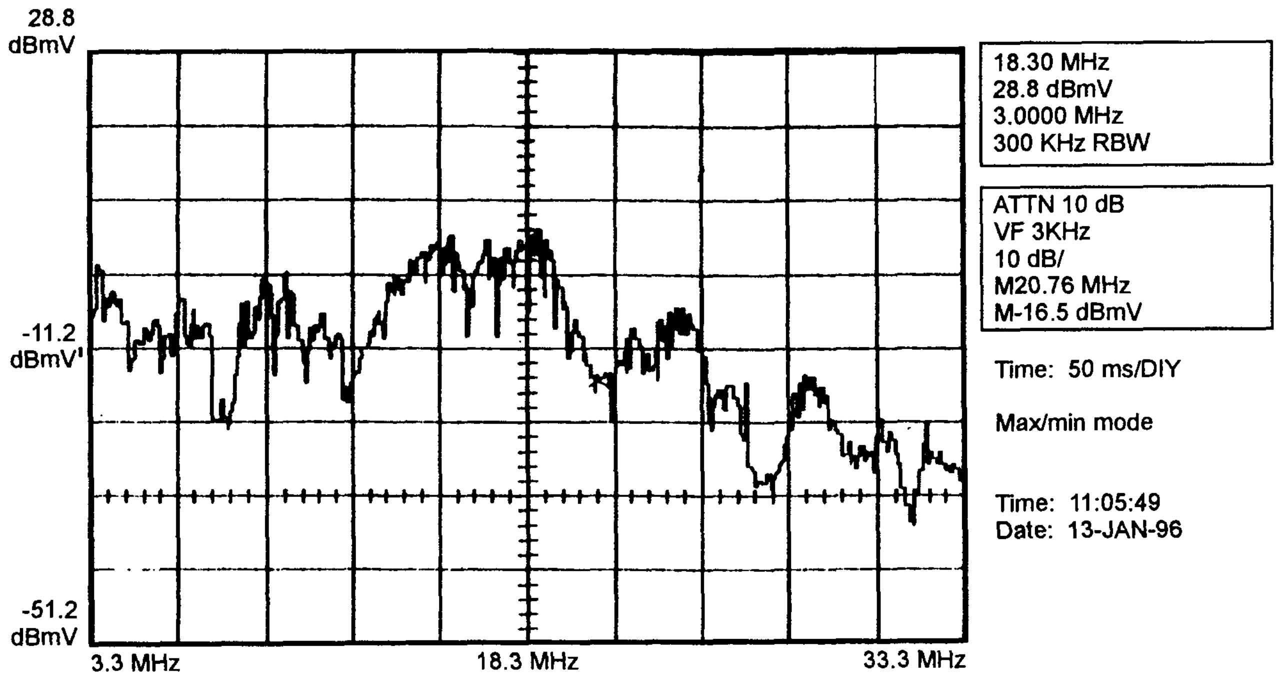


Figure 12

MANAGING THE HIGH-SPEED CATV DATA NETWORK

Esteban Sandino, P. Eng.
Rogers Cablesystems Ltd.

Abstract

Setting up and running high-speed data services over cable will require changes in the way services are currently provisioned and supported over a CATV network. Such changes will involve evaluating the way different groups within the cable organization interact with each other, as well as generating efficient service provisioning and problem resolution procedures. The overall goal will be to maintain the highest level of service availability through efficient service support and fast resolution of any network related problems. The building of the required expertise to provide and support the next generation of high-speed data services may involve the creation of new technical groups, and also the modification of existing operational procedures to accommodate needs specific to the new products. This paper describes the work Rogers Cablesystems has done to date to achieve these objectives.

INTRODUCTION

The support of high speed data services over two-way cable plant requires a cable operator to commit to offering the highest levels of quality and reliability. Potential users of high-speed data services will have a very low tolerance to service outages. Even a few seconds of service interruption during high-speed data delivery, which may not be considered serious by a typical cable TV customer, can induce unacceptably high loss rates during data transfers. High loss rates in turn can lead to unacceptable delays in the delivery of desired information as perceived by the end data customer.

This paper will discuss the service goals Rogers has initially defined as the minimum requirements that must be met to properly support high-speed data services. The Rogers™ WAVE™ service recently launched in the Newmarket, Ontario licensed cable area is the first one of such services. Next, the roles of the different groups within the Rogers cable organization in supporting this new service will be described. Lastly, the different processes that have been defined for service provisioning, problem resolution, and change control will be explained. Finally, next steps planned to improve the overall process will be discussed.

DEFINING SERVICE GOALS

Achievable targets for the desired level of service were defined from the outset. For Rogers™ WAVE™, these targets are as follows:

- Resolution of all cable outages within 6 hours of being reported. This applies for outages in both the forward and reverse directions.
- No reported field problems developing into outages.
- No detected data routing and modem response problems developing into customer calls.
- Achieve a 99.90% initial service

availability the first year after service introduction, or a maximum 8.75 hours of outage time per user per year as per the Canadian Cable Television Association (CCTA) service quality guidelines, with the eventual goal of reaching a 99.99% availability level.

These targets will evolve as the level of technical expertise increases, and all groups involved better understand their role in maintaining the desired levels of service quality and reliability.

ORGANIZING THE HIGH-SPEED SERVICE SUPPORT GROUPS

Service support is provided by a number of groups within the Rogers cable organization. Some of those groups have been recently created to specifically support Rogers™ WAVE™ services, and work closely with traditional cable TV personnel.

Figure 1 illustrates the various groups within Rogers and the division of responsibilities among them: the Technical Action Center (TAC), the data network group, and field technical support groups. Not shown in this picture, but still very much an integral part of the organization, is the Technical Service Representative (TSR) group. More specific descriptions for each of these groups are provided in subsequent sections of this paper.

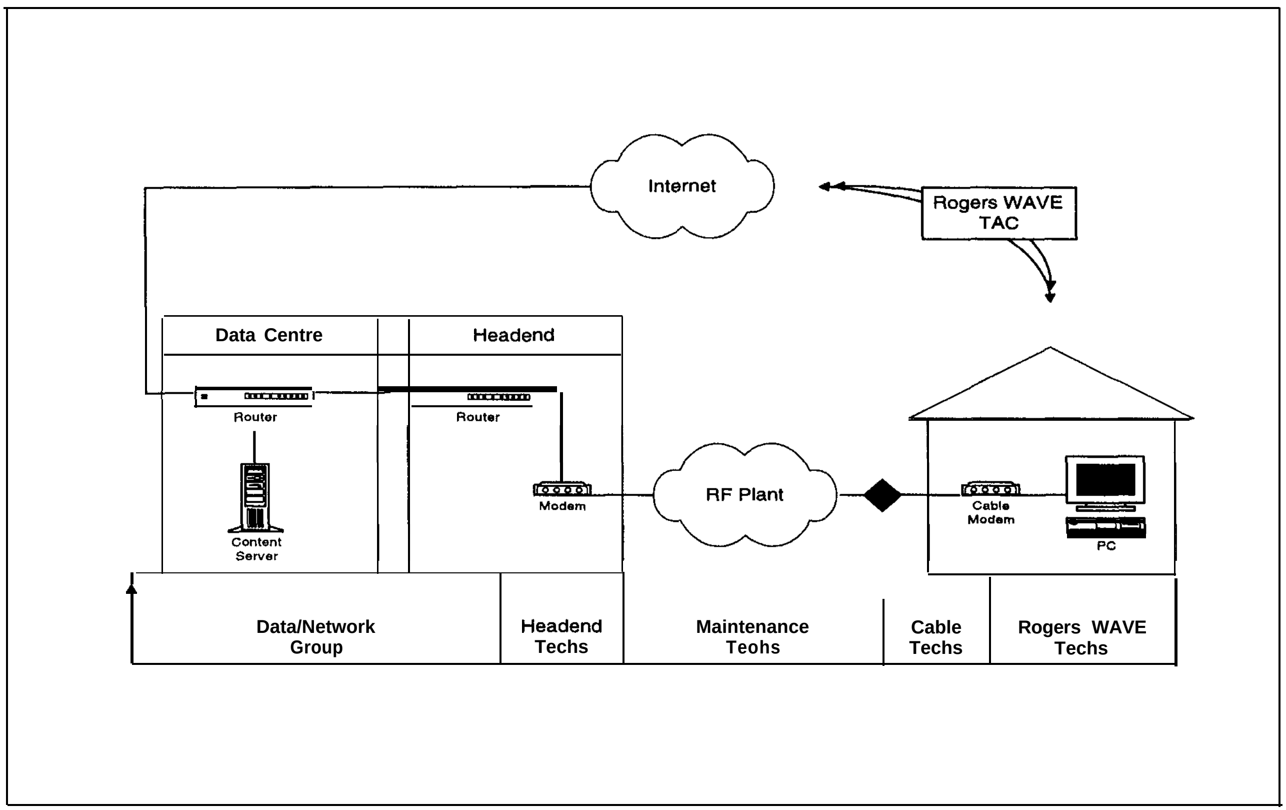


Figure 1. Rogers™ WAVE™ Service Support Groups

Technical Service Representatives (TSR)

The TSR is the first point of contact for all Rogers™ WAVE™ customers. Customers will contact a TSR for one of the following reasons:

- A general inquiry
- A Rogers™ WAVE™ equipment problem
- A software-related problem
- A third-party problem, e.g.: failure to connect to an on-line service provider
- Request for new service connection

TSRs are the focal point of contact between customers and the other technical groups. TSRs trained for the Rogers™ WAVE™ service have built an extensive knowledge base that enables them to quickly address the most commonly encountered hardware and software configuration problems. In addition, it is their responsibility to collect as much information as possible on reported problems they are unable to solve over the phone prior to escalating to the second level of customer support.

Technical Action Center (TAC)

The TAC primary functions can be summarized as follows:

- Continuous monitoring of both the cable and data portions of the Rogers™ WAVE™ network for outages
- Continuous monitoring of the of Rogers™ WAVE™ data connections. Information is derived from the round trip propagation time of test data packets sent to all data host terminals. The percentage of successfully returned packets and their average return trip time are captured in a log and plotted on a regular basis. The obtained plots provide a trend of the overall functioning of the Rogers™ WAVE™ connections, and the percentage of time when hosts have been unreachable
- Acting as second level support for the TSRs, and taking ownership of reported problems after escalation from other groups
- Isolating the cause of an outage, or any other reported problem, and escalating to the appropriate technical group responsible for its resolution, i.e.: On-call data analysts, cable headend supervisors, cable maintenance and construction, or other
- Tracking, in coordination with those cable divisions where Rogers™ WAVE™ services are offered, of all work done that can potentially affect the integrity of the cable and data networks supporting WAVE,.. This includes upgrades, physical network re-configuration, planned outages, etc.
- Updating appropriate logs to reflect changes in the Rogers™ WAVE™ network
- Maintaining an updated problem log of all pending trouble calls and their resolution

- Service provisioning support role in terms of overseeing the mapping of potential customers to the physical cable plant, and to the logical data network

- Maintaining an on-call schedule during off-hours of operation

To help TAC carry out these tasks, they have been given access to a number of resources:

Rogers Integrated Network Management System (INMS)

The INMS is the primary tool used for the remote monitoring and control of the cable distribution network, including RF coaxial distribution amplifiers. For each distribution trunk station in the Rogers cable plant, the INMS provides two important functions:

- Control over reverse bridger switching and thus control over which service areas can feed signal back to the main cable headend. This is extremely useful in isolating problem feeder areas from the rest of the return system

- Control over the use of a 6-dB attenuation pad to help find the approximate location of reverse noise sources on trunk and feeder in conjunction with bridger switches

The INMS is augmented by the availability at the TAC of a remote video feed from a headend spectrum analyzer, which allows constant monitoring of reverse noise levels for any service areas supporting Rogers™ WAVE™ services.

SunNet Manager

Currently, monitoring and control of all SNMP devices, including the current generation of Zenith cable modems, are performed through a SunNet manager station. The main function of this station is to perform polling of all network devices. Device availability status is constantly updated based on the polled information.

All devices monitored by the SunNet manager are grouped by level of functionality on a network map to allow quick visual alerting when unreachable network devices are detected and alarms are generated.

Plans are under consideration to evolve either the existing SunNet management platform, or an alternate one, to allow full access to control functions as well as monitoring of all SNMP manageable devices in the network. Extensive work is still required to reach the level of a truly intelligent system that not only allows collection of network statistics, but can also perform continuous analysis and generation of appropriate traffic analysis and other network performance reports.

Billing and Trouble Call Tracking System Terminals

Customer billing and service information is stored in Supersystem which is the billing system developed by Rogers for CATV services. The TAC can quickly access the above information through Supersystem terminals, and can also look at notification screens providing CATV outage information for all system service areas.

In addition, the TAC has developed its own database and trouble log system to keep track of various reported problems and their resolution. Use of a database assists in the generation of trend analysis and problem summary reports. This database has also evolved into a repository of all information pertaining to any customer terminal equipment deployed in the field or at cable headends.

Additional Tools

Two other tools extensively used for troubleshooting are data network analyzers and vendor-supplied utilities for the control of various operational parameters of deployed cable modems. The latter allows the TAC remote control over certain cable modem functions such as frequency tuning and power output level setting. Remote tuning is an extremely valuable feature as it allows the switching of field operating frequencies should the original reverse frequencies become impaired by reverse noise problems.

Data Network Group

A network group has been established to look after all data networking and service issues related to the Rogers™ WAVE™ service. Their responsibilities fall within the following three areas:

Data Routing and Internet Access

The main function in this area is the design and implementation of a high-speed data network architecture for the provision of Rogers™ WAVE™ service. This includes establishing minimum performance levels to assist with proper network service provisioning, and to define a standard network design that can properly scale to large customer bases, e.g.: 100,000 homes passed from a single cable headend.

Activities in this area also include the provisioning of Internet access links and the monitoring of the traffic over those links that will determine bandwidth expansion strategy. A related task is the provisioning of high-speed data links between individual cable headends and their corresponding Internet access nodes.

Network Management

The major responsibility in this area is the ongoing tracking of network performance by cable data segment. This is achieved through the collection and analysis of cable modem and other network elements' SNMP statistics such as broadband segment utilization, collisions, errored packets, transmitted and received packets, etc.

In addition to collection and analysis of SNMP statistics, additional scripts have been created to track the availability of cable modems and other network elements. This is done through pinging which involves transmission of test packets to all active devices and keeping track of the number of packets successfully returned by the device.

Rogers™ WAVE™ Network Servers

Major responsibilities in this area are the setup, operation and maintenance of Web, E-mail, and Internet News servers. This role has expanded to include any additional servers required to offer new services under Rogers™ WAVE™ such as remote server content mirroring, personal Web space, multiple E-mail accounts, and others. Server maintenance includes the monitoring of server resource usage required to properly provision for additional hardware to accommodate any increase in the number of customers accessing a service.

These responsibilities also include the generation, issuing, and enforcement of server usage guidelines for Rogers™ WAVE™ technical maintenance personnel.

Field Technical Support

Field technical support for Rogers™

WAVE™ services is provided by the following groups: headend technicians, who look after headend cable modem equipment installation and maintenance as part of their daily responsibilities; maintenance technicians, who look after regular maintenance and troubleshooting of two-way cable plant; cable installation technicians, who also look after installation of the cable modem and drop at the customer premises; and lastly, the Rogers™ WAVE™ technicians.

The Rogers™ WAVE™ technicians look after installation and configuration of both PC Ethernet adapter cards and installation of Internet access software. In addition, they perform on-site troubleshooting of cable modems and PC hardware when necessary, and are notified by the TAC or TSRs through cable dispatch of pending service orders for new installations or equipment replacement.

DEFINING REQUIRED SERVICE SUPPORT PROCEDURES

The establishment of procedures to facilitate the provisioning and management of the new generation of high-speed data services is critical to the success of the cable operator. Having a proper process in place may even be considered by some to be more important than having a working cable modem technology. It is not enough to be able to transmit data bits between the customer and the headend. The cable organization must also be capable of provisioning the new services in a fast and efficient manner, of troubleshooting and resolving any potential problems before they affect data customers, and of managing network changes in a way that minimizes any service disruptions.

The model that Rogers has followed to achieve the above objectives is based on having a central entity whose main task is to coordinate the activities of the various groups responsible for the support of data services. The Rogers Technical Action Center (TAC) has become such an entity. Figure 2 illustrates this model and the central role the TAC plays. The remainder of this paper takes a closer look at what Rogers has done to address each of the above requirements.

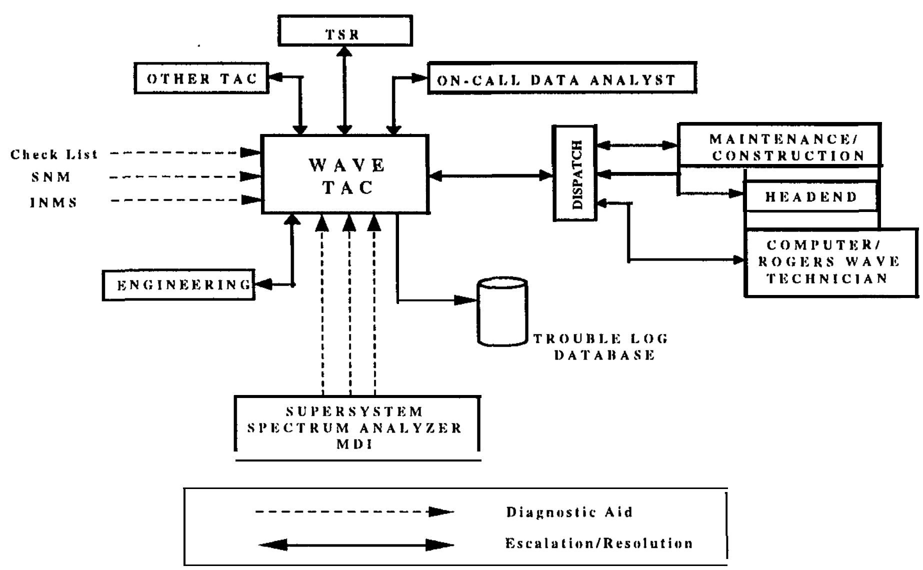


Figure 2. Organization Task Flow

Service Provisionine Procedures

Service provisioning involves a number of steps from the initial customer service request to the cable modem installation and start of the service. Figure 3 depicts the entire process and the tasks performed by each of the support groups involved.

Requests for new service are initially handled by the TSR group whose first function is to confirm the customer lives in a Rogers™ WAVE™ service area, and to verify minimum hardware and software requirements are met. Once the above information has been verified, the next step is to provide the customer with a service ID and any necessary passwords depending on the selected service options. TSRs then generate the necessary service and work orders, and request to the TAC to add the new customer to the service database. Adding the customer to this database results in the assignment of a start-up operating frequency for the cable modem, IP addresses for the customer PC, and any other information required to enable access to the Rogers™ WAVE™ service.

The final step, cable modem installation and service activation, is coordinated with the TAC. The TAC performs an initial test using the start-up frequency initially assigned to the cable modem to verify proper connectivity to the cable headend. Upon successful completion of this test, the modem is assigned its new operating frequency, and the installer can then proceed with the hardware and software configuration process to complete the installation.

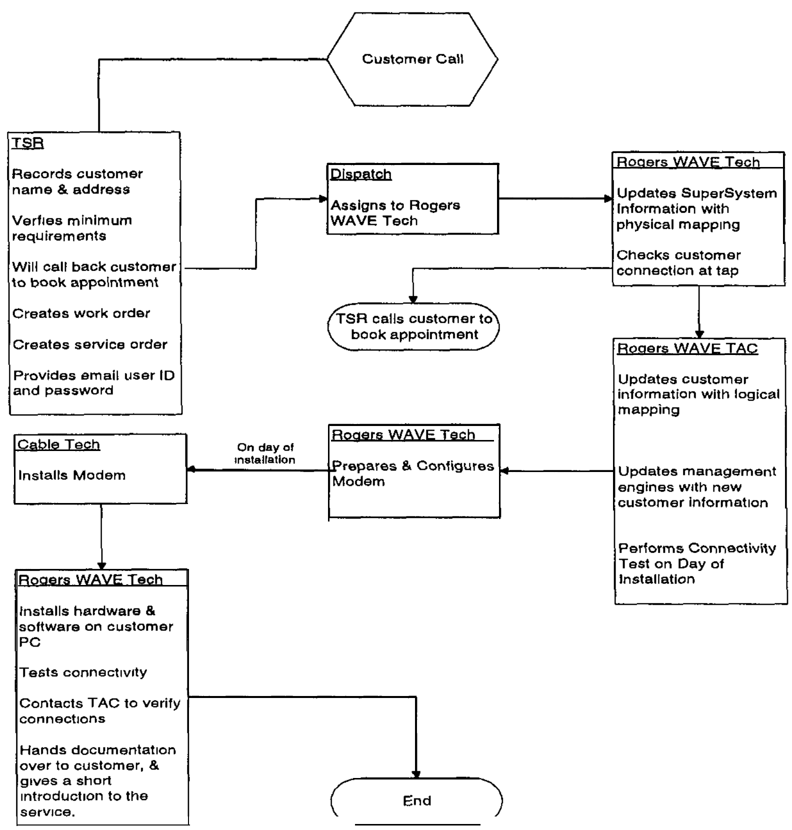


Figure 3. Service Provisioning Process

Problem Resolution Procedures

First Level Troubleshooting and Support

Individual Rogers™ WAVE™ customer reported problems are initially handled by the Technical Service Representatives (TSRs) who are the first level of telephone support. The TSR's role is to determine whether the reported problem is cable related, PC related, or "other", and to resolve or escalate to the Technical Action Center (TAC) which constitutes the second level support.

TSRs access Rogers™ WAVE™ customer information which has been entered into the CATV billing system, and have access to PCs supporting the same data services offered to customers. They can find out about any signal outages through special notification screens built into the existing trouble call tracking system for CATV operations. TSRs also have the ability to conduct echo testing to any hosts on the cable data network to verify network response times and connection quality. In addition, a background-running utility in the TSR PC allows the TAC and TSRs to exchange notes and updates regarding any activities affecting Rogers™ WAVE™ services.

Once the TSR has diagnosed the situation, he/she enters the appropriate fault and clearing codes available in the trouble call tracking system and which were created specifically to deal with faults affecting the new data services. A description of the problem is entered as well as troubleshooting steps taken on special fields in the trouble ticketing application.

Second Level Troubleshooting and Support

The second level Rogers™ WAVE™ customer support function starts when a TSR refers the reported problems to the TAC after determining that he/she is unable to resolve them. The TAC then takes ownership of all reported problems and is ultimately responsible for their resolution. It also has access to the billing and CATV trouble call tracking systems for customer and problem information, and to Rogers™ WAVE™ services as seen by the customer.

In its monitoring capacity, the TAC looks at both the cable and data components of Rogers™ WAVE™ and alerts TSRs and cable dispatch of signal outages and other problems through the CATV trouble call tracking system. Furthermore, the TAC also has the ability to book new trouble calls which can in turn generate pending work and maintenance orders at cable dispatch. Cable dispatch can then prioritize and assign those orders to the on-call CATV technicians. The TAC thus becomes responsible for all trouble tracking and for ensuring proper problem resolution. After successful resolution, the TAC can clear a pending order by entering the appropriate resolution code in the trouble call tracking system. Figure 4 illustrates the current problem resolution flowchart.

If a referred trouble call is resolved without further referrals to other Rogers™ WAVE™ technical support groups, the TAC enters the appropriate clearing code that best describes the action taken and closes the pending call. Should the TAC require further clarification or assistance from a customer during troubleshooting, then the TSR is asked to contact him/her directly. The TAC does not have direct contact with the customers.

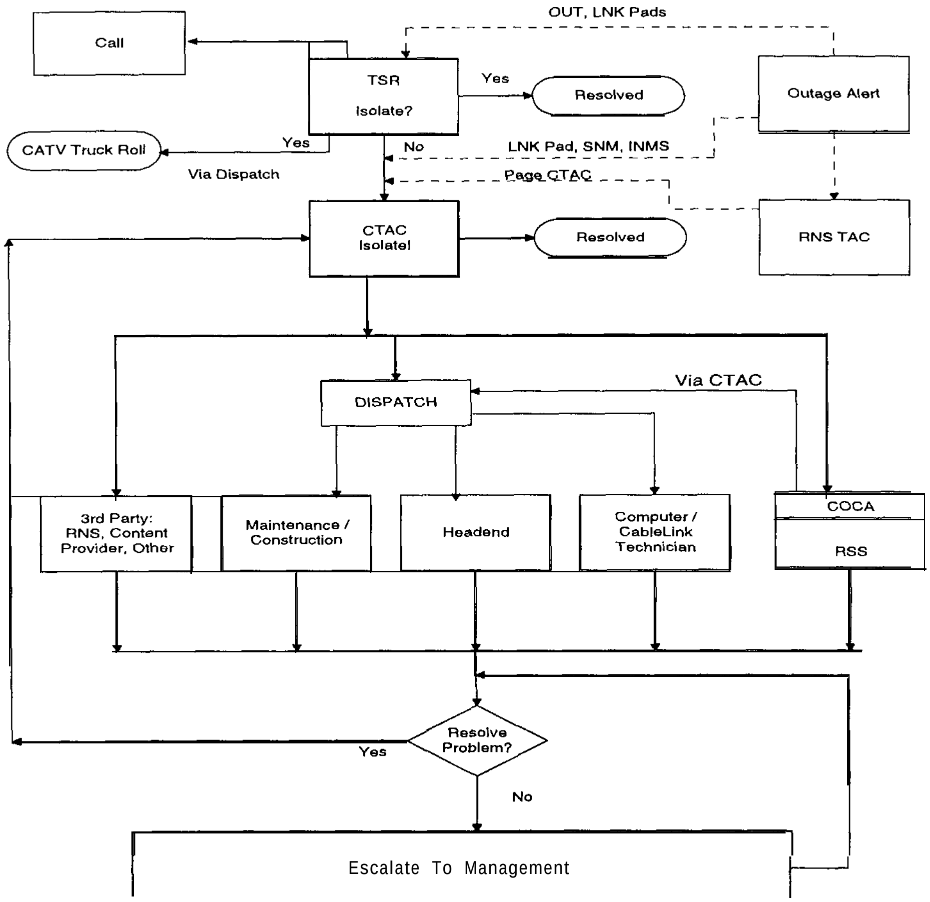


Figure 4. Problem Resolution Flowchart

Change Control Procedures

Change control procedures are currently under implementation whose main goal is to ensure all network changes, whether affecting the physical cable plant or the overlaid data distribution and switching network, are communicated to the TAC and eventually authorized by it. The TAC will evolve from strictly a monitoring and troubleshooting entity to a clearance center responsible for the management of the entire end-to-end network.

Change control is being introduced in two phases. Phase 1 is implemented with the help of the various cable groups responsible for provisioning of Rogers™ WAVE™ services. Phase 1 requires that each of these groups notifies the TAC 24 hours in advance of any network changes which have the potential to affect service. All involved groups are required to provide the TAC with the following information:

- Name of the originator
- Location(s) affected
- Date and time of the change(s)
- Expected duration of the planned activity
- Type of work and justification for it
- If potential for an outage exists, provision for a backup plan and

potential number of customers that could be affected

Phase 2 will require that any changes on the network be authorized by the TAC, and that request for changes be made to the TAC at least 72 hours in advance. The TAC will keep track of all changes, and will authorize or deny change requests. Eventually, a maintenance window will be enforced for all networks and all changes will take place within that window.

Upon notification and approval of any changes, the TAC will also be responsible for informing all customers of the planned activities and of the potential for any problems affecting their service.

CONCLUSIONS AND NEXT STEPS

Rogers has implemented new procedures to support the introduction of high-speed data services, specifically, the Rogers™ WAVE™ service offering.

In order to offer high levels of service quality and reliability, the need to have end-to-end visibility over all the network components has been clearly identified. This visibility is required over both the physical cable plant and data network operation to ensure that potential problems are addressed before they lead to service degradation, and to allow prompt resolution of any service outages. To help achieve these goals, Rogers has set up a Technical Action Center (TAC) with the mandate to monitor operation of two-way plant in general, and Rogers™ WAVE™ services in particular. The TAC constitutes the second level of Rogers™ WAVE™ product support after the traditional Technical Service Representatives (TSRs). TSRs continue to be the first contact with cable data customers and perform preliminary problem troubleshooting. Should TSRs fail to achieve proper resolution at this initial stage, the TAC takes ownership of the problem and becomes responsible for its proper resolution. The TAC has access to a wide variety of tools to rapidly diagnose the problem, and to escalate it further if necessary to other technical support groups. The support groups encompass traditional cable technical groups, i.e.: construction and customer service, or any other groups specifically created to support other aspects of the new service, i.e.: data network administration and maintenance groups.

Although initial service goals have been defined, these requirements were generated within the context of Rogers™ WAVE™ . The need has already been identified to expand these requirements to cover all two-way cable services, and to slowly migrate current TAC operation into a global Network Operations Center (NOC).

To achieve the latter objective, current network monitoring procedures and tools will need further refinement. The migration towards a fully empowered TAC to enforce change control procedures is one of the first steps already taken in that direction.

ACKNOWLEDGMENTS

The author wishes to thank everyone at Rogers who is part of the Rogers™ WAVE™ team, and whose efforts and tremendous enthusiasm have taken this new service to its current level of success.

Meeting the Needs of the Headend-of-the-Future Today--The Structured Headend

Al Swanson, Gary Mayer and Bill Hartman
ADC Telecommunications

Single media CATV networks are rapidly evolving into two-way, multi-media networks. They will be expected to reliably support a much broader range of services and technologies. Digital and analog signals will work in the same network to deliver telephony, data, and video to residential, business, public, educational, and government customers. The timing of the roll-out of each of the services to these markets may vary tremendously, and customer churn will certainly be a part of each market. There may be a smaller number of larger headends that will still need to deliver picture-perfect video. All this on a network that is expected to have unprecedented reliability, and is expected to be there "yesterday".

So there is a lot of confusion, but also some conclusions.

These needs are driving service providers to expect much more from their network infrastructure. They are looking for full connectivity in the fiber and coaxial networks that will be linking digital and ATM switches, as well other digital, RF, and analog elements. To successfully achieve this network performance means planning for advanced services and deploying effective expansion strategies. A cost effective plan that incorporates the needs for flexibility, reliability, and growth, will deliver:

- Service evolution that is fully supported by circuit access and protection, service assurance monitoring, and disaster recovery provisions.

- A reliable network ready for expansion.

- A strategy for upgrading existing systems, regional headend consolidations and completing new-build situations.

- And perhaps most important, modular, growth-oriented headends matching capabilities and incremental investments with customer requirements.

Before significant resources are invested, the plan should consider impact of the following:

- Services timing
- Facility restrictions
- How to get to the "next" services
- Equipment types
- Necessary space
- Technical performance and equipment standards
- Maintenance and test provisions
- Design limitations

ADC Telecommunications is able to bring practical hardware solutions as well as system expertise to the needs of the multimedia service providers. ADC is one of the first telecommunications suppliers to develop and deploy systems capable of delivering integrated video, voice, and data over a broadband network to residential and business subscribers. Our strong background includes cable management, access systems, and high speed data and wireless transmission. This knowledge has been combined with CATV and systems integration industry experts. We already have several years experience building and turning up systems that meet today's rapidly emerging competitive environment.

This experience leads us to the conclusion that headends being modified or built today must be thoroughly planned to accommodate this radically new environment. We have developed a headend design philosophy that structures capabilities to incrementally meet service and technology needs. This solution will allow you to change or add service offerings quickly, and will allow you to offer those services geographically targeted to meet demand.

What The Future Means Today

Traditionally, the bulk of the design effort expended in a CATV system design has been for the distribution systems external to the headend. CATV providers study their customer's demographics and service territories in great detail. The CATV customer base is well understood. The focus is on maximizing plant performance and reliability, as well as minimizing noise and distortion. Fiber optic technology has moved further out into the cable plant reduing amplifiers in cascade,to improve reliability while reducing noise and distortion.

In contrast, headend floor plans and layouts reflect the same basic principles common ten years ago. The focus has been on improving signal quality and channel capacity. These headends may be severely limited in meeting new plant demands.

The time to build more capable headends is now. Their designs must serve as platforms to deliver advanced broadband services, profitably expanding as new customers and services are added, yet with minimal cost and disruption. The new design model should consider the following:

1. The CATV provider faces the challenge of successfully and cost-effectively adapting to become a provider of Broadband Services while competing against other telecommunications service companies. The new service possibilities include telephony, wireless personal communications, and Internet gateway access, as well as two-way video and targeted commercial insertion.

2. The ability to add and delete channels in the analog video line up with minimal disruption, lowest cost, and highest system reliability. Line-up changes in the past have traditionally meant rewiring major portions of the headend.

3. Service demands and available technologies will continue to evolve, often in unpredictable ways. The headends that are built today must be able to grow with these demands while protecting invested capital.

4. The take rate of new services may initially be small. As the popularity of the new service grows, a migration plan is required that will allow

smooth expansion at the lowest possible incremental cost. Moreover, the take rate may vary from area to area, and even from neighborhood to neighborhood. A headend design is required that will allow for geographically "spotty" growth with a minimal capital investment in new equipment. Service equipment should only need to be added to those pieces of the network that are supporting customers and generating revenue.

The ability to flexibly accommodate this migration requires new thinking about headend design. Traditional design principals and advanced hardware alone will not solve these challenges. While today's designs are delivering better and better signal quality, they are severely limited in meeting added near-term, and certainly longer-term, requirements The issues of flexibility, service migration, and protection of investment must be addressed by connecting available advanced hardware and new, tailored services, in innovative ways.

A New Philosophy In Headend Design

The headend of the future must accommodate not only video services, but also telephony, data, and wireless communications. The result is a facility which is not just a telephone central office or a video distribution office, but is a Broadband Services Distribution Office. The key elements in this concept include:

- Modularity of service provision, both in terms of services and where they are offered. Pre-designed equipment arranged in module bays focused on specific types of service or size of service area. They include Narrowcast modules for commercial insertion and Public, Education, and Government (PEG) services, as well as Advanced Services modules for telephony, Internet access, digital video, and PCS.

- Structured design and design rules that establish in advance how to accommodate change. The future design must allow you to insert new advanced services at the appropriate point in the network with minimal disruption to ongoing operations. This will result in a more reliable facility, lower maintenance and upgrade costs, and financial protection of existing equipment investment.

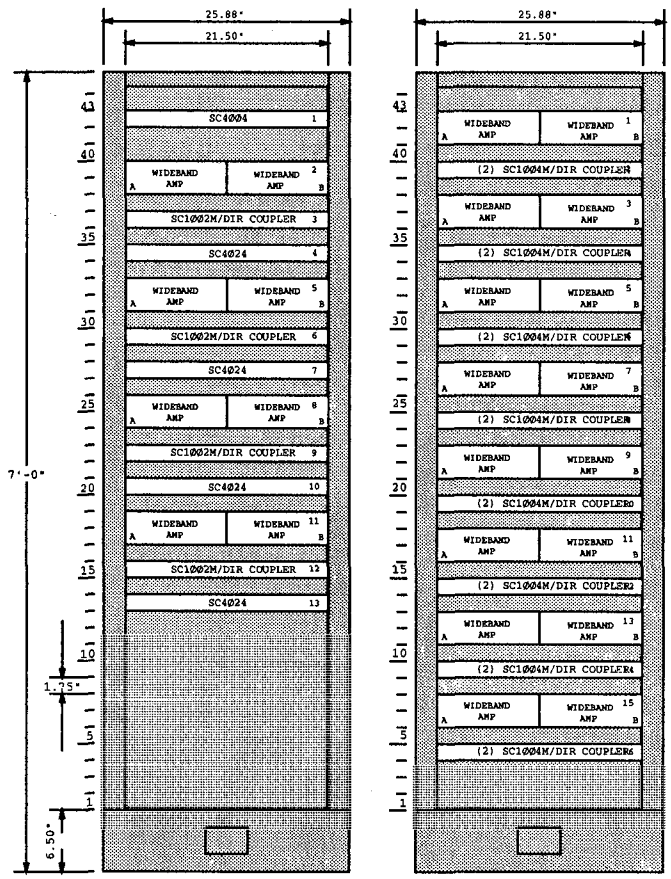
- Combining CATV know-how and RF expertise with the best quality and performance standards from Telco networks. This includes planning likely network requirements for test, access, and system reconfiguration. Achieve maximum service reliability through redundant circuit routing for lost satellite feeds and off-air signals, the failure of video and audio processing equipment, and the loss of commercial power.

- Begin with the end in mind. Reasonably plan for the full services and customer base the facility will be called upon to support. Design, but do not necessarily initially equip, to accommodate a full channel and service line-up. With full feature needs allocated within the scope of the modularly structured facility, future floor space, power and other limitations can be recognized. Actual near-term building and installation is scaled back to a portion of overall capabilities. The service provider is now positioned to meet change with matched, and planned, investment.

- Include open racking and bays for efficient maintenance and system reconfiguration. By eliminating cabinets and doors in the headend, today's complex headend can be packaged in a smaller facility. Technicians will be able to add and re-arrange hardware and locate trouble much more quickly.

- Plan for signal and cable growth and routing using racking systems. New video offerings require many more pieces of equipment, and thus much more cabling than a few years ago. The use of under-floor systems often results in a tangled mess of power, video, audio, and fiber cables. This means tracking problems in locating cables, kinks and microbends in fiber cables, and crosstalk. Tugging and pulling on cables as the means of circuit tracing stresses connections on other equipment, further hurting network reliability. Finally, open floor panels can be a safety issue in facilities that are undergoing a constant configuration change.

- Identify and focus on critical details, such as grounding systems. Careful attention to grounding, for example, reduces the incidence of ground loops, resulting in higher audio quality with less AC hum, and ease of trouble shooting. Poor grounding



SPLITTER/COMBINER BAY
FOR LEVEL 1
7'-0" UEF EIA RACK
FRONT VIEW

SPLITTER/COMBINER BAY
FOR LEVEL 2 & 3
7'-0" UEF EIA RACK
FRONT VIEW

Figure 1

practices may work in a traditional headend, but will prove to be a disaster when broadband telephony or digital services are added to the CATV network. In addition, poor grounding practices cost the service provider money as electrostatic discharge transients reduce equipment reliability.

- Integrate the splitter/combiner networks into the overall system philosophy of planning for service and technology insertion (figure 1). The network is placed close to related equipment and the module insertion point to achieve easier access. The typical field experience with this has shown labor savings of a factor of 4 in installation time, while maintaining 0.33 dB frequency response variation from 5 MHz to1 GHz, and 35 dB isolation port-to-port up to 450 MHz.

Group over several actual projects. This experience shows that the investment in the structured Broadband Services Distribution Office is 5% to 7% higher than recent, but less capable, industry designs. It will typically deliver savings of at least 20% annually when compared with Headend operations and maintenance costs of today. Of course, the demand on headend facilities is not static, but is increasing significantly. As a result, structured plant savings of greater than today's 20% are likely.

The Signal Flow diagram

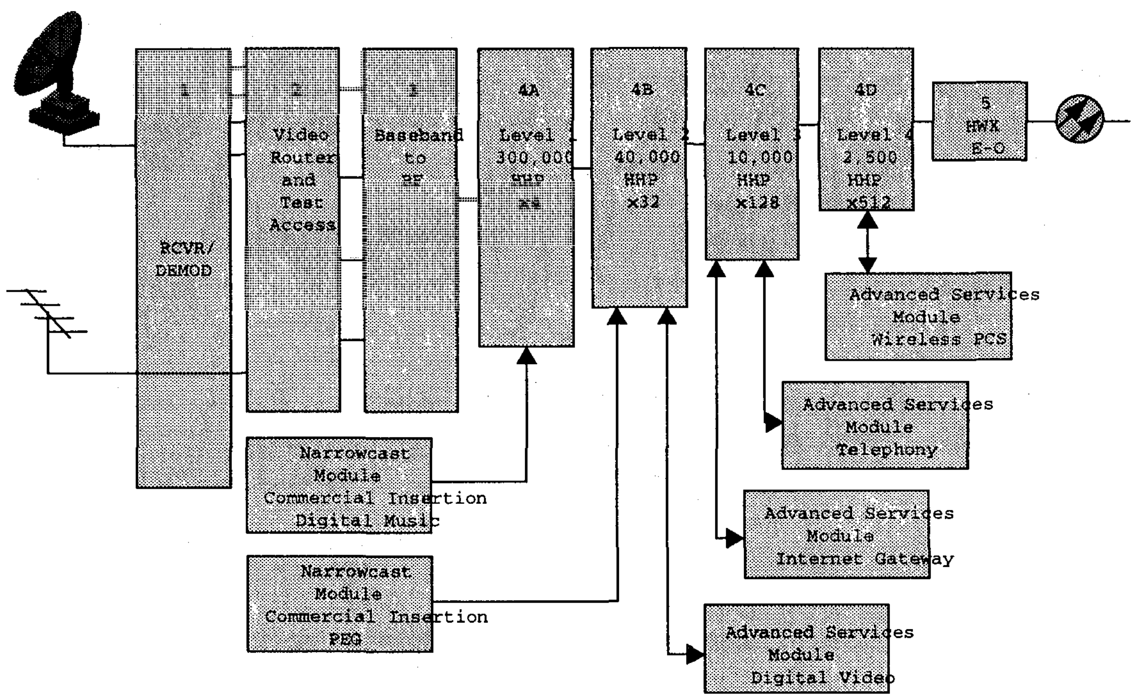


Figure 2 (Signal Flow)

What Does It Cost? How Does It Work?

This design approach has evolved within the ADC Systems Integration

(figure 2) represents the design of the Broadband Services Distribution Office network. It is divided into sections, or levels, that represent increasingly smaller geographical areas. The capacity of each of the modular levels is:

Level 1 - 1.2 million homes passed.
Level 2 - 300,000 homes passed.
Level 3 - 40,000 homes passed.
Level 4 - 10,000 homes passed.
Level 5 - 2,500 homes passed.

Satellite channels and off-air channels are converted into baseband signals and routed through an Audio/Video Patch Panel. The functionality of this panel allows for high density packaging for testing and alignment procedures. It also provides the capability for manual patching of circuits around devices and to reroute circuits during equipment failures.

From the Audio/Video Patch Panel, channels are directed to a baseband crosspoint routing switcher. The purpose of the switcher is to automate channel line up changes and provide an insertion point for system-wide services. For example, the Emergency Broadcast System will be added to the network at this point, as would commercial insertion sources, and PEG sources that are common to the entire serving area. From the routing switcher, the baseband-to-RF conversion section, consisting of video encoders, scramblers, BTSC encoders, and modulators, converts each channel to an individual RF channel, which is combined into one line for splitting into geographic levels.

The first split breaks the signal into Level 1 and Level 2. Level 1 signals are common to all subscribers. These common channels are the analog broadcast channels from the satellite feeds, off-air broadcast stations, and the Emergency Broadcast System. Up to 1.2 million homes may be served by this point in the network.

Level 2 splits the channels into 4 parts, each of which can serve up to 300,000 homes. This level would typically serve a metropolitan area. Commonly, this is where narrowcast services, such as commercial insertion, would be injected.

The Level 3 split breaks each of the Level 2 channels into 32 parts, each leg serving up to 40,000 homes. This level would typically serve smaller cities or areas within a metropolitan area. PEG channels are typically inserted at this level.

Level 4 splits Level 3 into 4 parts, serving approximately 10,000 home serving areas on each leg.

Level 5 splits each of the Level 4 legs into 4 parts, and can serve up to 2,500 home neighborhood areas on each of its legs.

At each point in the splitting network, combiners are installed which provide insertion points for advanced services. For example, the first deployment of telephony into a CATV network might serve 500 homes spread throughout the entire serving area in a "friendly" field trial. In this case, one Host Digital Terminal (HDT) would be required, and could be installed on a leg at the level 2 split through a 1:8 combiner, taking into account transmission delay limits on the telephony system. Additional HDT's would be added at this level during the early stages of commercial deployment. Once the telephony offering gains a "significant" customer base and sign-up rate, the decision would be made to move the HDT's down to the level 3 split, where many more combining ports are available. The same HDT's can be reused and additional units added. Since the insertion ports were designed into the network from the beginning, there is no major rewiring of the Broadband Services Distribution Office and no unexpected impact on end-of-line performance. With further growth in the telephony service take rate, the HDT units would be moved down to level 4, continuing to link costs to revenue. Again, rewiring and disruption of the facility is minimal.

An alternate scenario might be that the majority of initial Internet customers can be served at Level 3 for Internet access. However, suppose that one of the 2,500 home neighborhoods has a much higher take rate than other neighborhoods. Because of system flexibility, an Internet Advanced Services Module can meet this neighborhood demand at Level 4. This minimizes the capital investment while growing the system, even if the growth is not geographically even.

In addition to growth flexibility, the combining networks at each level permit insertion of other advanced services. In addition to telephony, many carriers have indicated a requirement for wireless Personal Communication Services. The roll-out and growth pattern of PCS services may be similar to that for cable telephony. The Remote Antenna Signal Processor (RASP) will be connected at a combiner panel for communication with Remote Antenna Drivers (RADs). A typical PCS Module (figure 3) will serve approximately 10,000 homes passed. Similarly, downstream Internet traffic from data bridge/routers will be inserted into the combiners using the Internet Advanced Services Module. Other services, such as SEGA, Digital Music Express, and

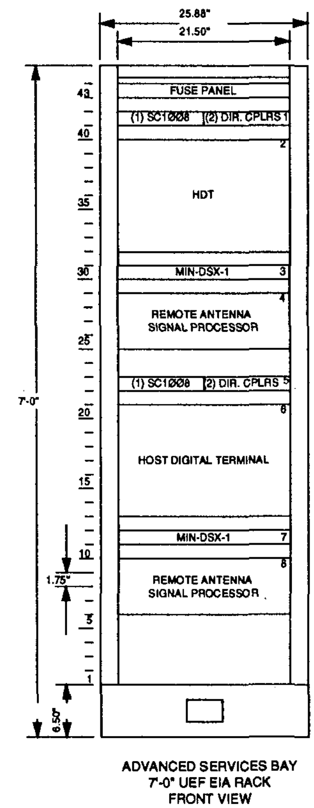


Figure 3

NVOD are accommodated in exactly the same way. This tailoring of equipment provides maximum flexibility while minimizing investment.

Location

The Broadband Services Distribution Office may be in a single physical facility, as represented in figure 2, or the levels may be in different geographic sites. For example, the TVRO function and Level 2 split may be in one central facility. If we assume digital super trunking, with a 30 dB optic loss budget and 0.4 dB/km total loss, the Level 2 channels may be transported up to 85 km at 1550 nm without optical repeating. The Level 2 split can serve up to 4 smaller facilities in a metropolitan area. Some systems may require further transport down to a Level 3 facility. Using available equipment with analog DFB lasers, a transmit level of 8 mw and an optic path budget of 9 dB, lower level facilities can be extended a further 36 km. The flexibility in the dispersal of facilities is appropriate for the CATV provider that serves many areas or cities that are not adjacent to each other.

Performance

Of course the end-of-line requirements are a critical element of the overall design. We recommend that the Structured Headend/Broadband Services Distribution Office be designed to better performance levels than currently required by the FCC. The limits are shown in the following:

FCC Requirements	Recommended Goal
C/N = 43 dB	C/N = 49 dB
CTB = 51 dB	CTB = 53 dB
CSO = 51 dB	CSO = 55 dB

With standard headend and outside plant amplifiers and fibertransmitters/receivers, achievement of these goals is possible. Calculations show that in an overall CATV network, with end-of-line performance standards equal to the recommendations, the headend contributes less than 1.5 dB for CTB, and less than 0.2 dB C/N. So it is possible to design a headend that is flexible and expandable without sacrificing quality and performance.

Summary

The competitive telecommunications environment is filled with opportunity for new services and profits. Building headends and networks to meet the range of possible services means levels of system complexity and uncertainty never before seen in this industry. Competitors, however, continue to make their own decisions and investments to drive the new telecommunications markets.

While operating in this new environment is a challenge, the Broadband Services Distribution Office and its structured, modular approach makes laying the infrastructure for these services today not only possible, but optimal. Rather than a "solution of boxes", this standard design offers pre-planned integration that serves now and in to the future. This solution maximizes the return on headend investment by allowing service offerings to be added or changed quickly, with maximum reliability and performance.

MINIMIZING MEMORY COST IN SET-TOP BOXES

Jeff Mitchell
Rambus Inc.

Abstract

Memory is the largest contributor to the cost of a set top box, typically comprising up to half or more of the component cost. This paper examines methods of minimizing this cost using commodity memory parts consumed in large volumes by the PC industry and by taking advantage of new developments in high density, high bandwidth DRAMs.

Architectural methods of extracting bandwidth from conventional DRAM are examined, along with performance requirements of emerging set-top box architectures. Alternatives are presented for meeting product performance goals while preserving low cost.

THE EFFECT OF MEMORY ON SYSTEM PERFORMANCE

Digital set top boxes require tremendous amounts of bandwidth in order to deliver the interactive experience desired by consumers. MPEG–2 decoding, stereo audio, computational 3D graphics and interactive user interfaces demand performance levels previously unheard of in consumer products. Yet the cost constraints of the consumer market demand that inexpensive, relatively low performance Dynamic Random Access Memory (DRAM) be used for memory. In the past, computer designers have used many DRAMs in parallel in order to increase memory bandwidth. This approach works well in PC's or workstations where there are many megabytes (MB) of memory, but is not effective in a set top box where the total memory size may only be as little as two to four MB. Two or four MB requires only one or two DRAM devices, whereas four or more devices are typically needed in order to provide sufficient bandwidth from conventional DRAMs. This makes a difficult choice for the system designer - add more memory and increase cost or accept lower performance.

Recently there have been new developments in the DRAM industry that offer solutions to the memory size/bandwidth dichotomy. Two new types of DRAM devices, Rambus™ DRAM (RDRAM®) and Synchronous DRAM (SDRAM) offer improved bandwidth over traditional devices and yet still maintain an inexpensive DRAM cost structure. Both are designed to be used in personal computer main memory applications, the largest user of DRAM memory. Over half of the world's annual DRAM shipments go into PC's. Interactive set top boxes which use these new high bandwidth types of memory can meet their performance objectives and take advantage of the PC cost/volume curves. This allows the consumer to have the best of both worlds - high performance at low cost.

MEMORY ALTERNATIVES

DRAM has always been the memory type of choice in cost-sensitive applications. Although DRAM performance is poor relative to other types of memory devices such as Static RAM (SRAM), DRAM has the virtue of being both dense and cheap. System designers go to great lengths to get the performance they need from DRAM and avoid using other types of memory in order to keep product cost down.

The conventional method of increasing performance in DRAM systems is to put several devices in parallel. If a single device cannot provide the required bandwidth, two devices used in parallel will double the performance. This also doubles the total amount of memory and nearly doubles the number of interface pins on the memory controller. The effect on the system of using DRAMs in parallel for increased performance is shown in TABLE 1. This example uses 16 Megabit (Mb) page mode DRAM devices in a 16 bit wide organization (1M x 16).

DRAMs	*Bandwidth (MB/s)*	*Controller Pins*	*Total Memory*
1	66	40	2 MB
2	133	60	4 MB
4	266	120	8 MB

TABLE 1. SYSTEM EFFECTS OF OPERATING DRAMS IN PARALLEL

Memory Granularity

In order to achieve a desired level of bandwidth, some number of DRAMs must be used in parallel. The minimum amount of memory required to be used in parallel is called the memory granularity. As devices are paralleled to increase performance the total system memory grows proportionately. Large memory granularity can be a problem in applications where cost requires that the total memory be kept to a minimum.

In past generations of DRAM technology memory granularity was usually not an issue because DRAMs were less dense, with fewer total memory bits in each device. Four 4 Mb devices in parallel yields the same total bandwidth as shown in TABLE 1, but the memory granularity would only be 2 MB. This is a factor of four less than with 16 Mb DRAMs.

So why not just use lower density devices to get the needed performance? Besides the increased board space, power, and the sheer number of pins needed on the memory controller, DRAM economics dictate that the cheapest price per bit will be found on the densest device.

When designing a new product, the most cost effective memory to use is the densest DRAM technology available. This will yield the lowest total memory cost. FIGURE 1 shows the relative cost over time of various densities of DRAM. For new products it is best to target the densest device that will be cost effective in the time frame that the product reaches mass production.

For example, a product starting design today would target the 16 Mb generation while also giving consideration of how to use 64 Mb devices should the anticipated lifetime of the product extend beyond 1998 or 1999.

Considering that a 64 Mb DRAM is 8 megabytes of memory in a single chip, it is

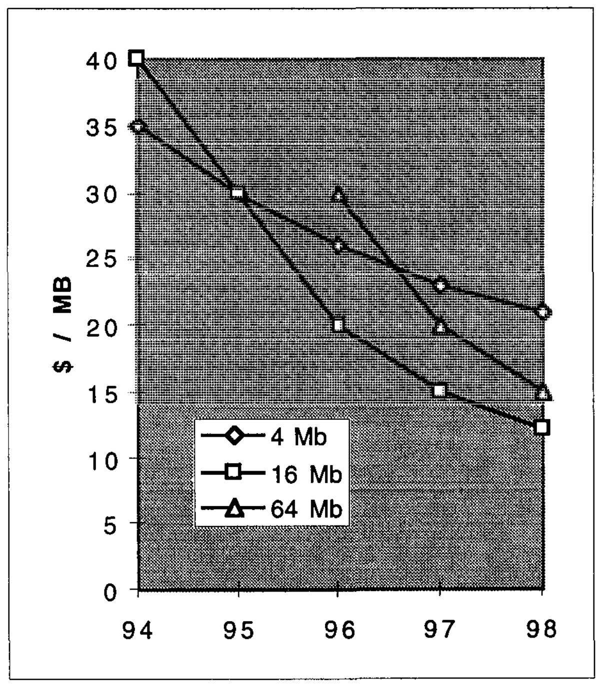


FIGURE 1. DRAM PRICING TRENDS

easy to see that the conventional practice of using DRAMs in parallel to obtain more bandwidth has become impractical to do in small memory systems. There are two alternatives to this practice - wider DRAMs and faster DRAMs.

Wide DRAMs

One alternative to using several DRAMs in parallel is for the DRAM manufacturers to simply make wider devices with data bus widths of 16, 32 or 64 bits. This provides the same benefit as parallelism but without increasing the total system memory size. This approach works up to a point, but then becomes both financially unattractive and technically difficult.

As a DRAM is made wider the die size increases and the package gets larger and more costly. With more I/O pins, it also becomes more expensive to test. These factors tend to negate the cost advantages of DRAM.

A wide DRAM also cannot be operated as fast as a narrower device. The increased number of output pins causes more noise and ground bounce. The remedy to this problem is to run the device slower, which offsets the performance advantage of being wider. Wider parts must also have more pins providing power and ground connections, which again increases cost.

Wider DRAMs are a partial solution to achieving higher performance, but at some width around 32 bits this approach reaches diminishing returns.

Fast DRAMs

An alternative to making a wider DRAM is to make a faster DRAM. Here the objective is to keep the device width down to a manageable size, but increase the speed at which it operates.

There are two types of 'fast' DRAMs, Synchronous DRAM (SDRAM) and Rambus DRAM (RDRAM). These two DRAM derivatives are similar in that they both use a conventional memory core and run the external interface at a high speed. This provides the economic advantages of a conventional DRAM while providing much higher performance.

Synchronous DRAM (SDRAM)

An SDRAM is a conventional DRAM mated to a synchronous interface. The synchronous interface aligns data transfers into and out of the part with an external clock reference.

Synchronizing the data transfer to a clock allows for tighter timing parameters and therefore a higher operating speed. SDRAMs can run in systems at speeds up to 66 MHz, about double the speed of a conventional DRAM. Doubling the interface speed means that only half as many devices are needed for a given bandwidth. This reduces the memory granularity to half that of a conventional DRAM.

Rambus DRAM (RDRAM)

As with SDRAM, RDRAM is a conventional DRAM mated to a synchronous interface. An RDRAM has a 64 bit wide internal bus running at 75 MHz. The RDRAM connects to a memory controller, which also has a 64 bit wide internal bus running at 75 MHz.

These wide internal busses narrow to only 8 bits externally without any impact on performance. This gives an RDRAM the performance of a 64 bit wide DRAM while retaining all of the cost advantages of a narrow 8 bit external bus.

Type	*Bus Width*	*Package Pins*	*Bandwidth (MB/s)*
DRAM	16 bits	42	66
DRAM	32 bits	100	133
SDRAM	16 bits	50	133
RDRAM	8 bits	32	600

TABLE 2. PIN/BANDWIDTH COMPARISON

TABLE 2 summarizes the number of signals, package pins, and relative bandwidth of DRAM, wide DRAM, SDRAM and RDRAM.

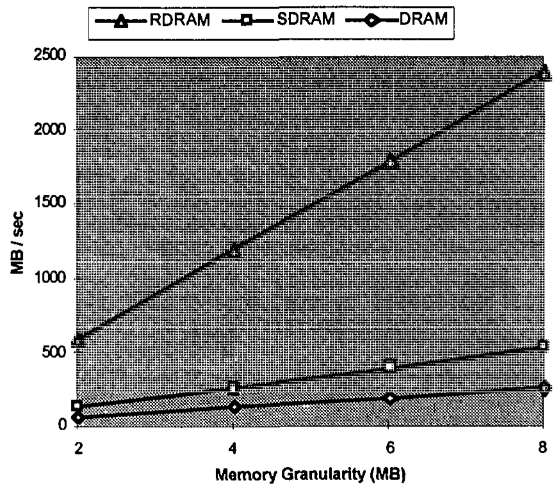


FIGURE 2. MEMORY GRANULARITY VS. BANDWIDTH

With each type of DRAM there is a straight-line relationship between bandwidth and memory granularity. This relationship makes it straightforward to approximate the memory granularity for a given level of system performance If the system performance requirement lies above the line shown in FIGURE 2 for a particular type of DRAM, either another type of DRAM will have to be used or the bus width will have to be increased, with a corresponding increase in memory granularity and system cost.

For example, an application which requires 250 MB/s of system bandwidth can be designed one of three ways, depending upon whether DRAM, SDRAM, or RDRAM is used. FIGURE 3 shows block diagrams of example systems comparing total bandwidth, memory granularity, and number of controller pins required for each of the three solutions.

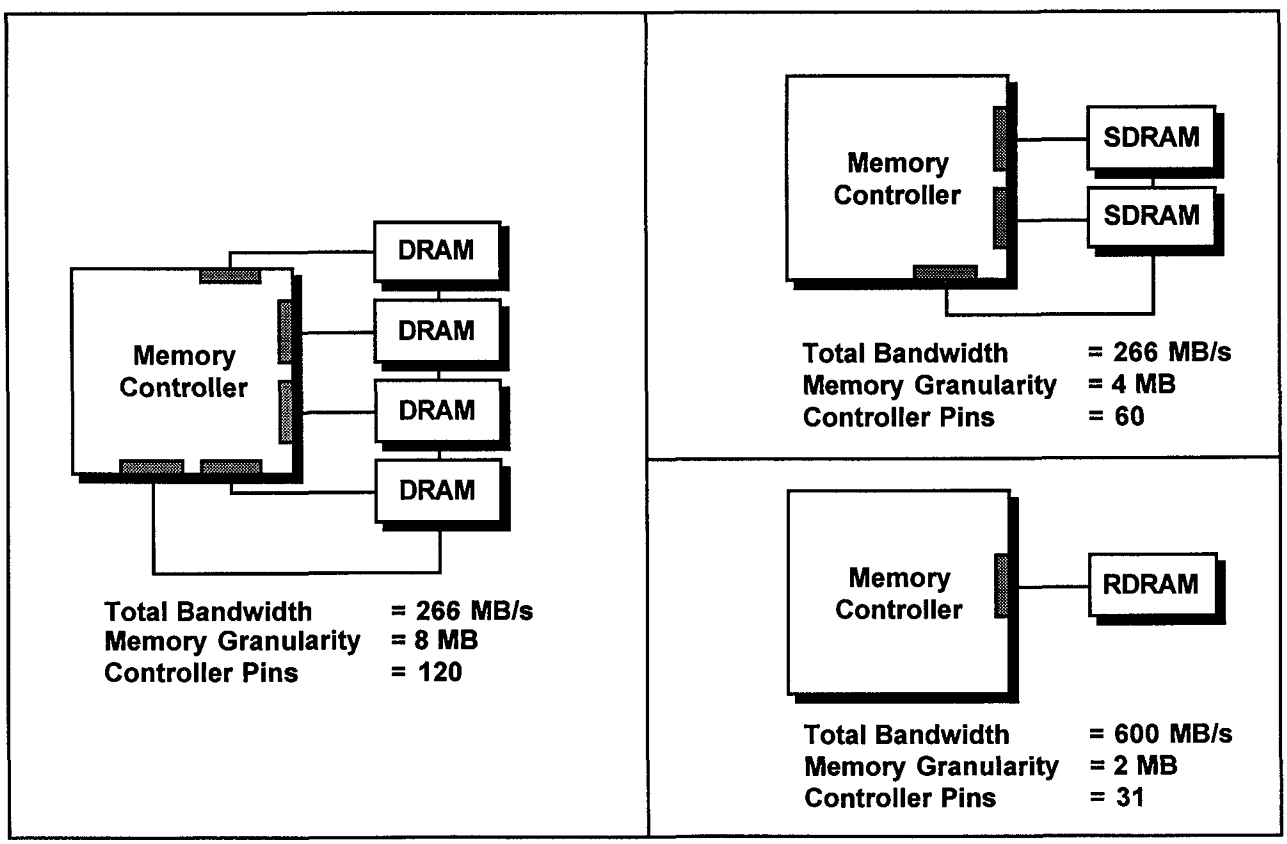


FIGURE 3. HIGH BANDWIDTH MEMORY SYSTEMS

SET-TOP BOX PERFORMANCE REQUIREMENTS

Depending upon the targeted application, a set-top box can have a broad range of memory bandwidth requirements. TABLE 3 lists the approximate bandwidth required for several common functions.

Function	*Min B/W*	*Max B/W*
Video	15	30
CPU	5	100
MPEG–2	100	200
2D Graphics	50	200
3D Graphics	100	300
Sound	10	50

TABLE 3. BANDWIDTH REQUIREMENTS (MB/s)

At one end of the spectrum a simple analog decoder has very modest bandwidth requirements. At the other end of the spectrum, a fully digital web-capable system with 2 channels of MPEG–2 for picture-in-picture and a fully interactive 3D user interface could easily require over 500 MB/s of memory bandwidth. This much bandwidth would require 16 MB of DRAM, 8 MB of SDRAM, or 2 MB of RDRAM. Clearly, it is difficult if not impossible to make a high performance, cost effective consumer product using conventional DRAM. Estimating system bandwidth requirements from TABLE 3 and comparing to FIGURE 3 gives an idea of what memory options are available for a cost optimized product.

Distributed vs. Unified Memory

There are two architectural methods to obtain the required system memory bandwidth. The conventional method has been to attach the required amount of memory to a chip performing a specific function. For example the microprocessor would have some memory connected directly to it, the video decoder would have some more memory separate from the CPU memory, and so on with each separate chip in the unit. This approach works well as long as the memory can be cleanly partitioned and there are no problems with memory granularity.

An alternative method is to unify the memory and have all functions operate directly out of the same block of memory. While this eases the memory granularity problem by combining many small pieces of memory into one large pool, it also increases the bandwidth that is needed from that one pool.

A move toward unified memory architectures is being motivated by the increasing integration of functions. Integrating previously separate functions into a single chip forces unification of the memory for all of those functions.

IMPLEMENTATION EXAMPLES

To demonstrate how system performance and memory granularity interact, two example systems are profiled below. These cover the entire range of set-top box functionality ranging from a simple dedicated decoder to a high end fully interactive system. The examples assume a unified memory architecture since that provides the potential for the lowest cost system by utilizing the densest memory devices.

Dedicated MPEG–2 Decoder

In this example a single NTSC MPEG–2 stream is being decoded. In addition there is a simple user interface generated by the CPU. From TABLE 3, the total system bandwidth can be estimated:

Video	15 MB/s
CPU	10 MB/s
MPEG–2	100 MB/s
TOTAL	125 MB/s

This bandwidth can be provided by either two DRAMs, one SDRAM, or one RDRAM. The memory granularity in this example is 4 MB if DRAM is used, or 2 MB if SDRAM or RDRAM is used.

Fully Interactive Terminal

An advanced interactive terminal may consist of a complete Internet web browser in addition to a multiple stream MPEG-2 decoder (for picture-in-picture or faster response to channel surfing) with surround stereo along with a fully interactive 3D user interface. Such a terminal would need a significant amount of memory bandwidth. Again estimating total system bandwidth from TABLE 3:

Video	30 MB/s
CPU	100 MB/s
3D Graphics	300 MB/s
Sound	50 MB/s
MPEG-2	200 MB/s
TOTAL	680 MB/s

To provide this bandwidth from DRAM would require over 20 MB of memory! Digital systems are generally designed to support memory systems in binary increments, so a 128 bit data bus with 32 MB would be the memory granularity for a DRAM based system. An implementation using DRAM obviously would be too costly to be a consumer product.

A memory subsystem built from SDRAM would require a 64 bit data bus and 16 MB of memory in order to achieve the needed 680 MB/s of bandwidth. Design compromises could get the bandwidth down to 533 MB/s which would require only 8 MB of memory. Not cheap, but getting there.

Using RDRAM would require only 4 MB of memory which would provide well over the required 680 MB/s of bandwidth. Over 500 MB/s of spare bandwidth would be available for other functions or future performance improvements. Alternatively, if it were possible to put all of these functions into 2 MB of memory, then as with the SDRAM example the system could be re-engineered to get the required bandwidth down to 600 MB/s. This bandwidth can be satisfied by a single RDRAM.

A set-top box with this kind of high end functionality is not likely to become commercially viable for several years, at which point the most cost effective DRAM will be a 64 Mb device. The higher density DRAM will exacerbate the memory granularity problem. Using 64 Mb devices the minimum DRAM system would be 128 MB, SDRAM 32 MB, and RDRAM 8 MB (a single RDRAM device).

COST REDUCTION THROUGH INTEGRATION

Electronic products become less expensive every year. This is due to improving manufacturing yields of the electronic components and higher functional integration. The key to cost competitiveness in consumer products is taking advantage of increasing levels of IC integration.

Integration and unified memory architectures are complementary. Functions that require several chips, each with their own memory space, are becoming integrated into a single IC. When this is done the separate memories must also be integrated into a single space.

Integration of components has several benefits. Functions that completely reside in a single chip do not have to communicate with each other through I/O pins. An integrated device has a smaller total die area and fewer package pins than the same functions spread across several chips.

However, it may still have too many pins to be in a cost effective package if excess pins have to be used to get the needed bandwidth from the memory system. This again points to the benefit of high pin-bandwidth memory devices.

The cost benefits of integration can be very compelling. A consumer product that is similar in functionality and implementation to a set-top box is a video game console. In this marketplace there is an excellent example of the advantages of integration and unified memory architecture.

Sega Saturn

FIGURE 4 is a block diagram of a Sega Saturn. This is a 32 bit game console designed to support high performance 3D graphics.

The Saturn has a very distributed system architecture. There are several microprocessors, each with their own memory subsystem. The 3D graphics subsystem is spread across two chips, each of which is connected to its own private memory. Even the audio subsystem has a separate dedicated memory.

Because each of these memory systems is small, they are implemented using older technology 4 Mb DRAMs. This is much less cost effective than using current generation 16 Mb DRAM. In a distributed system such as this, it is impossible to use more cost effective higher density DRAM without increasing the total memory capacity tremendously.

Adding to the system cost is the large number of interconnects between the components. Several four layer printed circuit boards are required for connecting the devices together.

Nintendo 64™

FIGURE 5 is a block diagram of the Nintendo 64, a 64 bit game console designed, as was the Sega Saturn, to support high performance 3D graphics. The component integration level in the Nintendo 64 is substantially higher than in the Saturn design. Except for the game cartridge, the only memory in the system is 4 MB of Rambus DRAM - two devices.

The high bandwidth and low pin count interface of the RDRAM allow all of the 3D graphics, sound generation, and CPU control to be integrated into the Reality Coprocessor ASIC. The only other components in the system are the RISC CPU and some small glue chips.

The Reality Coprocessor in the Nintendo 64 provides the same functionality as the Saturn's two video display processors and audio processor. This high level of integration allows all of the memory that has been distributed in small pieces throughout the several subsystems in the Saturn to be collected into a single pool of Rambus DRAM. The Nintendo 64 takes advantage of 16 Mb DRAM technology for maximum cost effectiveness.

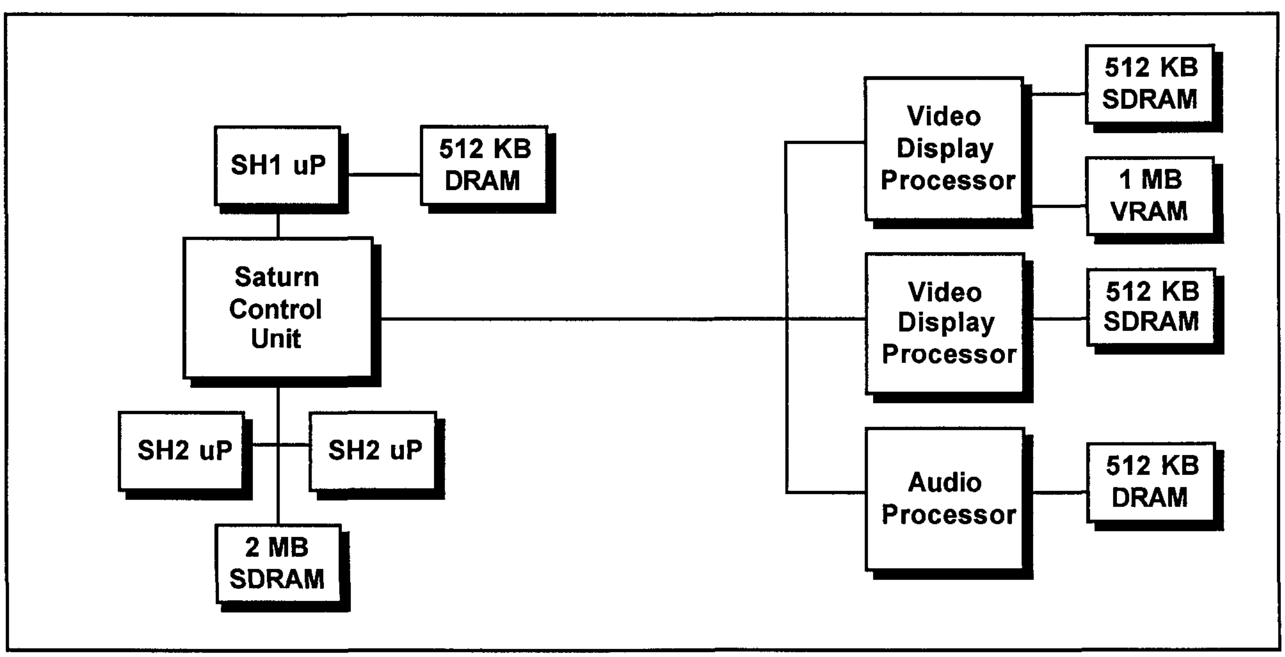


FIGURE 4. SEGA SATURN BLOCK DIAGRAM

The entire Nintendo system fits in a 6" x 6" form factor, which due to the simplicity of the design can be implemented on a single low cost two layer printed circuit board. The cost savings on the PC board alone is $5.00[1].

MINIMIZING SYSTEM COST

Consumer electronic products achieve cost reduction primarily through integration. This has two effects on system memory. The first is that there are physically fewer chips and pins to connect to the memory devices. The same amount of memory bandwidth therefore has to flow through fewer I/O connections. Second, the memory devices themselves become more integrated, packing more bits of memory into a single device. Again, the same amount of memory bandwidth has to flow through fewer I/O connections.

These two effects have a common result, that is fewer I/O pins connected to memory. This provides a cost savings, but can adversely affect system performance unless compensated for by using a higher bandwidth memory device.

The low individual component cost of standard 4 Mb DRAM can be deceiving. To minimize system cost in consumer products attempts should be made to take advantage of higher component integration levels and the lower cost per bit of 16 Mb DRAM. The high bandwidth Synchronous and Rambus DRAMs make such high levels of system integration technically feasible.

References:

[1] "The Ultra 64 Joypad", Interview with Genyo Takeda, "Next Generation" magazine, pp. 38-40, February 1996

Trademarks:

Nintendo 64™ is a trademark of Nintendo

Rambus™ is a trademark of Rambus Inc.

RDRAM® is a registered trademark of Rambus Inc.

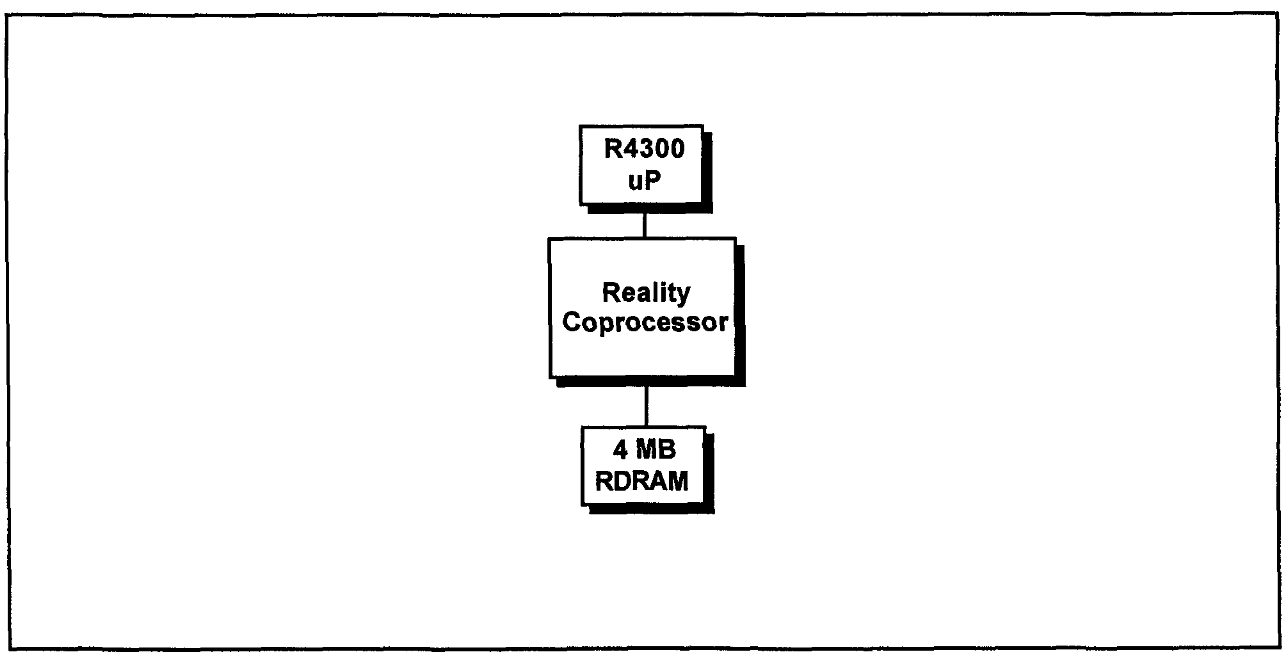


FIGURE 5. NINTENDO 64 BLOCK DIAGRAM

MORE EFFECTIVE OPERATION OF THE BROADBAND NETWORK

John C. Anderson, P. Eng., Chief Scientist, Project Engineering
Rogers Engineering
Nick Hamilton-Piercy, Senior Vice President Engineering and Technology
Rogers Cablesystems Limited

ABSTRACT

The traditional tree-and-branch one-way broadband networks have evolved into sophisticated hybrid-fiber-coax networks carrying advanced new interactive services. In order to realize the HFC network's full potential for increased quality and reliability, the operation and management practices of these networks need to evolve with them. Operational Support Systems (OSS) need to be put in place, concurrent with new work processes.

Rogers has gained significant experience from implementing an Integrated Network Management System (INMS) throughout its fiber and coaxial network. Rogers has also launched a high speed data service, WAVE, with enormous success, and gained considerable operational experience from the trial period. INMS was instrumental in the successful deployment of WAVE. This paper describes some of the initiatives and plans that Rogers Cablesystems is currently working on.

BACKGROUND

The cable television industry has been lagging behind other telecommunication carriers in adopting automation tools to operate and manage their networks. Since the inception of cable television in the 1950's, the industry has largely relied on the customer to be the network watch-dog and report any problems. Except for a few isolated ventures into status monitoring, few operators within the industry have any experience with comprehensive network management. Similar comments could be made regarding the operation of two-way systems. Computer-Aided Design and Drafting (CADD) systems have been in use for many years, but they have remained in the domain of the design and engineering departments, and consequently the information and data they contain has not been readily accessible to others. Computer-aided dispatch systems have been tried by a few operators, but not widely adopted as a customer service tool.

Most operators can share horror stories of uncoordinated construction and maintenance plant changes that remain undocumented and result in prolonged outages. Also typical are random interruptions of service for routine maintenance that go unreported and generate service calls. The recent frenzy of activity with high speed two-way data services for Internet access and work-from-home has raised the awareness within the cable TV industry that a new operations mindset and discipline are badly needed. A combination of competition in the marketplace that gives the consumer new choices, regulatory pressures for improved levels of service, and rapid technology changes are forcing operators to rethink the way they are operating their networks. New and innovative ways of managing the broadband network are required.

MANAGING THE NETWORK

Implementing network management systems is a first crucial step in this process. Software systems that manage network elements (i.e. trunk amps, fiber nodes, modems, etc.) are currently available from equipment vendors. The problem is that each manufacturer has developed proprietary protocols to communicate with their own devices, and standard interfaces between the systems do not exist. This does not mean that a "wait and see" attitude is necessary. At the very least, individual element management systems do provide a wealth of data about that part of the network. Different philosophies exist as to whether monitoring should extend to the ends-of-lines, to trunk amplifiers, or just to fiber hubs. Regardless of the decision, it is important to implement the element management systems early in network deployment. These systems will not only provide more immediate network information, but they also start to instill a new mindset about pro-active network management within the organization.

Most end-of-line and trunk status monitoring systems use two-way data modems to communicate with the amplifiers. The activation of these systems is not only a good training ground for two-way operation, but they also indicate the condition of the plant for carrying data. If the status monitoring system cannot operate reliably, it is almost certain that new data services will not operate as well. Once data services are deployed, status monitoring systems, and particularly those with bridger switching capabilities, are extremely useful in isolating return path problems and isolating faults so that they do not affect the rest of the network. Manpower savings using these tools can be substantial, and service restoration times are greatly improved.

New network management "umbrella" systems are also available which will facilitate the integration of several element management systems into a comprehensive network view. These systems also allow the scripting of user-defined "rules" that facilitate alarm correlation when numerous alarms flood in due to single fault. The "rules" sort through the alarms and report only one problem rather than a long list. "Expert" systems that "learn" from previous experiences are becoming available and will be able to shorten the diagnosis times for problems.

CADD systems typically provide only a graphic representation of the network and are simple tools to automate the drafting process. New Automated Mapping and Facilities Management (AM/FM) systems attach useful data to the graphic elements, and they provide a connectivity model of the network, from the primary hub out to each customer. This plant data can be sorted, queried, and dynamically linked to other systems. AM/FM technology, combined with network management systems, can provide a powerful set of tools to manage network performance.

MANAGING WORK PROCESSES

It is quite clear that the old "wait for the customer to call" scenario is no longer viable in the new competitive environment. Network management centers need to be created providing 24 hour, 7 day per week monitoring of the network. With modern software and network connectivity, these centers can be established regionally or even nationally. This network management center becomes the clearing house for all activity on the network. Not only equipment alarms, but physical changes, equipment configuration changes, network access security and network performance measurement fall under the responsibility of the network management center. Fault management starts with the detection of network events. These events are then correlated into a single fault which is then diagnosed and repaired. Centralized technical equipment experts can guide field crews to the specific cause of the problem and eliminate the tedious task of finding network problems from a truck. The network management system brings the problem to the expert, rather than dispersing scarce experts into the field.

During this process, network recovery may be taking place as redundant equipment or routes are switched into operation. Trouble ticketing and tracking systems log the sequence of events including the dispatching to field crews, right up to problem resolution and closure. Configuration and change management employs element management software and an outside plant database to log all reconfigurations, upgrades or physical plant changes during routine maintenance. Performance management is responsible for measuring key network indicators, compiling statistics, and generating reports of network performance against standards. Performance management is necessary to fine tune network operation and procedures. Security management involves the configuration and monitoring of security elements to prevent unauthorized network usage and restrict access to physical plant such as equipment enclosures and hub sites.

MANAGING NEW SERVICES

New services such as high speed data and video-on-demand are being deployed on microprocessor-based hardware with intelligent agents. While it is crucial that the physical network is managed effectively, it is equally important to manage the logical connections to the customer. In the case of high speed data services and Internet access this would include the file servers, bridges, routers, modems, that make up the connectivity with the customer. Similar functionality to physical network management is required, in order to manage alarms, configuration, performance and security. However, a whole new set of parameters need to be managed at the service level, including IP address allocation, session management, software distribution and version control, cable modem control, network usage metering, email ID allocation, bit-error testing, etc.

In the case of video-on-demand, the virtual channel to the customer needs to be established through coordination of video file servers, switches, channel modulators and box controllers. The session needs to be established and discontinued at predetermined times, with appropriate information sent to the billing system. It would be ludicrous to set up and bill for a video-on-demand session only to find out after the fact that the network feeding the customer had been down during the session.

Linkages between the physical network management system and the logical network management system must be established.

MANAGING INFORMATION

Work processes and software tools must be put in place to manage the gathering, storage, manipulation and access to network and customer information. A typical information relationship structure is shown in Figure 1.

Rogers Information Linkages

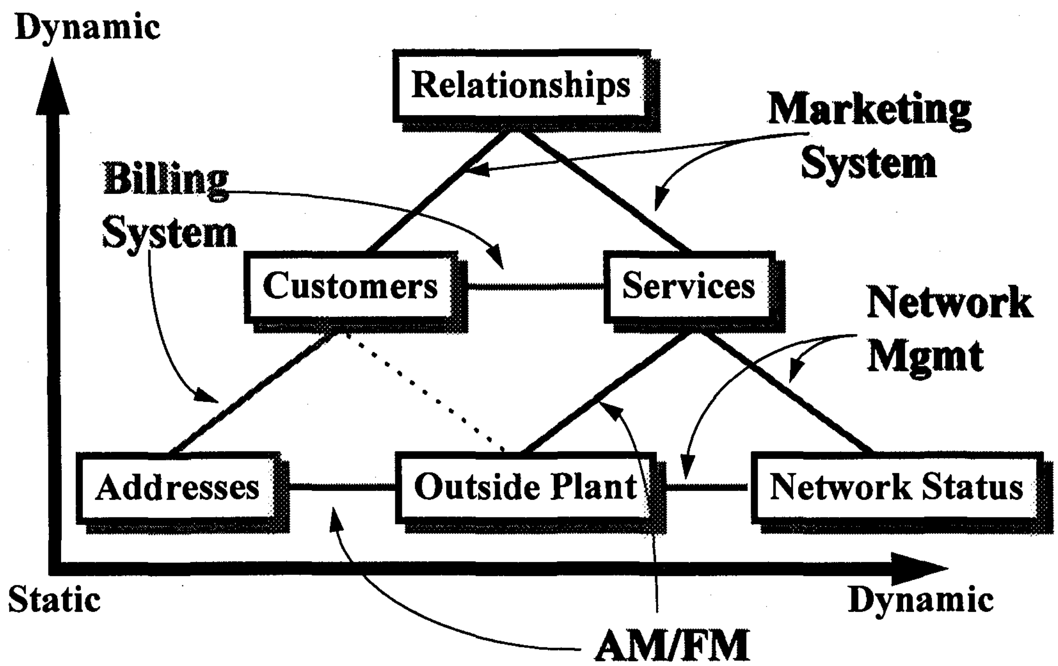


Figure 1

Different operational support systems are better suited to manage the different and diverse types of data. The key concepts are that data should be stored once, the data should be entered as close to its source as possible, and the data should be linked dynamically to the other systems that need it. Certain data such as customer addresses are very static and change infrequently, if at all. Customers occasionally change addresses, and physical outside plant changes occur due to rebuilds or maintenance. However, the most dynamic changes occur in network status and the provisioning of services delivered to customers. These are areas where real-time systems are needed to track and manage information. The ability to manage these data interrelationships effectively will translate into a significant competitive advantage in the deployment of new service offerings.

OPERATIONAL SUPPORT SYSTEMS

To support the work processes, a number of software tools will be required as depicted in Figure 2.

The tools will include a Network Management "Integrator" which bridges several element management systems into a comprehensive system. This is linked to the Workforce Management System which maintains an inventory of the workforce available at all times, the skill sets of each technician and the truck inventory of spare parts. The network management integrator has a facility for the user to script "rules" into the alarm messages. For example, if a string of sequential trunk amp alarms arrive at the same time as an alarm for their parent fiber node, the rule base would recognize the entire set of alarms as a fiber node outage and report it as such. The network management operator is buffered from the flood of alarms and presented with only one. Consequently, only one trouble ticket is opened and only one technician is dispatched directly to the problem. Without this capability, several technicians might have been dispatched to what appeared to be a number of different problems.

ROGERS CABLESYSTEMS OPERATIONS SUPPORT SYSTEMS ARCHITECTURE

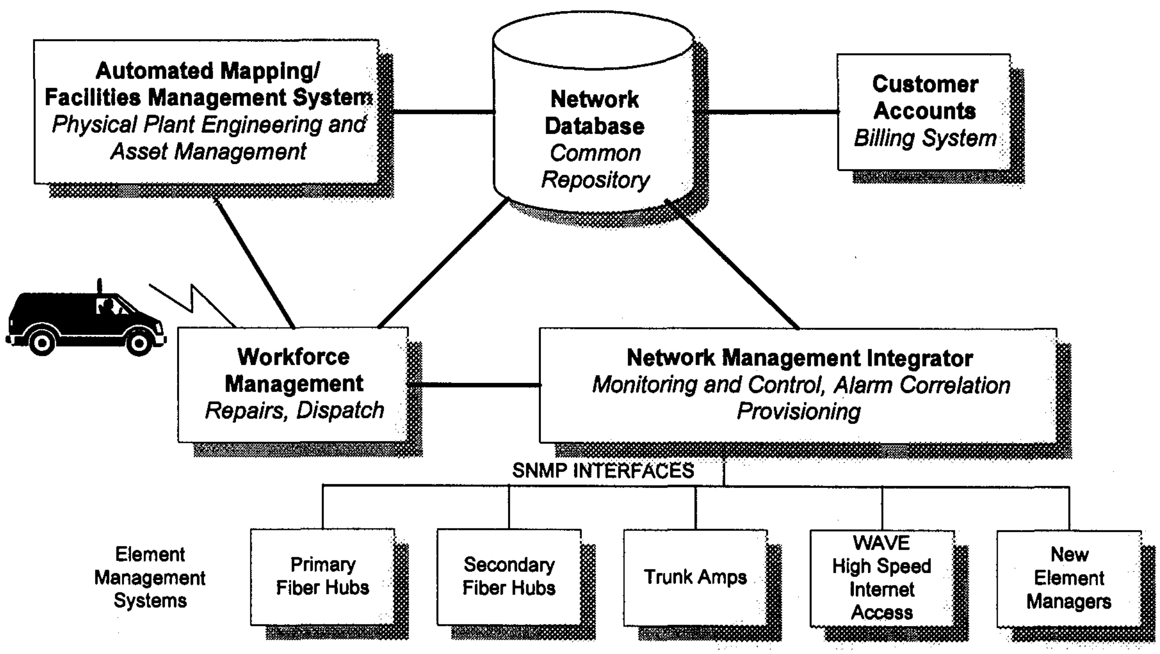


Figure 2.

As alarms are correlated and confirmed on the network management system, they are uploaded to the Workforce Management System. The appropriate fields are automatically populated on a new trouble ticket. The trouble ticket is assigned to the technician with the appropriate skills and spare parts who is closest to the problem. It is then dispatched electronically to the vehicle and appears on a laptop computer. The technician can query the network management system from his laptop for further information or access the facilities management system for schematics or wiring layouts. As the technician proceeds through the repair task, he updates the system regularly, and each update is automatically time stamped and reported back to the dispatch system. If the fault is not repaired within a predetermined time, the problem is escalated to a supervisor for intervention to determine if additional help or resources are required. Once the repair is completed, the entire sequence of transactions are stored in the database for future analysis and reporting. Equipment failure rates will be analyzed, and a knowledge base is built up based on past repair experience. The next time that similar problems occur, technicians can query the database for previous fixes and speed up repair time. Any changes that are made in the physical plant, either temporary or permanent, are uploaded to the AM/FM system and logged. This eliminates the problem of "phantom" changes to the network since all occurrences are logged, dated and time stamped. While outages are in progress, the customer database is updated with the affected equipment locations. Any customer inquiries during the outage can be dealt with efficiently since the CSRs are fully aware of the outside plant status at all times.

"RETHINKING" THE WAY WE MANAGE THE BROADBAND NETWORK

It is clear that if we are to offer new and innovative services, with better quality and reliability than ever before, the industry must evolve to a more pro-active customer-centric approach to managing the broadband network. A number of initiatives need to be taken to achieve this vision:

- provision network equipment with monitoring and control functionality
- initiate centralized network management centers that operate 7 days a week, 24 hours per day
- deploy new services with terminal devices (i.e. modems, DVC boxes) that can be remotely managed and configured under software control
- "re-engineer" work processes to layer on new operations support technology and to look for weaknesses in currently practices
- plan for and build an information infrastructure to support the new work processes
- start training both field and office staff to think like telecommunications providers

The provisioning and delivery of high speed data services, video-on-demand, and telephony will demand that the broadband network be operated and managed with more sophisticated surveillance, workforce management, and database tools. Operators need to start long range planning for these tools and, more importantly, start deploying the infrastructure for them now. The software and hardware systems are a challenge in themselves, but the bigger challenge will be the establishment of new work processes and attitudes within the workforce that will make the systems pay off in better levels of quality and reliability.

MPEG STANDARD CONFORMANCE TESTING FOR INTEROPERATILITY

Mukta Kar, Majid Chelehmal, Rich Prodan
Cable Television Laboratories, Inc.

Abstract

An important objective of the MPEG standardization effort is to define a set of video and audio compression techniques and to increase interoperability among various digital audio/ video devices made by different manufacturers. To further this effort, the MPEG committee has defined a set of normative compliance requirements in a part of the MPEG standard which sets forth procedures and bitstreams to test MPEG-2 related products.

Many manufacturers, particularly small ones, do not have enough resources to test their product thoroughly for compliance. To help cable operators and the cable industry, Cable Television Laboratories, Inc. (CableLabs®) has set up a conformance testing laboratory. These efforts should minimize concern about incompatibility among MPEG products and should assist in the development of a multivendor base of interoperable equipment.

This paper reviews some of the background of the MPEG-2 Standard and its compliance requirements. Discussion about non-realtime and realtime conformance testing is also included. The current conformance testing capabilities at CableLabs and the results of bitstream analysis from several manufactures realtime encoding system are discussed.

INTRODUCTION

As digital information delivery to the consumer is becoming reality, millions of devices will be interconnected to handle telecommunications and entertainment services. Interoperability among the equipment of various manufacturers will increase the speed of deployment while decreasing cost. For this reason, it is particularly important for cable companies to adopt interoperable standards.

In the wake of the Telecommunication Deregulation Bill recently passed by Congress, one can expect competition in the future encompassing various service providers. Without sufficient interoperability among equipment, the opportunities for timely, cost-effective mass deployment of video services may be impeded in a highly competitive marketplace.

Interoperability is a key issue. It helps stimulate strong competition for product innovation in the marketplace. Competition ensures more reliable products at lower prices. With the tremendous advances in hardware and software technology, the analog video and audio equipment makers are changing to digital design. In the analog world, most of the audio/video equipment manufactured by various vendors was not interoperable. Some of the audio/video equipment (e.g., analog settop boxes) cannot be moved and plugged into different cable systems in different regions of the country.

The interoperability of digital television utilizing the MPEG-2 audio-visual compression standard is examined in this paper. A brief history and overview of this standard is provided. The role of conformance testing to determine compliance to the definition of the standard and to allow interoperability at the baseband compressed program multiplex level is discussed. The MPEG-2 conformance laboratory established at CableLabs facilities is described. Some results of conformance testing of several vendors MPEG video encoder are provided. Future work and conclusions are summarized.

Compression Standards

To promote interoperability, equipment manufacturers, broadcasters, cable operators and government entities have expressed a desire for broader standardization. Various standardization committees have been formed under the administrative umbrella of the International Organization for Standardization (ISO). These standardization committees are know as JPEG (Joint Picture Experts Group) and MPEG (Moving Picture Experts Group).

The JPEG committee created a document for compressing still pictures. A picture can be compressed and decompressed without any reference to other pictures. So a picture can be compressed in non-realtime and then be stored or transmitted. The first document from the MPEG Committee is known as MPEG-1. MPEG-1 is optimized for moving pictures of SIF (352X240) format with a bitrate of about 1.5 Mbits/s used in single program application (e.g., CD ROM). Later, the MPEG committee extended its MPEG standardization work to include CCIR-601 broadcast quality images. The extended standard is known as MPEG-2. Also, it added a new system layer including a transport stream layer for realtime delivery of multiple programs over error-prone channels. Using MPEG-2 compression and transport syntaxes, interoperability among equipment of various vendors is possible. This paper deals with the MPEG-2 standard and the conformance testing of encoded bitstreams.

MPEG-2 OVERVIEW

In the second phase of work, the main goal of the MPEG committee was to extend the MPEG-1 standard such that it could be applied to a wide class of applications involving audio, video and other data with broadcast resolution and bitrates up to 15 Mbits per second. Major requirements were (a) forward compatibility with MPEG-1, (b) good picture quality, (c) flexibility in input format, (d) random access capability, (e) inclusion of VCR controls (fast forward, reverse play and slow motion), (f) bitstream scalability, (g) low delay for two-way isochrouous communication, and (h) the capability of multiplexing several programs in a single bitstream with resilience to transmission errors. Bitstream scalability is defined as the ability to discard a portion of the bitstream, but decoding the remaining bitstream will produce a recognizable picture of lower quality. Soon it was realized that a bitrate restriction of 15 Mbits/s is not required and the MPEG-2 standard can be applied to higher bitrate applications such as high-definition television.

The MPEG-2 standard is divided into three principle parts; video compression, audio compression, and system multiplexing and transport. MPEG-2 is basically a collection of tools defined in such a way as to satisfy the requirements of a wide variety of applications. All the tools are not needed for a specific application. A subset of tools is enough to implement a particular application while maintaining full interoperability among the equipment manufactured by various vendors that support that application.

MPEG video compression is based on discrete cosine transform with motion compensation (DCT-MC) algorithms. DCT is a block orthogonal transform which reduces spatial redundancies inherent in a picture by compacting most of the energy into lower order coefficients of a block. Coefficients are quantized and higher order terms are truncated. The technique of motion compensation provides further compression by taking advantage of temporal redundancies of successive frames. Only motion compensated (spatially offset) differences between pictures are transformed. The resulting motion vectors and quantized DCT coefficients are variable length coded using entropy coding including Huffman coding and run-length coding.

The MPEG video coding support is divided into profiles and levels. A profile is defined as a subset of the entire bitstream syntax. Within a profile, there could be various levels (picture resolutions). There are five profiles: simple, main, SNR scalable, spatially scalable, and high profile. There are also four levels: low (SIF), main (CCIR601), high-1440, and high (HDTV resolution). Using the combination of these profiles and levels, 11 different types of decoding are defined. A decoder could have one or more of these capabilities. MPEG audio compression uses the MUSICAM (masking pattern universal sub-band integrated coding and multiplexing) algorithm and supports mono, stereo, and surround quality audio.

In the MPEG-2 system document, the syntax and semantics are defined for multiplexing multiple programs in a single bitstream. In the system/transport stream layer, bitstreams are split into 188 byte packets (184 byte payload and 4 byte header) as shown in Figure 1. The header carries various information fields. The PID (packet identifier) is the most important one. The PID is a unique integer number associated with an elementary stream in a single or multi-program transport stream. The packets can carry various video and audio channels and other information, such as synchronization and timing, encryption, program information, access control, etc. The PID numbers help sort the packets into specific streams to which they belong.

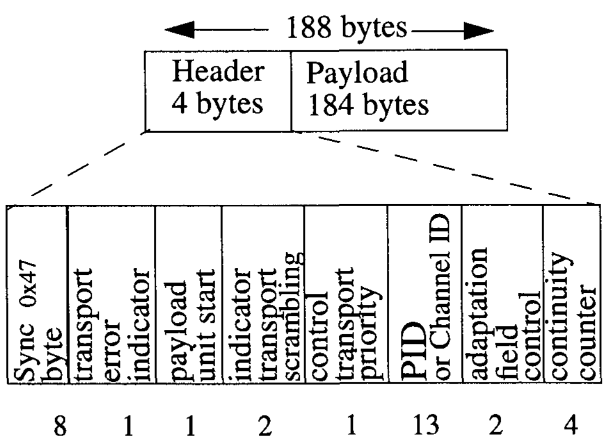


Fig. 1 MPEG-2 Transport Packet

Conformance Testing

Interoperability is the main goal behind creation of MPEG standard documents. It is assumed that equipment designed complying with the rules stipulated in the three documents (audio, video, and systems) will be interoperable, at these levels.

The MPEG-2 committee realized there were questions on how to test compliance and the degree of compliance, so it added a fourth document known as Part 4, Compliance. Part 4 specifies tests for ensuring compliance with video, audio, and systems specifications. The document recommends that a number of bitstreams be used for the test. These bitstream files have been carefully designed to test various characteristics of MPEG equipment. They will be especially useful for testing settop (decoder) boxes. These bitstream files are stored in various FTP (file transfer protocol) sites around the country and can be downloaded by anyone using FTP or similar software. CableLabs is an active participant in the MPEG standardization process and has provided the FTP site for system-related bitstream files.

Equipment manufacturers designing MPEG-2 related equipment (encoders, decoders, etc.) should test their equipment using the procedures specified in Part 4 so that they can specify their equipment as MPEG-2 compliant. Many companies do not have sufficient resources to test their equipment extensively. Such equipment may run into interoperability problems. This situation could be avoided by submitting equipment to a neutral laboratory that would test MPEG-2 equipment and check the degree of compliance. To assist cable television operators in evaluating MPEG-2 equipment, CableLabs has set up a conformance testing laboratory to test encoders and decoders. The main goal of conformance testing is to broaden the scope of interoperability among the equipment produced by various vendors.

CableLabs plans to test bitstreams both in non-realtime and in realtime. A set of software tools has been designed and developed to analyze bitstreams captured from an encoding system (video encoder and system multiplexer). A schematic block diagram of the hardware configuration for the CableLabs non-realtime conformance verifier is shown in Figure 2. A set of software analysis tools checks the syntax and semantics of such bitstreams at the system and video layers. The second approach employs realtime testing of encoders and decoders using specific hardware. Currently, video bitstreams can be stored over a short time only (several seconds, maximum). If an encoding system generates bitstreams with infrequent problems, observing such problems using a non-realtime method will be difficult. In that case, the encoder bitstream has to be analyzed in realtime over a considerable time interval.

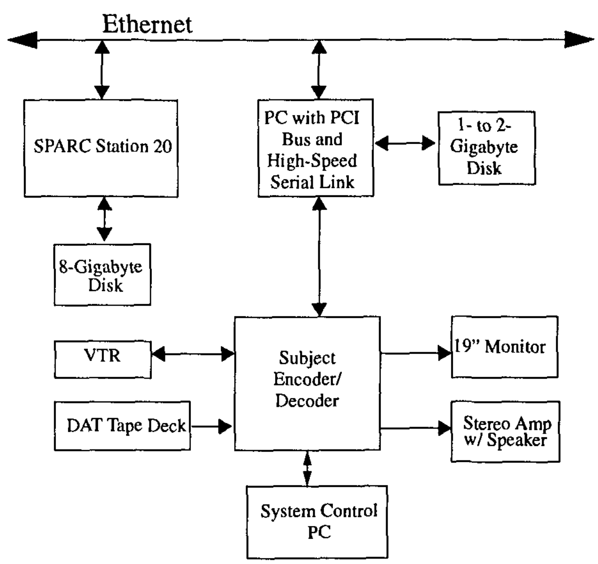


Fig. 2: Non-realtime Conformance Verifier

Non-Real-Time Conformance Verifier

At present, the non-realtime version of the conformance verifier is operational. A brief description of various components shown in Figure 2 is given below.

A Unix SPARC Station 20 serves as the main computer for the various MPEG-2 software tools used to create and to analyze MPEG-2 bitstreams. The 8-Gigabyte disk stores large bitstream files as captured from encoders or provides input to decoders. A PC has a PCI bus, a fast Ethernet card, and a specialized serial card. This serial card is used to send/receive data to or from an encoder or decoder, respectively. The 1-2 Gigabyte hard disk, connected to the PC provides a fast SCSI link for fast storage/retrieval of data. The SPARC-20 and PC are connected to Ethernet for networking. The VTR is digital D-5 or a D-1 tape recorder, or could be a video realtime disk. A DAT tape deck is used as the audio source material.

In a non-realtime configuration, a bitstream from a realtime encoder is captured using the PC with a high-speed serial link and stored on the hard disk. Then the bitstream is transferred to the SPARC-20 and analyzed using a set of conformance checking tools. The software tools are designed such that each tool reads a parameter file where various flags, such as buffer-verification, PSI (program specific information) checking, etc., are set to desired values. The analysis output can be displayed on the screen and be logged into a file at the same time, if desired.

The software tools have been primarily divided into two parts: transport stream multiplexer/demultiplexer and video elementary stream encoder/decoder. The demultiplexer tool reads PSI (Program Specific Information) from the transport stream and separates the elementary streams into individual stream files for video, audio, and private data. The video decoder tool checks the syntaxes and semantics of the video elementary stream file while reconstructing the picture. It also displays a frame once it is complete, and can send the decoded frames to other devices.

There are a significant number of syntax elements and parameter values to be checked in a transport stream. CableLabs demultiplexer checks about 300 syntax elements, and important parameters like buffer level verifications for all three buffers specified for T-STD model, bitrate, PCR frequency, etc., applicable to MP@ML. The demultiplexing tool has the capability of reading PSI information from a bitstream or can read a user-supplied one (if the stream has no PSI or erroneous PSI). Features are incorporated such as setting initial offset into the bitstream, logging program output to a file for later investigation, information display level for diagnostic output level of detail, etc. The video decoder checks 73 important syntax elements including timing parameters (validity of Presentation Time Stamps [PTS] and Decoding Time Stamps [DTS]), motion vector search ranges, valid temporal references, etc. The decoder handles both Main Profile and Simple Profile at Main Level, including Dual-Prime prediction mode. Additional software tools have been created to support conformance test activities. For example, the software video encoder and transport stream multiplexer may be used to create transport stream.

CableLabs has been performing conformance testing of bitstreams from a number of encoder vendors. Errors and warnings are tabulated and provided to the vendor company for appropriate action to correct the problems. The bitstreams captured from several realtime encoders under test have been analyzed and sample results presented below.

The conformance verifying decoder found the following problems in several bitstreams: 1) In the Dual Prime mode, forward prediction vectors were not pointing to the adjacent previous picture. 2) In the field prediction mode for a B-picture, the motion vectors violated the picture boundary limit. 3) Video access unit (AU) was not aligned with the PES packet in some transport stream packets. 4) The continuity counter did not increment in some transport packets. 5) The PTS time-stamp value was different from the expected value.

Future Research

With the current conformance testing software tools, syntax and semantics can be checked methodically in non-realtime. But if problems occur infrequently or intermittently, much time and storage will be required for post-analysis. A realtime verifier would alleviate these bottlenecks in such a situation.

CableLabs is designing and developing a realtime conformance testing tool. The realtime verifier will not make the previously developed non-realtime software tools obsolete. Checking syntax and semantics in detail is easier to do in non-realtime. A thorough checking of syntax and semantics will be very computationally intensive. Hence, we believe that realtime and non-realtime tools will complement each other. Filtering bitstream violations using hardware without detailed analysis will greatly reduce this computational burden.

Conclusion

Digital television utilizing the MPEG-2 audio-visual compression standard facilitates interoperability of equipment for cable distribution of entertainment programs. The CableLabs conformance testing laboratory has been established to foster a multiple vendor universe of interoperable products. The current non-realtime and future realtime components can assist in solving ambiguities and verify adherence to the established MPEG-2 standard, as well as industry established standards unique to the cable environment.

References

1. CODING OF MOVING PICTURES AND ASSOCIATED AUDIO, ISO/IEC JTC1/ SC29/WG11, N0742 (Compliance), N0801 (Systems), N0802 (Video) and N0803 (Audio).

2. Bhaskaran, V. and Konstantinides, K., "IMAGE AND VIDEO COMPRESSION STANDARDS, Algorithms and Architectures," Kluwer Academic Publishers, Boston, 1995.

3. Stubbs Graham S. and Wechselberger, Tony, "DIGITAL COMPRESSION - INTEROPERABILITY FOR THE CABLE INDUSTRY", - paper presented at NCTA conference, May 1994.

4. CableLabs internal documents on conformance testing.

The authors are employees of Cable Television Laboratories Inc., Louisville, Colorado. The conformance tools have been jointly developed with DiviCom Corporation of Milpitas, California.

MPEG-2 and SONET

Michel Fortier, Ph. D., eng.
Northern Telecom,
9300 TransCanada Hwy, St. Laurent, PQ, H4S 1K5 (Canada)
Tel.: (514) 956-3473 ♦ E-mail: michel.fortier@nt.com

Abstract

MPEG-2 can be used in many services such as video trunking, pay-per-view (PPV) and Near Video on Demand (NVOD). As MPEG-2 technology becomes commonplace in the cable headend, SONET technology plays a key role in backbone transport of digital content.

With SONET, MPEG-2 can be transported using ATM or non-ATM modes. Both in the core and access network, SONET equipment will be standard. End-to-end connectivity is possible on SONET networks that support both modes. Thus SONET can be used as the transport technology of choice to carry MPEG-2 to interconnect service providers to users with exciting new services.

1 - Introduction

Over recent years, the means for handling and delivering video signals has changed dramatically. Digital video signal processing is now common practice, and is part and parcel of all aspects of producing video programs. Until recently, programs to be used in distribution have been kept in uncompressed format, analog or digital. However, it is now possible to store video programs in digital compressed form. The degree of compression is usually based on the end use - either for long term storage and use in future productions, in which case the resulting picture quality must be of high quality and the compression ratio lower, or for use in server-based services where storage and eventual bandwidth delivery requirements will put constraints on the compression level, and hence, the resulting picture quality.[1]

In today's digital world, a gamut of services exist:

- Cable distribution — In upgrading networks, the cable operators are looking into digital access and distribution technologies to enhance the quality and robustness of the signals to be delivered over long distances, as well as investigating video compression technologies to enable new delivery mechanisms.
- Telco video distribution — With the new regulation in place allowing the telcos to provide service, operators are looking at ways to use broadband channels to carry digitized video signals in their networks.
- Video trunking — Whether for planned services, or for "just in time" services, video networks are being deployed to provide points of presence in many venues (studios, sports centers, convention centers, etc.) which allow the network operator to haul the video signals either to the processing centers, i.e., the broadcast studios, or to other end-points.
- Conversational video services — These services have been addressed until now by H.261 codecs and the higher bit-rate DS3 codecs. Cost and picture quality are of prime concern. MPEG-2 offers opportunities of providing better quality pictures with reasonable bandwidth for premium video services.

Accessibility to the network and the transport mechanisms are the key factors for the operators to establish networks which achieve end-to-end connectivity for video signals in their many formats. This paper presents a view of this connectivity from the perspective of SONET because of its emerging pervasiness as a standard in digital telecommunications networks. The discussion will also show the interworking with digital video, with particular emphasis on MPEG-2 video.

1 Note that the decoded picture quality can be quite good; degradations can even be undetectable. This will depend on the amount of compression used.

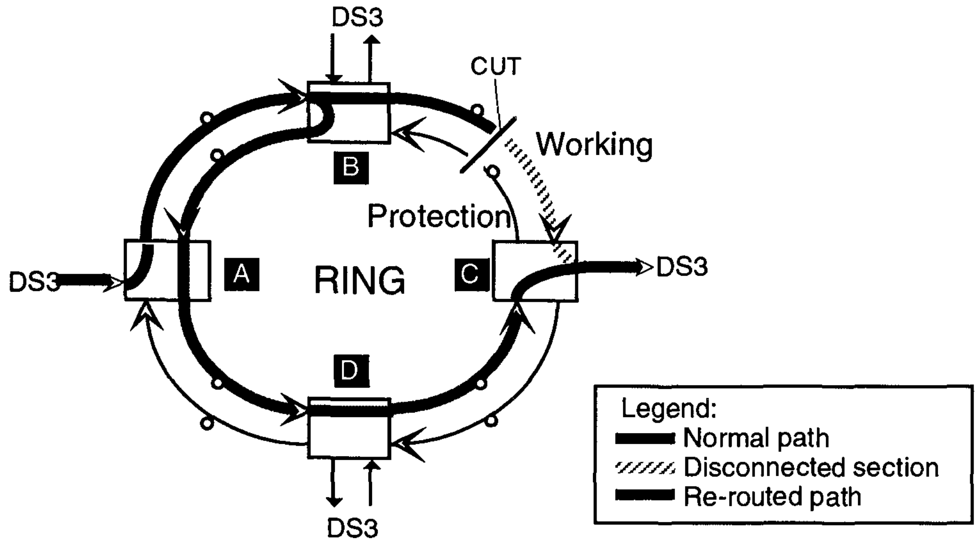


Figure 1 – Example of Bi-directional Line Switched Ring

2 - End-to-end Connectivity What does SONET offer?

SONET (Synchronous Optical NETwork) is a digital optical transmission infrastructure which has a hierarchy of transmission speeds, or bit-rates – i.e., OC1, OC3, OC12, OC48, OC192. It supports many kinds of channel protocols at the data level. At the physical layer, ATM uses SONET transport at various bit-rates, e.g., OC3c, OC12c, etc. Furthermore, DS3, DS1, and DS0 signals can be mapped into the SONET infrastructure.

Within the SONET infrastructure, there are built-in mechanisms to perform network management as well as carry the related commands using standardized protocols. This network management covers all aspects of the OAM & P process.

SONET being an optical and digital network, the transmission is inherently robust and it can be carried over much longer distances when compared to analog circuits. Moreover, extra fibers can be laid in for redundancy and the reliability of the network can be made very high. The SONET equipment provides for switching time slots on the terminal points, and automatic reconfiguration on failures of ring architectures. For instance, using techniques like bi-directional line switched rings, the network can be made tolerant to single fiber cuts. When activated, this technique re-routes traffic from normal working links to protection links which carries data around the ring in the opposite direction to reach the otherwise disconnected network element. This is shown in Figure 1 where traffic was going from node A to node C. A cut between nodes B and C results in traffic then going from node B to node A via the protection path, then to node D, and finally to node C.

This type of network is becoming all widely accepted in the world of telecommunications. The transport hierarchy equivalent to SONET in countries other than Canada, the United States, Japan, Korea and Taiwan is called SDH (Synchronous Digital Hierarchy). SONET and SDH are compatible with the exception of minor channel maintenance differences. Hence, video signals can then be distributed across continents, notwithstanding the issue of transcoding the video signals from one standard to another, whether compressed or not.

3 - MPEG-2 technology – Technical issues

Most transport of digitized video signals requires compression at some point. This stems from the need to optimize investment in back-bone infrastructures. MPEG-2 is a compression technology which can be applied in many parts of a distribution system. Its applicability will be governed by the user's quality of service criteria. Key technical issues are:

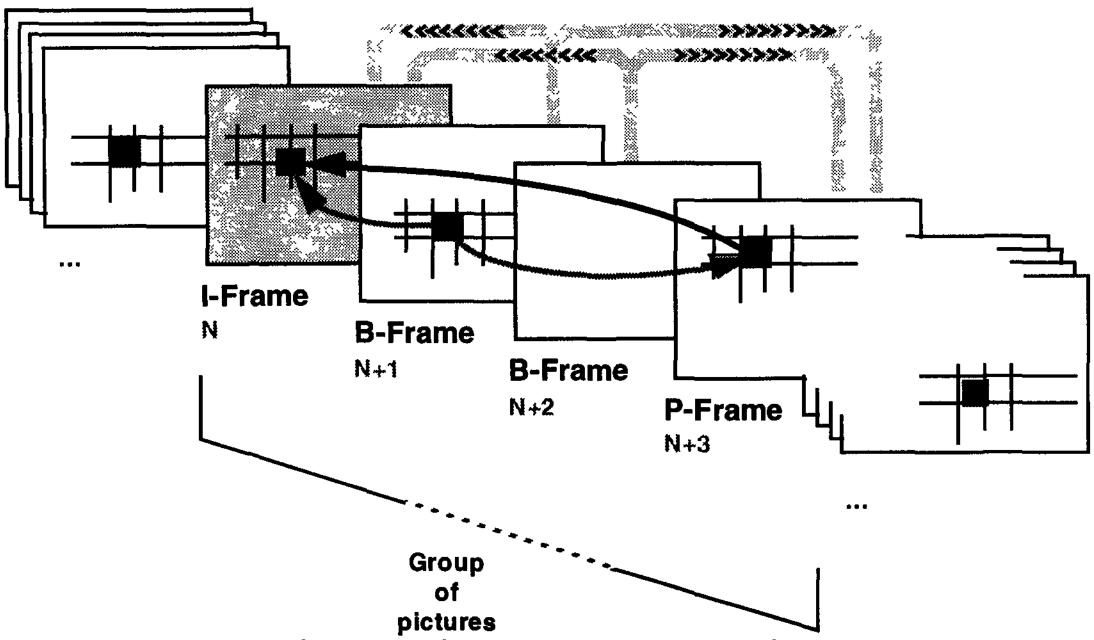


Figure 2 - I, B, P frame relationships

- Bandwidth versus Picture Quality
  MPEG-2 spans a range of bit-rates between 3 and 15 Mb/s. Picture quality in this range can vary between average to very good, and this will depend to some extent on the types of pictures being compressed (e.g., sports scenes, "regular" movies, training material, etc.). Typically, many MPEG-2 streams use the 4-9 Mb/s range, where 4-6 Mb/s may be used for regular and training material, and the higher end, 6-9 Mb/s, will be used for sports, as well as advertising material for insertion into programs.

- End-to-end Encoding-Decoding Latency
  In one-way video services, this aspect is not usually a factor. This latency appears as a combination of the codec's (coder-decoder) use of an IBP coding mode along with the so-called video buffer found in the codecs. The IBP mode refers to the way coded pictures are referenced. This process can be easily visualized with Figure 2.

  I-frames, or intracoded frames, refer to frames that use only information within the current frame being encoded. P-frames, or predicted frames, are differentially coded frames using prediction from previous frames in the coding process which may be I- or P-based frames. This prediction uses motion estimation to optimize the encoding. The B-frames, or bi-directional frames, operate on bi-directional interpolation. This process uses reference information in past and "future" frames to form the best interpolation possible. The overall technique is such that I-frames when coded produce more bits than P-frames which, in turn, produce more bits than B-frames. As a rule of thumb, the I:P:B ratio can be seen as 20:5-to-10:1 but it is far from static. Factors affecting the bits-per-frame production include scene changes, highly textured areas, moving object, etc.

The combined use of all frame types produces the best compression (e.g., the lowest bit-rates) for a given decoded picture quality. However, the IBP mode produces the greatest delay. In general, there are two B-frames between reference I-P, P-P, or P-I frames which accounts for an intrinsic delay of two frames; from the point of reference of a B-frame, the future I- or P-frame must be delayed appropriately before it can be used. This is then combined with the video buffer which must be filled to an operating level to avoid overflows/underflows. The size of the buffer translates into an equivalent delay at a given bit-rate. This can be of the order of ~200 ms. This is further compounded by the video processing delays which can then produce overall delays that are much longer than 200 ms. — in some cases several hundred milliseconds.

As indicated before, one-way video services would not have problems with this type of delay other than synchronization issues where multiple streams need to be integrated from real-time sources. Again, this would be taken care of via buffering mechanisms. For two-way conversational video services, I-frame only coding as well as IP-frame coding provide the best delays because they do not need information from "future" frames and the video buffers can be made to introduce as little delay as possible.

MPEG-2 has techniques for timing recovery to help reconstruct the stream at its original bit-rate. This allows for display and synchronization of the far end operating frequency to be linked back to the source. This timing recovery can be done in ATM and non-ATM working environments.

An MPEG-2 Bitstream can contain audio, video, and other ancillary streams. These are initially formed into packetized elementary streams (PES). Then, a multiplexer engine creates a stream from these PES streams in one of two formats: program streams (PS) and transport streams (TS). The PS format is to be used in "error-free" contexts while the TS format is to be used mostly for error prone environments (e.g., SONET transport or ATM). However, there are no specific rules in their use, only their intrinsic adaptation to certain environments.

While it is true that the TS format and AAL5 can be used in ATM environments, other means are possible. Carriage of streams can also be done using DS3/E3-level signals. Unfortunately, the standards bodies have not addressed such transmission channels. Users will want interconnecting networks, i.e., streams coming into, and multiplexed by one vendor's equipment into a DS3 for instance, should be capable of being demultiplexed by another vendor's equipment. Flexible multiplexing and demultiplexing will be required.

Nevertheless, SONET inherently transports either mode of transfer, non-ATM and ATM. The ATM mode offers advantages with respect to giving operators the possibility of transferring streams composed of single or multiple program streams into one transport stream. With the latter, the end-user can switch from one program to another within the same transport stream. The user can also switch from one transport stream to another. Non-ATM transfers can be used in point-to-point connections, and also in drop-and-continue scenarios. Point-to-point refers to users who have a link established with a remote point, whether permanent or temporary. Drop-and-continue refers to the capabilities offered on SONET rings to drop the signal at the network element and transport the same signal on the ring to be dropped at the next network element, and so on. This capability offers a mechanism for broadcasting around the ring. Resources can also be reallocated as needed with switching functions. The maximum number of network elements in the ring is sixteen to preserve switching times in case of failure as set by Bellcore standards.

4 - MPEG-2 and SONET

Integrated video networks – The bandwidths of channels required to transport data of all kinds are becoming larger. Voice, data, and video networks can be merged together and managed as one, whether internally the channels are mixed or segregated. Hence, the network operator need not contend with overlay networks which, in turn, provides efficiencies, e.g., economies of scale. All services are carried on one SONET network therefore reaping the benefits of economies of scale, maintenance, staff training, system integration, and others. MPEG-2 sources can therefore be distributed across networks in a transparent fashion. This applies whether transporting PS or TS formats in non-ATM or TS in ATM environments.

• Hierarchical Accesses to Servers

In setting up networks for services like video on demand (VOD), or near video on demand (NVOD), or access to multi-media servers, a hierarchy will be set up in the network. Let us address one view of a NVOD scenario. In the example shown in Figure 3, a two-level hierarchy is displayed. In the first level, source material can be gathered and stored on global servers, holding program and advertising contents which are to be used in later programming. In the second level, the local servers get data from the global servers and are replenished according to scheduled downloads which provide the contents for daily programming. The output of the servers goes to the home via the RF distribution plant in one of two ways:

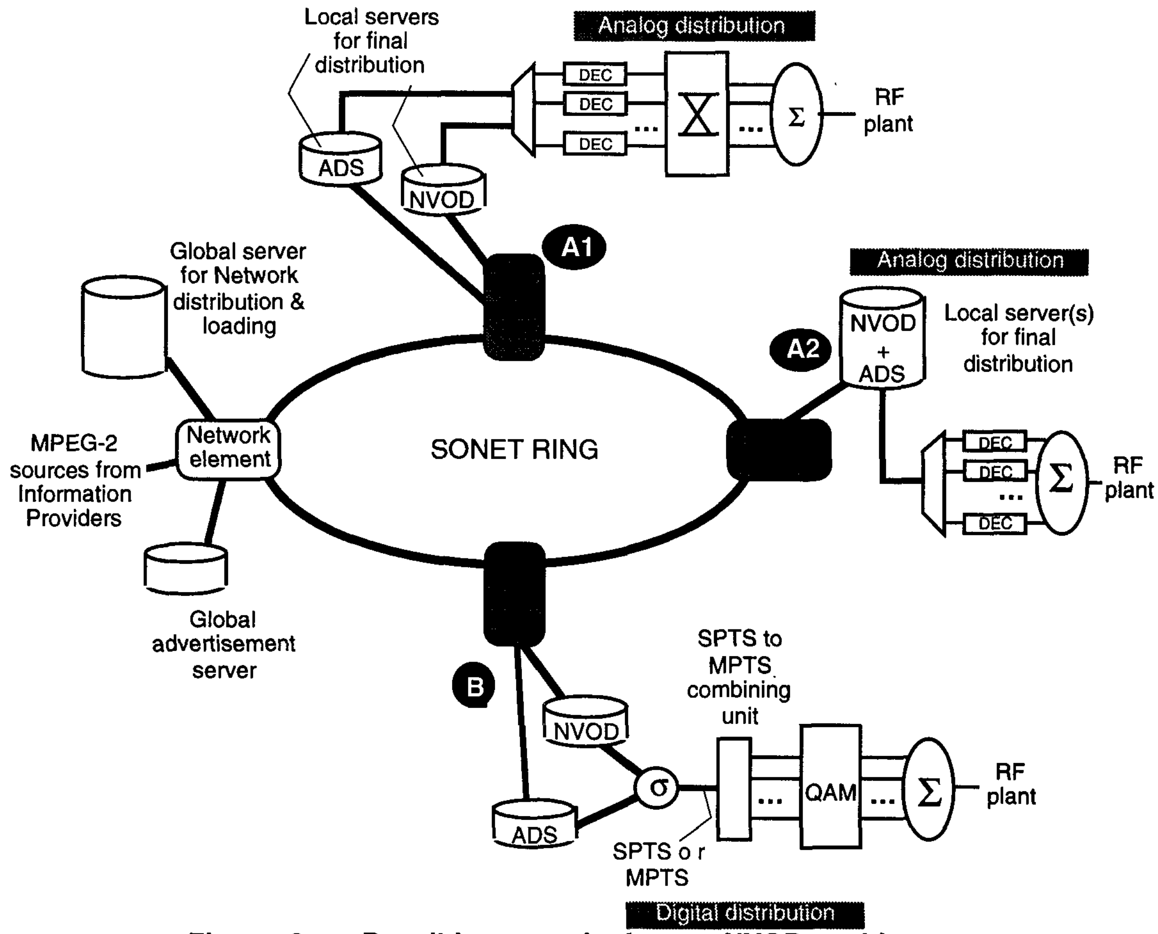


Figure 3 — Possible scenario for an NVOD architecture

A - analog video signal to the set-top box, or

B - a digital stream to the set-top box.

Method A can be done in two ways. One, shown by the reference mark A1, would be to have separate servers for the program and the ads. Both servers send data to a demultiplexer for distribution to a bank of decoders. The decoders' outputs then go through a switch to assign video signals to a modulator before going to the RF plant. In this way, a program runs until a commercial needs to be inserted. Then, the appropriate signal from the commercials' server is routed by the switch to the right modulator.

The second method, marked A2, stores the programs and the ads on the same server, eliminating the need for an analog video switch. In this technique, a program runs until the time for the commercial. Then, the server is instructed to get the appropriate file for the ad.

In Method B, similar scenarios can be used, separate or combined program/ad servers. The difference is that the streams are first put through a stream combining unit which multiplexes multiple MPEG-2 streams to form a multiple program transport stream. The output of these stream multiplexer units are used as input to QAM modulators before entering the RF plant, i.e., the streams are sent in digital format to the end-user. At the user's home, a set-top box receives the QAM-formatted signal, extracts the bitstream, and then, demultiplexes the appropriate program for decoding by the set-top box decoder.

5 - Interworking of non-ATM and ATM video networks

Presently, there is a need for interworking scenarios between ATM and non-ATM networks. The current networks have not all converted to ATM operations and there will be a transitory phase where streams will either originate in non-ATM format or will cross network boundaries.

Two kinds of migration scenarios can exist to migrate from non-ATM to ATM transport: (i) using ATM to emulate non-ATM circuits, and (ii) taking the non-ATM bitstream into an adaptation interface to convert it into ATM cells. In the first case, the SONET ATM network transports the constant bit-rate data without mapping the MPEG-2 stream into ATM. At the other end of the network, it then reconstitutes the constant bit-rate stream. In the second case, an MPEG-2 bitstream is converted into ATM cells using an AAL5 protocol. At the other end, the cells' payloads are demultiplexed to form a bitstream to be delivered to the end user, or to continue on in the network.

A method is also needed to provide ATM to non-ATM adaptation between networks. Transporting ATM cells via DS3 UNI circuits is one method to adapt between networks. At the receiving end, the cells can then be extracted and put back into an ATM network to continue on to their destination. An alternative is for the cell's payload to be converted with interface units into a constant bit-rate stream. This alternative can only be utilized if the bit-rates selected for transport on the ATM network are well adapted to the bit-rates used in the non-ATM networks. In other words, the non-ATM network may well use interfaces for MPEG streams that handle only specific bit-rates, for instance based on DS2 granularities. Then there will issues in trying to convert an MPEG-2 stream at say 4.5 Mb/s. Either, the interfaces will not accept such bit-rates, or stuffing bits will be needed and a varying degree of network inefficiency will result depending on the proportion of stuffed bits relative to the payload. Flexible interfaces will be required, but more importantly, a standardized way of multiplexing into one or various non-ATM networks will be needed.

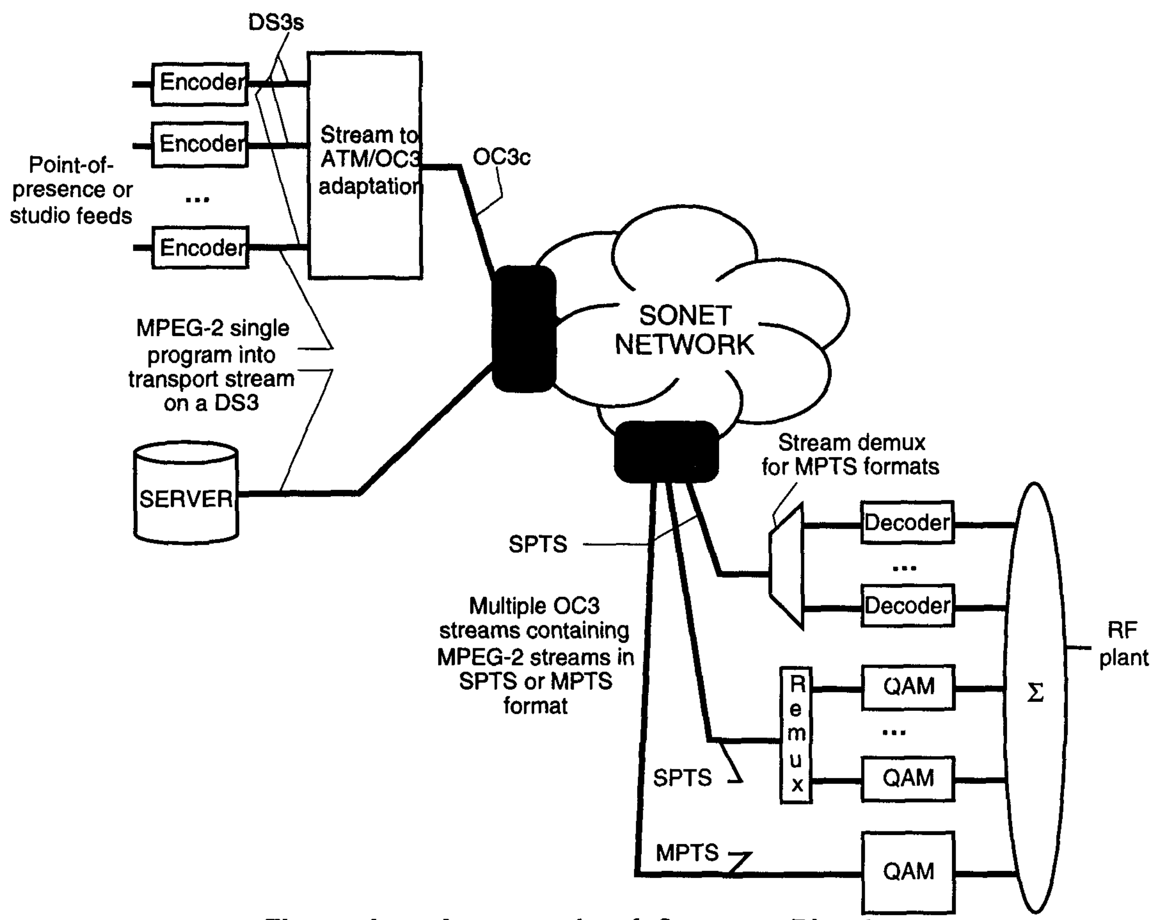


Figure 4 — An example of Constant Bit-rate Streams Adapted to ATM/OC3

Figure 4 shows one of the adaptation scenarios just discussed from non-ATM to ATM transport. Essentially, a number of encoders perform compression on sources coming from venues where major events are taking place, or from studio feeds for instance. The compressed signals are transmitted over non ATM channels such as DS3. The diagram only shows DS3 in seemingly the access portion, however, the access could be followed with a network part that is DS3-based until it gets to the portion that is ATM-based.

At this juncture, the signals can be converted for ATM transport over SONET using AAL5 stream multiplexers. The streams are placed then into OC3c connections and directed to the network element for entry into the SONET network where they get routed to appropriate destinations. At the network exit/destination, the single program transport streams (SPTSs) or multiple program transport streams (MPTSs) are directed to equipment that will either:

- convert the streams back to analog video using decoders, or
- have the SPTS streams remultiplexed into MPTS to be put through QAM modulators, or
- have the MPTS go directly to QAM modulators.

The resulting signals can then be combined as the final output for the RF plant.

■ 6 - Conclusion

The previous presentation has discussed the issues of MPEG-2 video on non-ATM and ATM networks. Inherently, MPEG-2 has been designed as network independent. It can be transported by ATM as well as non-ATM networks. SONET can be seen as a fabric which can carry both types of networks as well as providing links between the two.

The previous presentation has discussed the issues of MPEG-2 video on non-ATM and ATM networks. Inherently, MPEG-2 has been designed as network independent. The standards have addressed transport of MPEG-2 with ATM cells and the ATM networks can be based on SONET infrastructures. However, MPEG-2 streams can also be transported via non-ATM circuits, like DS3 channels, in a non-ATM format. This has not yet been addressed for standardization. At this moment, SONET is a widely accepted standard in the digital telecommunication field. Networks across the country, across the continent, and between continents can be connected and interwork. Moreover, the networks support voice and data, and will be the backbone of the multimedia highways. SONET is a network fabric for many markets: the public telephone network, the private enterprise networks, the CATV networks, etc. and MPEG-2 can benefit from either mode of transport, ATM or non-ATM.

■ 7 - References

[1] ISO/IEC, , IS 13818, Coding of Video and Associated Audio in the 3-15 Mb/s Range, 1995

[2] Daniel Minoli, Video Dialtone Technology, McGraw-Hill, 1995

Network Availability Consumer Expectations, Plant Requirements and Capabilities of HFC Networks

Tony E. Werner
Vice President of Engineering
TCI Communications, Inc.

Pete M. Gatseos, Ph.D.
Vice President of Strategic Research
TCI Communications, Inc.

Abstract

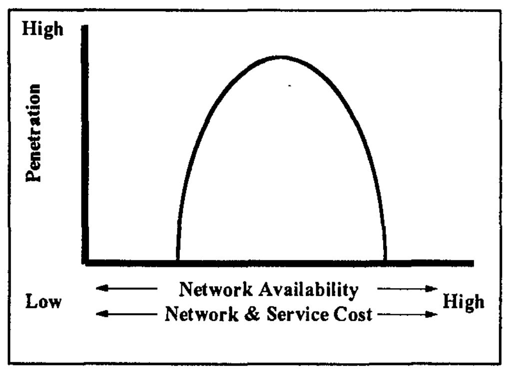
Within a consumer expectations framework, this paper investigates the value equation of cost versus network availability.

Clearly there are certain low levels of network availability that will be deemed unacceptable by customers. Equally as clear is that you can continue to increase network availability at a greater cost. Of course, this embedded capital cost in turn translates into higher service costs, ultimately becoming a price hurdle to the customer. Figure 1 provides a graphical view of this equation.

The network provider's objective is to satisfy customer expectations for a given service while providing a reasonable network cost —in other words, a reasonable investment that can be reasonably recovered through the price of the service or product.

This paper takes a wholistic view of key cost elements involved in network availability and reliability vis-à-vis customer expectations as they relate to cable TV services.

Figure 1
"Network Availability
The Cost-Value Relationship"



Introduction

Let's begin by defining the expectations of cable TV consumers with respect to network availability, *i.e.*, the ability of the cable network to perform at a given instant in time.

Very briefly, consumer expectation levels are set by their experiences with all other service providers that they have encountered. When a variety of services are typically delivered at a given level by most service providers, this level becomes the "norm", or in other words, the "expected". Of course, when expectations are met, consumers are satisfied that they have received delivery of services purchased. When expectations are not met, consumers become dissatisfied.

Further, in the free marketplace where competitors seek a means to generate greater revenues, service performance is continually improved (to gain a competitive advantage), thereby raising service level expectations over time. Indeed, as can be seen in Figure 2, holding all other variables constant and when services are at normative levels, (*i.e.*, customer expectations are being met), this results in parity with the competition. More importantly, when service levels exceed customer expectations, a competitive advantage is achieved.

Here in this discussion of cable network availability, consumers have electronic equipment such as their television sets, stereo systems, and telephones as a basis for comparisons. More specifically, because their TVs, stereos, and telephones perform error-free in nearly every instance that they are used, consumers have come to expect very high levels of service reliability from products/services that appear to be similar in nature, *i.e.*, "electronic" products and services. Therefore, cable services are intuitively measured/assessed by consumers against the reliability levels delivered by today's TV/stereo sets and telephone services.

It is also important to recognize that consumers have a higher probability of detecting a cable TV network failure. Indeed, according to Bellcore, cable customers are ten times more likely to experience a cable outage than phone outage. This is due to the fact that television sets are in use about 7 hours per day, while phones are in use only about 30 minutes per day on average. (See Table 1.)

Figure 2: Service Delivery and Consumer Expectations *vis a vis* Competitive Standing

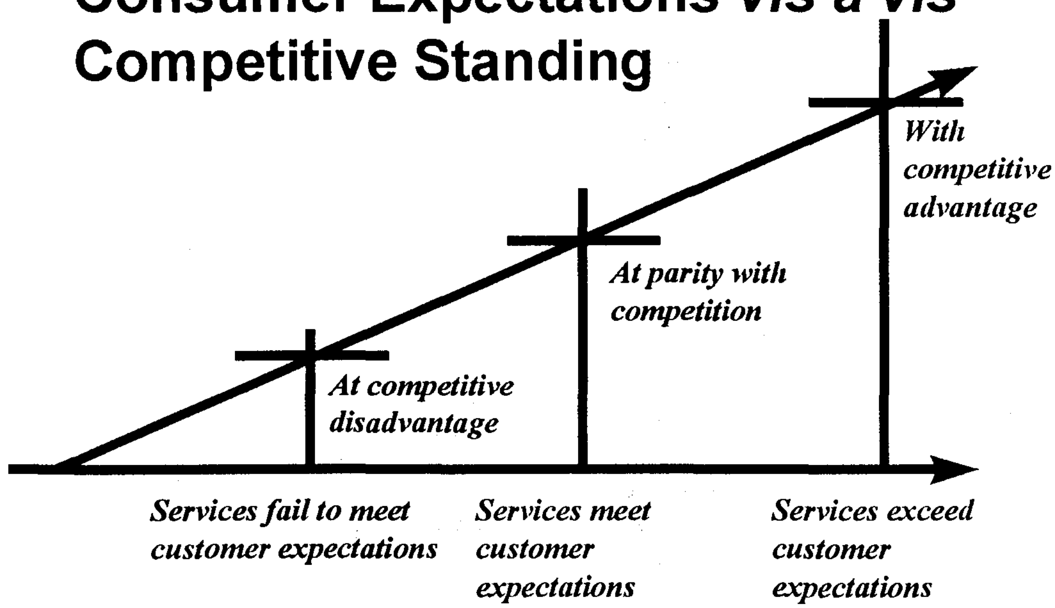


Table 1
Phone vs. TV Outage Perceptions[i]

	Usage per day Averages
Message Telephone Service (MTS)	30 minutes
Television viewing	***7 hours***

Definitions

By definition, network reliability is the probability that the system will not fail in a defined period of time. The frequency of failures over a defined period of time reflects the probability that the system will fail in a defined period of time, representing system unreliability. The well known term, Mean Time Between Failures (MTBF), reflects the inverted value of the frequency of failures and is usually used as a measure of reliability.

While network reliability is important to a network provider, a more important measure is the network availability. This measure is defined as the ability of a network or a unit to perform a required function at a given instant in time.

It can also be expressed as the percentage of time, within a given time interval, during which the network is capable of providing the service. This assumes that the external resources, if required, are provided.[ii]

In mathematical terms, the same can be expressed as MTBF/(MTBF+MTTR), where MTTR is the Mean Time To Repair with given external resources. This formula demonstrates that the network availability can be improved by increasing the MTBF, which represents reliability. Another way to improve network availability is to decrease repair time (MTTR) by providing more maintenance resources. Design, although not indicated directly in the formula above, is also a critical factor, since network availability can be greatly increased by limiting the number of cascaded devices.

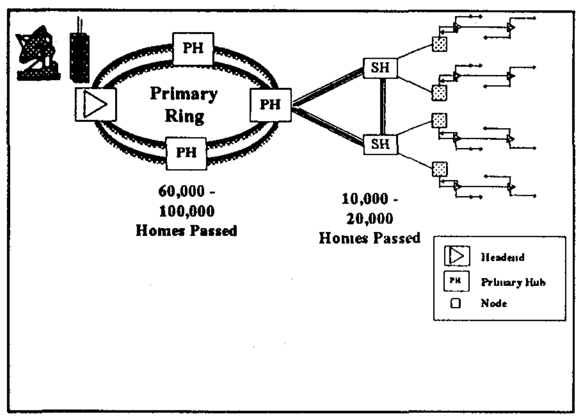
Elements of Network Availability

Network availability is a function of three basic elements, which are:

- MTTR
- MTBF
- Architecture

Mean Time To Repair is often overlooked in network availability, but can provide significant improvements to network availability at little or no cost.

Figure 3 — HRC



Using the Hypothetical Reference Circuit (HRC) in Figure 3 and empirically derived failure rates outlines in Table 2, we will evaluate repair time sensitivities.

Table 2
Failure Profile[iii]

	AFR (%)	MTBF	MTTR
FWD FO Tx	2.3	43[iv]	150
RET FO Rx	1.4	71[v]	150
FO Cable	0.44	227[vi]	270[vii]
FWD FO Rx	1.4	71	150
RET FO Tx	2.3	43	150
Cable	0.23	443	152
Connector	0.01	7,110	91
Power Supply	2.21	45	51
3rd Party Damage	3.56	28	99
Other	3.71	27	225
Passive	0.04	2,241	79
Fuse	0.66	152	44
Amplifier	1.75	57	63

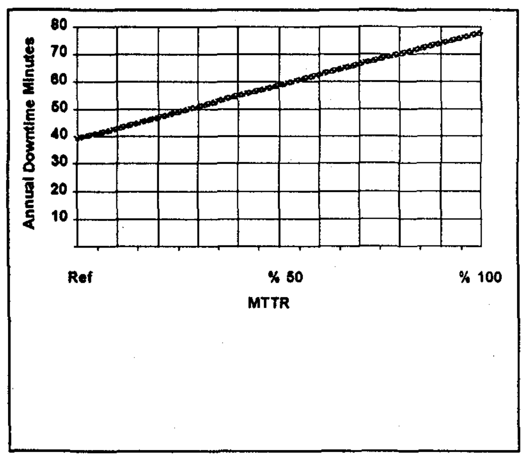
Based upon the figures in Table 2, the reference architecture would provide annual network downtime[1] of 39.01 minutes, based upon a statistical model. The Pareto of the downtime contributions are contained in Table 3.

Table 3
Downtime Analysis[viii]

	Redun-dancy	NDT (min.)
Primary Ring	Total	0
Secondary Ring	Total	0
Fiber to Node	None	15.94[ix]
Express Coax 1	None	7.64
Express Coax 2	None	6.46
Tapped Coax 1	None	3.35
Tapped Coax 2	None	3.40
Tapped Coax 3	None	2.23
Total		39.01

Figure 4 illustrates the impact of increased MTTR on network down time. The reference level at the beginning of the graph utilizes mean repair times as outlined in Table 2.

Figure 4
MTTR Impact on NDT



[1] Excludes simultaneous failures of standby and commercial power systems.

Table 3: Real Time to Repair Comparisons

(excluding power failures, drops, head end equipment, and network interface units [NIUs].)	*System A with aggressively managed repair times*	*System B with little emphasis on repair times*	*System C with little emphasis on repair times*
	With minimal fiber options (Minutes)	*With adequate fiber options (Minutes)*	*Very good plant (Minutes)*
Mean time to repair	62	135	124
Median repair time	48	105	95
Mode repair time	35	120	60
Average network availability	99.987	99.950	99.986
Average network downtime	66 minutes	243 minutes	72 minutes

Source: Computed from "Network Availability Requirements and Capabilities of HFC Networks" by Tony Werner and Oleh Sniezko, November 1995.

MTTR and Its Impact

Greater network availability begins with management, and the agenda and standards it sets for its employees.

In Table 3, three systems are compared along with several key measures. While System A's plant architecture was older, with minimal fiber upgrades, management's focus on maximizing network availability paid great dividends. System A, using any measure, out-performed the less well managed System B — despite System B's superior (but not atypical) plant. Note also how these two Systems compare with System C, which enjoys superior plant technology.

Very clearly, when management has employees focused on minimizing impacts of all the factors surrounding outages, great strides can be taken in meeting the demands placed on cable TV network availability by consumer expectations.

(This data should also be interpreted to mean that, when employees perceive network availability from "end to end" to be critical to meeting customer expectations, all manner of remedies will be pursued vigorously to minimize disruptions to the customer TV viewing experiences.)

To illustrate further, the repair times in Table 1 are actual repair times detected in a well maintained Midwest system. In reviewing several other systems, we have found several that exhibit repair times close to this system. We have also found several that have significantly longer repair times — some extending to over three times as long.

By having an MTTR of 150% over the reference, which several systems do, you will cause an annual network down time (NDT) of over 150 minutes versus the 39.01

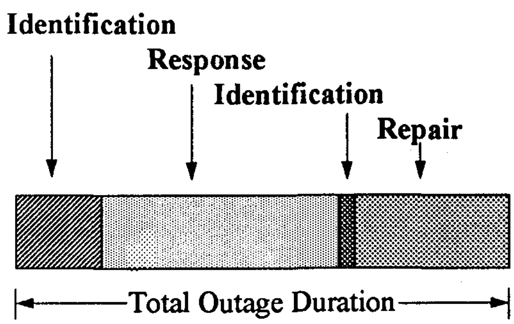
minutes, based upon the reference values. These two different availability outcomes are achieved with identical system performance for MTBF and architectural design. Managing outage response programs can significantly impact system availability at minimal or no cost!

The average repair time, often referred to as MTTR, is comprised of three major components affecting the total outage duration. They are:

- Identification;
- Response; and
- The actual repair.

Figure 5 diagrams the markup of a typical Hybrid Fiber Coax (HFC) outage.

Figure 5
Repair Time Components



In a typical HFC system, these contributions break down at: 3-10% for identification, 50-60% for response time, and 30-40% for the actual repair. These exact distributions vary, based upon several factors, but usually the major portion of the repair time is not the actual repair, but rather the identification and response time. This is because of the modular nature and short repair times associated with a bus network.

MTTR Opportunities

Taken from the consumer's perspective, even the most sophisticated plant performance tracking databases monitor only a small set of total sources of outages to the customer's TV set. As illustrated in Table 5, among five areas that consumers believe the network (*i.e.*, from "end to end") failed, only two are tracked by one of the cable industry's best plant monitoring systems.

Table 5
An Illustration of Customer Perception of Sources of Outages

Customer reported source of outage	*Actual plan performance reports*
Weather	Not reported
"Cable network"	Reported
Power company failure	Reported
Subscriber drop	Not reported
Converter	Not reported

Taken in this light, the conventional industry wisdom that consumers overestimate the actual number of plant outages is contradicted. From the consumer's "end-to-end" delivery perspective, plant monitoring measures are too limited in scope. Cable industry measurements, therefore, need to be more thoroughly articulated, standardized, and continually monitored.

Now let's look at a few select techniques to improve identification and reduce identification times. A tool that gets a lot of attention in this area is status monitoring. Historically, status monitoring in a CATV architecture was not significantly beneficial. With high penetrations of customers and trained phone staff, outages could be determined with 5 to 10 minutes of the occurrence via phone calls. Exact location identification was performed from continued phone calls while a technician was en route. The overall savings from status monitoring was typically the first 5 to 10 minutes that it took to determine an outage from phone calls. This would be traded off against the capital and operating cost of status monitoring systems, and in most cases, more value could be provided from investing that money in some other portion of the network.

It should be noted that, in the past, return networks would have to be activated and maintained purely for the status monitor, adding this capital and operating cost directly to the cost of status monitoring.

Today, status monitoring is going through a rebirth in HFC networks under the name of *network management*. While there are several reasons for the re-entry of network monitoring and management today, be cautious not to see this as a panacea. While we like to call it network management, it is still largely monitoring today. Some sophisticated networks may use it to enable backup fiber routes or even backup equipment, but mainly it is a monitoring tool. In the future, as systems offer significant levels of telecom services, the same upper level management systems may also provide provisioning, and then it will be graduating into network management.

With this said, there are situations in which monitoring is more practical today than in the past. First, computers have grown in power and user friendliness significantly in the last 5 to 10 years, allowing for a much more practical, useable implementation. Secondly, many operators are activating the return path of the network for new services, so this cost does not have to solely burden the cost of monitoring. The third reason is that several new services will go in at low penetrations, and due to this and their less-than-full-time usage pattern, return path failures may go unidentified for several hours without status monitoring.

Let's look at some of the other methods of reducing MTTR by focusing on repair times. Several operators enforce zone maintenance practices. This procedure has a particular maintenance technician responsible for a geographical area of plan, and as such is required under most conditions to be in that area.

This means that his or her drive time (response time) will be minimal. Another technique is to operate two shifts of maintenance technicians. The first shift works from 7:00 a.m. to 4:00 p.m., and a second shift works from 10:00 a.m. to 7:00 p.m. This ensures that staff is in the system for more hours of the day and ultimately insures a shorter response time for system failures during this portion of the day. The wider the spread of the work shifts, the greater the system coverage, but also the greater the cost for supervision and shift premiums (if applicable).

The final important point is to manage the outages in general. This means monthly reviews of every outage, looking for the opportunity to eliminate or minimize the occurrence in the future.

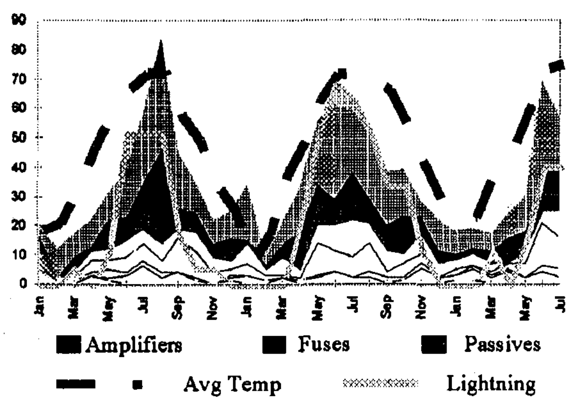
These opportunities include: training, truck inventory levels, pre-made optical cables, and individual employee performance levels.

The MTBF Opportunity

Now let's focus on the second component of network availability: MTBF. This is another area that can offer significant improvements in network availability at minimal cost. It should also be pointed out that, while it is important to minimize repair times, thus improving the overall network availability, it is better to have no interruption at all. This objective is largely affected by MTBF.

There are several basics for improving MTBF, but the main area is again system-specific analysis. Of all the systems that we have analyzed, they all have significant low cost opportunities for improvement. Whether it is poorly performing equipment, improper grounding or improper fusing policies, the latter is simple to correct, and in several of the systems reviewed, it can reduce the system outages by up to 30%. Figure 6 identifies a seasonal outage pattern that is characteristic of most of the systems analyzed.

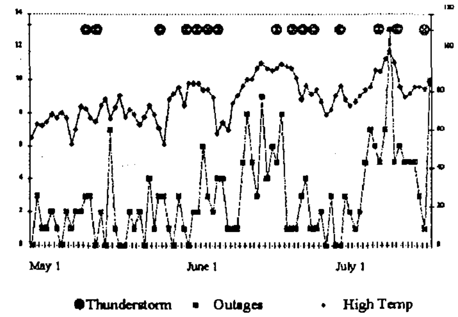
Figure 6
Outage Seasonally



The area graph shows outages from January 1993 through July 1995. The dashed line indicates average temperature, and the light solid line thunderstorms. After identifying the seasonal nature of outage activity, we obtained and overlaid weather[x] information for the same period. While somewhat intuitive, both temperature and thunderstorm activity obviously influence the system reliability.

To understand this better, we analyzed day-by-day activity for May-July of 1995. These results are graphed in Figure 7.

Figure 7
Day by Day Correlation



While there are times when the thunderstorms caused outages directly, the majority of the increased outages coincide with high temperature. This was demonstrated on July 13, 1995, when the high temperature hit 101 degrees Fahrenheit. This provided for the highest device failure of the entire year. Thunderstorms occurred on July 12 and 14, but not on July 13. As you review Figure 6, you can see several other points when thunderstorms occurred without significant increases in network failures. A close examination of the failures on those days often revealed water damage, not electrical damage. July 21, 1995, the last data point on the chart, revealed high electronic failures resulting from the electrical storm. Therefore, it is clearly a combination of effects, but the predominance in this system is temperature, not thunderstorms.

The device failures during the high temperature days were predominantly components burned by excessive amperage or shorted devices themselves. It is unclear whether these failures are the result of thermal instability of certain of the electronic devices or if the commercial power system becomes more unstable because of the increased loads for air conditioning.

The next step needs to be a careful power analysis to determine the exact cause or causes of the heat-related failure mechanisms. In either case it is likely that significant improvements can be made at minimal expenditure through selective power conditioning. If even the top 30% of these peaks can be eliminated, major improvements can be achieved in network availability. Furthermore, the savings in staff for technicians and phone personnel would likely go a long way toward paying for these improvements.

Network Architecture Opportunities

Now let's review architecture and its impact on network availability. This is often the element most focused upon for achieving high network availability. As demonstrated earlier in Table B, while architecture is an important element to network availability, even the best architecture will not achieve the desired results if not managed properly. Likewise, even classic cable systems can achieve remarkably good network availability with diligent management.

This is not to say that architecture isn't important, but rather that it is not the only factor in network availability. It can also be one of the higher cost methods of achieving high reliability, and if not carefully managed, can push prices beyond the ability of most consumers to pay (as illustrated in Figure 1).

If you are building a new system or upgrading an old system, network availability should be a primary factor in architectural decisions. Several architectural designs affect reliability. These include the following:

- Self-healing rings;
- Number of serial devices; and
- Redundant components.

All of these architectural designs help mitigate perhaps the most strategically important area of concern, losing customers to competitive TV delivery systems because consumers perceive competitors to provide higher levels of TV signal reliability.

For example, recurring outages frequently are confined to specific geographic areas in which network failures have proven elusive or resource limitations have precluded implementing a more permanent solution. In these situations, the probabilities of customers experiencing an outage greatly increase. Even more ominous, the groups of customers impacted will be repeatedly denied what they have already paid to receive.

Taking this state of affairs one step further, let's assume that 40% of the customers who watch TV daily are impacted by one outage and are angered by it (because they were very engaged in the program). If there is one outage per day, the laws of probability tell us that 16% will not only experience two outages, but be angered by them on *both* occasions. If there are three outages, 6% of our customers will have been moved to anger on all three occasions! Of course, the number of customers impacted can be quite great should these outages occur during prime time, *i.e.*, peak TV usage "dayparts" or times of day.

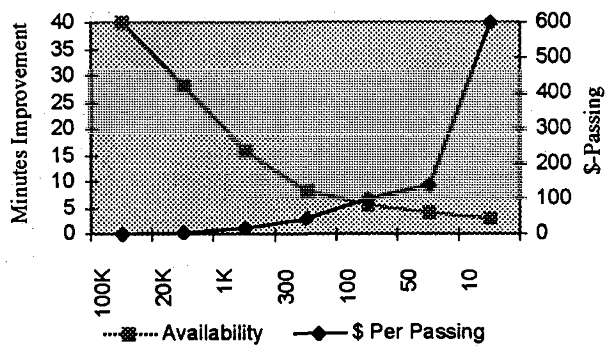
To help prevent recurring outage situations, architectural design alternatives now will be considered.

First, self-healing rings, which are typically used in fiber optic networks and offer significant improvements in network availability by providing virtually 100% network availability[2] to a portion of the network that has a higher MTTR and affects a significant number of customers. The question is how deep to run fiber rings.

[2] Redundant rings do not in theory provide 100% availability; however, with limited numbers of high MTBF, low MTTR, parallel devices, the reliability is practically 100%.

It is clear that fiber rings to hubs serving 100,000 to 200,000 customers can be implemented at a low cost per passing and provide significant improvements in network availability. It is also clear that fiber rings to the home with an alternate power network can provide virtually 100% network availability, but will also carry a ridiculous price. Determining the best value is again system-dependent. Figure 8 provides a cost versus reliability curve for extending fiber rings to various depths of the network.

Figure 8
Reliability versus Cost
(to close the ring)



The right-hand axis is cost per passing associated diamonds and the upward trend. The left-hand axis is annual minutes of improvement from ring implementation. The X axis is the size of homes passed grouping that is ring backed up. The exact relationship for cost versus availability improvement varies by system and exact implementation. Figure 8 provides a generic representation illustrating the diminishing returns as you extend the self-healing rings deeper into the network. At the large primary hubs, rings can be installed for as little as $3.00 per home passed[3] and provide availability improvements of nearly 40 minutes. Significant availability improvements are still maintained at the 20,000 homes and at the 1,000 homes passed levels. These hub or node sizes can be ringed without significant cost penalties. The exact point of maximum value is dependent upon the specific system (aerial/ underground/ density/ etc.) and the technology implementation. The further you move out on the curve, the smaller the improvement in availability and the higher the per passing cost.

Serial device or cascade units also affect network reliability. For each amplifier that you shorten your cascade, you improve network availability by between 3 and 7 minutes, based upon Table 2[xi]. Cost modeling based upon actual network design indicates that upgrade costs increase by $20 to $40 per home passed to go from 4-amplifier cascades to 3-amplifier cascades.

Limiting other devices such as passives also provides improvements, but they are quite minor in nature. The improvements that you obtain from limiting these devices is likely more pronounced in frequency response and other technical parameters than in network reliability.

A final architectural tool is to employ redundant devices such as lasers and receivers in serving area fiber systems. Based upon Table 2, this could reduce the 15.94 minutes to 4.75 minutes. This comes at a significant, although you now require redundant lasers and receivers for up-and-down stream operation at every node. Depending on node size, this can add over $10 per home passed to provide a theoretical improvement of 10.69 minutes.

We refer to it as a theoretical improvement based upon historic experience with fail-safe switching at low levels of the architecture. While redundant switching at primary and secondary hubs is quite effective, this technology has been less effectively implemented at lower levels of the architecture. Several reasons probably

[3] This is the incremental cost above a non-ring installation.

account for this. The first is that, in order to have economics prove in at low levels of the architecture, less reliable switching and detection circuits have been employed. In addition, just because of the sheer numbers, management and maintenance has been difficult. This is not to say that node redundancy cannot be implemented reliably, but it is just cautioning based upon previous experience with fail-safe switching in RF amplifiers. In this experience, undesired switches caused almost as many problems as they solved.

Strategic Implications

We began this paper by noting that service providers who exceed consumer expectations will enjoy a competitive advantage over their rivals. We also noted how cable networks can be compared to service reliability levels achieved in telephony. It is therefore logical to try to determine the service availability levels in delivering telephone services. This can be taken one step further as well — by determining the TV signal availability of direct to home satellite TV providers, *e.g.*, DBS/DSS competitors.

In this manner, benchmarks for reaching service parity or service superiority can become the target levels for cable operators — to meet or exceed — as technological advances as well as managerial/financial resources permit.

For telephony standards, Bellcore has published a telephony service availability standard of 99.99% or 53 minutes per year downtime maximum.[xii]

In regard to DBS/DSS, signal availability (excluding environmental factors such as "rain fade") averaged 99.90%.[xiii]

Unfortunately, overall cable industry norms are unknown — as such data is not nationally or systematically collected. At this point in time, it remains up to each individual cable system to assess their signal availability performance relative to current DBS/DSS competitors as well as to consumer expectations (created when making comparisons to comparable products and service at a given cost). It is sufficient here to restate what was said at the beginning of this paper: It is only when service levels exceed that of the norm and that of competitors that a competitive advantage may be achieved.

Minimally, it is clear that you must reach a certain level of network availability to not be perceived inferior to your competitors. The counterpoint is, of course, that you can use superior network availability as a competitive edge.

Caution should be applied, however, once you reach an acceptable level of network availability, because you can continue to increase network uptime, but at a higher capital cost. If this increased capital cost is translated into higher service cost, you can actually hurt your sales penetration levels and competitive position. The objective is to find the optimum intersection of capital cost and network availability.

[i] Bellcore, TA-NWT-000909, "Generic Requirements for FITL Systems Availability and Reliability Requirements", December 1993

[ii] TR-NWT-000418

[iii] Werner/Sniezko SCTE Emerging Technologies 1966, "Network Availability Requirements and Capabilities of HFC Networks"

[iv] Based upon manufacturer's AFR and validated with TCI 1994 field results. The number is thought to be conservative.

[v] Werner/Sniezko SCTE Emerging Technologies 1966, "Network Availability Requirements and Capabilities of HFC Networks"

[vi] Victor T. Hou (Bellcore), "Update on Interim Results of Fiber Optic System Field Failure Analysis"

[vii] Bruce Corrigan (Rogers Cable TV), "Regional Fiber Outage Report, July 1990 to June 1993)

[viii] Werner/Sniezko SCTE Emerging Technologies 1966, "Network Availability Requirements and Capabilities of HFC Networks"

[ix] Include up- and down-stream components for full duplex availability.

[x] National Oceanic and Atmospheric Administration (NOAA)

[xi] Werner/Sniezko SCTE Emerging Technologies 1966, "Network Availability Requirements and Capabilities of HFC Networks"

[xii] Bellcore, TA-NWT-000909, Issue 2, "Generic Requirements for FITZ Systems Availability Requirements" December 1993.

[xiii] *Multichannel News*, July 25, 1994, p. 103.

Nonlinear Distortions Distribution in HFC Networks

T.C. Tsao

Arthur Yang

AT&T Taiwan Telecommunications, Co. Ltd.

Abstract

This paper uses a new method to analyze the nonlinear distortions, CTB, CSO and CIN in the HFC Network.

Each carrier is treated as a vector, the amplitude, frequency and phase are set for each channel.

Both 550 MHz & 750 MHz all analog system and 750 MHz mixed analog and digital HFC system are analyzed. The level variations, random phase between channels and the output tilt of a real cable system are considered.

For the all analog systems, the distributions of CSO & CTB vs. tilt, the worst channels vs. tilt and the modulation effects are proposed.

For HFC systems with mixed analog and digital signals, the CIN noise distribution vs. tilt and worst channels vs. tilt are proposed.

This analysis provides some insight into the nonlinear distortions distribution in real cable systems.

Introduction

Cause of Distortions

The distortions are caused by nonlinear devices in the network, including amplifiers and AM fiber systems. [1], [2]

Distortion performance for the analog system includes Composite Second Order (CSO) and Composite Third Order Beat (CTB). Testing methods for these performance criteria, are well established. [3]

Distortion performance for the mixed analog and digital HFC system could be specified as Composite Intermodulation Noise (CIN), CSO and CTB. [4], [5]

Analysis method

The conventional "beat count" method results in the well-known beat tables for both CTB and CSO.[6] This "beat count" method only counts the beat number on the specified frequency locations relative to the video carriers. There is no consideration for level and phase variations among channels using the beat count method. The beat count method does not accurately characterize present day or planned cable systems which carry video modulated carriers, use tilted output or carry both analog and digitally modulated carriers.

A computer program was designed to analyze the CATV system with mixed analog and digital signals. The output tilt and modulation effects were also considered.

The program assumes that the nonlinear characteristic of any device is expressed by the transfer function. [7]

$$V_O(t) = K_1 * V_{in}(t) + K_2 * V_{in}(t)^2 + K_3 * V_{in}(t)^3$$

where

$V_O(t)$ *is the output signal*

$V_{in}(t)$ *is the input signal*

K1, K2 and K3 are the 1st, 2nd and 3rd coefficients

All the high-order terms are neglected. The time-variant input signal is assumed to be

$$V_{in}(t) = A1 * COS(W1 * t + \varphi_1) + A2 * COS(W2 * t + \varphi_2) + \ldots$$

Where

A1, A2 ... are the amplitudes

W1, W2... are the frequencies

φ_1, φ_2 ... *are the phases*

The amplitude, frequency and phase of every carrier can be set separately. Thus, the synchronized / non-synchronized carriers, flat/sloped output and CW/modulated carriers could be simulated by the program.

In order to analyze the relationships among all the discrete and composite beats in detail, all the beats within 0.25 MHz intervals in the 6 MHz NTSC channel are recorded. This feature can simulate the FCC frequency offset or count the total power in any interval just as the resolution bandwidth setting of spectrum analyzer.

The analyses are based on a hypothetical nonlinear device with specified nonlinear coefficients, K_2 and K_3. Any other device will have similar characteristics except the reference value.

In order to prevent confusion, the EIA Channel designation and the STD frequency allocation plan are chosen. CH-95 to CH-99 are not included in these analyses.

The tilt is defined as the level difference between the highest and lowest analog channels.

Detailed results of the analyses are shown in the following sections.

Analog System Results

CSO Distribution

The distribution of CSO vs. tilt and the worst case distribution CSO vs. tilt is shown in figures 1, 2 and 3.

There are two extreme groups, CH-5/CH-6 and the highest channel. In the case of CH-5/CH-6, the CSO distribution is caused by the special frequency assignment. In the case of the highest channel CSO distribution is due to the largest beat counts and the highest beat power.

There exists a transition point, about -6.5 dB tilt. Any tilt higher than this value, CH-5 is the worst channel. In this region, when the tilt is getting higher, the CSO is getting worse. When tilt changes, both the carrier power and the beat power of CH-5 decrease, but the carrier power decreases much faster than the beat power.

For the tilt lower than -6.5 dB, CH-78 (547.25 MHz) is the worst channel for 550 MHz system; CH-116 (745.25 MHz) is the worst channel for all analog 750 MHz system. In this region, when tilt increases from -15 dB to -6.5 dB, the CSO is getting better. That is because the carrier power of the highest channel increases faster than the CSO beat power accumulation.

If the same active device is used, when the tilt is higher than -6.5 dB, the CSO of 550 MHz system is about 4 dB better than that of the all analog 750 MHz system. When tilt is lower than -6.5 dB, the CSO of 550 MHz system is about 3 dB better than that of all analog 750 MHz system.

Typically, for typical cable systems, the negative tilt will not happen in the trunk lines, but in the distribution network and the input stage of the settop box negative tilt is likely.

CTB Distribution

The distribution of CTB vs. tilt and the worst CTB vs. tilt are shown in figure 4, 5 and 6.

When tilt increases, the CTB performance becomes better. At the same time, the worst CTB channel moves toward the low frequency band. For example, in the 550 MHz, 0 dB tilt system, the worst CTB happens on CH-41. In the 550 MHz, 9 dB tilt system, the worst CTB channel moves to CH-25, but the CTB performance improves about 7 dB.

When tilt increases, both carrier power and beat power decrease, but the carrier power decreases much faster on the low band than that of high band. So the worst case CTB channel will move toward the low band.

As the worst case channel moves toward the low band, the composite beat power of the worst channel becomes lower because the beat counts and the power of each beat decrease. Though the carrier power decreases when the tilt increases, the carrier to composite beat ratio continues to improve.

If the same active device is used, the CTB of 750 MHz all analog system is about 6.5 dB worse than that of all analog 550 MHz system.

Modulation Effects

Several papers discuss the modulation effects on system performance. [8], [9]&[10]

When the carriers are modulated, the CSO and CTB performance will be better because the average carrier power is about 6 dB down from the peak value. A 6 to 7.8 dB CSO improvement factor is reported and a 10 to 12.8 dB CTB improvement factor is reported as well.

In the analysis, the video information are assumed to be in Gaussian distribution.

The simulated modulation effects on CSO and CTB are shown on figure 7, 8, 9 and 10. Note that all these modulated curves follow the trends of the non-modulated (CW) curves.

For CSO curves, because the beat counts are very small, compared with that of CTB, the modulation effects are much more apparent. It follows the ***Law of Large Number*** of statistics.

Mixed Analog & Digital System Results

In mixed signals system, the digital carriers do not contribute any new CSO or CTB beat as the analog carriers. Instead, the beating process within digital carriers and between digital carriers and analog carriers generate the new impairment, Composite Intermodulation Noise, CIN. [5]

For convenience, 4 MHz is chosen as the noise bandwidth just as that in the analog channels.

CIN3 Distribution

CIN3 is the third order composite intermodulation beat noise. The beats involve at least one digital carrier.

The CIN3 Distribution vs. channel number and tilt is shown in figure 11. The worst CIN3 channel vs. tilt is shown in figure 12.

While CIN3 is distributed within both the analog and digital bands, most of the CIN3 is located in the digital band and the high channels of the analog band.

The impact of CIN3 to analog channels is greater than the impact to the digital channels.

In the digital band, the CIN3 performance is almost the same with either tilt or level difference changes. In the analog band, for increasing slope, the digital carriers will increase as well, same as the power of CIN3. But the analog carriers will decrease; thus, the CIN3 in the analog band will be worse for increasing slope.

When the level difference

increases, i.e. the digital levels decrease, the CIN3 in digital band remains the same because the composite beat noise decreases about the same amount. But in the analog band, CIN3 becomes better because the beat noise from digital band decreases. The limitation to the level difference will be the BER performance and the system noise level.

CIN2 Distribution

CIN2 is the composite of all the second order beats involving at least one digital carrier.

Compared with CIN3, CIN2 is about 20 dB lower than CIN3, the total CIN is almost the same as CIN3.

The total CIN should be considered in the analog band CNR calculation.

Conclusion

For all analog systems, CSO has two extremes, CH-5/CH-6 and the highest channel. If tilt is higher than the transition value, -6.5 dB, CH-5 is the worst channel. On CH-5, when the tilt becomes higher, the CSO becomes worse. With tilt lower than -6.5 dB, the highest channel is the worst channel for CSO and as the tilt decreases, the CSO will become worse.

The CTB performance for all analog systems will be better and the worst channel will move toward the lower frequency channels as the tilt increases.

For the mixed digital and analog signals system, the CIN noise in the digital band remains almost the same as the level difference or slope changes. But in the analog band, the CIN performance becomes worse when the slope increase, and improves when the level difference increases. The limiting factors of the level difference will be the system noise floor or desired noise performance and the BER performance requirement.

References

[1] "Broadband Network Technical Training Seminar Workbook ", C-Cor Electronics, 1992.

[2] James A. Chiddix, " AM Video on Fiber in CATV Systems : need and Implementation ", IEEE Journal on Selected Areas in Communications, Vol. 8. No. 7. Sep. 1990.

[3] "NCTA Recommended Practices for Measurements on Cable Television Systems", Second Edition Revised, NCTA, 1993.

[4] Lamar West, "A New Technique for Measuring Broadband Distortion in Systems with Mixed Analog and Digital Video", 1992 NCTA Technical Papers.

[5] Jeff Hamilton and Dean Stoneback, "The Effect of Digital Carriers on Analog CATV Distribution Systems", 1993 NCTA

Technical Papers.

[6] "Cable TV Reference Guide", Philips Broadband Networks, Inc. 1992.

[7] Keneth A. Simons, "The Decibel Relationships Between Amplifier Distortion Products", Proceedings of the IEEE, Vol. 58. No. 7, July, 1970.

[8] Ed Mitchell, "Composite Second Order Distortion and Distribution System Performance", 1985 NCTA Technical Papers.

[9] John A. Mattson, "Specifying AM Fiber System Performance", 1990 NCTA Technical Papers.

[10] Louis D. Williamson, "Laboratory vs. Field Measurement of AM Optical Links -- Reconciling the Difference", SCTE Fiber Optics, 1990.

[11] Jeff Hamilton and Dean Stoneback, "The Effect of Digital Carriers on Analog CATV Distribution Systems", 1993 NCTA Technical Papers.

Authors :

T. C. Tsao was born in Taiwan, Republic of China, in 1965. He received the B.S. degree from Feng-Chia University, Taiwan, in 1987.

From 1989 to 1991, Mr. Tsao was a development engineer at AT&T Taiwan Telecommunications. From 1992 to 1993, he joined the Optical Subcarrier Project of the National Science Council, responsible for CATV AM-VSB Optical Transmission System. Since 1993, he has been involved in CATV system applications, including headend and distribution systems engineering.

T. C. Tsao can be reached at email: internet!ms1.hinet.net!tctsao.

Arthur Yang was born in Taiwan, Republic of China, in 1967. He received the B.S. degree from Chung-Yuan Christian University, and the M.S. degree from National Taiwan University, both in electrical engineering, in 1991 and 1993 respectively. During 1992 to 1993, he also joined the Optical Subcarrier Project of the National Science Council, responsible for CATV FM Optical Transmission System.

Mr. Yang has been with AT&T Taiwan Telecommunications since 1995, involved in CATV system applications engineering, including optical transmission system and distribution electronics engineering.

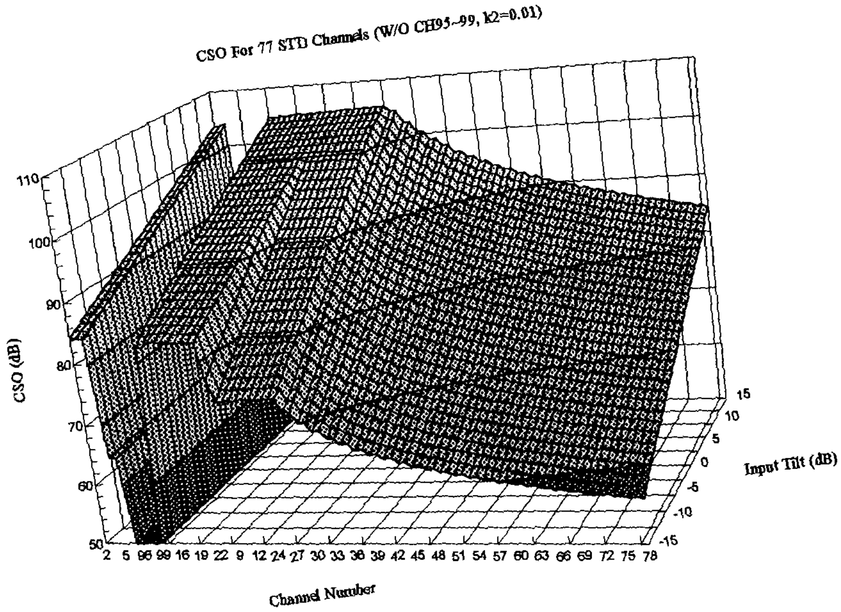


Figure 1

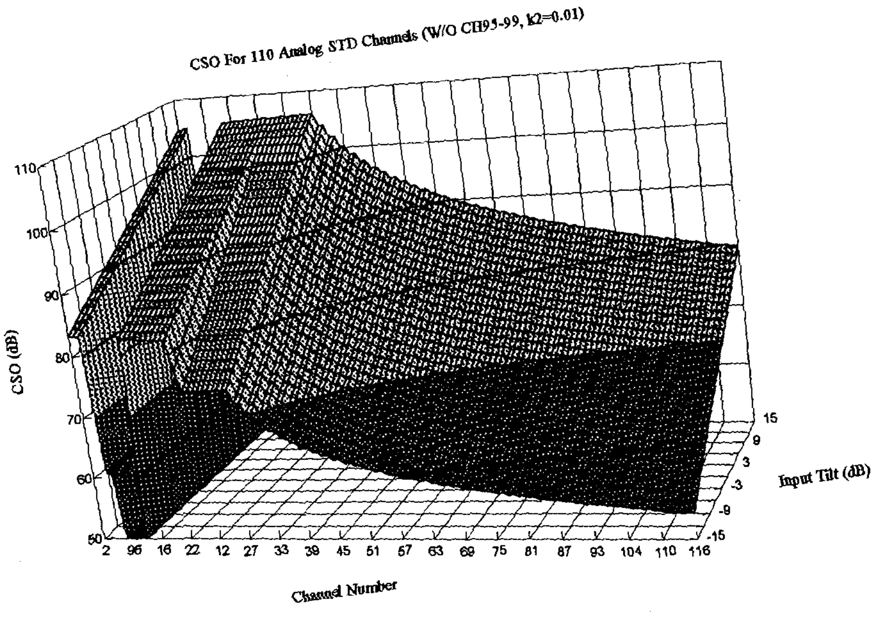


Figure 2

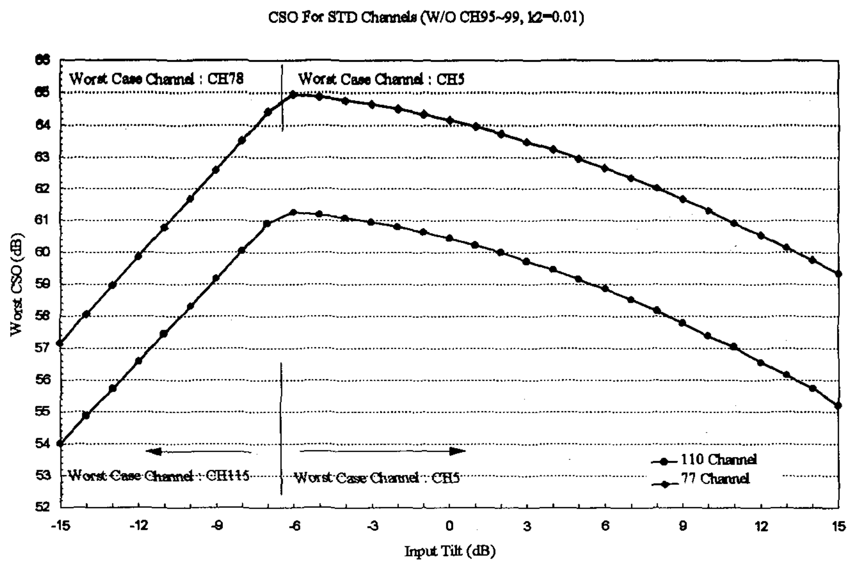


Figure 3

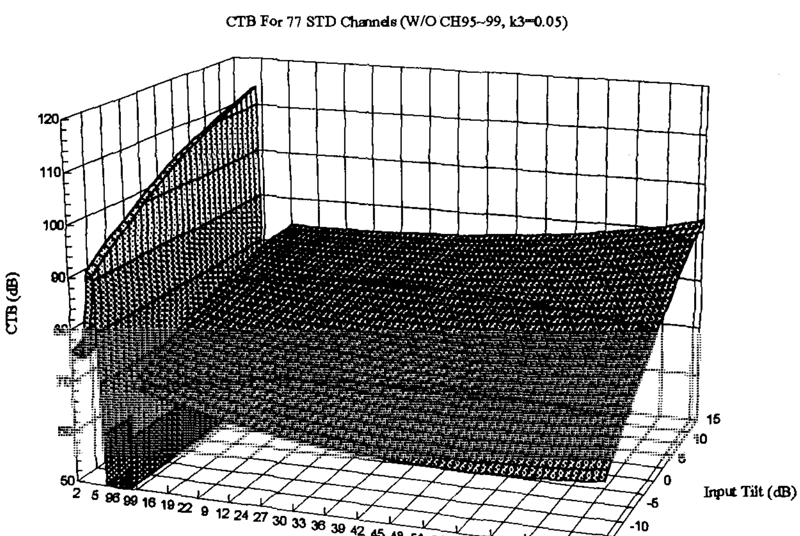


Figure 4

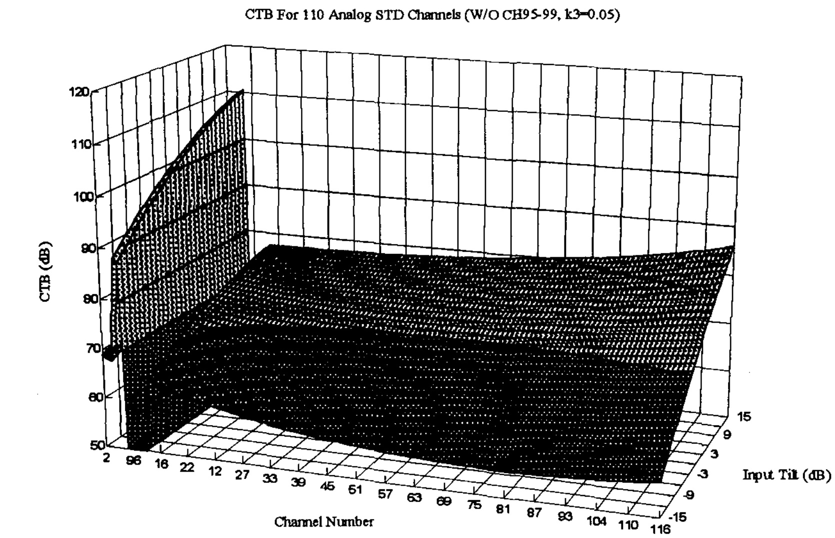


Figure 5

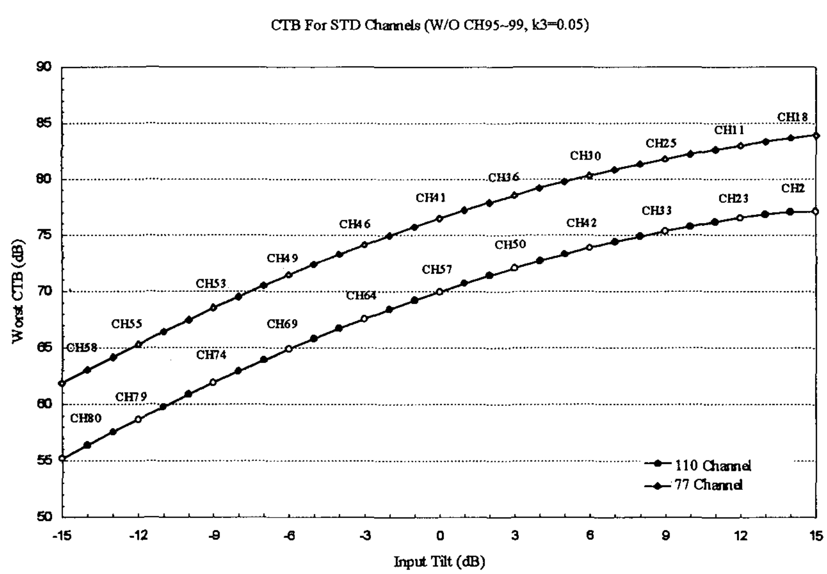


Figure 6

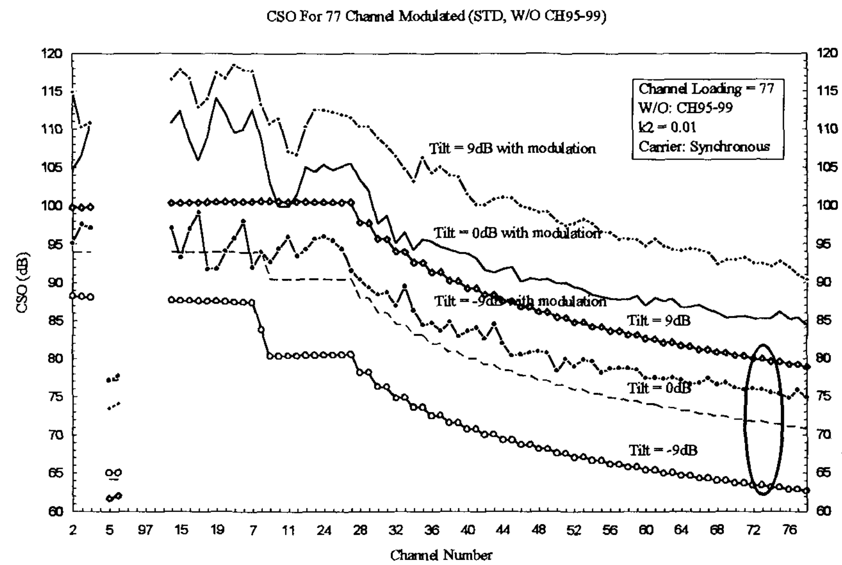


Figure 7

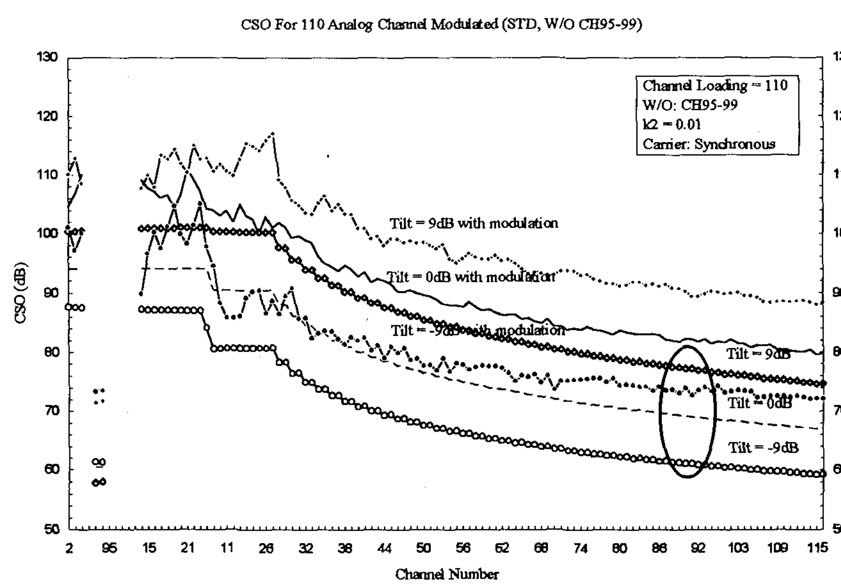


Figure 8

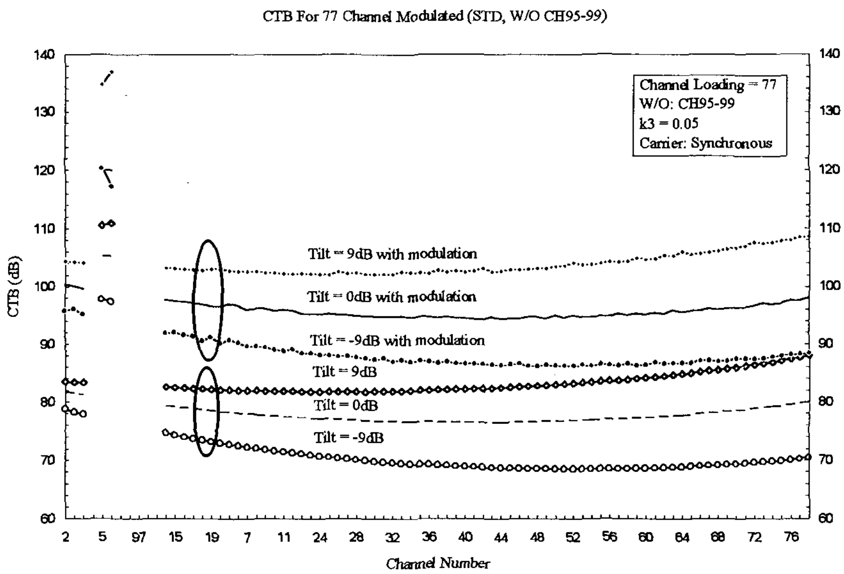


Figure 9

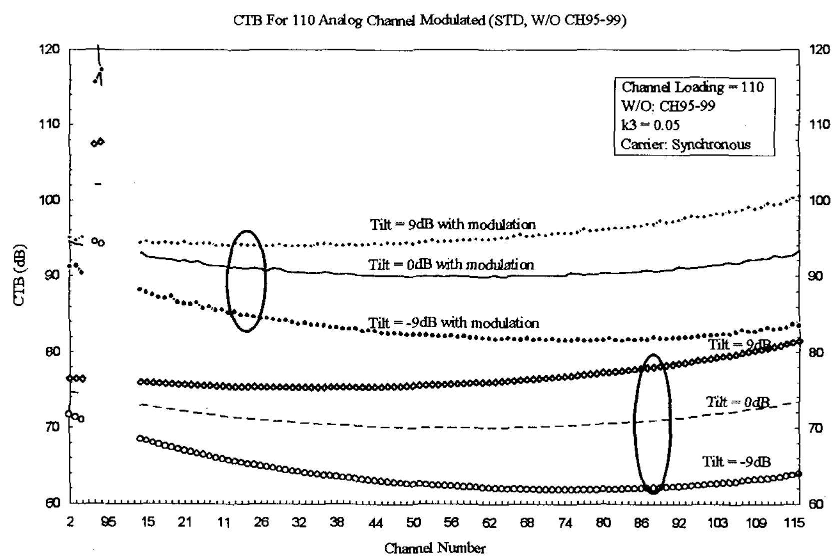


Figure 10

77 Analog + 33 Digital STD Channels (W/O CH95-99, k3=0.05, Digital Level Lowered 8 dB)

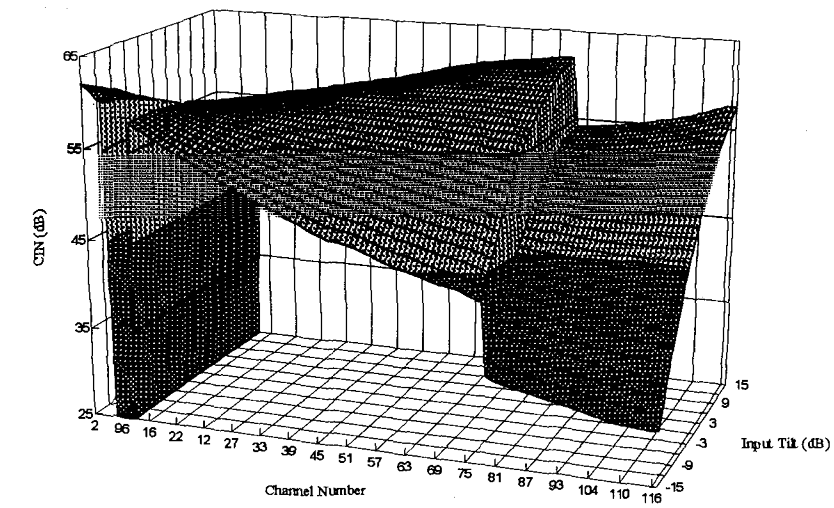


Figure 11

77 Analog + 33 Digital STD Channels (W/O CH95~99, k3=0.05, Digital Level Lowered 8 dB)

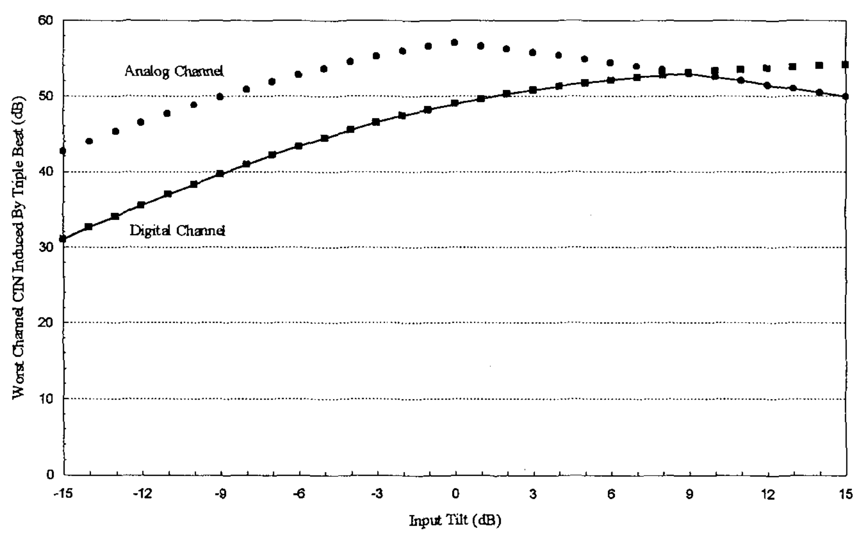


Figure 12

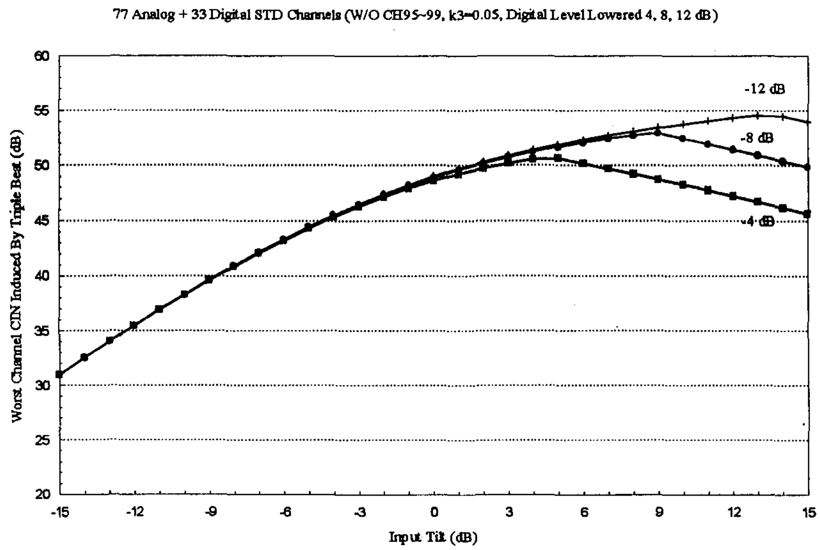

77 Analog + 33 Digital STD Channels (W/O CH95~99, k3=0.05, Digital Level Lowered 4, 8, 12 dB)
Worst Channel C/N Induced By Triple Beat (dB)
Input Tilt (dB)
-12 dB
-8 dB
-4 dB
60
55
50
45
40
35
30
25
20
-15
-12
-9
-6
-3
0
3
6
9
12
15

Figure 13

On-Demand Network Transport Architecture

Louis D. Williamson
Sr. Project Engineer, Advanced Engineering
Time Warner Cable

Abstract

The paper explores the architectural requirements to support an on-demand interactive digital network within the hierarchy of a hybrid fiber coax network. A traditional entertainment network is fundamentally a broadcast network. As new services, such as telephony, high speed modems and interactive televisions are introduced, can the same physical transport support them? What are the differences in their requirements? What are the behavioral differences of a broadcast network versus an on-demand network? The paper reviews the current distribution hub-master hub based HFC network and discusses available options to accommodate new digital switched services.

INTRODUCTION

All HFC networks are not created equal. Time Warner Cable's Master Headend, Transport Hub, Distribution Hub architectures were designed to accommodate various new businesses proposed by the industry. These new services can be categorized into three businesses; Entertainment, PC Data, and Telephony. They all have specific requirements that necessitate different implementation strategies on the Hybrid Fiber Coaxial (HFC) network.

Entertainment services require a high bandwidth, constant bit rate, highly asymmetrical communication path to deliver MPEG-2 programs to the television. PC data requires a high bandwidth, bursty, variable bit rate, communication path to delivery data to the personal computer. Telephony requires a low bandwidth, symmetrical, isochronous communication path for delivery of voice to the telephone. Equipment developed to deliver each of these services is optimized to satisfy their delivery requirements. All must be layered upon the existing HFC plant that has been optimized for the delivery of broadcast analog television channels. This paper will discuss how the on-demand interactive entertainment signals are added to an existing HFC plant.

GOALS

The traditional cable television system has been a broadcast network. The headend was the origination point for all services delivered throughout system. All of the services flowed from this point to all subscribers. Occasionally there maybe some local signals that were inserted at different areas to fulfill franchise requirements, but for the most part all customers receive that same channels. In addition, most of the services delivered were analog. Lately, digital services such as video game download channels, and digital music services have been deployed, but they remain as broadcast service. The primary delivery mechanism for these services has been amplitude modulated fiber-optic links and analog coaxial cables optimized for delivery of broadcast services.

On-demand interactive entertainment services bring major changes to this method of delivery. No longer can you deliver the same channel mix to all the customers since a customer in an area may request a specific program at any time. An on-demand interactive entertainment system must be designed to handle these on-demand requests simultaneously. The number of simultaneous requests that the network can service is referred to as the 'peak load' of the network.

Designing for peak load is common for telephone companies, and they have a wealth of historical data to base their design upon. For an on-demand interactive network, peak loading behavior is not very well known. No one really knows how many customers will be requesting the new release of a blockbuster hit the first Saturday night it is available on the network, or what happens to a movie's demand over time. The network must be designed to be able to deal with a variety of situations that may require that more channels be available in one area than another. This focusing or narrowcasting of channels is a new concept that cable television engineers must understand to be able to design cost effective on-demand networks.

It is imperative that an on-demand interactive network be integrated into our existing broadcast based system, if only to preserve the investment to date. On-demand interactive services by nature are sent to a specific customer in a specific area. The network must now support the routing of a channel to an area. But, only a portion of the available channels are on-demand, the majority of channels are still broadcast to all customers. The introduction of new on-demand interactive entertainment services must consider the existing infrastructure and to work harmoniously with it.

New services must be introduced incrementally and cost effectively. Even though we all believe that on-demand interactive entertainment services will be a great business someday, it would not be prudent to design or build a network to handle those anticipated peak loads today. The network must be designed so it can grow as the demand for services increases.

In review, the goals of an on-demand interactive entertainment system are:

1.Capability of handling peak loads.

2.Ability to grow as demand increases (scalability).

3.Fit into existing broadcast infrastructure.

4.Cost effective deployment.

Let review the components necessary to support an on-demand interactive entertainment system and TWC's Hybrid Fiber Coaxial design to understand how to implement a network that meets these goals.

INTERACTIVE COMPONENTS

The required components for an on-demand interactive television network are: media server, ATM switch, cable gateway, QAM modulators, and a data channel gateway. These components and their interfaces will be briefly discussed to determine the requirements that each places on the HFC network.

Set-top terminal

The set-top terminal is the consumer interface device. The set-top terminal is designed to receive both analog channels and MPEG-2 digital channels. The set-top terminals are usually designed to receive both analog and digital channels with one 6 MHz tuner.

Since it is necessary to be able to communicate with the set-top terminal while the consumer is watching an analog program, most set-top terminals have an out-of-band data channel. The out-of-band channel is referred to as the Forward Data Channel (FDC).

For on-demand interactive entertainment services, the set-top terminal must have a real-time Reverse Data (RDC) to communicate back to the headend. Both the FDC and the RDC use standard IP protocol for encapsulating messages.

Media Servers

Media servers are large computing platforms that are optimized for delivering MPEG-2 compressed video and audio streams. Proposed servers can deliver from 100 to 250 different 3 megabits per second video and audio streams. OC3 and OC12 interface have been proposed to connect these servers to the ATM switch.

Media servers are designed to be installed in an environmentally conditioned room. They also require much more maintenance than typical cable television equipment. The maintenance required includes: loading and removing assets, replacing disk drives, and cleaning fan filters.

ATM Switch

An ATM switch is required to route the MPEG-2 streams that originate from the media servers. Although the routing of the stream can be done with the media servers, the ATM switch provides for redundancy of media servers and helps tremendously with load balancing and scalability. The input ports on ATM switch are OC3c or OC12c to accommodate the outputs for the media servers. The outputs of the ATM switch are OC3c. The output feeds the interactive cable gateway.

Cable Gateways

The primary purpose of an Interactive or Broadcast cable gateways is to create 'Funny Cable Rates' (FCR). Today there are 2 FCRs, FCR1 that accommodates 64 QAM and FCR2 that accommodates 256 QAM. These rates are needed because they are the rates that can be put into a 6 MHz channel spacing using QAM modulation. FCR1 is approximately 27 megabits per second FCR2 is approximately 36 megabits per second. The input to a cable gateway is a SONET OC3c. The output from the cable gateway is electrical TAXI.

Cable gateways also perform some additional functions that are required including: PID re-mapping, stream encryption, and video and audio re-synchronization.

QAM modulators

Quadrature Amplitude Modulation (QAM) is the preferred modulation method for transportation over the HFC. The QAM channels that are created by the modulators are referred to as Forward Application Transport (FAT) channels. These channels carry MPEG-2 video and audio to the set-top. FAT channels can also be used as high speed download channels to dynamically send new applications to the set-top that the consumer request.

Data Channel Gateway

The Data Channel Gateway (DCG) is used for sending commands and other information to the set-tops terminal. The DCG consists of two primary units, the Forward Data Channel (FDC) modulator and the Reverse Data Channel demodulator and an IP router. Together these two units send and receive all messages for the set-top using standard IP protocol. Additionally, the DCG must provide the timing necessary to synchronize all of the set-tops attached to the coaxial plant serviced by the DCG.

Together these pieces are the primary communication elements that are needed to support an on-demand interactive entertainment system. There are many ways that these elements can be assembled to create an on-demand interactive entertainment system. Next, we will briefly describe Time Warner Cable's HFC network and how we are planning on assembling these components to form a scaleable, cost effective system that can handle high peak loads.

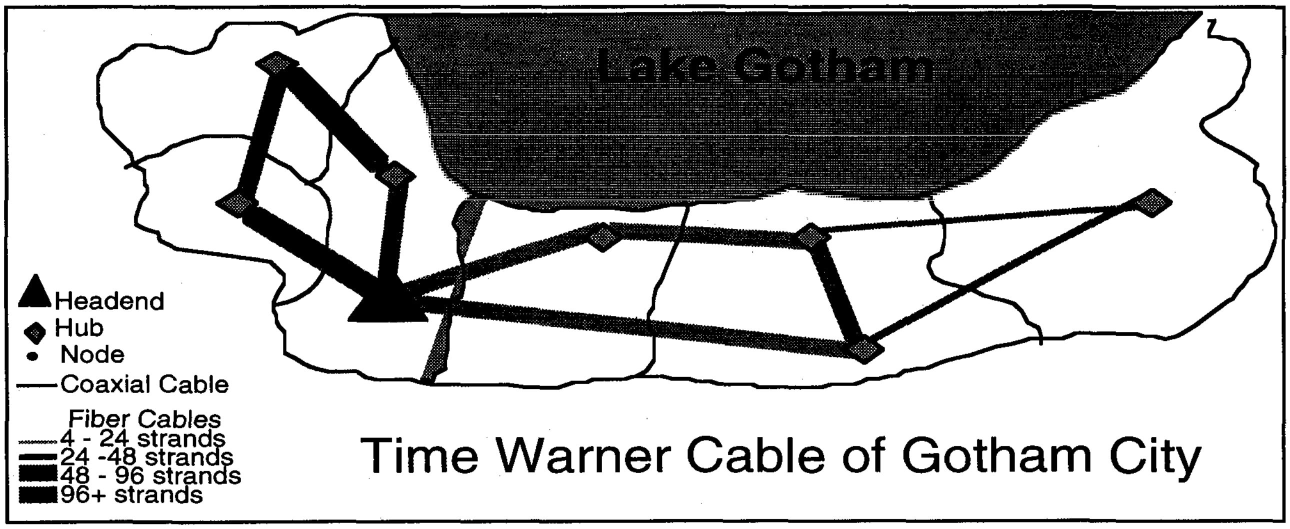


Figure 1 Headend and hub transportation architecture.

TWC'S HFC DESIGN

There are 3 primary parts to the TWC's HFC design, the headend, hubs, and HFC plant. (See Figure 1).

Headend

The headend is the main facility in a cable television system. The headend is the origination point for most of the analog services offered to our customers. This will also be the facility where most digital services will originate, especially early in the deployment stage where penetration of digital set-top terminals is low. This is the facility where the servers, media storage, ATM switches, and QAM modulators are located.

There is fiber optic cable continuity between the headend and all of the hubs in the cable system. Analog and digital channels are transported over these fiber optic cables from the headend facility to the hubs in the network.

The headend is a staffed facility. The headend is an environmentally conditioned location. These are important considerations for equipment like media servers and ATM switches that require frequent maintenance and stable environmental conditions.

Hubs

The Hub in a TWC system is the facility that is used to serve approximately 20,000 home passed. The hubs are connected to the master headend by a ring of fiber optic cables. There are typically 5 hubs on each fiber optic cable ring. Six fiber strands are dedicated to each hub for on-demand interactive entertainment services. There are typically 40 nodes served from a hub. A node typically serves 500 homes passed.

The hubs vary in size from an Optical Transport Node (OTN), which is a large pedestal, to a small building that may be up to 500 square feet. This facility is typically not staffed. There is standby powering at all hubs. All equipment that is located at this facility must be capable of remote monitoring and configuration.

Hybrid Fiber Coaxial Plant

The plant architecture described here is commonly called the Hybrid Fiber Coaxial (HFC) plant. Four parts compose an HFC plant; the transportation fiber optic cables, the distribution fiber optics cables, the node, and the coaxial cables.

The transportation fiber optic cables interconnect the headend to all of the hubs in a logical ring network configuration. The ring provides redundancy to all hubs for analog broadcast signals and the telephony signals. Redundancy may not be provided for on-demand interactive digital services. Typically, five hubs are on each transportation fiber optic cable ring. The distance from the headend to a hub is typically less than 20 miles. If this distance exceeds 20 miles, the signals may be repeated at an intermediate hub, called a transport hub, along the fiber optic cable path.

The Distribution fiber optic cables connect the hub to the nodes in a logical star network configuration. Up to 4 nodes may share one fiber strand in the forward direction. All reverse nodes are served by individual fiber strands. Typically, four fibers are available at each node, but all services are planned to be delivered on one fiber strand. No redundancy is provided at the node. (See Figure 2.)

A node is the optical-to-electrical conversion point. The node converts the forward optical signals into electrical signals for distribution over the coaxial portion of the plant. The node also converts the electrical signals from subscriber set-top terminal equipment in the home into optical signals for delivery back to the headend. The node typically serves 500 homes.

Coaxial cables connect the node to the homes in a logical tree network configuration. Coaxial cables are preferred in this portion of the plant because the electrical signals can be easily repeated with RF amplifiers. Repeating is necessary to overcome the losses of the splitter that are used to feed the homes. Typically, six amplifiers or fewer are in series between the node and a home.

ON-DEMAND TRANSPORT ARCHITECTURE

Figure 3 shows the TWC on-demand transport architecture. The media servers, ATM switches, cable gateways, and QAM modulators are located at the headend. The headend provides an environment that is much more suited for these components in the early phases of deployment. As the demand for service increases, the components at the headend may be replicated to provide for redundancy. A data channel gateway is located at each hub.

Each hub will have six optical fibers dedicated to on-demand services. Four optical fibers are used transport the FAT channel to the hub. Two optical fibers are used to transport signals to the data channel gateway. TWC has chosen to use amplitude modulated fiber-optic links to distribute the MPEG-2 FAT channels to the hubs, but has chosen to use SONET OC3 links to distribute the control signals to the data channel gateway located at the hub. This 'hybrid' transportation approach provides many of the requirements necessary to achieve the goals stated earlier.

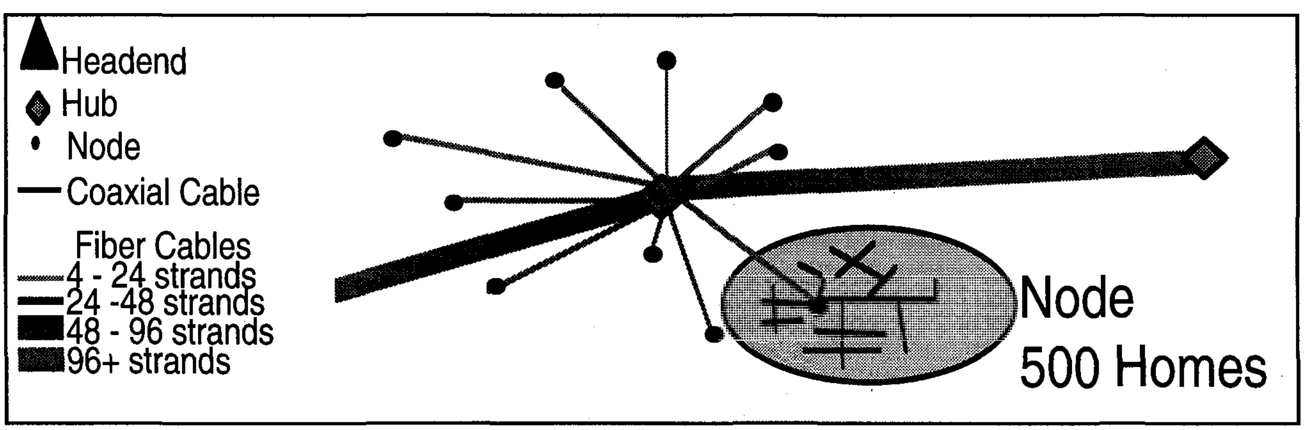


Figure 2 Hybrid fiber coaxial distribution architecture.

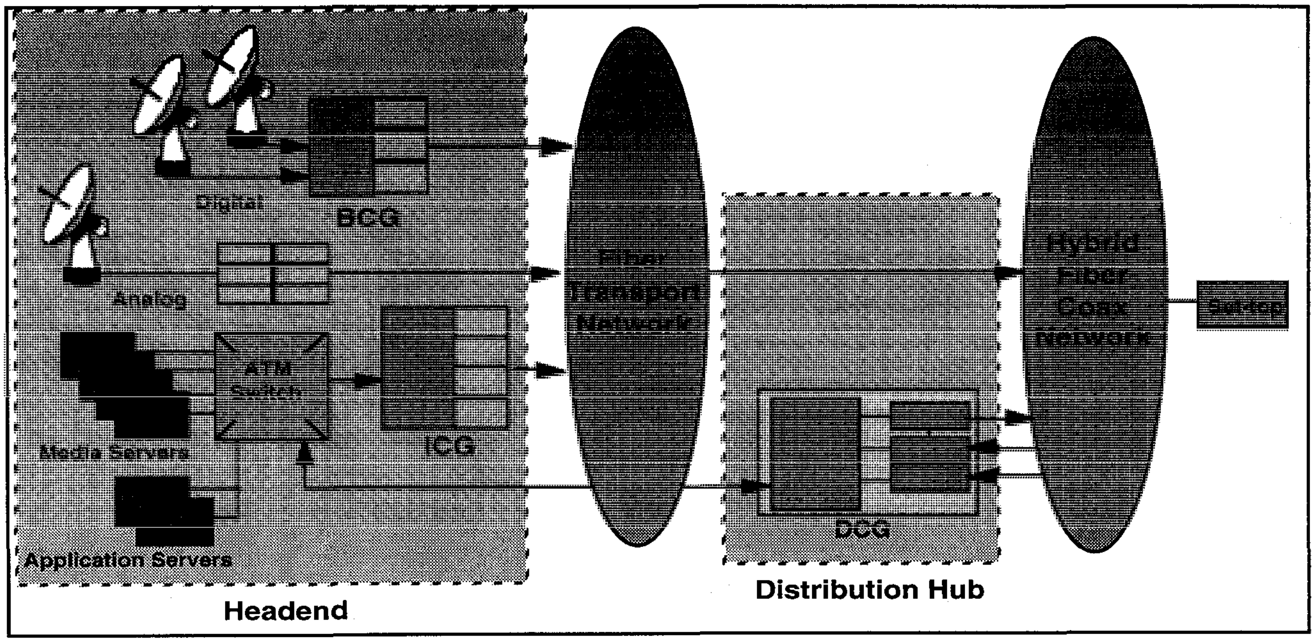


Figure 3 On-demand interactive entertainment architecture.

FAT Channels

When you contrast SONET transport versus amplitude modulated transport on the fiber optic cables, the amplitude modulated fiber optics is the most economical. On a single fiber we are able to transport 621 megabits per second to the hub using frequency division multiplexing. The FAT channels are transported to the hub on the frequencies that they will be carried on over the coaxial cables. At the hub they can be combined passively with the analog channels before distribution to the nodes. If we had chosen SONET OC12, then we would have needed 2 fiber to transport 622 megabits per second of information to the hubs. SONET transport would also have required add drop multiplexers and QAM modulators located at the hubs. Since the hubs are small and unstaffed, this was impractical in many situations.

It is very easy to introduce an on-demand service using amplitude modulated links. You can literally broadcast it to all the hubs, or to some number of hubs. As demand increases, you can segregate the hubs and begin to supply each with dedicated channels. With SONET, you must install the entire infrastructure before you can begin to offer on-demand services. You could start with OC3, but even this is more expensive than starting with amplitude modulated links.

Currently, we are constructing 750 MHz bandwidth plant. This gives us forward channel capacity from 50 to 750 MHz, and reverse channel capacity from 5 to 40 MHz. Current frequency allocation is for analog channels to be in the 50 to 550 MHz range. This gives us capacity for 80 analog channels. The 200 MHz of spectrum from 550 to 750 is reserved for all future services including on-demand interactive entertainment service, PC data services, and telephony services. It is anticipated that 138 MHz of this spectrum will be used for on-demand interactive entertainment services. This provides for 23 FAT channels on the coaxial portion of the plant, or 621 megabits per seconds using 64 QAM modulators. Since there are 4 optical fibers available to a hub, this provides almost 2.5 gigabits per second to a hub that serves 20,000 homes, or enough capacity to simultaneously supply a 3 megabit stream to 4% of the homes.

Data Channel Gateway

Even though we chose to use amplitude modulated links for the FAT channel transportation, we chose SONET for transportation to the data channel gateway.

The Forward Data Channel (FDC) originates at the hub. The Reverse Data Channel (RDC) terminates at the hub. These devices are located at the hubs for the following reasons:

1.Reverse bandwidth can be reused by segmenting the returns from different neighborhoods.

2.Thermal noise power is reduced by segmenting the returns from different neighborhoods.

3.FDC and RDC are typically tied together to synchronize the reverse transmission slot timing for set-tops operating in the neighborhoods.

4.The reverse bandwidth efficiency is increased by minimizing the round trip propagation delay.

The input of the FDC and the output of the RDC at the hub are ethernet. Ethernet is easily remote through SONET. Hubs are typically unstaffed. We need to be able to remotely control and monitor the equipment located at the hubs. Simple Network Management Protocol (SNMP) over ethernet was the most economical choice. Since this is IP-based it can easily be transported over the SONET links also.

There are also some economies when all 3 new businesses are deployed that become available for this relatively low speed data. Once PC data and telephony businesses are deployed, a SONET infrastructure will be installed to the hubs to support them. The bandwidth needed to support the DCG can be piggybacked to these businesses with minimal additional cost. Adding enough SONET equipment to support the FAT channels in a fully deployed entertainment business would still cost more than using separate amplitude modulated fiber-optic links.

SUMMARY

By using a combination of amplitude modulated fiber-optic and SONET links, an on-demand interactive entertainment system can be added to an existing hybrid fiber coaxial network. Amplitude modulated fiber-optic link provides a low cost entry into these new services and can grow as service demands increases. The network infrastructure scales from relatively small peak demands to very high demands, making the modern HFC network very suitable for the evolving future of interactive television.

Open Architecture for Multi-Vendor Compatibility of Element Management and Status Monitoring

Kiran Babu
Design Engineer
Philips Broadband Networks Inc.

Abstract

This paper addresses the software architecture for an Element Management System which can be developed keeping protocol dependency, device specifics and management platform dependency transparent. The unique aspect of this architecture is that it provides the seamless integration of any kind of device from any manufacturer and any kind of communication at the Element Management System. This paper provides an Object Oriented Design technique of such an Element Management System which can be easily migrated to Distributed Object Oriented Network Management of the future.

Introduction

Element management and status monitoring products currently offered by manufacturers are exclusively for the devices and networks developed by those manufacturers. Desirable management systems will be flexible enough to monitor products from many manufacturers with various communications protocols. As networks become more complicated, element management software based on an open architecture makes the most sense.

History

Initially, cable monitoring systems merely verified that a device was operating. Gradually, monitoring systems evolved to report on the status of key operating parameters. A typical status monitoring application had four components for each class of device monitored: user interface, communication protocol, performance analysis, and database storage.

To monitor the operational status of their equipment, manufacturers developed proprietary status monitoring applications that used unique protocols to communicate with each type or class of device. User interfaces, performance analysis, and database storage components were also unique to each device class. Developing specialized status monitoring applications was not only time consuming, but also resulted in a limited, single-manufacturer monitoring system.

Proprietary communications protocols are the major obstacle to developing a flexible system that is capable of monitoring the devices of various manufacturers. This problem was identified not only in the cable industry, but also in the data communications and telecommunications industries.

Migration to Better Concepts

As networks grew in size and capability, simple status monitoring evolved into hierarchical network management. This network management hierarchy, shown in Figure 1, has four layers, each with a specific function.

The **element management layer** accepts information from all devices (or elements) in a network, translates the information into a form understood by all other layers, and passes it along to the network management layer. This layer consists of a third-party open platform running multiple manufacturer-specific applications.

The **network management layer** manages multiple element management layers.

The **service management layer** manages the services of multiple networks.

The **business management layer** manages the services layer below it and implements business strategies.

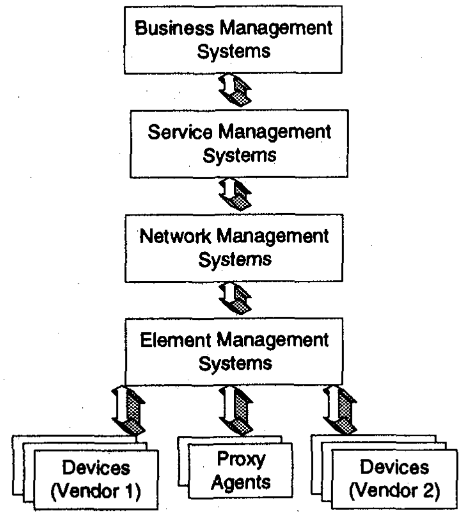


Figure 1.
Network management hierarchy has four layers

To achieve interoperability, the devices need to have a standard interface protocol instead of proprietary protocols. Specifications have been developed for a standard management protocol and a standard way of describing the management information of the devices.

Although new device development will be based on these standards, legacy devices still present the problems of multiple protocols. The element management layer plays a critical role in solving these problems. Proxy agents in the element management layer act as translators between the proprietary protocols used by the legacy devices and the standard management protocols used by the other layers in the hierarchy. Once this translation has occurred in the element management layer, the other layers need not address this issue.

This paper focuses on the element management layer, where the protocol translation and multi-vendor interoperability takes place. The three higher layers are generic for any organization and can be handled by third-party vendors.

Third-Party Open Platform for Element Management Applications

Third-party developers provided an open platform with basic features. An element management platform is open when any number of manufacturers can develop the applications specific to their devices under the same platform. Basic features include creating networks, triggering device-specific user interfaces, managing databases, and communicating with agents through standard protocols.

To manage a manufacturer's device from this open platform, a separate interface called a manufacturer's specific application must be created. These manufacturer-specific applications should be integrated with the platform's core software using application protocol interfaces and may include such things as device-specific user interfaces, databases, and performance analyses. This integration of software will be transparent to the user.

The following figure shows the relationship between the open management platform and a manufacturer-specific application.

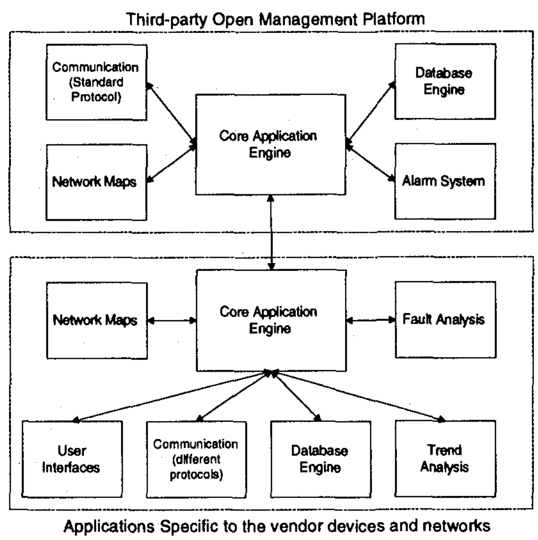


Figure 2.
Open platform supports manufacturer-specific applications

Problems and Solutions

Most third-party open element management platforms assume that the manufacturer-specific applications use the basic features provided. This would allow manufacturers to concentrate on the portions specific to their devices. Most manufacturer-specific applications, however, bypass the third-party open platform and are developed from scratch for the following reasons.

- **Problem:** No open management platform in the market supports all of the standard and proprietary communications protocols. The open management platform may provide communication support for one standard protocol, but there are many standards (for example, SNMP, CMIP, CORBA, and TL1) and more being developed.

  Solution: This problem can be resolved if a third-party platform can be customized to handle any communication protocol.

- **Problem:** Network map creation is not standard, and different management applications use different methods. Most of the open management platforms support a geographical organization of devices; however, there is no standard defined in this area.

  Solution: It is possible to define standards for network map creation common to cable, data communication, and telecommunication networks. For example, headend or central office devices can be grouped in racks and shelves and field devices can be grouped geographically in different regions.

- **Problem:** Open management platforms provide limited support for device-specific user interfaces because information and graphical representation of any type of device is unique. Although third-party open management platforms can provide some support for vendors to describe and graphically represent their specific devices, device-specific user interfaces are often created from scratch.

  Solution: Third-party, open platforms can be customized to include any device from any manufacturer.

- **Problem:** Open platforms provide very limited support for fault analysis. Once the manufacturer-specific application receives a fault notification from the device, it performs a root cause analysis using device dependency information to track the fault. The open platform applications do not recognize device dependencies.

  Solution: The open platform can be customized to include device-dependency information. This is possible only if the devices are maintained as generic objects with dependencies defined for their parameters.

- **Problem:** Not only do open platforms provide limited support for fault analysis among products from the same manufacturer, but they also cannot track faults among products from different manufacturers. Each manufacturer-specific application is maintained as a separate entity in open management platforms. A device managed by one application may be dependent upon a device managed by another application. These dependencies must be communicated between applications to track trends and resolve faults.

  Solution: Open platforms must be expanded to keep track of the dependencies of the parameters within a device, between different device types from the same manufacturer, and between devices from multiple manufacturers. This would allow the open platform to handle root cause fault analysis for all manufacturers' devices.

- **Problem:** Third-party open platforms have limited access to information about the devices being managed by manufacturer-specific applications. Therefore, database storage, reporting, trend analysis, and trouble ticketing are implemented by the manufacturer-specific application.

  Solution: If the device information is maintained by the open platform, it is possible to have a generic implementation of reporting, trend analysis, and trouble ticketing functions.

Resolving these problems will lead to an element management system that takes maximum advantage of the features provided by an open platform. The complete element management system requires an open platform which can be customized to allow element management of devices from multiple manufacturers.

An Object-Oriented Solution

Managing many devices from various manufacturers can be accomplished when the element management system uses an object-oriented approach. An object-oriented approach uses a generic object created in the element management system to represent an actual device in the network. The user communicates with the object, and the object communicates with its counterpart in the network to get and receive information. For this to occur, the object must know about the device it represents, as shown in Figure 3.

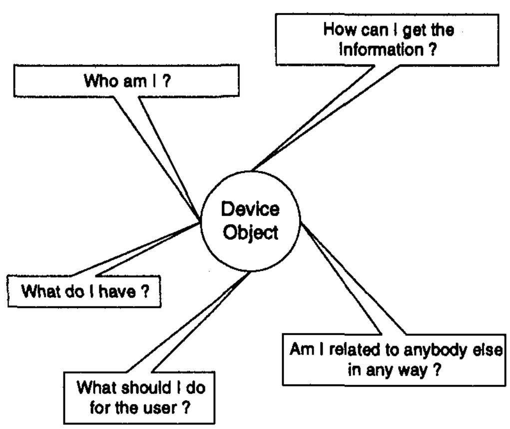


Figure 3.
The object must know about the device it represents

Each object is identified by its device class or type. Then, it is further distinguished from others in its class by key parameters. Some parameters are static, while others are dynamic. Dynamic parameters can be automatically updated by the application or manually updated by the user.

Thus, to manage a device, an object representing the actual device has to be created in the element management system. Through an interface for each device class, the user will interact with the object to initialize, update, monitor, detect and analyze faults, and analyze trends.

Initializing Objects

The user initializes the object by defining the device it represents. This definition specifies the type of device and values for key parameters which uniquely identify the device in its class. To do this, the object has to know what interface to display and the key parameters that require user input. Once the object gets enough information from the user to identify the actual device it represents, the object can communicate with the actual device to get values for additional parameters.

Updating Parameters

The user may want to change some device parameters. To allow such changes, the object must know which interface to display and which parameter values the user can change. The user communicates the changes to the object, and the object then communicates them to the actual device.

Monitoring Parameters

An object must know which parameters to monitor and their acceptable ranges. In addition, it must know which user interface displays the parameter values. Then, the object gets the parameter values from the actual device and shows them to the user.

Detecting Faults

If a parameter value falls outside of the acceptable range, the user is notified of a *critical condition* by either the actual device or the object. The user defines the acceptable range for the object, and the object communicates the range to the actual device.

Objects regularly poll actual devices to verify that monitored parameters fall within the acceptable range. If an actual device or its agent has intelligence to check the parameters being monitored, it can detect critical conditions and notify the element management system directly.

Analyzing Faults

In a network, a fault in one device may depend on a fault in another device. This may be because both devices are on the same signal path. For example, an abnormal parameter value in device A may depend on an abnormal parameter value in device B. Similarly, device B may be dependent upon device C, and so on. These dependencies may be within the parameters of a device, within devices of the same class, and even within devices from different classes. The list of dependencies in a network is called the inference engine. Using the object, an inference engine can be defined for each device in the network .

Individual objects can be formed into a network representing the actual network being monitored. If an object has an inference engine, it knows the other objects in the network on which it depends for a fault. Hence, a fault for an object can be analyzed and the root cause of the actual device's fault can be identified. This method of analyzing faults is generic, but the inference engine for each device is specific.

Analyzing Trends

To analyze trends an object compares its performance to a history database of parameter values. Each object class can have templates defined by the user to support trend analysis. These templates can define the database storage, report, and real-time graphic structures for each device class. Trend analysis can be a separate process analyzing the performance of devices and networks. This process can be totally independent of the objects, since it is based on information already stored.

The catch is that the trend analysis process learns about the actual devices being managed from the database. To do this the database storage must be generic, so that reporting functions like trend analysis can learn about the devices from the database itself. This is feasible because objects are generic and learn about themselves from templates.

Communicating between Objects and Actual Devices

Communication between the object and the actual device it represents is not generic because the actual device may use any protocol, standard or proprietary. But a generic approach can be developed to a certain level by using specific protocol handlers, also called communication handlers.

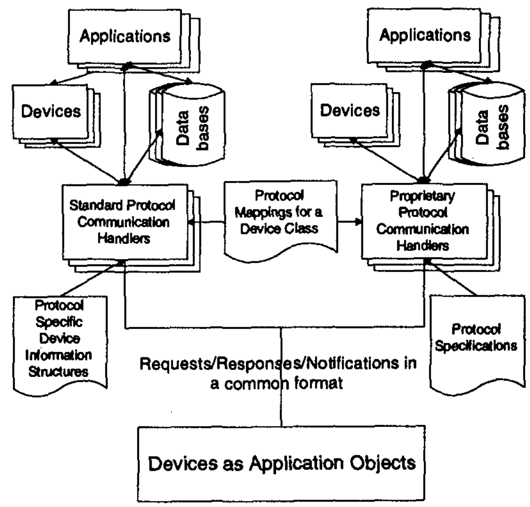


Figure 4.
Specific protocol handlers aid in communication between objects and actual devices

A device with a standard protocol maintains the device information in a certain format. A device with proprietary protocols also maintains the device information in a certain format. A communication handler for each protocol extracts the device information structure and maintains it in a common format. The communication handlers for proprietary protocols should have protocol specifications. A common protocol should be defined between objects and communication handlers.

Typical command structures between the objects and communication handlers to get or set a parameter value in the actual device or collect responses and notifications are shown below.

<Set/Get><DeviceClass><Key Parameters to Identify Device Instance><Request Parameter ID>[Value]

<Response/Notification><DeviceClass><Key Parameters to Identify Device Instance><Request Parameter ID><Value>

Communication handlers will map the requests to protocol-specific requests and handle the communications for each object. Each device class has to be associated with a corresponding communication handler.

The Total Solution

Adding a new device class to the element management system requires that device management information, dependencies, user interfaces, and communication be customized.

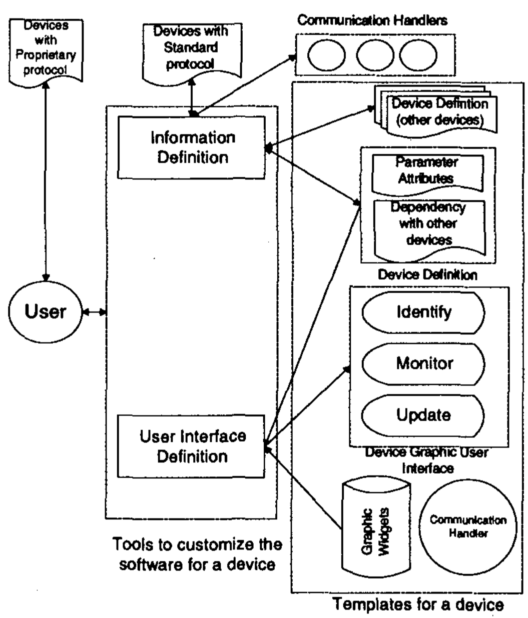


Figure 5.
Customizing to add a new device class

Figure 5 shows how to customize using tools in the element management system.

The information definition tool displays information about the other device classes, so the user can assign dependencies. These dependencies will be defined for each class. This tool should understand the device information specifications written for different standard protocols like SNMP, CMIP, CORBA, TL1, and SQL. Device information structures defined with proprietary protocols can be manually entered by the user or read from files with an intermediate standard defined for proprietary protocols. This tool can prepare a common structure for device definition, which describes the device class identification, protocol identification, and parameter attributes. In addition, the information definition tool should identify a specific protocol communication handler for a device class.

The user interface definition tool provides the user interface screen definitions for each device class. With this tool, the user defines graphic widgets for each parameter of the device class. The object associates parameters to these graphic widgets. The user interacts with the graphic widgets to monitor and update parameter values. User interfaces defined for a device class are shared by all objects in that class.

These tools help the user create a template for each device class. An object will be created from this template whenever the user adds a device to the network map. Specifics of the device class will be read from this template. Implementation of the object for any class will be almost generic and can be applied to any device from any vendor. In addition, whenever a object is added to the network, a corresponding communication handler is defined in the template and attached to the new object.

The graphical organization of the managed devices is a network map. Each device can be represented by a specific icon, and the icons can be arranged to represent the actual network topology. The network map should be created using passive elements like icons for each device, geographical maps, and racks and shelves with different configurations. These elements help the user create a graphical representation of the network with devices as groups.

Figure 6 shows a run-time version of an element management system.

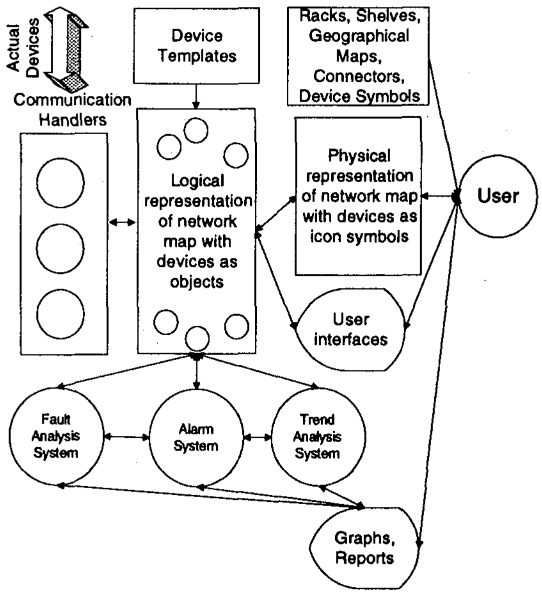


Figure 6.
A run-time element management system

Since the manufacturer-specific information can be customized and retrieved from templates, the whole element management system can be generic, and an open platform can accommodate multi-vendor interoperability.

This type of architecture will save users a significant amount of time, since adding new devices can be done easily with tools provided by the element management system.

However, a major effort is required to implement communication handlers for each protocol. This effort could be avoided if the industry agrees upon one standard management protocol.

Distributed Object-Oriented Management

The architecture proposed so far is based on object modeling of a device. For an object to communicate with its counterpart in the network, huge amounts of time, effort, and money are being devoted to implementing a protocol wrap-around. If the management system object is implemented as a client object and the corresponding server object is implemented as an actual device, a time-consuming effort can be avoided. This is possible with CORBA (Common Object Request Broker Architecture), and it is ideal for network management of massive networks like cable which contain many elements.

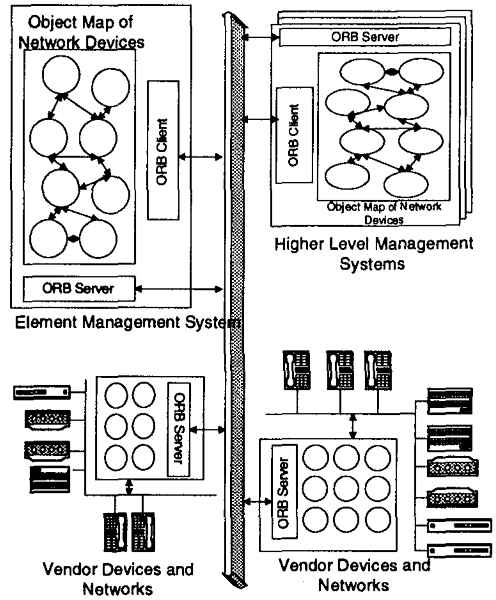


Figure 7.
Distributed object-oriented management using CORBA

With CORBA, the object exists in the device or agent itself. A mirror image is maintained at the management system. The communication wrap around is invisible to the developer. This approach has a significant impact because objects can be distributed to any level of management. As a result, any level of management can be applied at any level. If a distributed, object-oriented protocol is used, the architecture proposed in this paper will plug and play for any device a manufacturer can come up with.

Conclusion

Designing an element management system based on an open platform, as proposed in this paper, will allow operators to customize the system to accommodate any new device from one or multiple manufacturers. A management system based on this open platform architecture allows the operator to add new devices without developing new device-specific applications. Instead, the operator takes advantage of the open platform and customizes it to incorporate management information, fault patterns, user interfaces, and communications protocols of the new devices.

The Realities of The Retail Sale of "Navigation Devices"

Walter S. Ciciora, Ph.D.
Technology Consultant
Southport Connecticut

Abstract

The 1996 Telecommunications Act included SEC. 629. COMPETITIVE AVAILABILITY OF NAVIGATION DEVICES, *also called the "Bliley Amendment" after its author, Congressman Bliley. The potential of substantial hazards has been created. The danger exists that a) consumers will be confused and purchase devices that fail to meet expectations, b) the technical integrity of services will be impaired, c) reception of signals on other receivers in the home will be degraded, d) interference will be caused to other subscribers, e) signal leakage will hamper legitimate uses of the over-the-air spectrum, and f) economic waste will result.*

With adequate care, many of these hazards can be minimized and the consumer offered choices. Accomplishing this will not be easy.

INTRODUCTION

The 1996 Telecommunications Act included a section on the commercial availability of Navigation Devices. This section creates a requirement that the FCC establish rules. It is important for all involved with in-home equipment to become familiar with this section of the law and to participate in the formation of the FCC's rules. The financial and operational consequences of this law and the rules that will be issued are substantial. Subscribers will be affected in many ways. The likelihood of subscriber problems has been increased dramatically. The devil is in the details. The fine points of the FCC rules will dictate the outcome. As is often the case, Congress leaves the hard work to the regulatory agency.

Subscriber education will be more important than ever. Cable operators will have to engage in extensive information campaigns to ensure that consumers make informed choices which maximize the value of their cable subscriptions.

THE LAW

The section of the law reads as follows:

SEC. 629. COMPETITIVE AVAILABILITY OF NAVIGATION DEVICES

"(a) COMMERCIAL CONSUMER AVAILABILITY OF EQUIPMENT USED TO ACCESS SERVICES PROVIDED BY MUTICHANNEL VIDEO PROGRAMMING DISTRIBUTORS. – The Commission shall, in consultation with appropriate industry standard-setting organizations, adopt regulations to assure the commercial availability, to consumers of multichannel video programming and other services offered over multichannel video programming systems, of converter boxes, interactive communications equipment, and other equipment used by consumers to access multichannel video programming and other services offered over multichannel video programming systems, from manufacturers, retailers, and other vendors not affiliated with any multichannel video programming distributor. Such regulations shall not prohibit any multichannel video programming distributor from also offering converter boxes, interactive communications equipment, and other equipment used by consumers to access multichannel video programming and other services offered over multichannel video programming systems to consumers, if the system operator's charges to consumers for such devices and equipment are separately stated and not subsidized by charges for any such service.

"(b) PROTECTION OF SYSTEM SECURITY. – The Commission shall not prescribe regulations under subsection (a) which would jeopardize security of multichannel video programming and other services offered over multichannel video programming systems, or impede the legal rights of a provider of such services to prevent theft of services.

"(c) WAIVER – The Commission shall waive a regulation adopted under subsection (a) for a limited time upon an appropriate showing by a provider of multichannel video programming and other services offered over multichannel video programming systems, or an equipment provider, that such waiver is necessary to assist the development or introduction of a new or improved multichannel video programming or other service offered over multichannel video programming systems, technology, or products. Upon an appropriate showing, the Commission shall grant any such waiver request within 90 days of any application filed under this subsection, and such waiver shall be effective for all service providers and products in that category and for all providers of services and products.

It is not clear where the term "Navigational Devices" comes from. The impact of the law is much broader than just electronic program guides which is the usual meaning of the term. All set top boxes are included in the scope of this law.

Set Top Boxes vs. Consumer Electronics

In the past, the Consumer Electronics Industry has positioned set top boxes as evil things, the work of the devil. From their previous point of view, these nasty things get in the way of all of those wonderful TV and VCR features and functions (whether you want those capabilities or not!). They believe that the reason consumers don't use TV and VCR features and functions is not because they don't care for them, but because the evil set top box gets in the way.

Sometime in 1995, there was an apparition from heaven that told the consumer electronics industry that they were not seeing things clearly. The set top box is in fact a wonderful thing that can be sold at retail! Rather than trying to kill off set top boxes, the consumer electronics industry now wants to make and sell them. We've entered a new era.

It may be that the consumer electronics industry came to the realization that the cable set top box has nearly all of the same characteristics as a VCR when considered from the perspective of the interaction with a TV. The VCR has a tuner, power supply, remote control, and a remodulator. These elements duplicate the corresponding elements in a TV. The VCR requires the TV to be tuned to channel 3 and left there. Alternatively, both VCRs and set top boxes are available with base band video and audio outputs. Many consumers use the VCR's remote control to tune channels and the TV's remote control to adjust volume. These consumers don't miss Picture In Picture (PIP) since they've forgotten they have it. Other TV tuner features such as sleep timers are rarely used anyway. VCRs have automatic bypass switches which convey the full spectrum to the TV's tuner when the VCR is off. The same feature is available on set top boxes. The VCR is in almost every way a "set top box" which has as much interference with the features and functions of a TV as does a cable set top box. Yet almost no one complains about VCR interference with TV features!

Now that nearly all TV households have more than two TV's and most have at least two VCRs, the consumer electronics industry needs something else to sell. Since the VCR is really just one type of set top box, the expansion of this category to include other types of set top boxes is attractive.

A Cable Operator's View

Very little in life is all bad or all good. Most things have advantages and disadvantages. From a cable operators' perspective, the advantages to subscriber ownership of set top boxes are primarily economic and potentially customer satisfaction. Certainly getting the subscriber to pay for the hardware and its maintenance is a real blessing. Also, there is a potential that the subscriber who owns the set top box will see it as a benefit rather than as something that gets in the way of TV and VCR features and functions. There is something about a pride of ownership that brings a "halo" effect.

Commercial Availability

Very clearly, the law of the land requires that the FCC make regulations, in consultation with appropriate industry standard-setting organizations, to assure the commercial availability of set top boxes for all kinds of cable services. However, it appears that the FCC is moving directly to a Notice of Proposed Rule Making (NPRM), bypassing its usual step of a Notice of Inquiry (NOI). This is very troubling since this is a very complex issue involving many affected parties. While it is true that Congress placed unreasonable demands and burdens on the FCC with the Telecommunications Act, the taking of shortcuts only increases the likelihood of serious problems later. Since Congress placed no deadlines for the FCC's completion of this difficult task, this haste seems unwarranted. Congress also did not place a deadline for the commencement of commercial availability. Clearly doing it right is more important than doing it early. The hazards of ill-conceived or inadequate rules will be with us for a long time since these navigation devices are durable goods with a decade or more of lifetime.

It isn't obvious how the FCC intends to comply with the requirement to consult with appropriate industry standard-setting organizations. Certainly, the NCTA Engineering Committee or the Society of Cable Telecommunications Engineers should be involved in any such consultation. Standard-setting organizations are experienced in the difficulties of reviewing all sides of issues and resolving them to a common position. There is no way to do this quickly. It takes time. Rushing to judgment will jeopardize the ultimate solution.

It is clear that the commercial availability of set top boxes will not limit the cable operator's right to also offer these devices. The law, as written, seems to offer the option of sale or lease by the cable operator. The important point is that the subscriber must be aware of the charges and that these charges are not subsidized by any service offerings.

Protection of System Security

Congress recognized the importance of protecting system security and the rights of service providers. There are only two ways to accomplish this. The first is to create a signal security system which is so secure that it cannot be broken. The second way is to separate out the signal security system and not allow subscribers to own and access it. In the latter case, the service provider is free to replace a breached security system without imposing financial loss on the consumer.

It is impossible to guarantee that a signal security system cannot be breached. There is no way to prove such a claim of invulnerability. The only alternative is for the entity wishing to sell the signal security system to consumers to guarantee the signal security system by placing sufficient funds in escrow to cover the full costs of recovery from a security breach. Since these costs will include the labor to replace the breached hardware as well as replacing the hardware itself, these costs are prohibitive. None the less, when cable equipment suppliers make large sales of set top boxes to cable operators, they pledge certain limited signal security guarantees. The cable operator makes a business judgment regarding these guarantees. Then the cable operator takes on the ultimate financial responsibility for fixing a security breach or replacing the defeated equipment without direct financial loss to the consumer. The retailer wishing to sell signal security directly to the consumer must also protect the consumer against direct financial loss in the event that the security system must be replaced.

The second solution is to separate out the security element and not sell it to the consumer. The service provider retains ownership and control of the security element. In the event of a compromise of the security, the service provider replaces the security element. There are problems associated with this. The analog security systems in use at present were not designed for such a partition. Forcing the partition may result in a loss in economics, a reduction of security, and a decrease in user convenience. Adjustments in the way analog security is done will be required. This will take time.

Waiver

Service providers or equipment manufacturers may request a waiver of the rules for assuring commercial availability of new technology still in its formative stages. For example, while standards are being settled for the hardware required to deploy a new cable service, non-standardized implementations used for market research or early introduction of the service may be granted a waiver. This is important because the standard-setting process is long and complicated. If every new service had to be standardized before market trial, new services would cease to be introduced. A further disadvantage of this approach is that the benefits of market trials of different approaches would be lost. In the current approach, multiple vendors supply a variety of implementations which are tested against consumer wishes. The implementation which best fits consumer needs and desires usually wins. In the event that it is learned that the service itself has insufficient appeal to justify a launch, a great deal of time and money is saved by the market trial. In this situation, no standards are needed and those who would like to provide hardware at retail are saved the losses involved in providing for a failed service.

The FCC must respond to the request for waiver within ninety days. Any such waiver applies to all service providers and to all equipment suppliers for that type of service.

SERIOUS HAZARDS

Serious hazards arise from subscriber set top box ownership. Significant extra costs and a loss in convenience can result from direct ownership. If the consumer fully understands the trade-offs and makes an informed choice, his satisfaction should increase. If the trade-offs are not clear, frustration, anger, and financial loss will likely result.

Servicing the Hardware

An important issue is the resolution of in-home problems. Most subscribers do not realize that the leasing of a cable set top box brings with it in-home pre-paid service. This service is generously provided and covers not only technical difficulties but even assistance in usage and user errors such as not checking to see if the set top box is plugged in. The lease fee covers the costs of providing this service. This is taken for granted. It will be important to advise subscribers who purchase their set top boxes that this assistance must come from the retailer, the manufacturer, or it will be available at an extra charge when provided by the cable operator. *Clearly* *managing subscriber expectations is critical to informed consumer choice and to minimizing and properly directing subscriber frustration and anger to its appropriate source*.

New Service Hurdles

A serious drawback to subscriber ownership of set top boxes comes from the hurdle this creates to taking new services. If the set top box owned by the subscriber does not have the technology to provide access to a service, the subscriber will have to choose between a) not taking the service, b) adding a second set top box to provide access to the service, and c) replacing the existing set top box with a new one that combines the old capability with the new. Lacking the technology for a service can arise from a) the service provider introducing a new service, or b) the subscriber moving from a cable system which did not offer the service to a cable system which does offer the service. Clearly this situation presents financial, convenience, and "just doing it" hurdles to subscribers who might otherwise try a new service.

Perhaps the simplest example of a technology change which impacts the consumer is the cable system's move to higher frequencies. It is reasonable to assume that over the next five to seven year, it will become technically practical to extend a cable system's upper frequency limit of 1 GHz or beyond. When this happens, a cable operator will replace the set top boxes which are leased. Those subscribers leasing the set top boxes will not be disadvantaged. However, those subscribers who have elected to purchase their set top boxes will have to make some choices and perhaps another purchase.

Launching New Services

Another aspect of this problem is that a new service needs to get subscribers quickly. If the cable operator provides the set top box, it will be installed in a large number of homes simultaneously. The service provider quickly adds subscriptions if the service is attractive. When the service provider has to rely on consumers purchasing set top boxes, there will be a significant reduction in the speed of penetration. Many service providers will not be able to survive long enough under these conditions. This raises yet another concern: what happens when a service fails? Under current conditions, a service which fails and leaves hardware useless does not directly impact subscribers. If the subscriber purchases a set top box for the new service and it then fails, the subscriber will bear the loss. Not only will this cause unhappiness, but it will further reduce subscriber willingness to try new services in the future.

Management of Expectations

While these problems are not yet upon us, the potential is high that they will be in the near future. It is worthwhile giving them some serious thought now. Probably the most important action to be taken by cable operators is subscriber education. The issue here is "management of expectations". If the consumer is well aware of the advantages and disadvantages of his purchase decision and proceeds with an informed decision, the likelihood of later problems is reduced.

It is unfortunate that the cost of advising consumers of their options will fall mostly – or even entirely -- on cable operators. This cost eventually raises the price of subscriptions. Consumers will ultimately pay for facilitating commercial opportunities for retailers.

Cable operators have the tools to make extensive and effective contact with subscribers. These tools must be utilized to minimize the problems that arise from failed expectations. Subscribers will see advertising for set top boxes available at retail. Some of this advertising will be misleading. It will certainly de-emphasize or omit the problems associated with such purchases. The cable operator must set the record straight. The cable operator needs to provide the other information required by the subscriber for making an informed choice.

THE "CABLE READY" REQUIREMENT

The cable and consumer electronics industries have struggled for many years arriving at technical specifications for "cable ready" TVs and VCRs. This effort intensified after the 1992 Cable Act and its Leahy Amendment concerning compatibility between consumer electronics equipment and cable services. That law required the FCC to establish a technical definition for cable ready TVs and VCRs. The FCC issued a Report and Order in May of 1994 which took the first step in that direction by including partial technical specifications on RF performance of cable ready TVs and VCRs. The process will be completed when the FCC incorporates the Decoder Interface specification in its rules. That, of course, will happen after the Electronic Industries Association's (EIA) and the National Cable Television Association's (NCTA) Cable, Consumer electronics Compatibility Advisory Group (C^3AG) makes its recommendation on the Decoder Interface to the FCC.

When the definition of cable ready is complete, it must apply to set top boxes sold at retail as well. All of the long and difficult work done leading to cable ready TVs and VCRs could be undone by the retail sale of set top boxes which fail to meet the FCC's technical definition of cable ready. It is reasonable to expect that the sale of set top boxes at retail will exceed the sale of cable ready TVs and VCRs if only because the prices of these set top boxes will be lower than most TVs and some VCRs. Additionally, consumers already have multiple TVs and VCRs and are more likely to buy cable ready set top boxes at retail rather than new TVs or VCRs.

Certainly if compatibility between cable systems and TVs and VCRs really is important, than that same degree of compatibility must be important in the retail availability of set top boxes. If the latter is not true, we must question the true importance of the former.

RF Requirements

The navigation devices sold at retail require the radio frequency requirements of a set top box rather than just those of a cable ready TV or VCR. The reason is simple. Both set top boxes and navigation devices sold at retail are connected in front of the TV or VCR. They add another tuner, intermediate frequency amplifier, and remodulator to the chain. If the noise or distortion contribution of the set top box or navigation device sold at retail is to be transparent, it must be much less than that of the TV or VCR itself.

It is a principle of radio physics that the first tuner dominates the noise performance of the system. If subscriber picture quality is to be maintained, then the navigation devices sold at retail must be held to the same standard as the set top box provided by the cable operator.

Arguments can, and will, be made that navigation devices sold at retail are available in the marketplace and marketplace forces should prevail. The marketplace is the most efficient force only when the consumer is well informed. It is difficult – near impossible – to have well-informed consumers on such technical and subtle issues. The opportunity for error is much too great. Since price is such a major factor in consumer purchase decisions, the pressure to reduce costs is difficult to resist. Inexpensive single conversion TV tuners in plastic boxes with inadequate shielding sold at retail would be a picture quality disaster.

Hazards to Broadcast Reception

The availability of navigation devices sold at retail can create serious hazards to the reception of broadcast signals both on cable and off-air. If the shielding of the tuner is inadequate, direct pick-up of off-the-air broadcast signals will mix with the signals conveyed by cable. Since signals travel through cable at a lower speed than through the air, the over-the-air signal will reach the tuner first. Inadequate shielding will allow this signal to penetrate the tuner and mix with the over-the-air signal arriving later. The result is a degraded signal including ghosts and moving diagonal bands. In some cases the degradation is sufficient to render the picture unwatchable. Broadcast reception on-channel is jeopardized.

A further difficulty arises if the navigation devices sold at retail leaks broadcast signals into the environment. TVs and VCRs in the same residence, not connected to cable may experience interference. Receivers in adjacent apartments may also suffer. If large numbers of inadequately shielded navigation devices sold at retail are installed, even portable TVs may experience problems.

Jeopardizing FCC Cable Technical Rules

The FCC has created strict technical rules for cable system operation. These rules include detailed specifications on set top boxes provided by cable operators. This will all be for naught if set top boxes available at retail are not held to the same specifications.

The chain is only as strong as its weakest link. If the second from the last link -- the subscriber-owned set top box -- is deficient, there is little point in imposing strict technical rules on the rest of the system.

It is reasonable to expect that cable operators will seek to be relieved of strict technical specifications on their set top boxes if subscribers connect lower grade set top boxes obtained at retail. Cable operators incur substantial costs in purchasing well-shielded double conversion tuners with excellent linearity specifications and strict restrictions on the spurious signals back fed into the cable system. All of this is pointless if consumers purchase lower grade set top boxes in significant quantities.

The Decoder Interface Solution

The cable and consumer electronics industries are reaching completion of the Decoder Interface specifications. The Decoder Interface consists of an Intermediate Frequency (IF) link of modulated signals selected by the TV's or VCR's tuner and a twenty six pin connector. The twenty six pin connector can have up to four bi-directional video twisted pairs, four bi-directional audio twisted pairs, a signal reference twisted pair, a control line twisted pair, and three bi-directional twisted pairs satisfying the video specification but reserved for future use. The control line determines the direction of flow of audio and video signal. In its full implementation, the Decoder Interface complies with the specifications for the Audio / Video Bus (AVBus) for interconnection of consumer electronics products. In its abbreviated implementation, the Decoder Interface has just one uni-directional video and audio twisted pair along with the reference and control twisted pairs and the IF connection.

The Decoder Interface and the AVBus connect to a cable which can support up to ten devices and span a distance of thirty feet. Expansion modules allow these numbers to increase. This system allows the addition of modules which decode subscription signals and provide other features and functions. The signal security modules are intended to be owned and provided by the system operator and the "feature modules" can be either purchased at retail, purchased from the system operator, or leased from the system operator at the consumer's option. Both analog and digital signals are accommodated. This comprehensive system allows for maximum consumer choice and facilitates competition in the provision of in-home equipment. When subscribers wish to try a new service, they plug in the appropriate enabling module. Their investment in equipment is preserved and they are protected against economic loss. Hurdles to trying new services are reduced or eliminated. Duplication of hardware elements is minimized.

THE LAME TELEPHONY ANALOGY

The technically unsophisticated and those who know better but wish to distort the facts to make a point continue to use the telephony analogy. Consumer ownership of telephone Customer Premises Equipment (CPE) has brought significant benefits and only a few problems. In particular, the telco claim that the network will be harmed has not occurred. By simple-minded analogy, the same is argued for cable.

Neglected in this are several significant technical facts. The spectrum used for telephony is just a few thousand Hertz (Hz). If these frequencies leak out into the environment, little harm is caused because these frequencies are not used for other purposes. The same is not true for cable. Cable uses frequencies in its closed environment which are used for many other purposes in the over-the-air situation. Aircraft navigation and communication and emergency services are just a few of the critical applications which must be protected from interference. If unregulated "Navigation Devices" are allowed to be connected to cable, their leakage of signals into the environment will be detrimental.

The telephone system involves individual circuit paths to the subscriber. If the subscriber's equipment emits interfering signals back into the telephone line, it will have limited impact on other subscribers. Since the cable network is a shared network, any signals put back into the system will cause signal degradation for other subscribers. This cannot be allowed.

In the case of telephony, the individual circuit path makes the theft of service very difficult. The same is not true of cable. The shared structure means that the same signal enters all homes in the neighborhood. Signal security is much more important in this case.

Allowing the retail sale of set top boxes for cable is a much more demanding situation requiring more care and regulation.

LEASE VS. PURCHASE

There are important reasons to lease or rent equipment rather than own it:

- can't afford it
- have other uses for funds
- aren't sure what they want
- wish to avoid maintenance and up-keep
- need continued support
- want protection from change

While home ownership is part of the American dream, many simply can't afford it. For these individuals, renting is the only alternative to homelessness. Without the rent or lease option, these individuals and families would be in a difficult situation. In many cases, the monthly rent is about the same or even more than a monthly mortgage payment.

However, home ownership requires an up-front down-payment which is beyond the means of many. Similarly, the availability of cable equipment at lease allows access to services, functions, and features which many could not afford to purchase and own. To restrict the lease to only security functions would deprive these consumers of the availability of features and functions they may desire.

Others might be able to afford the purchase of the in-home equipment but only if they gave up some other purchase or use for the funds. The choice should not be denied them. If there is a better use for the funds, that options should be available. Perhaps it is more important to buy a new TV or VCR and lease the set top box until some later time.

Families who are re-located frequently rent a home even though they can afford to purchase one because they are not sure where they want to live. The same principle applies to hardware at lease. By renting for awhile, a more informed purchase decision can be made later.

If a consumer purchases a navigation device sold at retail which has features closely matching the offerings of one service provider, say the cable company, and later decides to change to a service provider who uses another media, say satellite, fiber optics, or twisted pairs, the ownership of the navigation device will impose a cost to the change that would not be present if the consumer had chosen to lease all of the features and function in a set top box available from the service provider.

Some consumers choose to lease a condominium, not because they can't afford a home, but because they wish to avoid the continuing maintenance and up-keep. Either they don't have the aptitude for these tasks, or they have better things to do with their time, or they simply don't want to be bothered. The same choices should apply to the lease of a set top box.

Subscribers who want continued support from their signal provider would choose an equipment lease rather than take on the cost of service calls.

Finally, there are those who will want the latest services and want the protection from a change in hardware requirements needed to keep up with new services.

There is a growing trend in the lease of automobiles by consumers for much the same reasons. Automobile leases are not restricted to the basic models. All varieties of features and functions are available in this leased equipment.

THE DIGITAL ENVIRONMENT

There is a simple minded assumption that when digital television arrives, security will be high enough to allow subscriber ownership. This is a naïve proposition. There is no reason to believe that digital security will be any better in the long run than analog security.

Signal security is a running battle between the engineers creating the system and those who wish to defeat it. This is not an even match. The designers have a limited budget, a restricted staff, a short time to design the product, and the mature technology of the day to utilize for implementation. Those who would attack the signal security system have unlimited time, arbitrarily large numbers of participants, and an evolving technology. The power of personal computers continues to grow at a phenomenal rate. Both speed and processing power are accelerating. The ability to network these computers together also grows. As we look forward to cable modems with megabit per second speeds linking multiple Pentium Pro processors, truly massive computing power will be available to "hackers" at very affordable prices. Hundreds of millions of transistors interconnected at high speed will be available to attack the security systems devised at an earlier time under the constraint of being offered at affordable prices. It is ironic that cable's offering of these high speed connections to "the information supperhighway" will facilitate attacks on digital video security.

The vulnerability of a software security solution far exceeds that of a digital hardware solution. With software, the bank of interconnected Pentium computers is especially powerful. No circuit modification is required.

The only test for the conviction of some one who wishes to sell security systems to consumers is an escrow account covering the full cost of recovery from a breach. Anything short of that puts the consumer in danger of having his investment destroyed. Someone needs to be fully responsible for the replacement of the security elements when the security system is breached. If the cable system operator is to be replaced in this role, the retailer or manufacture must take on this serious responsibility.

CONSUMER PROTECTION

The consumer must not be forgotten in the rush to create sales opportunities for retailers. There are significant and legitimate consumer needs which must be protected.

Most importantly, the consumer must not be put in the position of having to abandon an investment in hardware because someone else has breached the signal security system.

Service and replacement parts must be available for the subscriber who wishes to own set top boxes. He must not be subject to "orphaned" hardware. Assistance in installation and operation is important too. If that task falls to the cable operator, the cable operator will have no choice but to charge for the assistance.

The consumer must not be hampered in his choice of services by a previous hardware purchase. It must be emphasized that the hardware is the means to the end. The consumer's principal need and desire is for the services which come through the hardware. The hardware is a facilitator; it must not be allowed to become a barrier.

The consumer who sees value in leasing equipment should not be precluded from doing so by the provincial desires of an industry which wishes to sell hardware.

Of course, the subscriber who continues to lease set top boxes should not have his service degraded by interference from deficient hardware purchased and installed by his neighbor.

These issues are extremely complex, difficult to understand, and change with time. The consumer must be provided with enough information to make an informed choice. There must be remedies to errors and situations where the wrong equipment was purchased.

UNANSWERED ISSUES

There are a number of unanswered issues. Just a few of them are considered here. The first issue concerns the rights of the cable equipment manufacturers. They seem to have been left out in the cold on this. The law applies to cable operators, not equipment manufacturers. But the business of the equipment manufacturer may be potentially impacted in a negative way. The design of set top boxes involves patents and intellectual property. It is not clear what happens when some of that intellectual property is involved. There is no legal requirement for the cable equipment manufacturers to license intellectual property. This may inhibit a cable operator's ability to comply with what ever rules are created.

Another issue involves equipment which is no longer manufactured or whose manufacturer is no longer in the cable supply business or any other business. What are the cable operator's burdens in these cases. For example, the Cube system or the Oak scrambling systems haven't been manufactured for many years but are still in use.

It is important to recognize that the "look and feel" of a service is an important part of its make up. The on-screen look and the methods of interaction with the remote control are an integral part of the service influencing its attractiveness and ease of use. This necessarily has implications on how the in-home equipment is packaged and offered. It is naïve to think that just a security module is all that is required.

The rights of artists, creative contributors, and holders of copyrights must be protected. If theft of service on a large scale becomes possible, programming will be withheld from this medium. The creators of this programming have no choice but to protect themselves. Without the ability to be fairly compensated for their work, they will not be able to continue. Consumers will end up being the ultimate losers with a loss of access and convenience.

CONCLUSION

The law of the land requires the FCC to create rules to assure the availability of set top boxes – "Navigation Devices" – at retail. There are important protections in that law for

- assuring industry standard-setting body participation
- assuring signal security
- waivers for developmental technology

There are many hazards in this law for consumers and for service providers. Care must be taken to minimize the difficulties.

It is critical that the FCC's technical definition of cable-ready apply to set top boxes sold at retail. If this is not the case, all of the careful work done after the 1992 Cable Act will be undone. In addition, the technical integrity of cable systems will be impaired, consumer signal quality will suffer, other subscribers' services will be degraded, and interference will be caused to other users of the spectrum including aircraft communications and navigation and emergency service communication.

The Decoder Interface can be a powerful tool in meeting the needs of this law.

Clearly managing subscriber expectations is critical to informed consumer choice and to minimizing and properly directing subscriber frustration and anger to its appropriate source.

THE AUTHOR

Dr. Ciciora is a technology consultant specializing in Cable Television, Consumer Electronics, and Telecommunications.

Most recently he was Vice President of Technology at Time Warner Cable. Walt joined American Television and Communications, the predecessor to Time Warner Cable, in December of 1982 as Vice President of Research and Development. Prior to that he was with Zenith Electronics Corporation, starting in 1965. He was Director of Sales and Marketing, Cable Products, from 1981 to 1982. Earlier at Zenith he was Manager, Electronic System Research and Development specializing in Teletext, Videotext and Video Signal Processing with emphasis on digital television technology and ghost canceling for television systems.

He has nine patents issued and several more pending. He has presented over one hundred papers and published about fifty, two of which have received awards from the Institute of Electrical and Electronic Engineers (IEEE). His papers have been translated into Japanese, Chinese, German and Spanish. Walt wrote a monthly columns for Communications Engineering and Design (CED) magazine and for Communications Technology (CT) magazine for three years each.

He currently serves on the Executive Committee of the Montreux Television Symposium. He was a member of the board of directors of the Society of Cable Television Engineers (SCTE) for six years. He was Chairman of the Technical Advisory Committee of CableLabs for four years and Chairman of the National Cable Television Association (NCTA) Engineering Committee also for four years. He was president of the IEEE Consumer Electronics Society for two years and is a past chairman of the IEEE International Conference on Consumer Electronics. He chaired the Joint Engineering Committee of the NCTA and the Electronic Industry Association (EIA) for eight years. He has served on several industry standard-setting committees. He currently co-chairs the Cable Consumer electronics Compatibility Advisory Group and its Decoder Interface subcommittee.

Walt is a Fellow of the IEEE, a Fellow of the Society of Motion Picture and Television Engineers (SMPTE), and a Senior Member of the SCTE. Other memberships include Tau Beta Pi, Etta Kappa Nu, and Beta Gamma Sigma.

Current interests center on competitive technology, the consumer electronics interface with cable, Digital Video Compression, Interactive Television, Multimedia, and High Definition Television.

Walt received the 1987 NCTA Vanguard Award for Science and Technology and was named "1990 Man of the Year" by CED magazine. CED also named him "1993 Man of the Year". He was the Fall 1994 Levenson Memorial Lecturer at the National Cable Television Center at Penn State.

Walt has a Ph.D. in Electrical Engineering from Illinois Institute of Technology (IIT) dated 1969. The BSEE and MSEE are also from IIT. He received an MBA from the University of Chicago in 1979. He has taught Electrical Engineering in the evening division of IIT for seven years.

Hobbies include helping his wife with her horses, reading, wood working, photography, skiing, and a hope to someday become more active in amateur radio (WB9FPW).

Transmission Technologies in Secondary Hub Rings — SONET versus FDM Once More

Thomas G. Elliot
Senior Vice President
Engineering & Technical Services
TCI Cable Management Corporation

Terrace Tower II
5619 DTC Parkway
Englewood, CO 80111-3000
(303) 267-5222
FAX (303) 488-3210

Oleh J. Sniezko
Director
Transmission Engineering
TCI Communications, Inc.

Terrace Tower II
5619 DTC Parkway
Englewood, CO 80111-3000
(303) 267-6959
FAX (303) 488-3210

Abstract

The choice of architecture and transmission technology for signal delivery from primary hub rings to clusters of coaxial buses with 125 to 500 homes passed per bus is a subject of an ongoing debate. This paper presents an analysis of the several transmission technologies suitable for these links, and namely:

- *dedicated fiber delivery system (simple FDM technology) from the primary hub ring to the coaxial bus (ring-star-bus);*
- *FDM technology with block conversion for targeted services;*
- *hybrid FDM (for all broadcast signals) and SONET (for targeted services);*
- *hybrid FDM (for analog broadcast signals) and SONET (for all digital signals).*

The impact of the following factors on the results of the analysis are considered:

- *service requirements;*
- *architectural requirements;*
- *capital and operating costs;*
- *future-proof level of the architecture and transmission technology.*

The results of the analysis can be used in selecting a solution suitable to a particular set of requirements and a size of the system. The authors also present their vision of the future and TCI's technology of choice at this point in time.

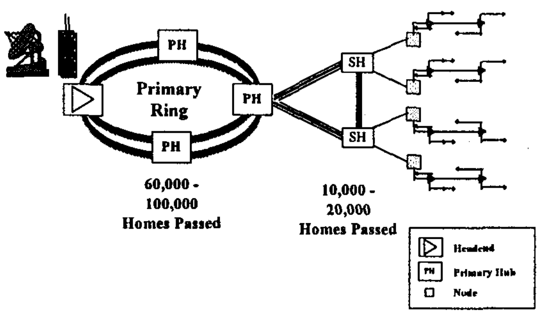
INTRODUCTION

As the world is rushing into the telecommunication age and becoming even more a 'global village' prophesied by McLuhan, we, telecommunications engineers, try anticipate future requirements and design a network that will not become obsolete before it is completed. Of course, we are restricted by the equation diametrically different from the equation used by some of the telecommunications service providers. Ours is that "profitability equals revenue minus expenses". Hence, our effort is full of circumspection. The equation and its implications are described by Dan Pike in his address to the Western Communication Forum[1].

HFC ARCHITECTURE

In most cases, we decided on a transmission technology in major parts of our network. Let us review the HFC network architecture as an example (see Figure 1). This architecture closely resembles CableLabs' Active Coaxial Network Architecture. Other MSO's architectures are similar in many solutions to the CableLabs' network. In the largest metropolitan areas, numerous headends are most likely to be connected in a ring to provide a fully redundant and survivable platform. In many cases, primary hubs are established to maintain signal quality delivered to the distribution network. These hubs serve from 60K to 100K homes passed.

Figure 1: HFC Network — Dual-Ring / Star / Bus



The other end of the HFC network is configured as a bus. To provide all the services anticipated, we have to share limited resources of the bus in our network among a limited number of potential subscribers. The number of customers that can be share the resources will depend on modulation schemes and other technical solutions employed as well as on demand for services The industry estimates this number between 125 and 500 homes passed. Usually, several of these buses are fed from a central location (nodes) to share the resources. These nodes serve clusters of 500 to 2000 homes passed.

The looming question is how to deliver signals from the primary hub ring to the clusters of customers served by the bus. This question encompasses two issues:

1. Architecture, and
2. Transmission technology.

Both issues are interrelated but are treated separately by network engineers for several reasons:

- lack of the optimal transmission technology for the range of services to be delivered;
- unpredictability of the future requirements; and
- the limits set by the equation ruling our industry (profit equals revenue less expenses).

This paper deals with the second issue — transmission technology to be used. However, to deal with this issue, we have to analyze possible architectures for the links between the primary hub ring and the clusters. The analysis is conducted for two geographical areas with 20K homes passed each, fed from a primary hub ring. The areas are located so that they can be connected into a ring to provide backup feeds. Each area can be logically arranged into twenty-two 900-home passed nodes. Each node feeds three buses of 300 homes passed. Several solutions are analyzed in a historical order as they were considered and championed by TCI.

These solutions are included within the wide range of architectural choices. One side of this range is the physical star architecture in which all nodes (in the extreme cases, all homes) are fed via a dedicated link. This architecture is deployed by telephone companies (switched star).

On the other side of this range is the bus structure or a ring structure feeding the nodes directly from the primary hub ring.

Most of the evolving HFC architectures deploy some intermediate solutions — redundant secondary hub rings to deliver signals and services over optical fiber from primary hubs (central offices) to the neighborhood of 5,000 to 40,000 potential customers. From this point on, the signals are distributed either over fiber or coaxial cable to clusters of 500 to 2,000 customers to allow for segmentation and targeted services delivery.

The HFC network in its completeness described above would be implemented only in large metropolitan areas. It will be scaled down in areas with fewer than 100K homes passed. In areas with fewer than 20K homes passed, the choice of architecture will be different (probably direct feed to the nodes) than in areas with 60K to 100K homes passed.

TRANSMISSION TECHNOLOGIES DEPLOYED

Primary Hub Ring

The primary hub rings utilize technologies suited the best for long distances and introducing the lowest level of impairments. The most common deployed so far are listed in Table 1.

The choice of the transmission technology in the past was based mostly on the requirements for the quality of signals and on availability of the technology (including its cost). For these reasons, no standard technology has been selected by the industry. Several studies by major MSOs and the industry[2 3 4] concluded that, in large metropolitan areas with high demand for wide array of services, it is cost effective and beneficial now to deploy SONET-based systems. It is the authors opinion that, as the quality of the video codecs improves[5] and the prices fall further, this technology will prevail in competition with proprietary technologies in primary hub rings, even for video distribution. This will probably happen in the next several years. There will be a niche for analog technologies (1550 nm optical links for example) in small markets with limited demand for competitive access provisioning for obvious reason of better cost-effectiveness.

Table 1: Transmission Technologies in Primary Hub Rings

Technology	Description	Comments	
		Pros	Cons
Analog FM	Baseband video signals FM modulated	• good overall signal quality	• problems with different analog scrambling systems • proprietary systems • non-standard network management elements • limited cascading • not suitable for system interconnects • high cost of interfacing to RF
Analog AM 1550 nm window	RF signals with higher-power, externally modulated lasers with optical amplification.	• transparent to any technology • low cost of interfacing to RF	• ring length is limited due to cascading noise of amplifiers (repeaters) • less reliable than 1310 nm lasers • network management proprietary • few vendors for analog video quality systems
Linear PCM	Baseband or IF video signals digitized and transmitted for long distances	• very good video quality • SONET-like features • numerous interfaces available (DS3 and subsets) • 16 to 32 high quality channels per wavelength • 8 to 16 IF channels per wavelength • QAM and VSB signals at IF can be digitized and only up-converters are required • drop/add capability • good reliability record	• non-standard, proprietary systems • limited number of vendors (two major vendors only) • non-standard network management system • high cost of interfacing to RF
SONET	Digital TDM optical hierarchy system	• reasonable video quality • potentially interfacing with any digital service • standard systems • standard network management • limited number of fibers • survivability • perfect for system interconnects • many vendors • drop/add capability • good reliability record • system prices lowering	• high cost of interfacing to RF

Optical and RF Coaxial Bus

Signals distributed over the final part of the distribution network must be compatible with terminal devices or network interface units, and must be suitable for bus type architecture and shared coaxial cable distribution network. These requirements practically define the type of transmission technology that can be deployed between secondary hubs and customers. This must be FDM with the type of modulation compatible with the terminal devices: AM modulated NTSC signals of EIA channel assignment; FM radio; QPR Sega channels, QPR digital radio services, QPSK telephony, QAM MPEG video channels, QAM and QPSK high speed two-way data, QPSK voice, etc. The alternative to this technology would require a placement of expensive signal conversion centers shared by few users.

Primary-to-Secondary Hub Links

Primary-to-secondary hub link engineers exercise a significantly higher degree of freedom while selecting a transmission technology for this part of the HFC network. Beside the paradigm of the signal compatibility with the remaining elements of the HFC plant, the following factors must be considered:

Service Needs

1. Cable television needs:
   - broadcast services,
   - NVOD services,
   - VOD services;
2. Telecommunications services.

Architectural Considerations

1. reliability:
   - redundancy,
   - network management;
2. quality;
3. interoperability and level of standardization:
   - comfortability of the potential end users with the technology selected;
4. cost of the final network per home passed and per active customer;
   - lay-out cost in light of time value of money,
   - operating cost;
5. technology availability; and
6. cost of becoming obsolete.

All these factors are dynamic in nature and change continuously with the progress in technology and declining prices of yesterday's novelties.

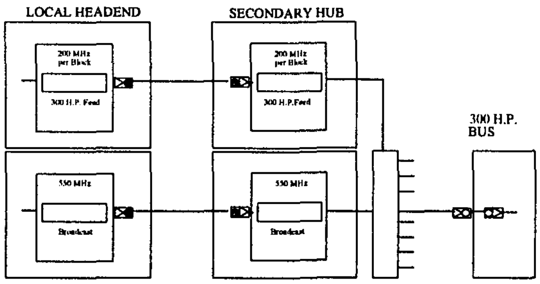
AVAILABLE CHOICES

The following technologies for connecting the primary hub ring with the nodes are the most popular among HFC network engineers:

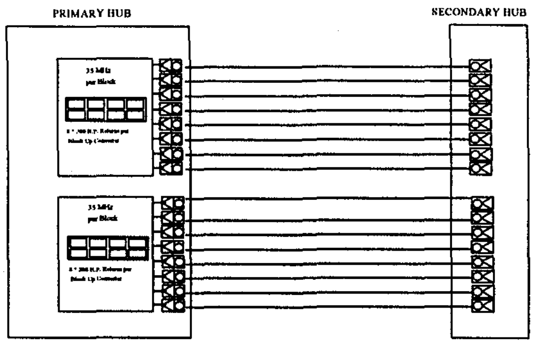
- dedicated fiber delivery system (simple FDM technology) from the primary hub ring to the coaxial bus (ring-star-bus);
- FDM technology with block conversion for targeted services;
- hybrid FDM (for all broadcast signals) and SONET (for targeted services);

These choices are presented in Figures 2 through 4.

Figure 2: Ring/Star/Bus Architecture

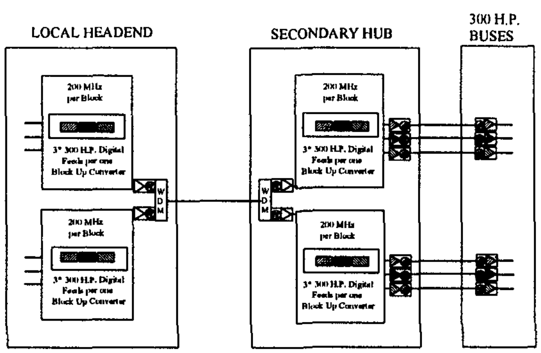


TSD LINK SERVES 300 HOMES PASSED
COMBINED WITH BROADCAST

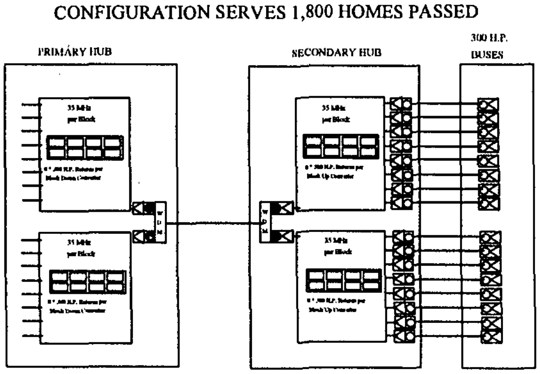


SECONDARY TO PRIMARY HUB REVERSE LINKS

Figure 3: Ring/Ring/Star/Bus Architecture with FDM Transmission

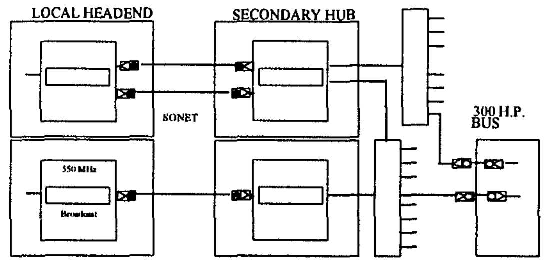


TSD BLOCK CONVERSION AT THE HEADEND
CONFIGURATION SERVES 1,800 HOMES PASSED



RETURN PATH BLOCK CONVERSION AT THE OTN
CONFIGURATION SERVES 4,800 HOMES PASSED

Figure 4: Ring-Ring-Star-Bus Architecture with Hybrid FDM/SONET Transmission



SONET DROP/ADD FETAURE PROVIDES
OUPTUT/INPUT PAIR FOR EACH 300 HP BUS

All three solutions are analyzed below.

ARCHITECTURAL CONSIDERATIONS

Ring/Star/Bus

This architecture provides the highest level of flexibility in the selection of the transmission technology. In this case, the FDM technology is deployed throughout all the network between the primary hub ring and the customer. The FDM technology can work in combination with a digital or analog transmission technology (and multiple access protocols or schemes) compatible with the terminal devices. There is, however, a major drawback to this technology: the number of fibers required to facilitate this architecture for two secondary hubs with twenty-two nodes, each feeding three 300-home passed buses (see Figure 1) would reach 154 without redundancy or 308 with redundancy. For rings with more secondary hubs, the number would be higher. If better quality were required in the forward direction, the number of fibers would have to be increased further. Moreover, to provide protection against single point of failure to the level of 4,000 homes passed, a significant number of optical switches or repeaters with RF switches would have to be deployed at the secondary hub. The cost of such an arrangement and the manageability problems are prohibitive.

Ring/Ring/Star/Bus

To limit the number of fibers in the links between primary and secondary hubs, some way of sharing the fibers had to be developed. We were considering two options:

- using FDM technology with frequency block conversion, and
- using dense WDM technology.

Since WDM technology was not mature at the time (and is not widely deployed in cable TV today), we concentrated on the FDM arrangement with block conversion. This approach limited the number of fibers in our reference secondary hub links to 27 fibers without redundancy or 54 fibers with redundancy. This is a big gain in fiber network simplicity traded off for more complex electronics.

In recent months we reviewed our experience from Hartford where we dealt with this problem somewhat differently. We applied SONET technology in links between primary and secondary hubs. This allowed for limiting the number of fibers even further. For our test area, the number of fibers would equal 7. This number would include forward analog (and digital) broadcast signal fibers (two fibers for dual-fiber configuration plus one back-up fiber), two fibers for SONET ring (OC3 for both forward and reverse voice an data signals), and two fibers (principal and backup) for reverse signals that cannot be easily interfaced to SONET at this time.

The only remaining issue is whether to use SONET for transmission of all digital signals or only targeted digital signals. This issue will be analyzed later.

Let us qualitatively compare the FDM block conversion technique with SONET technology (Table 2).

Table 2: Qualitative Comparison between FDM Block Conversion and SONET

Desirable Feature	Block Conversion	SONET
Low cost interface to RF	✓	
Many vendors		✓
Standard system		✓
Standard network management		✓
Limited number of fibers required	✓	✓✓
Interfaces with digital systems	✓	✓
Good reliability record		✓
Survivability	✓	✓✓
Same system for forward and reverse		✓
Drop/add capability		✓

Undesirable Feature	Block Conversion	SONET
Possible problems of instability	✓	
Potential of becoming obsolete	✓	
Potential of becoming single-vendor product	✓	
Fixed system frequency bandwidth split between broadcast and targeted services	✓	
High cost of RF interfaces		✓

Table 2 clearly shows that the SONET solution is technically superior to the FDM Block Conversion. Moreover, many of our partners on the manufacturing side are uneasy about possible unknown impairments that block-conversion technology can introduce (frequency instability, jitter, etc.). How do the two technologies compare based on cost? To judge this issue, we had to analyze requirements for different services.

SERVICE NEEDS

Let us start with the traditional cable TV services: broadcast entertainment, common at least for a single community. These services can be characterized by common collection points (signal importation, direct feeds, and local origination), centralized switching, and delivery of regional specialty services. Our old architecture (tree & branch) was perfectly suited for this type of service. However, with the advancement of more precisely targeted services and transactional services, this topology outlived itself. Beside such services as VOD, targeted to an individual customer, there is advertising targeted to a neighborhood (possibly as limited as single node or bus). Some may argue whether VOD will become a successful service and whether so narrowly targeted advertising will bring the revenue expected but we, engineers, consider these services as potentially viable and design the network with their delivery in mind. It does not mean that we provision for these service today (remember the equation under which we operate), but we try to make choices that are future- and service-proof.

Although VOD viability is questionable at present, no one questions the fact that the telecommunications services and other transactional services (for example, voice and videophony) will be targeted to an individual customer and that these services have the potential of bringing a sizable revenue to offset the expenses and be profitable. Signals for these services are usually distributed in a digital form. More importantly, they require two-way communication and significant bandwidth availability per customer in forward and reverse directions to be successful.

Interface Equipment Requirements

Both the architecture and the transmission technology selected must accommodate the distinguishable characteristics of the broadcast services and targeted services. In our analysis, we assumed that the analog broadcast signals will be transmitted using existing FDM technology. We also decided that all digital broadcast signals will be transmitted the same way for the foreseeable future. Any local (community or optical node specific) advertising will be injected to a dedicated channel in the set-top boxes. At this point, we did not account for these services. Table 3 lists the interface requirements for the two technologies analyzed for a series of services considered.

Table 3: Interface Equipment Requirements for SONET Based Transmission Network

Service	Existing Multiple Access Protocol (most common) or Interface	Existing Interfaces	Preferable Interfaces	Preferable Multiple Access Protocol	Additional Interface (Existing multiple access protocol or format and RF network)
Voice	TR303 or TR08 interface, DS0 or DS1	Direct to DS1 VT	NA	ATM	Bandwidth Manager
High Speed Data	10BT, 100BT, or FDDI	Router or Protocol Translator from LAN protocols to DS3 or OC3	Direct from a modem or bandwidth manager to DS3 or OC3	ATM	Fast Ethernet Switch, Servers, Bandwidth Manager, Frequency Translators
VOD	MPEG2	Mappers or Groomers into DS3 or STS1	NA	ATM	QAM or VSB Modulators, Demappers, Degroomers

The table indicates the equipment needed at the secondary hub to interface with the SONET ring. This equipment, together with SONET equipment for secondary hubs, replaces equipment required for FDM/Block Conversion scheme. The list of the replaced equipment would include the same interface equipment in primary hubs (lower quantity by 1:3 ratio in TCI), optical transmitters and receivers for forward and reverse (of 1,000 MHz bandwidth), and block converters. This assumes that the primary hub ring is SONET based.

Quick table analysis indicates that the voice and data services transmission over SONET may be already viable from the cost point of view, or viable in the near future. Unfortunately. VOD distribution over SONET would increase the number of QAM or VSB modulators significantly and at the current cost estimate ($3,500 per 6 MHz channel) would not make it viable. At this point in time we have to conclude that VOD services must be delivered in a broadcast mode with sufficient bandwidth dedicated to meet 20K homes passed area demand. Alternatively, separate fibers (saved by deploying SONET) will have to be redeployed with some additional transmitter for forward (some forward transmitters of higher quality and all reverse links will be saved). To avoid the confusion in costing, we excluded VOD services from costing analysis.

Cost Comparison

The tables indicate that the SONET solution may be already viable. The cost of the SONET option does not include the cost of SONET-RF interfaces (except for the network management signals and reverse video - codecs). For the required interface equipment, refer to Table 3. The cost does not include the cost of interface equipment between different protocols and SONET either. These cost are of very dynamic nature. Moreover, the level of interface will depend on each service business case and rate of success. The availability, cost, size, and power consumption issues of the interface equipment poses the greatest opportunity for progress.

Table 4: Reference System Characteristics

a) Block Conversion

	Service Area	Unit Numbers	Block Converter	He-SH Fiber #
Headend	100,000	1		110
Secondary Hub (SH)	20,000	4		34
Fiber Nodes (FN) per SH	900	22		
Broadcast Services (50-550 MHz)	300	67		2
TSD services (550-750 MHz)	300	67	3 TO 1	22
Reverse Services (5-40 MHz)	300	67	16 TO 1	4
Number of Fibers per Node	9	255		

b) SONET

	Service Area	Unit Numbers	Block Converter	He-SH Fiber #
Headend	100,000	1		8
Secondary Hub (SH)	20,000	4		12
Fiber Nodes (FN) per SH	900	22	DFB	
Broadcast Services (50-550 MHz)	300	67	DFB	2
TSD services (550-750 MHz)	300	67	DFB/SONET	5
Reverse Services (5-40 MHz)	300	67	SONET	incl. in TSD
Number of Fibers per Node	9	205		

Table 5: Total Cost for Secondary Hubs

	Headend		Secondary Hub	
	Space (RU)	$	Space (RU)	$
FDM / Block		$1,819,358	912	$2,920,003
Conversion			**TOTAL**	**$4,739,361**
Hybrid: FDM for Broadcast,		$178,910	1,314	$2,338,530
SONET for Transactional			**TOTAL**	**$2,517,440**

OTHER CHOICES

Only one choice at present can challenge the SONET solution from the point of view of features and advantages, and most likely cost. This is a high density WDM technology. When applied in 1550 nm wavelength window, this technology can results in a completely passive network between the primary hub rings and optical node. Only WDM equipment (passive) would be located at the secondary hub level (if required) to maintain network redundancy. It would certainly require high density WDM techniques to stay within the reasonable limits for fiber count. The biggest disadvantage of this technology is the fact that it is not commercially available. Combination of this technology and some reverse block conversion in optical nodes can lower requirements for the density of wavelength-division multiplexing.

INDUSTRY CHALLENGE — SUMMARY

To make the SONET solution in secondary hub ring a viable option, the industry must move towards developing standardized interfaces between SONET and other digital transmission techniques. A preferable solution would be to select a multiple access protocol capable of supporting most of, if not all, the services we want to provide over coaxial RF network (ATM) and develop interfaces with an ATM protocol on both sides and with SONET and RF interfaces respectively. Of coarse , this standardization should aim at lowering the prices. TCI will work with the potential vendors on target prices.

Equally important is lowering power consumption and size of the interfaces (and other equipment such as modems, routers, servers) to lower our power and real estate requirements for secondary hubs. Current anticipation of 200 square ft and 400 square ft huts for secondary hub significantly increases our problems with site acquisitions.

We think that, given we make progress on the issues outlined above, SONET technology can be successfully applied in secondary hub rings and the same driven deeper into our network.

ACKNOWLEDGMENT

The authors want to express their appreciation to Steven Dukes and Ken

Hoguta of TCI, Jeff Tokar of ANTEC, and Milo Medina, Jamie Howard, and other members of the engineering staff at @Home for their help in data gathering and their discussion about different matters expressed in this paper.

---

[1] Dan Pike, Cable TV Strategies, Western Communication Forum. The other equation described in the paper states that "revenue equals expense plus a profit guaranteed by law".
[2] Rogers Engineering, A Comparison of Transport Technologies for the Rogers Intercity Network, Internal Rogers Report
[3] Tony E. Werner, Regional and Metropolitan Hub Architecture Considerations, SCTE 1995 Conference on Emerging Technologies
[4] James Farmer, Issues in Handling Cable Signals within Digital Optical Interconnect Networks, 1994 NCTA Technical Papers
[5] As in [3]

Upgrade of 450/550 MHz Cable Systems to 600 MHz Using a Phase Area Approach

Robb Balsdon
Vice President, Engineering Services
Rogers Engineering

ABSTRACT

This paper reviews the issues that Rogers considered when it decided to adopt a 600 MHz two-way upgrade strategy for its cable systems that were a combination of 450 MHz and 550 MHz plant (450/550 MHz). The costs of upgrading existing 450 MHz and 550 MHz amplifiers to 600 MHz are summarized. The main advantages and disadvantages of implementing a 600 MHz upgrade are discussed. A cost comparison of upgrading 450/550 MHz plant to 550 MHz, 600 MHz and 750 MHz fiber to the feeder (FTTF) is presented. Lastly, the phase area approach that Rogers is using to manage the 600 MHz upgrade of its plant and deploy new services on the upgraded plant is discussed.

INTRODUCTION

There were five main objectives that contributed to the Rogers' decision in 1994 to utilize a 600 MHz upgrade:

- the Rogers cable systems must support the delivery of 77 analog channels and DVC services with additional spectrum above the last analog channel (550 MHz). .
- where ever possible, capital should be deferred based on a present value (PV) of capital costs including risk factors.
- any upgrade must maximize the reuse of existing assets where possible, in order to reduce total upgrade costs.
- any upgrade must not preclude the addition of fiber nodes to reduce serving areas to 500 homes passed in the future.
- the end of line (EOL) must be of acceptable quality (better than -55dB , -51 dB CTB, 47 dB CNR).

The plans of the Rogers marketing group included a total of 77 channels of: basic service channels, premium tier channels, pay TV channels and expanded PPV channels (20 channels) in 1995/96. The marketing objective was to not remove analog channels below 550 MHz to allow DVC services to be launched. Both analog PPV and DVC services("Digital Tier") would be offered coincidentally. Also, additional spectrum above 550 MHz was required for expanding other non-video services such as PC access (cable modem) services and other future data and voice services.

Similar to most cable operators, Rogers implemented a plant upgrade rather than a full rebuild because of the high cost of placing new cable in system rebuilds. Rogers investigated several plant upgrade alternatives for the upgrade of various Rogers cable systems. There was a wide variation in the existing bandwidth and plant quality across the Rogers owned cable systems.

The upgrade alternatives that were investigated are:

1) 550 MHz upgrade:
- upgrading the 450 MHz sections of plant to a minimum of 550 MHz.
- reducing all trunk cascades to a maximum of 10 amplifiers (plus 3 line extenders) using optical hubs with sufficient fibers installed to allow future segmentation to 500 home nodes.

2) 600 MHz upgrade:
- upgrading all plant to a minimum of 600 MHz, premised on reusing existing amplifiers were possible and purchasing 750 MHz distribution actives and passives to replace units that could not meet 600 MHz as a minimum.
- reducing all trunk cascades to a maximum of 10 amplifiers (plus 3 line extenders) using optical hubs with sufficient fibers installed to allow future segmentation to 500 home nodes.

3) 750 MHz fiber to the feeder upgrade:
- reducing amplifier cascades to 5 actives by installing optical nodes.
- installing 750 MHz actives and passives.

600 MHz Equipment Upgrade

The existing feedforward trunk amplifiers and line extenders resulting from 550 MHz upgrades in the Rogers systems in the early 1990's were analyzed to determine if the bandwidth of the amplifiers could be cost effectively extended to 600 MHz. The performance of the amplifiers were tested with analog channel loading to 550 MHz and DVC loading (8-10 dB below from video) from 550 MHz to 600 MHz. Only certain models and vintages of amplifiers were selected to be upgraded to 600 MHz, with approximately 50% of the 450 MHz equipment and 100% of the 550 MHz equipment being upgradeable. The resulting upgraded equipment met the 600 MHz specifications with DVC loading above 550 MHz. Additional spectrum above 600 MHz is available, however there is a gradual frequency roll-off after 600 MHz where the flatness specification cannot be consistently met. This spectrum is capable of carry more robust modulation formats such as 16 QAM and QPSK. The 600 MHz upgrade makes use of the excess gain available in the amplifier hybrid ICs that is typically reduced by interstage attenuation adjustments by the manufacturer. The 600 MHz amplifiers can be dropped into plant currently spaced at 550 MHz. In most cases a high percentage of the passives and multitaps used in the 550 MHz plant can maintain specification at 600 MHz. All 450 equipment was respaced to 600/750 MHz spacing. The cost savings of upgrading existing amplifiers over purchasing new is shown in Table 1.

Cost Saving Of Amplifier Upgrade To 600 MHz

Table 1

Amplifier BW (Type)	Saving Over New 600 MHz Trunk Amp	Saving Over New 750 MHz Line Extender
450 MHz	44%	57%
550 MHz	44%	80%

Note: Includes salvage of existing amplifier when purchasing new amplifier.

Since additional trunk amplifiers are required when respacing trunk from 450 MHz to 550/600 MHz, new 600 MHz trunk amplifiers were purchased to augment the upgraded 600 MHz trunk amplifiers. New 750 MHz line extenders were purchased since only about 70 % of the line extender requirements could be filled with upgraded 600 MHz units. These new 750 MHz line extenders were concentrated (installed) in specific systems (areas) to make any future upgrade to 750 MHz more cost effective in these locations.

<u>Advantages Of An Upgrade To 600 MHz</u>

The main advantage of implementing a 600 MHz upgrade are:

- minimum of 50 MHz of DVC spectrum can be cost effectively added for less than a 10% premium, typically less than $6 U.S. per home passed.
- 600 MHz can be used to offer 77 analog channels plus 72 digital services at 8:1 digital to analog channel (6 MHz) ratio.
- the cable operator can defer or eliminate the high capital cost required to install all the fiber/nodes required for 750 MHz FTTF (especially if a high percentage of the fiber cable installation is buried).
- if necessary, further fiber and nodes need only be installed into areas where additional capacity for new services are demanded.

If the requirement for 750 MHz capacity is uncertain and an upgrade to 600 MHz is adequate based on current analysis then a risk factor must be added to the company's current cost of capital. The cost of capital combined with the risk factor now can be used to compare upgrade alternatives. Table 2, compares the capital for a 750 MHz FTTF upgrade with a 600 MHz upgrade and adds the capital related to a later decision to expand the bandwidth of the system to 750 MHz when additional capacity is needed. The expenditures shown use a 10% cost of capital and the uncertainly about the requirement for 750 MHz was factored into the comparison by using a 10% risk premium. Depending on the risk factor used and the planned expenditure rate (to achieve the desired upgrade completion), it can be economical to upgrade to the capacity that meets expected requirements, but may be interim depending on changes in plans to offer new services and changes in available technologies, such as DVC.

Comparison Of Capital Expenditure Rate For Upgrades
Table 2

Upgrade Alternative	Disc. Rate	Capital Expenditure by Year (millions of $)						Total Capital	PV
		1994	1995	1996	1997	1998	1999		
450 MHz to		7.4	14.8	14.8	0.0	0.0	0.0	$36.9	--
600 MHz	10%	7.4	13.4	12.2	0.0	0.0	0.0	--	$33.0
Total: 450-600 MHz								**$36.9**	**$33.0**
600 MHz to	0.0	0.0	0.0	0.0	2.0	8.0	11.8	$21.8	--
750 MHz	20%	0.0	0.0	0.0	1.5	5.5	7.3	--	$14.3
Total: 600-750 MHz								**$21.8**	**$14.3**
Total: 450-600-750 MHz Upgrade in Two Steps								**$58.7**	**$47.3**
Upgrade Alternative	**Disc. Rate**	**Capital Expenditure by Year (millions of $)**						**Total Capital**	**PV**
		1994	**1995**	**1996**	**1997**	**1998**	**1999**		
450 MHz to		11.0	16.5	16.5	11.0	0.0	0.0	$54.9	--
750 MHz	10%	11.0	15.0	13.6	8.2	0.0	0.0	--	$47.8
Total: Direct Upgrade 450-750 MHz								**$54.9**	**$47.8**

Disadvantages of 600 MHz Upgrade

The main disadvantages of the 600 MHz upgrade are:

- the spectrum available for usage for 256 QAM type signals is limited to 600 MHz.
- non-standard 600 MHz amplifiers are used in the upgrade.
  the cable operator must manage the upgrade of the equipment
- through the OEM vendor or third party contractor.
- the maximum cascade of 10 trunk amplifiers and 3 line extenders provides an availability of approximately 99.95% (259 min. of downtime, at a 4 hr MTTR) in contrast to an availability of 99.98% (128 min. of downtime, at a 4 hr MTTR) for 750 MHz FTTF with 5 actives in cascade

Cost Comparison of Alternative Upgrade Strategies

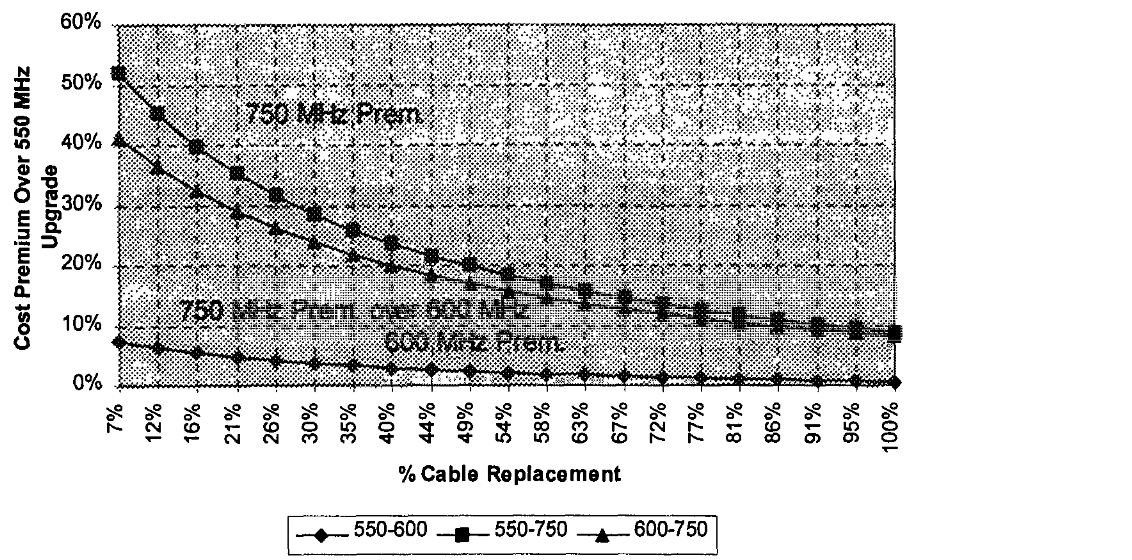
The upgrade costing used for this paper was completed based on a number of Rogers systems that were investigated in 1994 to determine the viability of the 600 MHz upgrade program. The costing was adjusted to provide a sensitivity analysis on the premium to upgrade a typical Rogers system to 550 MHz or 600 MHz or 750 MHz FTTF from an existing 450/550 MHz system. The analysis compares the premium in construction cost for a 600 MHz and a 750 MHz FTTF upgrade to the cost to upgrade a current 450/550 MHz system to full 550 MHz with fiber backbone to fiber hubs that could feed 500 home passed optical nodes in the future. The construction cost of these three alternatives was investigated relative to a number of variables, the main ones discussed here are:

- the percentage of coaxial cable that needs to be replaced as part of the upgrade alternative.
- the percentage of the existing amplifiers that can be upgraded to 600 MHz.
- the percentage of the existing plant (typically a 450/550 MHz mix) that is 550 MHz.

Figure 1, shows that the cost premium required to upgrade 450/550 MHz plant to 600 MHz is less than 10% and this premium falls as the amount of cable that needs to be replaced increases and the total cost of each type of upgrade increases as shown in Figure 2. As more of the existing cable must be replaced in an upgrade and as this percentage of cable increases beyond approximately 50%, the cable operator is effectively in a rebuild mode. While the premium to upgrade the existing 450/550 MHz plant to 600 MHz falls from 7% at 12% cable replacement, to 2% at 50 % cable replacement, the cost to upgrade to 750 MHz FTTF falls below a 20% premium beyond a 50% cable replacement threshold. Figure 1, illustrates why most cable operators consider a 750 MHz FTTF rebuild when most of the cable must be replaced.

As a result of the shorter amplifier cascades at 750 MHz FTTF and the higher gain 750 MHz amplifiers the physical spacing at 600 MHz and 750 MHz can be designed to be the same. Subsequently, a significant amount of feeder plant that was upgraded to 600 MHz is usable at 750 MHz, after exchanging 600 MHz line extenders for 750 MHz units.

Cable Placement vs Cost Premium Over 550 MHz Upgrade
FIGURE 1



Cable Placement vs Total Upgrade Costs
FIGURE 2

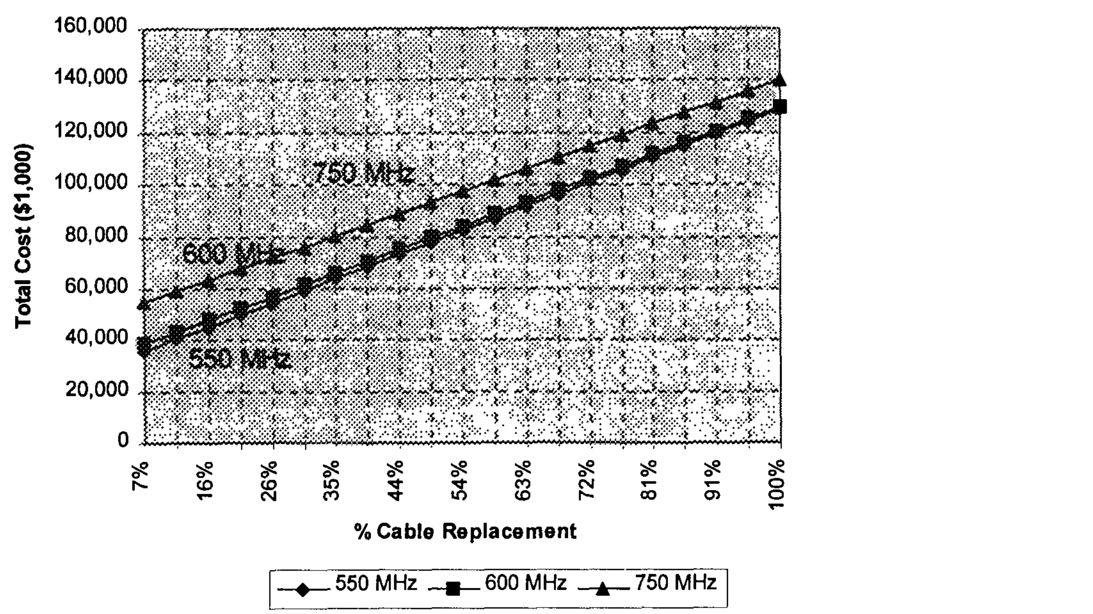
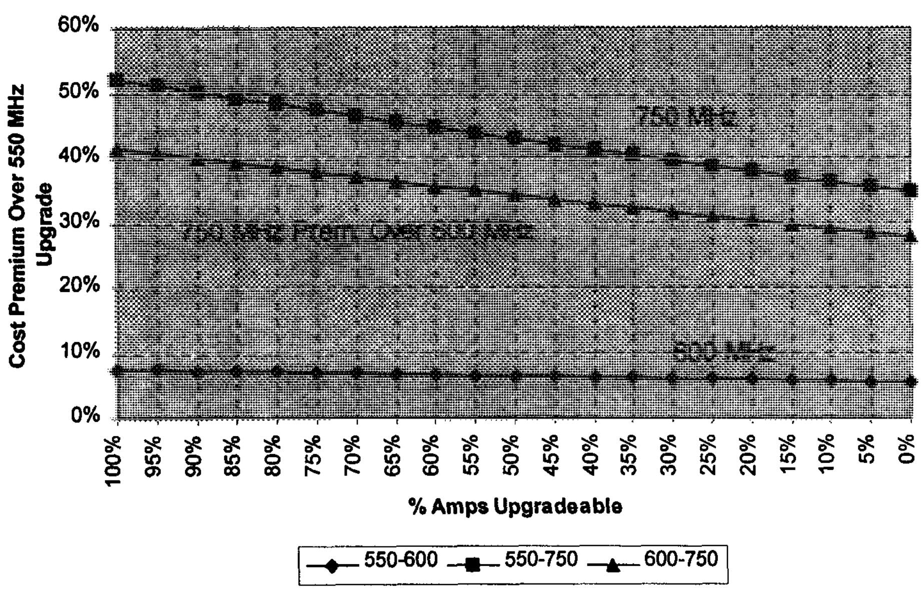
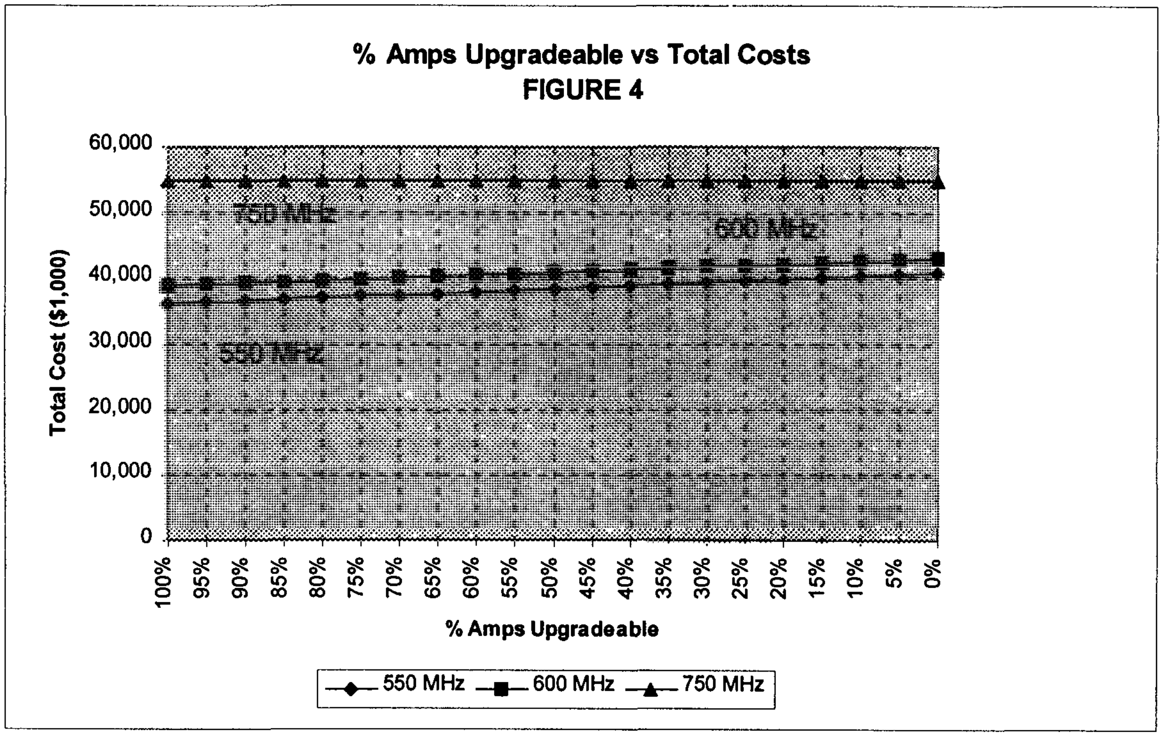


Figure 3, shows that the premium for 600 MHz doesn't change significantly if fewer amplifiers in the cable system are upgradeable. If fewer of the existing 450/550 MHz amplifiers are upgradeable to 550 MHz or 600 MHz, the cost of both the 550 MHz and 600 MHz upgrades increase at approximately the same rate, refer to Figure 4. The premium for 750 MHz FTTF falls faster than the premium for a 600 MHz upgrade since fewer 450/550 MHz line extenders are upgradeable and more new 750 MHz line extenders must be purchased. As a result, the total cost of the 550 MHz and 600 MHz upgrades increase faster than the total cost for 750 MHz FTTF upgrade. If very few of the amplifiers are upgradeable, it can still make sense to buy 600 MHz feedforward trunk stations, install 750 MHz line extenders and defer the placement of fiber and FTTF nodes. If significant (>30%) of the cable must be replaced the case for a 600 MHz trunk upgrade becomes weaker.

% Amps Upgradeable vs Cost Premium Over 550 MHz Upgrade
FIGURE 3





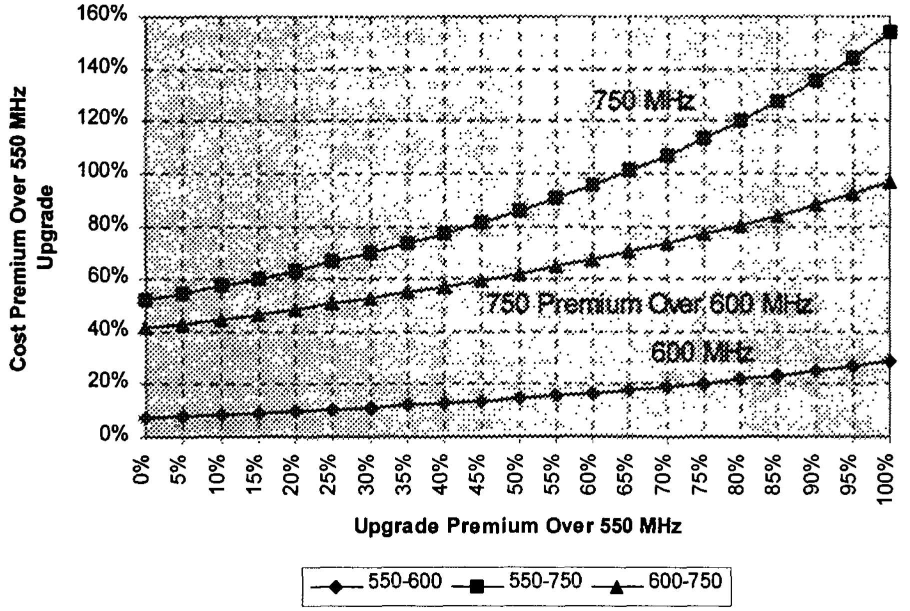
If a large percentage of the existing plant is 550 MHz (450 with 550 MHz areas), the total cost to upgrade to 550 MHz, 660 MHz or 750 MHz falls, refer to Figure 5. The total cost for 550 MHz plant falls faster than the cost for 600 MHz or 750 MHz FTTF plant as shown in Figure 6. As the percentage of 550 MHz plant in a system increases the cost shown for a 550 MHz upgrade is effectively the cost to install the fiber backbone, fiber hubs and nodes to reduce all trunk cascades to a maximum of ten. The premium to upgrade existing 550 MHz plant to 600 MHz is 20% to 30% higher than the cost to simply install fiber hubs and nodes in the existing 550 MHz plant.

Phased Approach

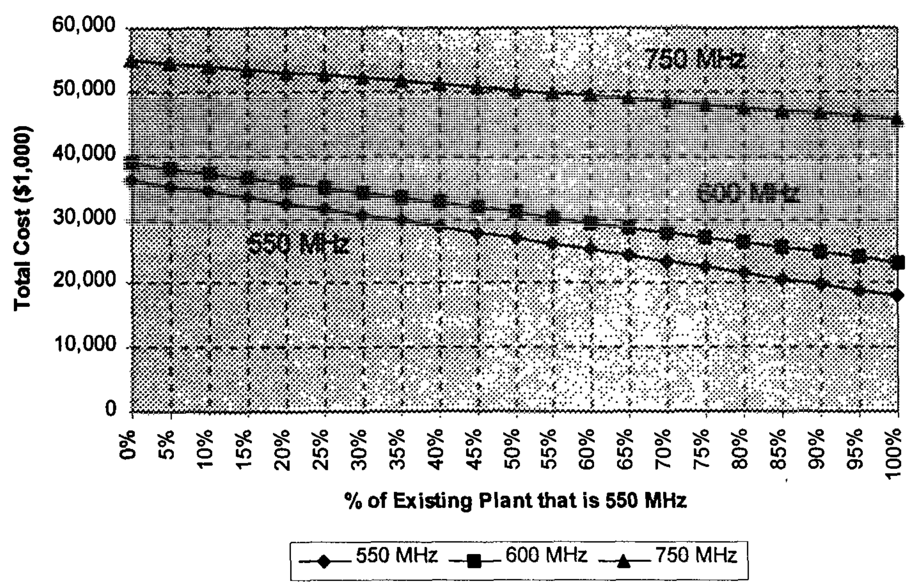
The phased approach to plant upgrades used by Rogers involves dividing a system into phase areas that typically have a similar: number of amplifiers, amount of coax/fiber cable and are served by a similar number of fiber hubs or nodes fed from the headend or regional hub. The use of fiber hubs or nodes allows a particular phase area to be treated independently from the other phase areas. Each phase area has its own database detailing:

- the existing equipment.
- the bill of materials requirements for upgrade.
- all other resources for the construction and testing of the phase area.
- costs per home passed, costs per subscriber and associated upgrade efficiency measures.
- characteristics of the plant or subscribers in a phase area, such as: line extenders per trunk station, home passed per active

% of Existing Plant that is 550 MHz vs Premium Over 550 MHz Upgrade
FIGURE 5



% of Existing Plant that is 550 MHz vs Total Cost
FIGURE 6



- and subscriber demographics within a phase area.

The ability to treat the phase areas independently provides several benefits:

- each phase can be tested and certified for: construction quality, end of line performance, operation of plant status monitoring and system control completely independent of the other phase areas.
- variations in construction cost and upgrade progress for phase areas can be identified early in the upgrade and factored into the completion schedule and cost estimates of the other phase areas and the entire upgrade plan.
- trouble shooting and clean up of two-way return plant can be implemented based on the planned launch of two-way services for each phase area.
- new services or products can be launched phase area by phase area depending on plant characteristics or subscriber demographics.

SUMMARY

Three main factors determine the viability of a strategy to upgrade existing 450/550 MHz plant to 600 MHz plant with a fiber backbone and fiber hub architecture. These factors are:

- quality of the existing plant (percentage of cable that must be replaced).
- the channel capacity required for expected new services.
- the cost of capital and the risk of stranding capital by building 750 MHz FTTF plant substantially before it is a proven requirement.

If in fact more capacity is required in the future, the 600 MHz plant can be upgraded to 750 MHz FTTF on an phase area by phase area and a system by system basis.

The phase area approach to upgrades is a way to dimension upgrade areas, collect important information on the plant and subscribers in that phase area to assist in a phase area by phase area project management of the upgrade and launch of new services to the subscriber.

VIDEO COMPRESSION PERFORMANCE: SUBJECTIVE EVALUATION OF PICTURE QUALITY

Bronwen Lindsay Jones, Independent Contractor
Richard S. Prodan, Ph.D., Cable Television Laboratories, Inc.
David A. Eng, Cable Television Laboratories, Inc.

Abstract

An experimental design incorporating psychophysical test methods for use in the subjective assessment of picture-quality variations due to compression artifacts is described. Standard-definition television pictures are used as test material, including video, film and some extensively utilized MPEG (Moving Picture Experts Group) test sequences.

This study is unique and may be the first of its kind. The effect of compression artifacts on picture quality is very subtle in nature. There are many trade-offs between and among bit rate, compression tools, resolution, picture content and recording format.

A combination of test methods is being utilized, based in part on a new International Telecommunications Union-Radiocommunication Sector (ITU-R) listening test procedure for testing very small differences.

Introduction

The advent of differing compression schemes and bit-rate combinations for video and audio information has caused confusion in the industry (over 100 combinations have been reviewed to date). Optimization or best-fit for viewer preference and expectation of picture quality is expected to vary with program content and demographics. Answers which will help guide industry standards are needed as are recommendations regarding variability with program content and source format.

Because standard-definition, compressed-digital picture artifacts have not been examined in any formal subjective manner, their effect on perceived picture quality to various viewer populations is not known. This proposal describes an initial small pilot study aimed at getting these answers, using industry experts as observers.

Method

The uniqueness of this study is in the new and very subtle nature of the manner in which compression artifacts manifest themselves in pictures. Objects in motion can exhibit blockiness, edge business, and a shimmery, twinkling "mosquito" noise that is visible in high contrast areas around sharp transitions. Test methods must be chosen carefully to fit experimental test conditions in order to avoid overly influencing the outcome. For example, compression artifact differences are small. A subjective scale with too little resolution will show large differences but not small ones. Small differences are perceived, but will not show up in the data.

Due to the expertise of the observers, the small differences in compared picture quality, and the internal nature of the initial study, a special combination of test methods is being employed. Unlabeled graphic scales are presented for recording viewer judgements in a continuous, proportional ratio-scale manner. In addition, viewers are encouraged to record comments in an information gathering technique which has become known in the industry as Expert Observation & Commentary (EO&C).

Observers

Expert observers from within CableLabs, including the authors/experimenters, set up the study and finalized the choice of appropriate test methods and test materials.

Expert observers from within CableLabs' membership (i.e., industry experts and video engineers) will be used as subjects in these pilot studies; non-expert viewers may be added later for more generalized test results. Experts are defined by the ITU-R (formerly CCIR) as "observers who have had recent extensive experience in observing picture quality or impairments, particularly of the type being studied."

Picture Source Material Selection

New motion-picture test material has recently been assembled by groups such as The Moving Picture Experts Group (MPEG), the Federal Communications Commission's Advisory Committee on Advanced Television Systems Planning Subcommittee Working Party 6 on Subjective Assessments (ACATS PSWP6), and some private parties. It generally consists of 10-to-35-second segments of video material (no audio) originated on film, on HDD 1000, on D1, and on Beta SP. This ensured excellent quality origination. Such quality origination is especially important for source comparisons.

Additional test material originated on film was provided by Lucasfilm Ltd. Test material that was representative of high-motion cable sports programming was provided by ESPN Engineering. Video sequences from the original CableLabs/Viacom compression test source were included.

The choice of which test material selections to use in an experiment is made by expert review and includes seeing all test material under all viewing conditions. Those sequences which are most sensitive to the impairments or artifacts being studied are chosen. The selected sequences are displayed according to system M 525-line component standards, after compression and expansion. As a result, a compilation of 17 selections from MPEG, Lucasfilm Ltd., ESPN, and Viacom are the primary test-material grouping. The material is arranged for presentation in several blocked and balanced pseudorandom orders.

Test Material Production

A studio-quality, Panasonic D-5 digital component VTR input source material to a digital video compression encoder connected to a decoder produced by the same vendor.

The encoder output fixed values of constant bit rate. The output of each decoder was tape recorded with a second studio-quality digital component VTR. A short segment of compressed digital information in MPEG transport stream format was captured and verified by CableLabs. The verification process included bitrate, resolution and profile elements, as well as MPEG-2 bitstream syntax at the system and video layers.

Test Equipment

Hardware included a studio-quality video tape recorder (Panasonic model AJ-D580P), the randomized tapes, and two or more Sony BVM-1911 CRT monitors. If viewer response-gathering becomes automated at a later date, computer keyboards may be included.

The study was conducted in a viewing environment closely matched to ITU-R viewing-room specification. One to three viewers participated at a time.

Experimental Test Design and Procedure

This study was segmented into two or three primary blocks:

1. Source and four compression techniques including full CCIR-601 resolution MPEG Main Profile (MP) using I, P, B frames; Simple Profile (SP) using I, P frames with MPEG-2 dual prime smart prediction; intra-coded and predicted-frame coded (IP)

MPEG-2 profile using I, P frames without smart prediction; and a proprietary profile using I, P frames with DigiCipher™ extensions. Processing was accomplished by a single manufacturer's encoder/decoder at bit rates of 8, 6, 4.5 and 3 Mb/s;

2. Source and full resolution MPEG MP, SP, IP, and proprietary processing as generated for comparison by two primary encoder/decoders at bit rates of 8, 6, 4.5 and 3 Mb/s;

3. Source, full, and 3/4 resolution MPEG MP, SP, IP and proprietary processing as generated by all encoder/decoders at bit rates of 8, 6, 4.5 and 3 Mb/s.

Block One is an evaluation of a single coding system operating at full resolution, but at different bit rates and with different profiles to determine just exactly how similar they are. There are as many as 32 possible operating conditions tested. Block two and three evaluate between and across manufacturer's systems.

Side-by-side pair comparisons, in parallel rather than in sequence, will be the primary response-gathering procedure. Unlabeled graphic scales will be used in a continuous, proportional, rather than discrete, manner. (This approach asks not only if something is better, but how much better; rating scales ask merely for a position on a continuum). Both members of each pair are judged and scored, in accordance with ITU-R Recommendation BS 1116. Expert Observation & Commentary (EO&C), as was used in the listening tests for the FCC ACATS in the US audio standards effort (recently concluded end of 1995, HDTV sound) will accompany the scaling procedure.

Test material will be made up of different groupings of the 17 test material selections previously described. The viewing distance from the screen (three or four times the picture height) was chosen by the expert observers and noted and reported.

Block one may be conducted repeatedly with any or all systems if deemed in need of a similar thorough evaluation. Blocks two and three will be used in the same manner for comparison across systems: ratio scales in side-by-side pair comparisons using EO&C.

Depending on the visibility of artifacts, it may be necessary to impose top and/or bottom anchors by including uncompressed source and/or very low bit rate compressed test material.

Presentation of Results

The presentation of results includes graphs and tables of means and standard deviations. Ratio scaling of data usually makes use of geometric, rather than authentic, means and standard deviations in order to lessen the influence of the spread of the members (proportions are of interest, not numerical values).

The size of this study is fairly small and manageable, therefore, viewer responses can be gathered with paper and pencil and the analysis of results can be done manually. If demonstrations become important or the number of observers grows a great deal, it may become desirable to automate viewer-response gathering. When this is done, the statistical processes are incorporated into the software and data can be plotted immediately upon completion of the test. This kind of automation also considerably speeds up data analysis and report writing.

References

1. Recommendation ITU-R BS.116, Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems.

2. Recommendation ITU-R BT.500-6, Methodology for the Subjective Assessment of the Quality of Television Pictures.

ISBN 0-940272-01-6; 0-940272-08-3; 0-940272-10-5; 0-940272-11-3; 0-940272-12-1; 0-940272-14-8; 0-940272-15-6; 0-940272-16-4; 0-940272-18-0; 0-940272-19-9; 0-940272-20-2; 0-940272-21-0; 0-940272-22-9; 0-940272-23-7; 0-940272-24-5; 0-940272-25-3; 0-940272-26-1; 0-940272-27-X; 0-940272-28-8; 0-940272-29-6; 0-940272-32-6; 0-940272-33-4; 0-940272-34-2; 0-940272-35-0; 0-940272-36-9; 0-940272-37-7; 0-940272-38-5; 0-940272-39-3; 0-940272-40-7; 0-940272-41-5; 0-940272-42-3; 0-940272-43-1; 0-940272-44-X; 0-940272-45-8; 0-940272-46-6; 0-940272-47-4; 0-940272-48-2; 0-940272-49-0; 0-940272-50-4; 0-940272-51-2; 0-940272-52-0; 0-940272-53-9; 0-940272-54-7