THE COMPLETE
TECHNICAL PAPER PROCEEDINGS
FROM:

A METHOD OF ANALYZING MPEG DATA IN ENCAPSULATED STREAMS

F. Eugene Rohling
DVA Group, Inc.

Abstract

This paper describes a method of analyzing encapsulated binary data streams for the purposes of performing detailed message analysis. This method evolved from a general purpose analysis tool used to analyze radar data. It is now being applied to the analysis of MPEG-2 content and access control data delivered both in-band and out-of-band. It is particularly useful for compartmentalizing the details of sensitive control and encryption information within the MPEG data strea of an access control system..

The method allows users to describe encapsulated framed data, parsing a binary data stream, and generating human readable output that can be used to analyze and resolve problems. The template files can be tailored and customized to reveal varying levels of proprietary and confidential data within the binary stream.

INTRODUCTION

This paper identifies a solution that helps test and field engineers analyze complex MPEG data streams. It uses the familiar NAS access control service as an example of data that has been encapsulated four times when it is received within a headend system. Finally, it discusses the need for these tools as new technologies emerge.

This paper specifically discusses access control data. Many off-the-shelf tools exist for analyzing standard MPEG-2 video and DOCSIS services. However, access control systems are by their nature proprietary, and tools for looking at stream usage of Motorola Broadband DigiCipher, Scientific Atlanta
Power Key, and other access control streams are usually held close. This makes it difficult for an MSO to find problems in his local system, especially when he is responsible for operating it.

Encapsulated MPEG Data

The National Access Control Service (NAS) owned by Motorola Broadband and operated by AT&T (now Comcast) is an excellent example of MPEG encapsulated data. Figure 1 shows the various layers of MPEG data. First, the DigiCipher OOB data is encapsulated into MPEG private data message packets. When it arrives in the headend, data is then sent from the satellite receiving device (IRT) across Ethernet to the out of band modulator (OM). That is, the OOB data is carried as an encapsulated MPEG data stream within a HITS multiplex through the satellite system. [1]

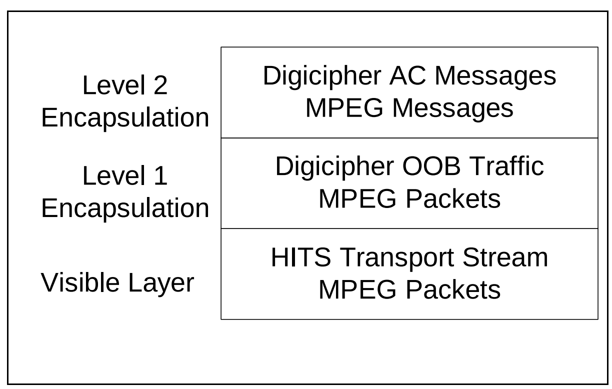


Figure 1 - NAS Encapsulation

A standard MPEG recording tool such as the DSTS by Logic Innovations allows you to record the data stream as it is received by the IRT. But if you want to recover only the data seen by the set-top, then you must remove the HITS transport stream.

Several one-off tools have been built to detunnel the data, but they are all considered proprietary by the AC provider. Sometimes an MSO has legitimate reasons to determine if his access control system is operating properly or if he is receiving all the data his contract with NAS provides.

Similar problems exist with Motorola DAC based local access controllers. In this case, the problem becomes more urgent because the MSO is responsible for the operation of the DAC.

In many systems, access control data is encapsulated on a TCP/IP network and sent to a modulating device. Rather than data being MPEG encapsulated in MPEG, it is now MPEG encapsulated within IP. While good Ethernet tools exist, they to not provide utilities to integrate with MPEG tools. [1]

Compartmenting Data

To give the MSO the tools Motorola originally used to develop DigiCipher would be giving away the keys to their access control kingdom. But to give MSO's tools that help identify if code objects are spinning, or if TV Guide data is still online, or to identify if channel maps are being provided to their facility are all reasonable requests.

A legitimate need exists to compartment the visibility of MPEG access control implementation so legitimate users can visualize it operationally without compromising the access control system.

Processing Binary Data

Many processing programs exist for processing text. Unix has a wealth of tools such as awk, sed, grep, lex, and perl. But converting a 100 MByte file from binary to readable text becomes unwieldy when the result can generate many Gigabytes of data and take significant time to sort through and filter that data.

It is significantly less time consuming for analysts to process binary data and extract only the information they need to do their task.

HISTORY

The problem of analyzing a complex data stream that has been multiplexed into many layers is not unique to the cable or MPEG industries. Instrumentation systems during the 1980 to 1995 time frame commonly mixed and multiplexed dissimilar data from many sources within a telemetry or tape recorded data stream.

The Link to Radars

A good example was a radar instrumentation system developed for the F-15, F-16, and B-1 aircraft by Lockheed Georgia under the Advanced Radar Test Bed (ARTB) program. The requirements for that system required it to visualize and record traffic from up to four MIL-STD-1553 data bus streams, up to four streams of telemetry data, several custom low, medium, and high speed data streams at an aggregate rate of up to 12 Mbytes/sec. This was a feat for the 1989 designed system. They also required the system to be versatile and instrument any of five radars on the three aircraft. The requirements finally required time stamping the data to +/- 10 microseconds.

High Speed Analysis Becomes Key

Instrumenting the aircraft, multiplexing data, recording data, and time tagging data was straightforward. Much of it was performed in hardware. But the system proved that reducing and analyzing the data became a significant labor intensive task. U.S. Air Force engineers likened the task of finding a needle in a hay stack.

This system evolved into the bench top Radar Instrumentation System (RIM-68) developed by Flexible Engineering Resources, Inc. (FER). This company developed a method of encapsulating the data in a common format and a method of parsing the data at high speeds so a small number of parameters could be visualized in both text and graphic format. The method was coined "MAcq" for Modular Acquisition.

MACQ FILTERING [3]

The "macq_filter" program performed the analysis side of this task was called the "MAcq_filter". It analyzed data for both real-time and post processing. It used "filters" that described the encapsulated nature of the data stream to both extract and process the stream into either human readable form or into derivative streams for off-the-shelf graphic programs to process as shown in Figure 2.

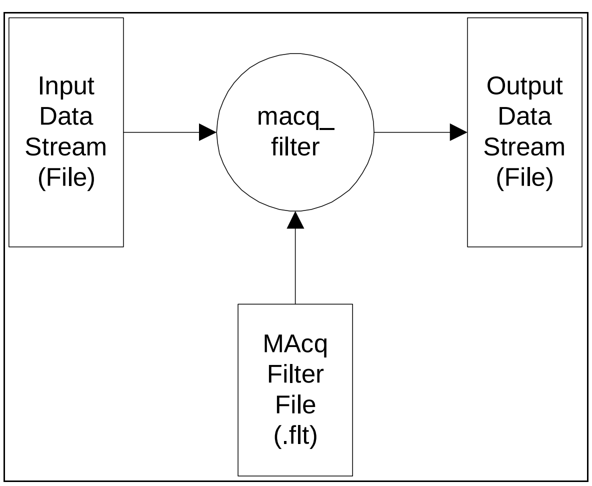


Figure 2 - MAcq Filter Process

Processing Frames of Data

The MAcq filter input description was designed to process nested frames of variable length data in a serial data stream. Figure 3 shows the format of the filter file. Note that the format of the filter file allows recursion. That is, optional filter frames can be nested within a top level scope frame to create the same data recursion effect often found with software recursion. This is the primary benefit of applying MAcq filters to encapsulated data problems.

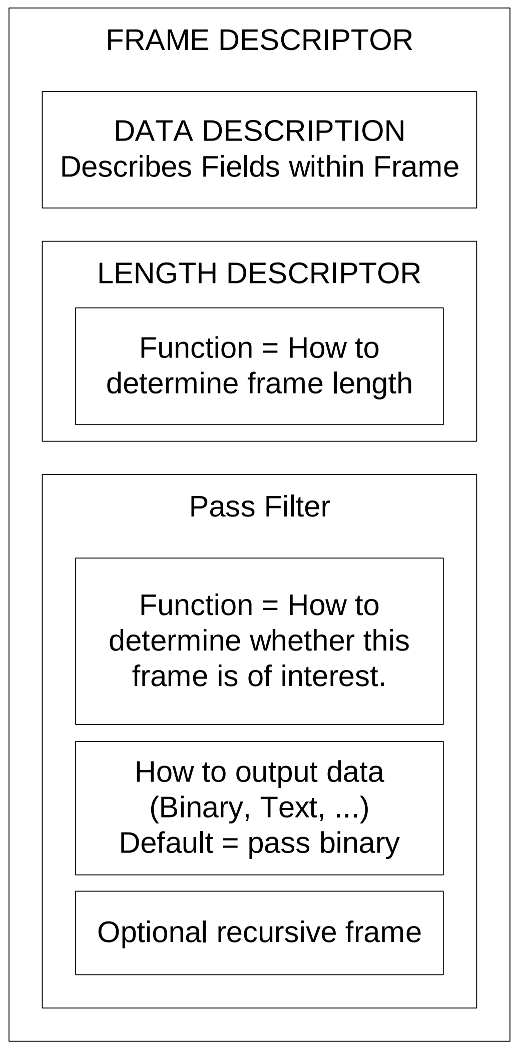


Figure 3 - Filter File Format

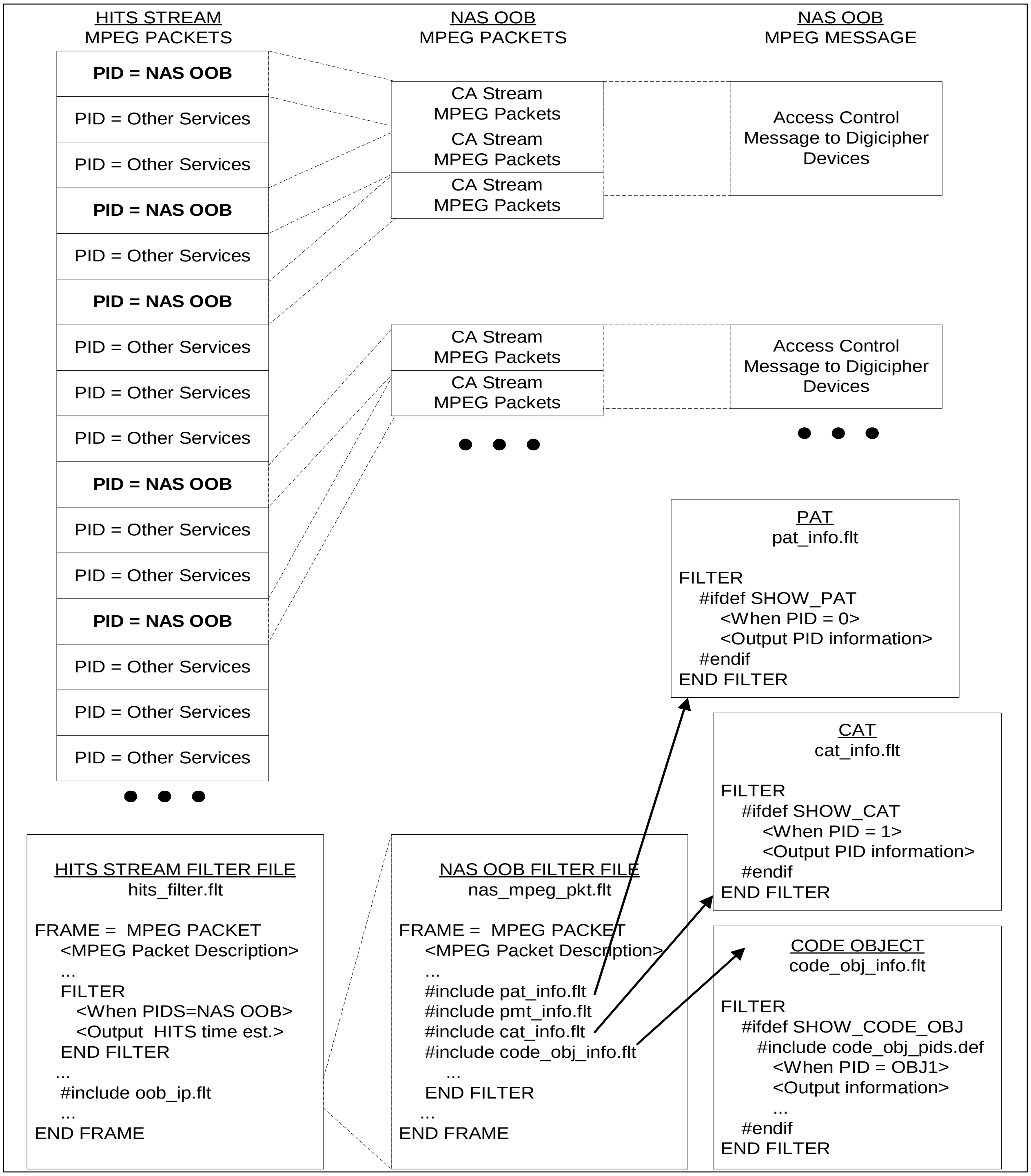


Figure 4 - Representing Encapsulation

Data Description

Each of the fields within a frame must be defined. The data description block within the filter identifies fields of data, such as the packet sync (47 Hex), the continuity counter, or the PID fields of an MPEG packet.

Variable Length Data

While MPEG packets are fixed format (188 bytes or 204 bytes), UDP / IP data is not. The length, however, can be readily determined from the contents of the UDP packet. Note the length clause contains a function used to establish the length of the arbitrary frame.

Selecting Data to be Processed

One or more pass filters look at frame headers and establish whether data needs to be passed. For MPEG data, the pass filter would likely select PIDS. For UDP data, it might select UDP source and/or destination ports.

Once data is selected, it is then processed. The output section defines what data is to be output. Output can be formatted text such as:

PID=234 TIME=88:99

or it can be binary data. Outputting binary data is quite useful for simple extraction of encapsulated data. That is, if all you want are the MPEG packets from a NAS IP OOB stream going to an OM-1000, you simply detunnel the UDP packets to that device.

Storing Data

The MAcq filter allows "scratchpads" to be used to temporarily store data. This initially became very useful when analyzing F-16 radar data.

APPLYING MACQ TO MPEG DATA

DVA Group began a research program in 2002 known as "Crown Royal" or CR to identify whether MAcq could be used to parse MPEG data and generate text output files.

Processing Frames of Data

Figure 4 shows how MAcq filter files can be used to describe and process the NAS satellite transport stream and extract the conditional access table (CAT). This shows a simple case of extracting OOB messages.

Need for Storage

Note that MPEG packets contain MPEG messages, and that MPEG messages can span multiple MPEG packets. When analyzing an MPEG stream in the general case, MPEG messages on multiple PIDs may interleave themselves in the temporal sequence of the MPEG stream. The MAcq scratchpad is useful for this case.

However, the MAcq implementation only allows statically defined scratchpads. This was fine for only detunneling OOB data, but was not adequate for cross PID correlation problems. As such, the general case of providing a general PID storage for detunneling MPEG messages was not adequate. Indexed scratchpads need to be added to the MAcq filter syntax.

Compartmenting Knowledge

In this context, compartmentalization refers to the Department of Defense (DoD) style security compartmentalization used during the cold war. That is, everything is on a "need-to-know" basis.

Access control providers have been reticent to only provide necessary information outside (and often inside) their corporate control. Providing MSOs and vendors with too much detail places the acces control provider at risk, and makes the MSO vulnerable to attack.

The MAcq filter provides a method of only providing information on a "need-to-know" basis. That is, filters that describe MPEG formatted information, or that simply announce the presence of a channel map, code object, or conditional access table may be appropriate for an MSO to obtain. However, the details of conditional access, especially key exchanges can be hidden by simply omitting the filters that are not needed.

PUTTING IT ALL TOGETHER

Engineers in the cable industry have many tools at their disposal. Many off-the-shelf products will parse Ethernet and IP packets, and others parse MPEG packets. Use of MAcq should take advantage of the strengths of existing tools.

Analyzing Local Access Control Data

Local AC data is often encapsulated on an Ethernet IP network. Off-the-shelf tools such as Etherpeek and the Unix tcpdump utility provide historical recording of Ethernet IP network in text or binary form. To make sense of the MPEG packets, however, requires the content to be detunneled.

The MAcq_filter can be used to detunnel the MPEG packets and put them in a form that MPEG analyzers can use. They can then be analyzed in native MPEG forms.

The same solution addresses instrumentation of systems in which video is transported across an Ethernet IP network. Many new MPEG re-multiplexors are being introduced that accept video streams across IP networks.

Using with Unix Pipes

Visualizing the delivery of code objects, VOD content, channel maps, and other necessary components of a cable system requires a tool that can output data in graphical form.

The macq_filter has been used in the radar community to visualize its effectiveness. The tool filters, processes, and then streams selected data in both real-time and playback instances into off-the-shelf 3-dimensional analysis tools.

The same can be applied to monitoring the OOB data within a headend. That is, MAcq can filter and process the access control stream and stream data into commercially (and sometimes free) third party software tools that display arbitrary bar graphs. This can be used to build tools that show code objects, channel maps, and other access control data as a percentage of bandwidth.

Work to Date

DVA Group has successfully used the original macq_filter program for simple tasks. The original program worked because 188 byte packets were long word aligned. It enabled analysis of PID distribution, continuity counts, and extraction of PIDS in binary form. It also allowed an encapsulated IP layer to be extracted from a given PID in an MPEG transport stream.

But extracting an OOB stream encapsulated within IP data could not be performed without being able to parse frames in byte word alignment.

SUMMARY

We have proven the underlying technology behind the macq_filter tool can help fill the gaps in commercial MPEG analysis tools. DVA Group continues to evolve the filter tool so it properly supports the needs of embedded cable systems in the future.

MANY THANKS

The author extends his appreciation to Michael Adams and all the people who helped write "OpenCable Architecture". By showing a top level view of how Motorola Broadband and Scientific Atlanta conditional access systems work, we can discuss real world cable industry applications in a public forum.

REFERENCES

1. Michael Adams, "OpenCable Architecture", 2000, Cisco Press.

2. Dr. Keith Montierth Jr, Todd Jahng, Lisa Chiang, Greg Rohling, Gene Rohling, "Three Dimensional Data Visualization System for the F-16 Program", JAWS S3 Conference, June 10, 1997.

3. Flexible Engineering Resources, Inc. "MAcq User Manual", September 13, 1997.

ADVANCES IN OPTICAL FIBER TECHNOLOGY FOR ANALOG TRANSPORT-TECHNICAL ADVANTAGES AND RECENT DEPLOYMENT EXPERIENCE

Andy Woodfin, Jim Painter*, Boh Ruffin
Corning Incorporated, *Comcast Corporation

Abstract

Performance assessments in analog video transport and distribution will be compared and analyzed, based on existing commercial optical and electronic equipment used with a variety of standardized optical fiber types. Particular emphasis is placed on comparison of capabilities with standard single-mode fiber to improve SBS thresholds on the order of 2dB, with associated increases in CNR, as well as improvements in CSO on the order of 8-9dB.

Experimental and simulated results will be presented, in addition to recent field data collected from actual physical links deployed by a major MSO. This is the first known commercial deployment of an alternate optical fiber type (i.e., not standard single-mode) expressly for the purposes of improving analog video transport capability.

INTRODUCTION

Among the primary telecommunications network architectures in use today, modern CATV designs provide unrivaled capability and capacity afforded by the hybrid fiber-coax (HFC) architecture. The underlying foundation of HFC networks is the optical fiber deployed primarily in the trunk/transport and distribution portions of the plant. By eliminating RF trunk amplifiers, increasing transmission bandwidth, enabling two-way transmission, and eliminating interference ingress, optical fiber has allowed CATV networks to transform into the pipes which now carry the full spectrum of voice, video, and data services. Undeniably, standard single-mode fiber has been the workhorse, and arguably the key element, in HFC design. Improvements in transmission capabilities have, as a result, historically been designed within the constraints of standard single-mode fiber characteristics. Appreciating the historical evolution of optical transmission over HFC architectures provides a useful perspective on these constraints, and the issues which consequently remain in nearly all modern HFC optical transmission systems.

AN ALTERNATE PERSPECTIVE ON TRANSMISSION TECHNOLOGY DEVELOPMENT

In early stages of HFC deployment, the benefits of transitioning from copper to optical fiber for CATV transport purposes were clear, with some of those advantages stated above. Significant development (and acceptance) was required in the optical transmission arena, however, to realize the large and powerful HFC networks of today. Optical transmission at 1310nm was typically viewed as sufficient where copper trunks were replaced with fiber, and the technology was relatively mature and economically feasible. The economics of system clustering and regional interconnection drove the need to adopt 1550nm transmission technology, where fiber loss is significantly less than at 1310nm and signals can be optically amplified with erbium doped fiber amplifiers (EDFAs). Standard single-mode fiber's chromatic dispersion at 1550nm is significantly higher than at 1310nm, however, which was a significant issue when the only sufficiently linear analog transmitters were high frequency chirp directly-modulated types[1]. The development of linearized externally modulated 1550nm transmitters addressed the issue of source chirp and interaction with fiber dispersion. However, fiber dispersion-induced self phase modulation (SPM) [2,3] was still an issue, in addition to exacerbation of the power-limiting impact of stimulated Brillouin scattering (SBS) (4,5) by the relatively narrow linewidth emitted by externally modulated sources. The severity of SBS was subsequently mitigated by integration of electrical pre-distortion and suppression techniques[6], although it is still a limiting factor in a number of system designs.

The transmission technology development summarized above can be viewed in the context of modification to optical fiber parameters, rather than working within the constraints of a fixed set of assumptions. While being an interesting academic exercise, it obviously does not address issues in the installed cable plant, where the fiber infrastructure is fixed. However, such an approach can indeed provide flexibility in designs for pending upgrades and rebuilds. Concerning the transition from 1310nm to 1550nm, significant reductions in attenuation at 1310nm would conceivably increase achievable transmission distances at the lower wavelength and enable wider application of lower cost transmitters, allowing enabling a broader application base for 1310nm. As illustrated in figure 1, however, Rayleigh scattering places a fundamental limitation on the minimum achievable loss at a given wavelength in current silica-based optical fiber, and current fibers closely approach that limit. While techniques exist to improve upon these limits through exotic materials and/or waveguide structures, they are not immediately adaptable into commercially viable fibers. The issue of high chromatic dispersion at 1550nm, on the other hand, has been addressed for some time in the long-distance telecommunications market with non-zero dispersion shifted fibers (NZDSF). NZDSF typically have dispersion on the order of 3 to 4 times smaller than that of standard single-mode fiber. Although designed primarily around the considerations of high capacity long distance networks, NZDSF can have direct benefit on CATV network designs by significantly reducing the impact of nonlinear and dispersion-related impairments such as SPM and composite second order distortion (CSO). Arguably the most significant limitation on analog transmission at 1550nm continues to be SBS, and the prevailing assumption has been that standard single-mode fiber best mitigates the effect. As SBS is directly dependent on the fiber's effective area (equation 1)[6], and standard single-mode fiber has a larger effective area (typically 80μm^2) than all NZDSF (typically 45-72μm^2)(7,8,9). However, it has been shown that some NZDSF are in fact superior to standard single-mode fiber in terms of SBS threshold, by as much as 2-3dB [10,11]. Fibers with this capability, coupled with optimally reduced chromatic dispersion, can show significant advantages over standard single-mode fiber to support real world analog transport network designs.

ASSESSMENT OF TECHNICAL ADVANTAGE

Details of technical capability

In simple terms, stimulated Brillouin scattering occurs in optical fiber due to a generation of acoustic waves in the optical waveguide, which create periodic variations in the fiber's refractive index. This periodic variation effectively reflects part of the original transmitted optical power thus diminishing the power seen at a receiver. The effect worsens with increasing launch power, so the signal reduction at the receiver cannot be overcome simply by increasing the transmitter output. The power threshold at which SBS begins to quickly deteriorate a signal is given by:

$$(1)\quad P_{th} \cong \frac{21A_{eff}}{L_{eff}\,g_B}$$

where P_{th} is the SBS-dictated optical power threshold (in dBm), A_{eff} is the fiber effective area, L_{eff} is the nonlinear interaction length, and g_B is the peak Brillouin gain of the fiber. As stated previously, standard single-mode fiber A_{eff} is larger than that of typical NZDF, but significant variability in threshold among different fiber types due to variation in Brillouin gain characteristics has been empirically explored. Regardless, this phenomenon is commonly overlooked and effective area dependence is typically the only consideration made. With the appropriate combination of reasonable effective area (>70μm^2) and Brillouin gain, some NZDSF can support higher SBS thresholds than standard single-mode fiber.

Figure 2 shows an SBS threshold comparison between several commercially available optical fiber types. Considering standard single-mode fiber as the presumed standard for SBS threshold, the most commonly deployed NZDSF varieties were evaluated in comparison. The three NZDSF variants considered were: large area NZDSF, characterized by a relatively high effective area (approximately 72μm^2)[7] in comparison to other NZDSF; high dispersion NZDSF, with relatively high chromatic dispersion at 1550nm (~8ps/nm*km)[8]; and reduced slope NZDSF, characterized by relatively low chromatic dispersion slope and very small effective area at 1550nm (0.045ps/nm^2*km and ~55μm^2, respectively)[9]. All fibers under test were at a nominal length of 50km, and tested in the configuration illustrated in figure 3, with backscattered signals detected through a self-heterodyne configuration. As indicated in the figures, large area NZDSF has a significantly higher SBS threshold than standard single-mode fiber, in spite of the fact that it has a lower effective area (72μm^2 and 80μm^2, respectively). A relevant point to consider is that the relative differences in SBS thresholds are nominally constant regardless of electronic-based SBS suppression techniques. In other words, a transmitter with maximum SBS-limited power of 16dBm on standard single-mode fiber could support approximately 18dBm over large area NZDSF, while a 17dBm standard single-mode fiber rated transmitter could accommodate a similar 2dB increase (to 19dBm) over large area NZDSF. Also significant is the fact that other NZDSF varieties can not support the SBS threshold allowed on standard single-mode fiber.

Aside from variation in SBS suppression capabilities among standard single-mode fiber and the NZDSF variants, an approximation can be made to assess the introduction of second order distortion in different fiber types. Second harmonic distortion for a chirp-free externally modulated source, as determined by fiber dispersion and nonlinear refractive index, can be expressed as:

$$(2)\qquad \frac{1}{4}m\ddot{\beta}^2 z^2\Omega - \frac{1}{2}m\ddot{\beta}^2 z^2\Omega^2 P\left(\frac{2\pi N_2}{\lambda A_{eff}}\right),[1]$$

where m is the modulation index, z is the fiber length, Ω is the modulation frequency, P is launched optical power, N_2 is the Kerr nonlinear-index coefficient, λ is the transmitter center wavelength, and A_{eff} is the fiber effective area. Also, note that

$$\ddot{\beta} = -(\lambda^2/2\pi c)D$$

is the second-order fiber dispersion coefficient , where D is the fiber dispersion coefficient. For standard single-mode fiber, large area NZDSF, reduced slope NZDSF, and high dispersion NZDSF, we can consider the chromatic dispersion at 1550nm (17, 4, 5.2, 8ps/nm*km, respectively) and effective area

(80, 72, 55, 63 μm^2, respectively. Assuming all other terms in (2) are constant, we can make a qualitative assessment of the relative magnitude of CSO impairment in each fiber by scaling the ratio of fiber dispersion to effective area, D/A_{eff}. As is evident from the table, all NZDSF should have a significantly reduced CSO distortion relative to standard single mode fiber.

Fiber Type	D/A_{eff} ratio ($ps/nm*km*\mu m^2$)
Standard single-mode	0.212
Large area NZDSF	0.056
Reduced slope NZDSF	0.094
High dispersion NZDSF	0.127

If only performance parity with standard single-mode fiber is desired, the comparative assessment of SBS thresholds carries significant implications in the choice of fiber type to deploy. While large area NZDSF can support any given single wavelength 1550nm transmission scenario designed around standard single-mode fiber constraints, a system design would otherwise require careful consideration and possible power budget de-rating to avoid significant signal degradation if deployed over other types of NZDSF (i.e., high dispersion NZDSF and reduced slope NZDSF). The true justification for a choice of fiber other than standard single-mode fiber would obviously come from a desire to achieve performance benefits, as opposed to simple parity with the effective standard (standard single-mode fiber). Therefore, the gains derived from exploiting an increased SBS threshold, as well as the merits of reduced chromatic dispersion and other optimal parameters, warrant exploration.

Taking Advantage of the Technical Benefits

A number of potential performance advantages can be identified considering the combined impact of increased SBS suppression and optimized chromatic dispersion. The most readily apparent benefit gained from an increase in SBS threshold on large area NZDSF is the capability to support higher optical launch powers, and consequently extend the distance over which in-line optical amplifiers (EDFAs) would otherwise be required. Assuming equivalent loss characteristics on large area NZDSF and standard single-mode fiber, and considering only SBS, a 2dBm increase in SBS threshold would translate to approximately 8km increased distance with equivalent end-of-line received power. Coupled with a reduced chromatic dispersion and optimal effective area, however, large area NZDSF can further increase capability by both supporting higher powers and mitigating distortions. Indeed, previous studies have demonstrated 100km transmission over large area NZDSF with no repeaters or in-line EDFAs[10], and at shorter distances (50km) with high launch power demonstrated significant CSO and CNR advantage with large area NZDSF (CSO<-65dBc, CNR>50dB), compared to standard single-mode (CSO<-52dBc, CNR>44dB) and reduced slope NZDSF (CSO<-37dBc, CNR>24dB). By extension, this capability could extend to supporting longer reaches or superior signal integrity over a fixed distance with large area NZDSF while remaining within the constraints of existing design rules (e.g., maximum allowable number of cascaded in-line EDFAs). Link engineering rules can also potentially be extended since the input power to cascaded in-line EDFAs can increase due to higher launch powers, thus improving EDFA output CNR.

Given the broadcast nature of analog video transport, another beneficial application of increased launch power capability with large area NZDSF would be the potential to increase the number of remote locations supported with a single transmitter. Particularly for those locations not immediately targeted for advanced services, the economics of basic service distribution from a single transmitter become appealing. As an example, as illustrated in figure 4, an additional 2dBm maximum launched power could scale from a 1x8 passive splitter (loss=9dB/output) to accommodate the additional 2dB loss encountered on each arm of a 1x12 split (loss=11dB/output).

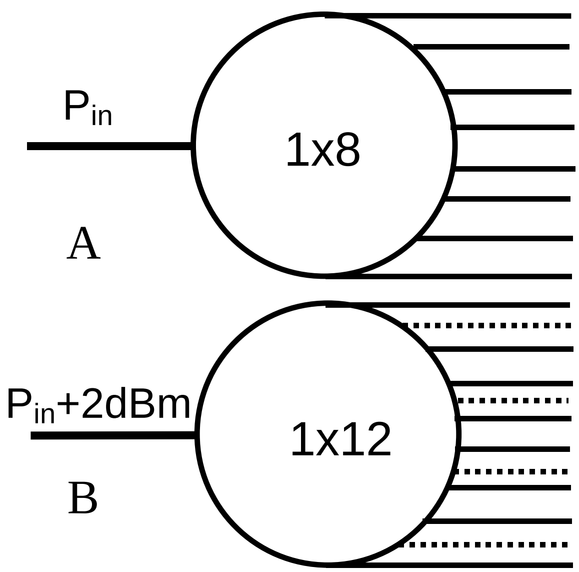


Figure 4: Splitter configurations: A-with standard single-mode fiber, B-with large area NZDSF

A promising possibility, again born from the coupled advantages of reduced chromatic dispersion and increased SBS threshold in large area NZDSF, is the ability to significantly increase the usable range of directly modulated 1550nm PEG transmitters. Typically characterized by significant frequency chirp and thus severely limited by dispersion-induced CSO, the introduction of large area NZDSF with reduced dispersion could potentially allow for PEG transmitter displacement of more costly externally-modulated sources to address trunking applications as opposed to simple signal insertion. The inherent SBS suppression resulting from modulation-induced spectral broadening, coupled with the improved power characteristics due to the lack of an attenuating modulator section, aids in drawing a significant comparison with conventional long reach externally modulated sources. This scenario is currently being experimentally evaluated at Corning.

FIELD DATA FROM DEPLOYED CABLE

Available commercial transmission equipment operating over a contiguous link of standard single-mode fiber was not capable of supporting internal CSO requirements of 68dB in the system link depicted in figure 5. As suggested previously, a reduction in total link chromatic dispersion could potentially mitigate CSO brought about by direct interaction between fiber dispersion and residual transmitter chirp, as well as CSO introduced by SPM-induced signal chirp (which also has some dependence on fiber effective area). Indeed, concatenating a 56.4km length of large area NZDSF to the previously installed 53.2km of standard single-mode fiber enabled a significant improvement in CSO. With an initial transmitter CSO of 76.9dB, the contiguous link of all standard single-mode fiber received 61.6dB and 59.4dB at channels 36 and 67, respectively. By introducing large area NZDSF into the latter portion of the total link, thereby reducing the overall accumulated chromatic dispersion, received CSO values with identical system parameters were 70.9dB and 71.4dB at the respective channels. For the two monitored channels, 9.3dB and 12dB improvements in CSO were realized over the total link, reducing the impairment such that it was well within the internal requirement. Note in addition the increased magnitude of improvement at the higher modulation frequency. Marginal improvements in CTB were also realized with the heterogeneous standard single-mode/large area NZDSF link, with 0.3dB and 1.1dB improvements at the respective monitored channels when compared with the homogeneous standard single-mode fiber link. Note again the slight increase in the performance delta at the higher modulation frequency.

CONCLUSION

Looking at the evolution of CATV networks and systems free from the technical constraints of the majority installed base of standard single-mode fiber allows for consideration of system solutions that can meet challenging performance requirements, extend the capabilities of existing transmission equipment, and provide opportunities to deliver significant savings in network flexibility and equipment cost. The capabilities of non-zero dispersion shifted fibers to significantly mitigate signal distortions are beginning to be explored in actual installations. Moreover, the large effective area subset of NZDSF allows for the broadest range of performance capability improvements among alternate fiber types, and in comparison to standard single-mode fiber.

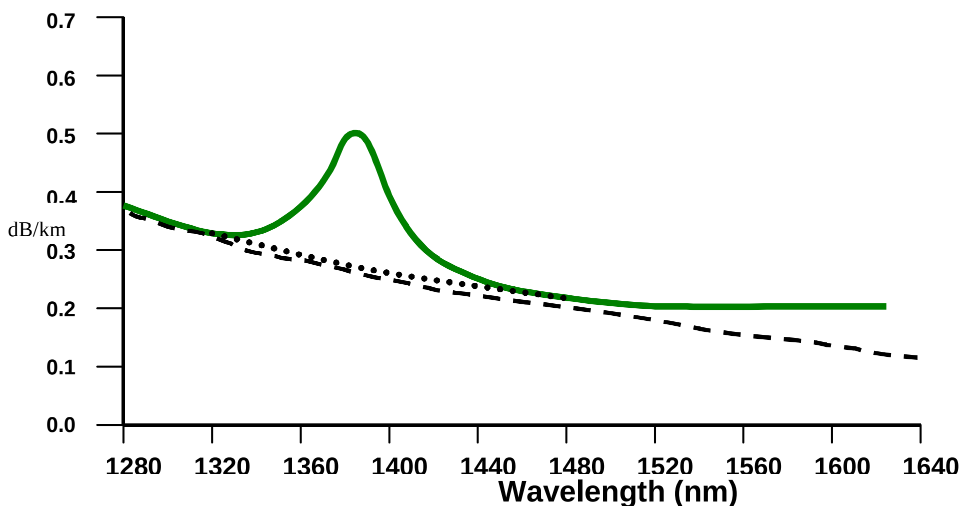


Figure 1: Typical attenuation curves for standard single-mode fiber (solid curve) and low water peak standard single-mode fiber (dotted curve), and fundamental Rayleigh scattering limit (dashed line)

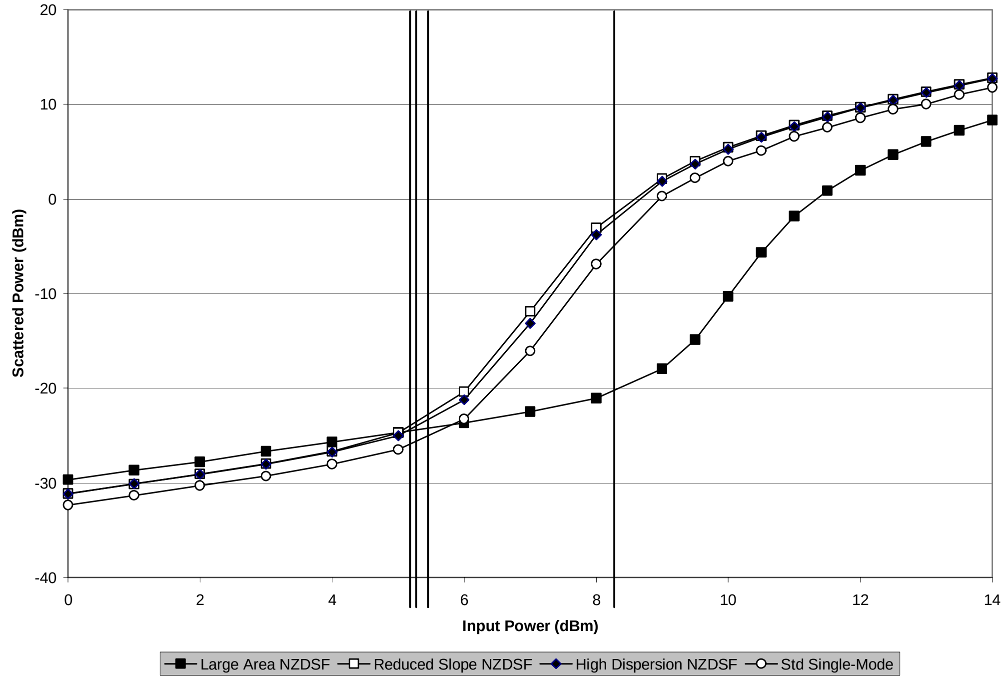


Figure 2: Comparison of input optical power and scattered optical power. Elbow of curve indicates approximate location of nonlinear onset of Brillouin scattering, values indicated by vertical markers for reduced slope NZDSF, high dispersion NZDSF, standard single-mode, and large area NZDSF, respectively.

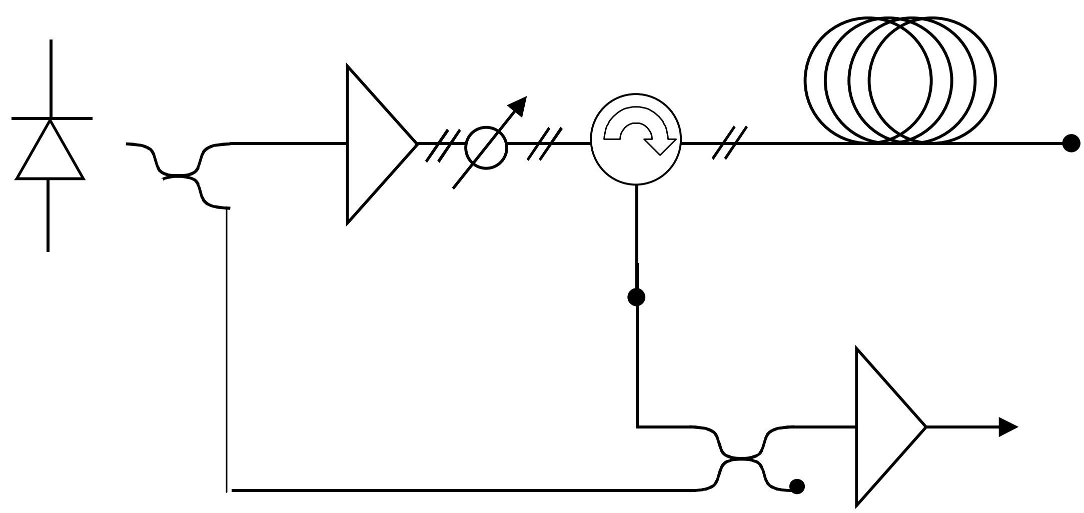


Figure 3: Experimental configuration for evaluating SBS threshold

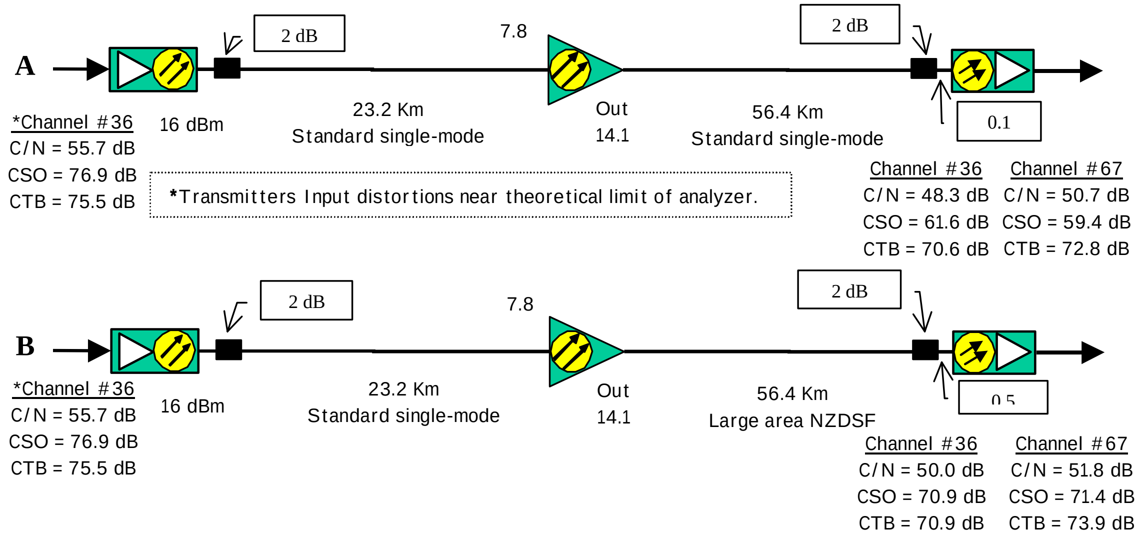


Figure 4: Configuration of installed links for comparison. A-Contiguous standard single-mode fiber link, B-Standard single-mode fiber extended with large area NZDSF.

REFERENCES

[1] M.R. Phillips, T.E. Darcie, D. Marcuse, G.E. Bodeep, and N.J. Frigo, "Nonlinear Distortion Generated by Dispersive Transmission of Chirped Intensity-Modulated Signals," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 481-483, 1991.

[2] F.W. Willems, W. Muys, J.C. van der Plaats, and R. Nuyts, "Experimental Verification of Self-Phase-Modulation Induced Nonlinear Distortion in Externally Modulated AM-VSB Lightwave Systems," Electron. Lett., vol. 32, pp. 1310-1311, 1996.

[3] R. Chraplyvy, "Limitations on lightwave communications imposed by optical-fiber nonlinearities," *J. Lightwave Technol.*, vol. 8, pp. 1548–1557, 1990.

[4] X.P. Mao, G.E. Bodeep, R.W. Tkach, A.R. Chraplyvy, T.E. Darcie, and R.M.

Derosier, "Brillouin Scattering in Externally Modulated Lightwave AM-VSB Transmission," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 287-289, 1992.

[5] R.G. Smith, "Optical Power Handling Capacity of Low-Loss Optical Fibers as Determined by Stimulated Raman and Brillouin Scattering,"*Appl. Opt.*, vol. 11, pp.2489-2494, 1972.

[6] G. P. Agrawal, Nonlinear Fiber Optics, 2nd ed., (Academic Press, San Diego, 1995).

[7] LEAF® Fiber Product Information Sheet, Corning Incorporated, September 2002.

[8] Teralight™ Metro Fiber Product Information Sheet, Alcatel, February 2002.

[9] TrueWave® RS Fiber Product Information Sheet, OFS Optics, February 2003.

[10] C. C. Lee and S. Chi, "Repeaterless Transmission of 80-Channel AM-SCM Signals Over 100-km Large-Effective-Area Dispersion-Shifted Fiber," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 341-343, 2000.

ANYTHING, ANYTIME, ANYWHERE: OPEN ADVANCED BANDWIDTH MANAGEMENT OF ON-DEMAND SERVICES

Michael Chen
Concurrent Computer Corporation

Abstract

The convergence of new technologies in an affordable fashion has given rise to new features that not only bolster customer demand, but also provide new revenue-generating opportunities. The ability to deliver the full promise of on-demand services -- "Anything, Anytime, Anywhere" -- is finally within reach, and customers are clamoring for their providers to deliver.

New features available now, and some of those envisioned for the future, are identified and investigated, as are the issues which face providers and vendors today. Observations and recommendations on the next generation of systems and their architectures are then offered in closing.

INTRODUCTION

Increased demand, competitive market forces, and technology advances have placed Gigabit Ethernet at the heart of new cable architectures offering additional revenue opportunities to the Multiple System Operator (MSO).

The adoption of standard Internet protocols has made the pervasive switching and routing capabilities which power the Internet available to these video delivery systems.

These capabilities provide a framework which, combined with new techniques such as network-based personal video recording (PVR), allow the MSO to deliver their customers the full promise of on-demand services – "anything, anytime, anywhere".

To deliver this, the MSO is faced with a bewildering array of challenges, from the selection and installation of compatible equipment to the configuration, management, and maintenance of this new infrastructure.

These issues facing both MSOs and equipment vendors today, as well as other looming issues, are further discussed below. The new features and capabilities of these systems, both at present and in future, are also identified and investigated. Finally, observations and recommendations are made for the design, procurement, and deployment of next-generation architectures and systems.

GIGABIT ETHERNET ON-DEMAND SYSTEMS

Current on-demand systems are largely being deployed using Gigabit Ethernet output. A typical video-on-demand (VOD) system employing Gigabit Ethernet looks like this:

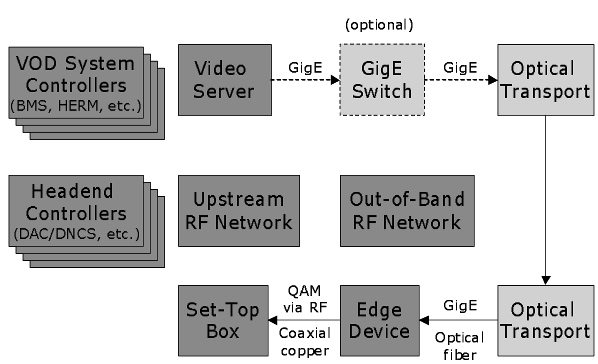


Figure 1. Typical Gigabit Ethernet VOD system.

The Gigabit Ethernet output of a streaming server is sent to an edge device, where it is combined and converted to a form suitable for display on digital cable set-top boxes. The server's output may be connected into a switch, and optical transport gear is used when necessary to transmit the signal across large distances.

The streaming server output is encapsulated within User Datagram Protocol (UDP) packets, defined as part of the Internet Protocol (IP) standards to provide low-latency data delivery, while taking advantage of the wide range of products and services the Internet explosion has produced.

Note that the output transmission is often implemented unidirectionally, since this allows the MSO to effectively double the amount of fiber bandwidth available. This one-way connection may require additional effort to configure systems initially, since many standards used with IP protocols assume the existence of a bi-directional network link for proper operation.

Gigabit Ethernet on-demand systems today are usually allocated dedicated network bandwidth for streaming. This often stems from the difficulty of ensuring sufficient quality of service to protect on-demand streams from being damaged by other data traffic. Having dedicated bandwidth for streaming, which the streaming servers then manage among themselves, greatly simplifies the overall system, and has accelerated the availability of Gigabit Ethernet solutions.

To Switch or Not to Switch?

Gigabit Ethernet switches were used in early deployments to aggregate the outputs of one or more streaming servers, when these servers were unable to generate enough traffic to fill an entire Gigabit Ethernet link.

Since streaming servers can now saturate Gigabit Ethernet links, a switch is no longer technically needed for deployment. However, the use of switches also provides new routing flexibility that was either unavailable or cost-prohibitive with prior output formats, and many of the new features which Gigabit Ethernet enables are built upon this functionality. For this reason, using a switched Gigabit Ethernet transmission framework is still quite advantageous for these on-demand services.

Asymmetric Deployment and Expansion

The division of labor between the streaming server and the edge device in the Gigabit Ethernet framework offers the MSO a new method for system deployment and expansion. Gigabit Ethernet's switching and routing functionality allows streaming servers and edge devices to be loosely rather than tightly coupled. The MSO can then deploy and expand edge devices separately from the streaming servers, allowing an asymmetrical buildout of the system.

A typical asymmetric buildout will overprovision the radio frequency (RF) edge with more edge devices than necessary to satisfy initial bandwidth demands. This is because installing new edge devices is often difficult to do without impairing the RF signal to a node, and requires more truck rolls to accomplish. The available granularities of optical transport equipment often will favor having more optical transport capacity than initially required, which may prompt the MSO to overprovision with edge devices at the same time.

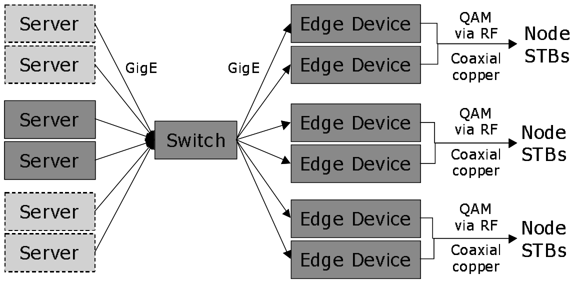


Figure 2. Asymmetric deployment and expansion.

Such an asymmetric buildout will generally add only as many streaming servers as required to meet current demand; as demand increases, more servers can be added at the headend and assigned to RF outputs of edge devices.

REALITY TODAY

What can today's Gigabit Ethernet on-demand solution currently provide?

"Something, Anytime"

Current solutions have limited on-demand content available. This is often not a storage capacity issue, but rather a rights licensing issue. The limited availability of on-demand content may force the MSO to select content for the on-demand system that is assumed to be more compelling than the broadcast digital cable offerings. This content typically includes movies, special events such as concerts, and popular sporting events.

Now Playing: "Anything, Anytime"

Personal video recorders (PVRs) such as Tivo can provide a wider selection of on-demand content to the home, but the limited availability of PVRs with integrated digital cable functionality curtails the overall benefit to the customer. PVRs also remove content storage control from the MSO at the home. This raises content protection issues, which tend to ripple back into rights negotiations.

However, successful trials of subscription video-on-demand content indicate that MSOs may not need to supply customers with DVR boxes to satisfy their desire for more varied on-demand content, as long as they can make desirable content available to their subscribers.

Network-Based PVR

In a network-based PVR approach, broadcast programming is recorded and stored by the MSO at the headend, rather than inside a consumer's set-top box, and is made available to on-demand streaming servers for transmission to customers upon request. Some implementations of network-based PVR allow a customer to pause a program in real time and use standard navigation features such as fast forward and rewind.

The advent of network-based PVR solutions levels the playing field with home PVR boxes, and allows the MSO to provide the full range of broadcast programming on demand, in addition to PVR functionality, without upgrading any customer premises equipment.

However, existing carriage agreements are likely to require renegotiation before broadcast programming will be allowed for on-demand viewing, so MSOs must aggressively pursue content rights to achieve the full potential value of network-based PVR.

"Many Streams, Each To There"

Current on-demand solutions can be scaled to meet the MSO's streaming capacity needs for their digital subscribers. However, these solutions often suffer from inflexible routing that dates from the previous generation of transmission technology such as

DVB-ASI and integrated quadrature amplitude modulation (QAM) and upconversion. Since this transmission equipment had little or no switching and routing capability, and the capability was often not cost-effective when available, each streaming session had a fixed route to its destination. This meant that only a smaller subset of on-demand servers could stream content to a given customer's set-top box.

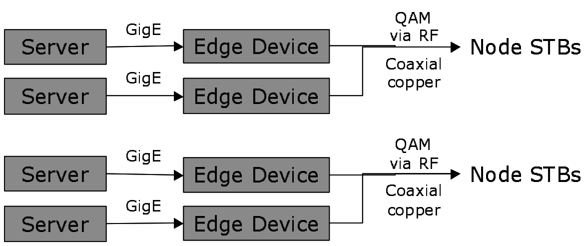


Figure 3. Fixed server-to-edge device routing.

For the MSO, these constraints meant that systems had to be designed for and sized to the peak demand expected at each hub, rather than the peak demand expected from the overall system. MSOs responded by defining system pricing in terms of cost per simultaneous stream, independent of service grouping or location. This pushed the cost of additional equipment to satisfy per-hub rather than overall requirements back on the equipment vendors, resulting in lower margins and profit from these sales.

This architecture is workable, but clearly not optimal for either MSOs or equipment vendors. MSOs must deploy larger systems that would otherwise be necessary, which impacts operational and maintenance costs, as well as complicating the issue of failure recovery. Equipment vendors must absorb costs imposed by sizing constraints at each hub, rather than at the overall system level. Performing asymmetric expansion of an on-demand system is further complicated by this routing inflexibility, since the expansions must again be performed at the hub level, not at the overall system level.

THE PROMISE OF TOMORROW

"Any Stream Anywhere"

"Any Stream Anywhere" is a phrase used to describe a system where any stream being sent from a streaming server can be directed to any set-top box. Looking from the other direction, this also means that any streaming server can satisfy a stream request from any particular set-top box.

A system with this property has many clear advantages. Since all streaming servers, not only a subset, can satisfy a node of set-top boxes, the total capacity provided by these servers can be sized against the demand of the overall system, instead of sizing each subset individually. This both eliminates unnecessary equipment, and also greatly simplifies the processes for installation and expansion. MSOs can set aside reserve streaming capacity to cover the entire system, rather than separate hubs or nodes.

A switched Gigabit Ethernet transmission framework can easily support the "Any Stream Anywhere" model, using the switching and routing functionality provided to direct traffic from any server to any edge device which transmits to a given set-top box. A conceptual diagram of "Any Stream Anywhere" for a system using a switched Gigabit Ethernet transmission framework is shown below.

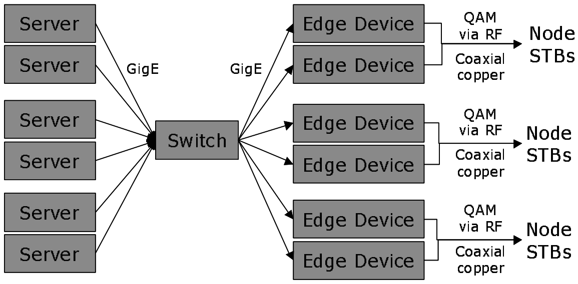


Figure 4. Gigabit Ethernet Any Stream Anywhere.

A basic implementation of "Any Stream Anywhere" using switched Gigabit Ethernet can take advantage of the fact that these systems are usually given dedicated network bandwidth. As long as the streaming servers can manage the available bandwidth properly, while taking into account the new switched infrastructure, few if any significant changes should be required to add the ability to support "Any Stream Anywhere" in an existing centralized Gigabit Ethernet system.

"Anything, Anywhere": Sharing Resources Between Multiple Services

A digital cable transmission system using Gigabit Ethernet has at least two distinct networks with resources to manage: the Gigabit Ethernet network used between streaming servers and edge devices, and the RF network between edge devices and set-top boxes. These systems also have an Ethernet network for command and control information, but that network is managed independently and falls outside the scope of this discussion.

RF Resource Sharing

Each RF frequency available has two separate but related resources to manage: the program numbers which can be individually tuned by set-top boxes, and the bandwidth which all programs using the same frequency must share.

A rudimentary level of RF resource sharing is easily achieved with a static partitioning of the available RF frequencies between the services sharing the RF network. This avoids most possibilities of conflict between services, but is clearly not optimal since resources unused by the assigned service are not available for reuse by other services.

An incremental improvement can be gained by changing the partitioning so that program numbers and their associated RF bandwidth can be assigned to services, instead of entire RF frequencies. However, the lack of mechanisms to guarantee quality of service (QoS) at this level makes it possible for an ill-behaved service to disrupt other services which share the same RF frequency.

Dynamic partitioning of these resources is clearly more efficient, but requires a resource management system to arbitrate requests. If the site in question uses the Scientific-Atlanta headend infrastructure, the Digital Network Control System (DNCS) is responsible for performing this function, using the DSM-CC protocol specified in the MPEG-2 standard. However, if the site uses the Motorola headend infrastructure, no such entity manages the RF resources. In this case, VOD system vendors have typically implemented their own internal management to handle resource sharing. Requests from other services for resource sharing can be accommodated by sending these requests to the VOD system for fulfillment.

At present, few services attempt to share RF resources with VOD systems, and the small number of involved parties makes solutions by private arrangement feasible. But as more potential services emerge, and providers begin to call for unified multiple vendor support, open standards should be adopted to define the interactions required for these services to share common resources.

Gigabit Ethernet Resource Sharing

Resource sharing for the Gigabit Ethernet network is simpler, thanks to both its inherent switching and routing functionality, and the suite of Internet protocols available for use. Like the RF network, Gigabit Ethernet networks have at least two separate resources to be managed: the addresses used to identify each device on the network, and the bandwidth available for data traffic.

Ethernet devices generally have unique Media Access Control (MAC) addresses, so only IP addresses generally need to be directly managed. The Address Resolution Protocol (ARP), part of the standard suite of Internet protocols, handles the matching of IP addresses with appropriate MAC addresses, and the Dynamic Host Configuration Protocol (DHCP) is often used to assign IP addresses to devices, whether on a static or dynamic basis.

Gigabit Ethernet network bandwidth can be statically allocated in a fashion similar to the RF bandwidth allocation described previously to provide a rudimentary level of sharing between services. Without any quality of service guarantees, an ill-behaved or misconfigured service can once again disrupt other services sharing the same network.

The effects of this disruption can be significantly worse for Gigabit Ethernet, since the vastly increased bandwidth available encourages a correspondingly higher number of sessions per link to share the network. But in this case, the Internet comes to the rescue, since mechanisms have been developed to ensure quality of service for IP and Ethernet traffic.

Gigabit Ethernet Quality of Service

There are several different methods, such as IP precedence, IP Type of Services (ToS), and Differentiated Services Code Point (DSCP), which can be used to specify which quality of service policy should be applied, if any, to IP traffic. Some of these methods overlap, and may conflict with one another if not configured and used carefully.

Fortunately, streaming servers are relatively immune to this problem, since the switch that receives their output can be configured to tag all incoming traffic on an input port with particular QoS settings. The streaming server is therefore not required to know how QoS will be implemented.

Edge devices are not so lucky, and so should be capable of receiving input with QoS tagging. QoS indications are not currently used to signal the relative priority of individual streams; therefore, vendors may note that it is safe for the edge device, as the last device in the chain, to ignore the QoS indications it receives.

Note that although lost data can sometimes be tolerated by other applications, within streaming video server output such losses are almost always clearly visible and objectionable to the customer. In light of this fact, best-effort queuing policies to enforce QoS are much more suitable for digital cable transmission than policies which result in lost traffic.

Gigabit Ethernet Bandwidth Reservation

The standard Internet protocol used to perform network bandwidth management is the Resource Reservation Protocol (RSVP). This protocol allows a receiver to establish a bandwidth reservation between itself and a specified source. Dynamic partitioning of network bandwidth between services can be readily accomplished with this protocol.

A crucial RSVP feature is its ability to accommodate portions of the network that are not RSVP-aware. This feature enables the gradual introduction of RSVP at sites with existing equipment that predates or otherwise does not support it. Although many existing Gigabit Ethernet switches support RSVP, and optical transport equipment is generally not required to do so, existing streaming servers and edge devices largely do not support RSVP. Even if direct support for RSVP is added, the encapsulation of these messages within UDP multicast packets may be required, as specified in Annex C of RFC 2205. This is due to various operating system and security issues regarding the use of raw sockets.

The fact that many Gigabit Ethernet switches provide RSVP support is again advantageous to streaming servers, since these switches may act as a sender proxy and hide the details of RSVP operation from the connected servers. In this situation, the switch maintains RSVP states, and generates required downstream "Path" messages in response to received streaming input.

Edge devices are, once again, not as lucky and may be required to directly support RSVP. The primary reason for this stems from the fact that unidirectional transport from streaming server to edge device is often employed to better utilize the available optical fiber. For RSVP, the receiver must initiate an upstream request for bandwidth reservation, but it is unclear what upstream path will be available to the edge device.

Bi-directional Edge Connectivity

Downstream video traffic requires much more bandwidth than upstream control traffic, which is why unidirectional transport from the streaming server to the RF edge is often implemented. However, having bi-directional connectivity at the edge would enable much simpler autodiscovery and autoconfiguration methods, and allow standard protocols used by the Internet such as ARP and RSVP to accomplish their tasks.

The establishment of bi-directional edge connectivity, with only unidirectional transport to the edge, requires a switch to exist between every edge device and the optical transport feeding it. This can become expensive, but an emerging new breed of equipment, combining switching and optical transport capability in the same device, may prove well-suited to this task.

As an alternative, devices may support methods such as the Unidirectional Link Routing (UDLR) protocols specified in RFC 3077 to logically create an upstream network path over a different connection, such as the command and control network.

Autodiscovery and Autoconfiguration

Automatic discovery and configuration methods are not strictly required for these systems to be deployed. However, for MSOs unfamiliar with the intricacies of these new systems, any automation that can help reduce the probability of misconfiguration, and also simplify system expansion, will clearly be of great value.

However, to implement autodiscovery and autoconfiguration, bi-directional connectivity and support for each device is required. Set-top autodiscovery schemes can use the upstream communications link provided by the RF network to perform these functions, but this makes open standardization difficult. Existing network equipment with full support for autodiscovery and autoconfiguration methods may require modifications to work with unidirectional links, as described above.

Complex Network Topologies

Up to this point, the discussion of "Anything, Anywhere" has been based on a simple centralized model, where streaming sources and switches are located at the master headend, and their output is distributed to hubs and nodes using optical transport. Although the simplicity of this model eases the discussion of issues which are not dependent on topology, real-world systems are much more complicated.

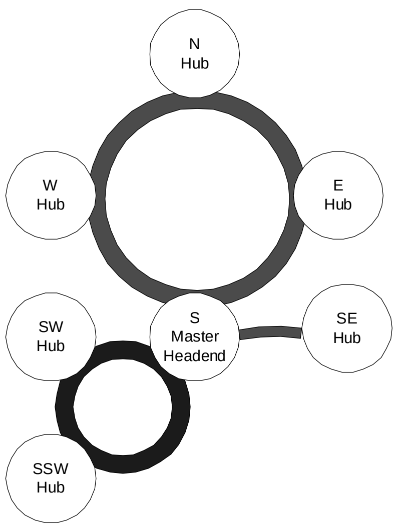


Figure 5. A slightly more complicated topology.

Since the cost of putting optical fiber in the ground is prohibitive, the topology of available fiber often dictates that of the services it carries. In other cases, the available space at headend locations may constrain the amount of equipment that can be installed. In addition, headends often also act as hubs to serve local customers. Lastly, redundant equipment is often used to provide failover capabilities. The net result of all this is that most real-world architectures diverge significantly from the ideal centralized model.

The multiple possible paths introduced by complex network topologies make routing and other management tasks much more difficult. However, complex topologies are generally chosen because they can be more flexible, and also more resilient if problems arise. Other possibilities which may drive MSOs to adopt more complex network topologies include the regionalization of functions such as broadcast feed generation, network-based PVR content ingestion, and reserve streaming capacity.

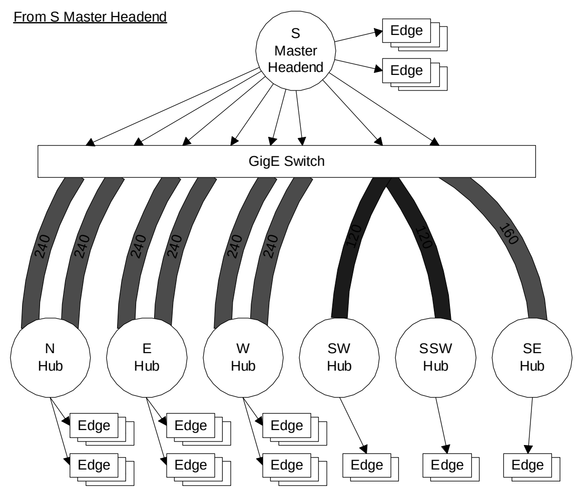


Figure 6. This example is centralized... almost!

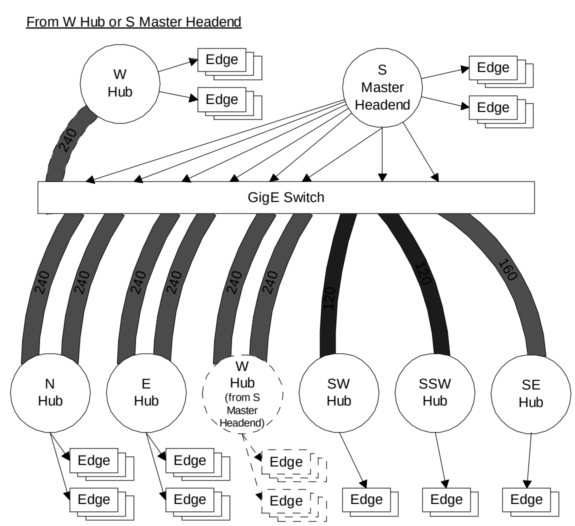


Figure 7. A server elsewhere on the main ring makes things more complicated.

Existing network devices and architectures have sturdy mechanisms available to handle failure detection and recovery, as well as other issues such as bandwidth reservation and quality of service. However, in some cases, the Internet solution does not quite fit the digital cable problem. For example, reserving bandwidth for a stream to a set-top box differs from the typical Internet case, due to the separation between the IP network and the RF network. A simpler problem thus becomes complicated in the digital cable space.

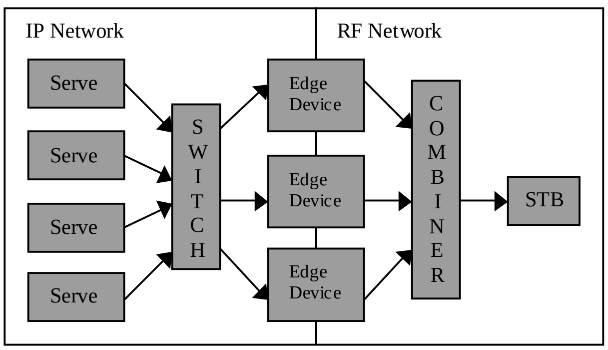


Figure 8. Gigabit Ethernet IP and RF networks.

The challenge here is to integrate network management functionality with the resource management of on-demand streaming systems. If this integration is performed at a high enough level, each piece can manage its own responsibilities and ask the others when external resources are required. But if the integration is performed at too low a level, the labor required to properly configure a system with complex topology may require that working autodiscovery and autoconfiguration methods be devised and implemented first.

High-Definition Video-on-Demand (HDVOD)

The advent of high-definition (HD) content for VOD systems is quickly approaching. In fact, HDVOD may have already arrived! HDVOD content differs from standard VOD content only in video resolution and bit rate, but support for these higher resolutions and bit rates can have ripple effects throughout an on-demand system. Care must be taken in both the underlying infrastructure and devices themselves to ensure that HDVOD content does not cause design limits to be exceeded.

RECOMMENDATIONS

Considerations for selecting and purchasing equipment for deployment:

Streaming servers should be able to fill a Gigabit Ethernet link so that switch ports and transport bandwidth are fully utilized.

Switches and/or routers should implement port queuing and QoS policy enforcement in a fashion compliant with streaming content requirements. Also, switches, routers, and transport equipment should only minimally modify the nature and timing of streaming media content. This is simplified by selecting equipment verified by vendors to interoperate correctly with other components, including both streaming server and edge device.

Edge devices should be upgradeable to support both variable bit rate (VBR) and high-definition (HD) streaming input. Devices with better buffering and dejittering capabilities are generally preferred over their competitors.

It must be decided up front whether asymmetrical deployment and expansion now merits the increased capital expenditure that it requires at initial rollout. Future cost projections for needed equipment will clearly play a significant role in this decision, as will the bargaining power brought by higher-volume and/or integrated purchases.

Considerations for designing or deploying a system:

The rollout and expansion of proven revenue sources such as on-demand services should not be delayed to wait for the promise of resource sharing with other services. The revenue to be gained now facilitates the expansion for these services later, and is a valuable hedge against the chance that other services may not end up as viable opportunities for additional revenue.

The network topology should not be complicated more than absolutely necessary, unless the benefits of doing so are tangible and compelling.

Component interactions should be kept at a high level when possible to accommodate differing implementation at lower layers. This avoids unnecessary problems that can arise from conflicting decisions made in the design and implementation of individual components.

The use of open standards should be encouraged for interoperability whenever feasible, but may not be required for existing or near-term deployments. This prevents unnecessary and unavoidable delays for acceptance and integration from impacting the timetables for these deployment.

CONCLUSION

Equipment vendors in this space hold an enviable position; they are poised in a market ready to explode with new business, and are positioned well to capitalize on that fact. The new features needed by MSOs are already being developed and deployed now, while open standards are being refined and proposed to allow smoother integration and interoperability for the future. The acceptance and adoption of these standards will allow vendors to focus on the development of next-generation features to drive the next wave of business.

For the MSO, this is an exciting time to be in the business, due to the convergence of several new technologies in an affordable fashion. This recent development has given rise to new features that not only bolster customer demand, but also provide new revenue-generating opportunities. The wide range of services available to customers has neven been more compelling. The ability to deliver "Anything, Anytime, Anywhere" is finally within reach, and customers are clamoring for the MSO to deliver this promise. The last remaining hurdle is to standardize rollout procedures to make them suitable for mass deployment, and then the MSO can let the good times roll.

REFERENCES

Chen, Michael, "Evolving the Network Architecture of the Future: The Gigabit Ethernet Story", SCTE Cable-Tec Expo Proceedings, May 2003.

Braden, R., Ed., Zhang, L., Berson, S., Herzog, S. and S. Jamin, "Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification", RFC 2205, September 1997.

Duros, E., Dabbous, W., Izumiyama, H., Fujii, N. and Y. Zhang, "A Link-Layer Tunneling Mechanism for Unidirectional Links", RFC 3077, March 2001.

Postel, J., "User Datagram Protocol", RFC 768, August 1980.

Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, March 1997.

Plummer, David C., “An Ethernet Address Resolution Protocol“, RFC 826, November 1982.

Nichols, K., Blake, S., Baker, F. and D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers“, RFC 2474, December 1998.

ISO/IEC International Standard 13818; "Generic coding of moving pictures and associated audio information", November 1994.

Gigabit Ethernet Alliance, “Gigabit Ethernet: Accelerating the Standard for Speed”, 1999. http://www.10gea.org/GEA-Accel1999_rev-wp.pdf

ARTIFCIAL INTELLIGENCE IN CABLE TV APPLICATIONS

Louis P. Slothouber, Aaron Ye
BIAP Systems, Inc.

Abstract

After many years and billions of dollars invested in a digital network infrastructure the cable TV industry finds itself unable to fully capitalize on that investment. Underpowered set-top-boxes, daunting integration issues, lack of standards, and huge capital costs hinder the roll out of new subscriber services at a time when competition from digital satellite providers is becoming acute. Fortunately, artificial intelligence (AI) technologies developed over the last forty years are directly applicable to many of the difficult technical problems faced by today's cable TV applications. Specifically, we describe how AI techniques can be applied to provide more personalized subscriber services, alleviate information overload, reduce backend server and human editorial costs, and to use available bandwidth more efficiently.

INTRODUCTION

The cable TV industry has invested huge sums of capital in recent years to upgrade both their networks and millions of consumer premises equipment (CPE) units from analog to digital. This has not only increased the quantity and quality of video that can be provided, but also placed a system controlled computing device, the digital set-top-box (STB), in every subscriber home. This high-speed, two-way network combined with a re motely programmable computer in the home provides the cable TV industry an opportunity to provide subscriber services that both Microsoft and the digital broadcast satellite (DBS) providers must envy.

Unfortunately, the evolutionary nature of the digital upgrade process has produced an architecture that is ill designed to support the multiple application services that are currently in development or on the drawing board. Initially, the STB was primarily intended to do little more than decode MPEG video. But the abundant bandwidth that the digital upgrades provided allowed for rapid growth in the number of video channels, far too many for the analog style scrolling guide to be practical. The need to overcome information overload caused by too many channels in a scrolling guide drove the development of the interactive program guide (IPG), a remote-control driven application that was squeezed into the confines of the STB. Today, a new set of business needs and opportunities drives the development of an array of new subscriber services, including video-on-demand (VOD), T-Commerce, information-on-demand (IOD), PC-like messaging, and games.

Clearly a STB that was originally intended to do little more than decode MPEG video is hard pressed to support all of these services. Further, the software architecture of the STB, which modified to support a single application (the IPG), typically requires costly integration to accommodate new applications and services. There are no standards for new services. As a result, most new services require costly servers to be deployed at the cable headend to perform much of the work, while the subscriber's STB acts as merely a dumb display device.

This current state of affairs is unfortunate. Because of the economics of the situation, the currently deployed STBs are likely to remain in the field for many years to come. Yet the technical limitations of both STBs and IPG applications impose significant integration and development challenges that impede the roll out of new, high-revenue generating services.

Surprisingly, there is an existing technology that could be applied to currently deployed STBs facilitating the full realization of the revenue potential enabled by a digital cable TV infrastructure. Even more surprising, this technology is neither proprietary nor a recent development. Rather it is the often misunderstood and under-utilized fruit of many decades of academic research: Artificial Intelligence (AI).

While popular understanding of AI revolves around jerky robots and giant, chess-playing super-brains, the true foundations of the science of AI consist of a cornucopia of techniques for performing complex tasks, such as user modeling, application of expert knowledge, dealing with uncertainty, etc. using limited computing resources. Many AI techniques are ideally suited to solving some of the most vexing problems in today's cable TV applications, and can often do so in the restricted computing environment of currently deployed digital STBs.

Carefully applied AI technology promises to revolutionize the subscriber services cable TV can offer, and to do so at a fraction of the cost of conventional, client-server systems.

EVOLUTION NOT REVOLUTION

The evolutionary development of today's cable TV infrastructures and applications is characterized by the *reactionary loop*, depicted in Fig. 1.

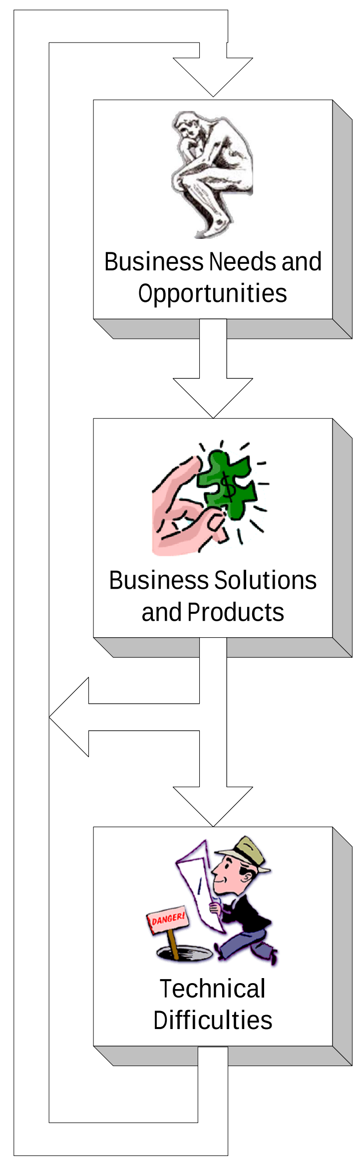


Figure 1. The Reactionary Loop.

Initially, a system operator or multi-system operator (MSO) identifies a business need or opportunity that solves a business problem (e.g., increases revenue, cuts costs, provides competitive advantage, etc.). The MSO then designs or purchases products or technical solutions that satisfy the need or capitalize on the opportunity. Finally, the product or solution reveals new opportunity, or technical difficulties, thus driving the next reactionary cycle.

Unfortunately, this mode of development produces ever more complex and costly, ad hoc solutions, as each cycle must accommodate the shortsighted decisions made on previous cycles. And the resulting solutions typically have little or no intercompatibility without costly software integration.

This state of affairs has led many MSOs to seek a “middleware” solution that provides a common foundation for future application and service development. Unfortunately such systems can never entirely overcome the inadequacies of a hardware and software architecture that has evolved via the reactionary loop. Middleware solutions isolate applications from the raw features available on the STB, forever limiting the role of such code to simple display tasks, and locking solutions into a client-server model. And while limitations imposed by underpowered STB hardware can be alleviated via backend processing, this only trades one problem for others, as this adds yet another layer of computation in an already tight STB environment, and server-centric backend solutions are notoriously expensive, and do not scale well for large numbers of subscribers.

THE CURRENT CYCLE

As of January, 2003 there are only a few markets in the US providing next generation services on modern STB hardware; the majority of cable systems, by far, are run on old technology.[1] To date, the majority of deployed cable TV STB software is not middleware based, but built around a single resident application, the IPG[2]. In some cases (e.g., DCT-2000) deployment of a new application requires direct integration with the IPG. In most systems out-of-band (OOB) bandwidth is a precious commodity, and little if any is available for use by third-party applications.

In this environment the cable TV industry faces a number of challenges, including high rates of digital churn, slowing digital penetration, shrinking subscriber bases, and the ever increasing competition of DBS. To continue to grow and survive the industry has entered the next cycle in the reactionary loop, as illustrated in Fig. 2. First, several critical business needs have been identified, including:

- Reduce digital churn rates, providing sufficient services to keep digital customers once they sign on.
- Compete with DBS by providing feature parity and significant feature differentiation, capitalizing on the two-way network.
- Create new features and services that can provide incremental revenue (e.g., pay-per-view).

---

[1] Primarily from the Motorola DCT-2000 and the Scientific Atlanta Explorer 2000 families.

[2] Mostly TV Guide/Gemstar or TV Gateway.

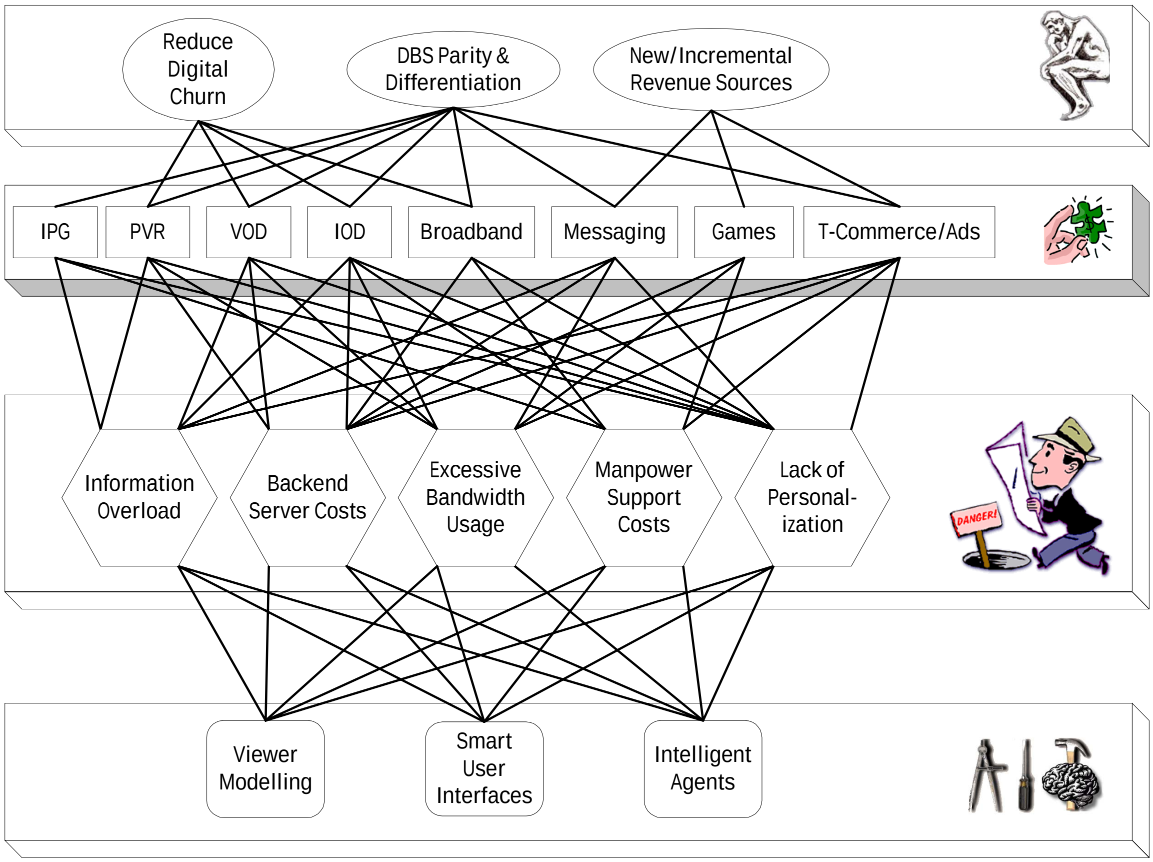


Figure 2. The current reactionary cycle, and the application of AI tools

To satisfy these business needs, a variety of new features and services have been defined and in some cases deployed. Among them are:

- **Video-on-demand (VOD)**. Providing TV programming of the subscriber's choice (initially PPV movies) anytime.
- **Information on Demand (IOD)**. Providing relevant news, weather and other information on TV anytime.
- **Messaging**. Providing instant messaging and e-mail services on TV.
- **Games**. Providing games and other interactive entertainments on TV.
- **T-Commerce/Advertising**. Providing custom adverstising and enabling online sales via TV.
- **Broadband Access**. Providing Internet service via cable modem.

However, we believe that such services, even if successful, introduce a number of new technical difficulties that must be overcome for these services to be adopted by subscribers, and to generate the revenue that justifies their implementation costs.

1. Backend server costs. As stated above, most such services are implemented via expensive hardware and software deployments at the headend. And while such solutions may work adequately when rolled out, they seldom scale well with the number of subscribers, and may fall victim to their own success.

2. Excessive bandwidth usage. It is the nature of such client-server solutions that information has to flow back and forth between client and server. Often this information (e.g., clicks on the remote control) must be transmitted out-of-band. But out-of-band bandwidth is a precious, contention-based commodity, and is often inadequate to the requirements of client-server solutions.

3. Information overload. Today's interactive program guides are useable for the several hundred channels available on digital cable. But how can they hope to cope with hundreds or even thousands of new programming titles made available by VOD. The subscriber will suffer information overload, hindering their ability to find and purchase VOD programming. Similar problems exist with the other new services that flood the subscriber with unprecedented quantities of information and numbers of choices.

4. Manpower support costs. Many new services, particularly IOD, games, T-Commerce and advertising require a significant number of people to provide content retrieval and editorial services.

5. Lack of Personalization. All of these services would be both more useable and more successful if they were personalized for each subscriber. Tailoring the information presented to the subscriber based on their interests and preferences provides a more efficient, and therefore more profitable, user experience. Research has also shown that systems that require personalization or learning are more sticky, retaining customers better than those without. [5]

Fortunately, AI can be applied to solve all five of these problems, and at a fraction of the cost of conventional, client-server systems.

APPLYING AI TO CABLE TV

Over the past forty years AI scholars have researched a variety of hard problems and developed a vast array of techniques, technologies, and tools for solving them. Serendipitously, most of this research was performed in an era when computing resources were scarce, so even a conventional cable TV STB is often adequate for their application. Table 1 lists several such technologies. [1,3,4]

Implicit in this discussion is that the judicious application of efficient AI technology allows much of the work that is currently performed by backend servers could be performed in a distributed fashion, directly on subscriber STBs, thus eliminating the need for costly backend servers. Such systems have been realized and are in operation today.[2]

AI Technology Class	Specific Techniques	Description	Cable TV Applications
Learning	Rote Learning, Inductive Learning, Neural Networks, Genetic Learning	Incremental improvement of task performance based on rote knowledge or examples.	Modeling the viewer based on previous actions to predict programming of interest for PVR or smart IPG.
Intelligent Agents	Information Retrieval, Knowledge Management, Commerce	Software that understands a complex task well enough to automate it, performing in a human role.	Automated content retrieval, reducing editorial staff for IOD, T-Commerce.
Expert Systems	Rule-based, Logic-based, Context-sensitive interfaces	Software that can apply expert domain knowledge to a problem.	Encoding knowledge about TV usage to provide smarter user interfaces.
Statistical Reasoning	Fuzzy Logic, Certainty Factors, Dempster-Shafer Theory, Baysian Networks	Reasoning with uncertain, incomplete, or noisy input	Widely applicable techniques useful in learning, agents, and expert systems.
Distributed Computing	Intelligent Agents, Edge-based computing, Peer-to-peer networking	Distributing pieces of a complex task among several distributed computers.	By pushing tasks down to the STB, obviates need for expensive servers.

Table 1. A sample of AI technologies applicable to cable TV applications.

It is beyond the scope of this paper to describe every possible application of AI technology to the cable TV industry. Instead, we will focus on three sample AI tools, each of which is directly applicable to many of the technical difficulties identified above. These tools are summarized in Table 2.

Intelligent Agents

Intelligent agents are small, active software components that understand a complex task sufficiently to assist a human in performing it, or to automate it entirely. One such task required of IOD and T-Commerce systems is the retrieval of content (text and pictures) to be displayed to the viewer, such as news, sports scores, stock quotes, current bargains, etc. Conventional solutions employ a staff of human editors who retrieve raw content from various network sources, revise it for display on TV. All such data is then broadcast out to the STBs for display. But it has been demonstrated that such tasks can be performed by intelligent agents running directly on the subscribers' STBs.

This approach has a number of advantages. First, it reduces or eliminates the need for an editorial staff. Second, it allows STBs to retrieve exactly the content that is appropriate for a given subscriber, and ignore everything else. This represents a significant reduction in the bandwidth requirements and often allows an on-demand request model to replace the broadcast model currently in use.

AI Tool	Applies to	Info. Overload	Server Costs	Bandwidth Usage	Manpower Costs	Lack of Personalization
Viewer Modeling	IPG, VOD, IOD, PVR, T-Commerce	Fewer programming or product choices	Runs on STB	No client-server network traffic	--	Customize views and content to target subscriber interests
Smart User Interfaces	All products and services	Automat or assist in obvious or repetitive tasks	Runs on STB	No client-server network traffic	--	Customize features & views based on subscriber abilities and context
Intelligent Agents	IOD, Broadband Portals, T-Commerce	Retrieve and display custom info., not everything	Runs on STB	No broadcast of generic info. Allows on-demand requests	Reduce or eliminate content editorial staff	Select agents that retrieve only desired content.

Table 2. Three AI tools and their applicability to the current reactionary cycle

Smart User Interfaces

Current cable TV user interfaces (e.g., VOD, IPG, IOD) tend to be static, providing the same set of capabilities to all subscribers at all times, regardless of the situation. The user interface, in this case, provides a means of operating a tool. However, significant AI research has been devoted to producing smarter interfaces, that operate more like an automated assistant. Rather than displaying a channel grid in numerical order, a smart IPG interface might order the channels based upon frequency of use. Or, by applying statistical reasoning and expert systems a smart IPG would "understand" the normal activities that a subscriber performs, either assisting or performing those activities automatically.

Viewer Modelling

Information overload is perhaps the most prevalent problem that results from the array of new subscriber services in the works today. Using an IPG to navigate through hundreds of channels is difficult enough. But add thousands of VOD titles and no subscriber is going to want to navigate through any static hierarchy to find a title worth paying for. However, by applying several statistical reasoning and learning techniques from AI, STB software would be able to monitor the programming viewed by a subscriber and construct a model of the tastes and preferences of that subscriber. Armed with this model, a smart IPG or VOD user interface could present the user with a small number of choices tailored to their preferences, and to provide a more dynamic navigation through the available titles based on subscriber tastes. By overcoming the information overload problem, and making it easier to find and buy programming of interest, such systems should allow VOD to realize its true revenue potential.

SUMMARY

The evolution of technical advances in the cable TV industry is the result of a reactionary cycle. As a result, current limitations of STB hardware and IPG software applications impose significant development challenges that impede the efficient roll out of new, high-revenue generating subscriber services.

However, judicious application of AI technologies, developed over the last 40 years, significantly enhance the range and quality of services that can be implemented via STB applications, and at a fraction of the cost of conventional client-server approaches.

AUTHORS

Louis Slothouber, Ph.D.,
Chief Scientist, BIAP Systems, Inc.,
lpslot@biap.com.

Aaron Ye, Ph.D.,
Chief Technical Offier, BIAP Systems,
aye@biap.com

REFERENCES

1. Barr, Avron, Edward Feigenbaum, eds., *Handbook of Artificial Intelligence*, vols. 1-3, HeurisTech Press, 1981.

2. BIAP Systems, Inc., "PiTV Overview", http://www.biap.com/pdf/Personalized Information Television.pdf, 2002.

3. González, Avelino, Douglas Dankel, *The Engineering of Knowledge-Based Systems: Theory and Practice*, Prentice-Hall, 1993.

4. Rich, Elaine, Kevin Knight, *Artificial Intelligence*, 2nd ed., McGraw-Hill, 1991.

5. Yahoo User Study, 2001.

BUILDING COMPETITIVE SYSTEMS: A PACKETCABLE PERSPECTIVE

Burcak Beser *
Juniper Networks

Abstract

PacketCable defines a network superstructure that overlays the two-way data-ready broadband cable DOCSIS 1.1 access network. PacketCable specifications define how PacketCable elements interact with each other and the protocols that are used between these elements. Since PacketCable certification/qualification only includes the protocol compliance of these elements, the certification/qualification does not suffice as the necessary means to provide a competitive service as provided by today's Public Switched Telephony Network.

This paper details some of the features that are necessary for competitive telephony service. Some of these features are not covered by PacketCable certification/qualification tests.

INTRODUCTION

The PacketCable project defines the protocols that are necessary for building competitive telephony.

Building a successful competitive telephony services over cable infrastructure requires the grade of service that is provided by Public Switched Telephony Network (PSTN) landline services to be met.

The grade of service provided by PSTN landline services has many dimensions that require different aspects of services to be engineered. Due to time and space limitations, this paper mainly focuses on the issue of perceived voice quality.

PUBLIC SWITCHED TELEPHONY NETWORK

Since the aim of the cable telephony is to match or exceed the service quality that is offered by the landline telephony systems, it is very important to understand the landline telephony of today.

Each phone call is carried as 64 kb/s bit stream, with bits flowing regardless of whether the sender talks or not. The speech signal is encoded at a sampling rate of 8 kHz, with eight bits per sample. The encoding is a simple lookup that maps sample amplitudes from 0 to 8159 to a 7-bit table entry, with a roughly logarithmic scale. There are two encoding schemes, A-law and u-law, where the former is found in European countries, while the latter is used in North America and Japan. The encoding is often also being referred to by its ITU Recommendation name, G.711 or called PCM coding [1]. The u-law encoding offers a signal-to-noise ratio of 39.3 dB for a full-range signal and a dynamic range of 48.4 dB. This is roughly equivalent to that of FM radio, except that the audio bandwidth is far lower.

The 8 kHz sampling period yields the basic clock period, 125 us, that is found throughout the digital telephone system, even when no voice is being transmitted. A number of these digital signals are then multiplexed into a single frame. For example, a T1 circuit consists of 24 voice channels, with one byte per channel. This packaging of channels into a single digital stream is called time-division multiplexing (TDM). A frame consists of these voice channels plus one or more synchronization bits.

Due to the TDM nature of the PSTN network the delay that is perceived by the users is mostly the propagation delay [2]. For North America the end-to-end delay worst case is calculated using maximum national distance of 6000 Km is found as 33 msec. In the same manner the long distance submarine fiber connection between San Francisco and Hong Kong can be found as 78 msec. It is important to note that even though the typical propagation delays are much less the impact of the PBX equipment, compression CODECs and multiplexers the given delays constitute good reference points.

GRADE OF SERVICE

The Grade of Service can be divided into two: the call connecting quality and perceived voice quality.

The call connecting quality depends on many factors including but not limited to call blocking, post-dial delay and accurate billing.

The perceived voice quality is generally more important than the call connecting quality, people tend to forget sporadic call connection problems, but when they have bad voice quality that they are paying for they tend to remember.

Perceived Call Quality

The voice quality in the PSTN networks was historically measured using 'mean opinion score'. The mean opinion score measures the subjective quality of a voice call. Historically the telephony providers invited people and used various call types (with delay, echo etc.) and recorded the results.

The MOS is a scale of 1-5 where the PSTN stands at 4.4 for local calls (perfect score). The score of national calls is generally above 4, which is considered as satisfactory. Anything below 4 may result with customer dissatisfaction with the service being received. Since the MOS is a subjective scale and requires subjective tests to be carried out which is not a good method of designing for a target. For design purposes ITU E-Model can be used [3].

The equation for the transmission rating factor *R* is:

$$R = Ro - Is - Id - Ie$$

Where,

- *Ro*, the basic signal-to-noise ratio based on send and receive loudness ratings and the circuit and room noise;
- *Is*, the sum of real-time or simultaneous speech transmission impairments, e.g., loudness levels, side tone and PCM quantizing distortion;
- *Id*, the sum of delayed impairments relative to the speech signal, e.g., talker echo, listener echo and absolute delay;
- *Ie*, the Equipment Impairment factor such as packet loss and CODEC loss, if CODEC being used is different than G.711.

CABLE TELEPHONY TARGETS

Cable Telephony competing with landline services aims to have E-Model R-value of 80, which corresponds to MOS scale of 4. The cellular services offer a much lower R-value than 80 but they have an advantage factor that adds to the total and improves the R-value. For new services, like satellite phones a correction value A is intoduced, to take into account the advantage of using a new service and to reflect acceptance of lower quality by users for such services. It is assumed that the Advantage Factor will be reduced over time as the service improves and the customers get used to the benefits of the new service. It is not recommended to include a non-zero Advantage

	Framing	Look Ahead	Coding	Decoder	Total Delay
G.711/10	10 msec	5 msec	1 msec	1 msec	17 msec
G.711/20	20 msec	5 msec	1 msec	1 msec	27 msec
G.729/10	10 msec	5 msec	10 msec	10 msec	35 msec
G.729/20	20 msec	5 msec	10 msec	10 msec	45 msec

Table 1 Delay introduced by various CODECs.

Factor for IP telephony because it is a replacement for existing services, rather than a completely new service.

CABLE TELEPHONY PERCEIVED CALL QUALITY

The perceived call quality of the cable telephony will be discussed using the e-model. As with the e-model the contributing factors of absolute speech delay and packet drop will be discussed.

Absolute Speech Delay

The absolute speech delay known as mouth to ear or one-way delay is a very important factor in the perceived voice quality. Figure 1 shows the drop in the voice quality in e-model with respect to absolute speech delay. To find the voice quality drop one has to calculate the absolute speech delay look up from the graph to find the delay related impairment in E-model [2].

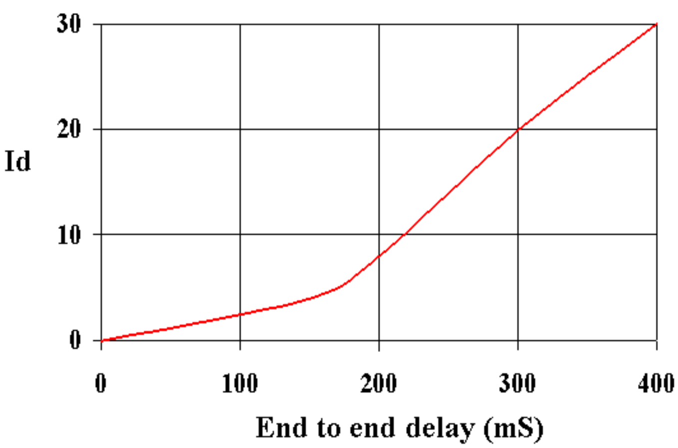


Figure 1 Impact of End-to-End delay using E-model

The absolute speech delay is contributed by many factors: Coding delay, Cable Access Delay, Network Side Delay and Jitter Buffer Delay.

Coding Delay

The MTA introduces a certain amount of delay due to framing, look ahead processing, and decoding. The delay introduced by the two PacketCable CODEC's are given in table 1.

Cable Access Delay

The delay introduced by the Cable Access network on the upstream direction depends on many assumptions. Some of the assumptions are listed below:

- The MTA's coding/de-coding clocks are slaved to CMTS DOCSIS master clock:

  If this assumption does not hold than the packets on the upstream-direction would experience a delay that is increasing/decreasing between 0 and 10 (the UGS interval) and then the same behavior will repeat as depicted in Figure 2.

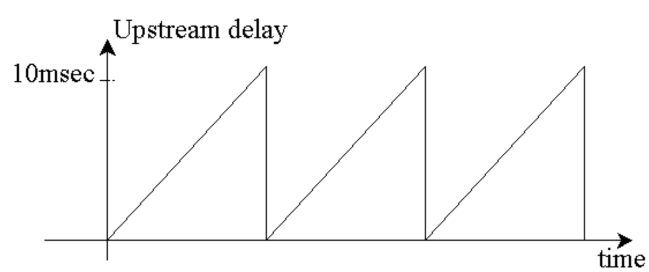


Figure 2 Upstream Delay when MTA's are not synchronized

- The MTA's framing interval will be aligned to UGS intervals:

  The first Dynamic Service Addition (DSA) message from the CMTS will include the time reference of the first UGS grant [4]. The MTA should align its framing interval such that the time between the end-of-framing and time to transmit upstream is minimized. If the time is not minimized or the framing is not aligned than there will be a constant value between 0 and 10 (UGS interval) upstream delay added.

If the MTA has implemented both the clock and framing synchronization than the delay on the upstream direction consists of:

Framing to transmit delay	<1 msec
Cable propagation delay	<0.8 msec
Cable receiving delay	<0.2
CMTS internal delay	<3 msec

A total of 5 msec is assumed for the upstream direction.

For the downstream direction cable delay consists of

CMTS internal delay	<3 msec
Interleaving/transmit delay	<1 msec
Cable Propagation delay	<0.8 msec
Reception to buffer delay	<0.2 msec

Making a downstream direction cable delay of 5 msec.

Network Side Delay

The network side delay depends on many factors such as the distance, the number of routers between end-points, the traffic and the connection technology. The end-to-end delay number for a national network is around 60-90 msec with jitter as large as 50 msec or more. When packet prioritization is being used the average delay remains about the same but the jitter reduces.

Jitter Buffer

The other delay factor is the jitter buffers on the de-coding section. The predicament with the jitter buffer is it is one of the items that is left for vendor differentiation.

The jitter buffer implementations can either be adaptive or static. On static jitter buffer implementations the buffer generally holds one or two packets, which spans one or two framing interval delays.

The goal of adaptive jitter buffer management is to remove the jitter while minimizing the amount of delay or incremental latency that is added to what's already been provided by the network. The adaptive buffer management schemes use interpacket arrival time variations, doing statistical analysis on it, then adapting the mean holding time of the packets or the jitter buffer length.

The problem with the static buffer management algorithms is that they tend to be conservative and assume the worst network cases. The MTA buffer management should be designed for end-to-end IP transport jitter values no matter where the call is connected. That means, if the worst-case jitter is 15 msec end-to-end then all the calls experience twice the jitter value delay of 30 msec.

The problem with the adaptive buffer management is twofold: First most of the adaptive buffer management statistical analyses schemes are not designed against the bursty nature of the network delay that is experienced today. Second, the buffer management is generally carried out during silence periods, which most probably does not coincide with the events that require jitter buffer changes.

The incorrect jitter buffer assignment has a two different impacts: When the jitter buffer is set to a value that is too low than the packets that arrive later then the buffered time will be dropped, and if the buffer is set to a too high value then the delay will be too high.

Depending on the network configuration and load the jitter on the VoIP packets may vary. In some cases some packets are so much delayed that setting the jitter buffer will result in an overall drop in voice quality.

Packet Loss

Almost all IP networks exhibit Packet Loss. Figure 2 shows the e-model impact of packet loss on G.711 CODEC, which is the only mandatory CODEC in PacketCable specifications [5]. As can be seen in figure below, the impact of packet loss is tremendous on the voice quality, if 1% of the voice packets are lost than the speech impairment due to packet loss is 25. Combining this with the starting point of the G.711 the quality of no-delay VoIP system with 1% packet loss results with a e-model rating of 70 which is equivalent to cellular phone quality.

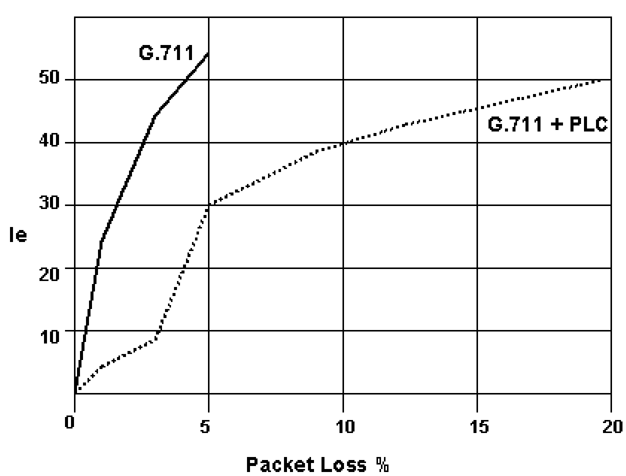


Figure 3 Impact of Packet Loss on E-model

The Packet Loss in IP networks can be attributed to several sources: queue overflow, synchronization, jitter buffer overflow/underflow, damaged packets. The impact of packet loss on the e-model impairments depends directly on the CODEC being used and whether Packet Loss Concealment (PLC) is implemented.

The impact of packet loss can be prevented if the packet loss concealment algorithms are implemented on the receiver side. As depicted in figure 3 the quality drops to 5 when PLC is implemented, which is one fifth of the non-PLC G.711 system. Unfortunately the PacketCable specifications do not mandate the implementation of PLC when using G.711 CODEC[1].

Packet Loss due to Queue Overflow

The queue overflow results when a certain interface receives more packets than it can send for a certain time duration. The solution for queue overflow is two fold: prioritization of packets, control of the queuing available to each priority.

The prioritization of packets is now a well-established standard in the IP world and is called Differentiated Services (DiffServ) [6,7]. The PacketCable standards already support the DiffServ packet marking. The DiffServ in general is the scheme that when a high priority packet is received that packet is being sent before the lower priority packets. Even when DiffServ marking is being used, it is still possible that at some funneling points, the router would receive a larger amount of high priority packets then it can handle and has to queue the high priority packets. In this case the packets either have to be queued for a long time or should be dropped. The issue with funneling is that it would cause all calls to be

[1] The term used for G.711 is 'RECOMMENDED' which is much weaker in terms of testing/certification viewpoint than use of the term 'MUST'.

impacted. If a certain link can handle at most 1000 calls and 1001st call is being connected all 1001 calls will be impacted not only the one call that is added.

The issue of packet dropping in DiffServ environment can be prevented by over-designing the network with no funneling points.

Packet Loss due to Jitter Buffer Overflow/Underflow

The jitter overflow/underflow is caused by the jitter experienced by the packets. As stated before every MTA has a jitter buffer, which can be static or dynamic. Since at any instant the jitter buffer is perceived as being static, only the static jitter buffer case will be considered.

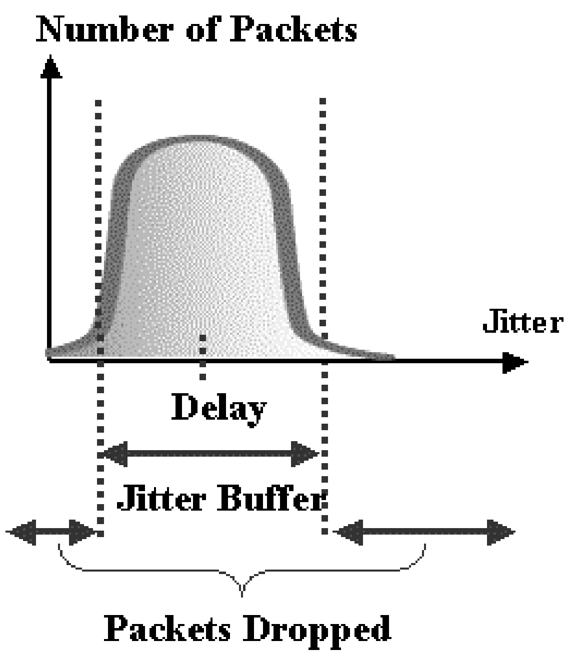


Figure 4 Packets Dropped and Jitter Buffer

Lets assume that the MTA has a jitter buffer of 20 milliseconds, which is actually a +/-10 milliseconds jitter buffer. What this means is that if a typical packet is received by time t0 then it will start to be played at time t0+20 milliseconds. Now assuming that the second packet is being received with a jitter of +10 milliseconds than it will be played without any jitter buffer induced delay. If a packet is early by 10 milliseconds (-10 millisecond jitter) than the packet will be delayed by 10 milliseconds. In short the jitter buffer regulates the delay variation on the packets and makes the impression that there is no jitter experienced by the packets only a constant delay of 20 milliseconds added.

The real issue comes to play when a packet experiences a jitter that is more than that of the jitter buffer. If a packet were earlier or later than the jitter buffer allocation, the MTA would drop the packet since it cannot handle the packet with in the jitter buffer. As shown in figure the jitter in a IP network can be depicted as a bell shape and the jitter buffer as a band within/engulfing this bell shape as shown in figure 4.

The jitter buffer may not be an issue if it can be set to a value that would always be able to accommodate the worst-case jitter in the network.

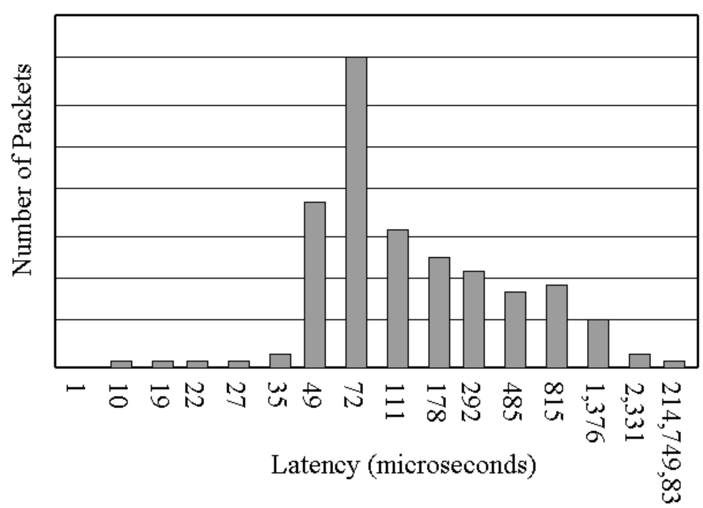


Figure 5 Example Core Router Jitter

Unfortunately this is not an easy task. Figure 5 shows a core router performance on a two-interface situation that no queuing is necessary. As can be noticed the spread of the delay is very wide [8]. If the delay from this one router is to be accommodated then a jitter buffer of 214 milliseconds will be required. The better approach is to set an acceptable packet-drop level and set the jitter buffer to this value around 2 milliseconds to get a packet drop around 1%.

Using similar analyses it can be shown that the best settings for IP network jitter buffer are around 1% packet drop. Any settings above the 1% packet drop will cause a larger delay introduced by the jitter buffer thereby dropping the voice quality faster than the voice quality improvements coming from the decreased packet drop rate. Any value above the 1% packet drop would result with a decreased voice quality due to the fact that the voice quality drop induced by the increased packet drop will not be compensated by the reduced delay.

Packet Loss due to Errored Packets

The packet drop due to errored packets is generally due to RF impairments in the cable plant. This is due to the fact that the modern IP transmission equipment provides reliable transmission with errored packets less than one in 10000.

In the downstream direction the packet-drop rate is in the order of 10^{-5} due to the fact that on the downstream the transmitter is more powerful and the bandwidth used for downstream transmission has better SNR (Signal to Noise) characteristics than the upstream direction [4].

On the upstream direction a provider has many possibilities that would impact the packet loss. Almost all of the countermeasures against the packet loss would have an impact on the perceived bandwidth on the upstream side. For example using a 1% packet loss a typical CMTS would be able to use a 2Mbps upstream channel whereas for a 10^{-5} packet drop rate a 612 Kbps will be achieved.

BUILDING AN EXAMPLE SYSTEM

Lets assume that the objective is to design a Voice over Cable system that would be competing against the PSTN landlines in North America.

Form the viewpoint of TDM based PSTN equipment the end-to-end performance of the PSTN network is impressive to say the least:

- Less than 1 in 1000 samples are dropped
- The end-to-end delay in North America is less than 35 msec.

When the cellular phones are taken as a base point these characteristics change as:

- As much as 3% packet loss.
- More than 200 msec of delay

The perceived quality of the cellular phone calls suffers from these characteristics.

CODEC Decision

Since the PSTN network will carry the voice as G.711, it is assumed that the G.711 will be used on the VoIP portion(s) of the networks. If this assumption is not valid than the initial coding loss and transcoding loss should be taken into account.

Absolute Speech Delay

Assuming that the aim is to be as close to PSTN landline services as possible the worst case has to be considered. The worst-case scenario for the Cable Telephony is that the call starts on Cable hops into PSTN and ends at Cable. As depicted in figure 1 the call starts in a user calling via a PacketCable certified MTA in Sunnyvale, CA to another user that has PacketCable certified MTA at the Providence, RI. The call goes to PSTN on the Gateway in San Francisco, CA, and then exits from PSTN on the Boston, MA. The delay on the PSTN segment between San Francisco and Boston is 30 msec.

Since the design is made for worst case, it will be assumed that the 10 msec packetization interval with G.711 coding will be used. Looking from the table 1 the coding and decoding delays of G.711 CODEC with 10 msec framing is 16 msec for coding and 1 msec for the de-coding, a total of 17 msec CODEC induced delay is found.

Since the coding is carried out on the originating MTA and de-coding on the ingress to PSTN, and coding on the egress from PSTN and de-coding on the terminating MTA, there two occurrences of coding-decoding in the Voice sections making 34 msec of CODEC delay.

The delay on the network side between the CMTS and the PSTN egress point is assumed to be 10 msec; when the cable access delay of 5 msec is added the total IP network side delay becomes 15 msec.

The Jitter Buffer delay is assumed to be 10 msec. The jitter buffer value of 10 msec should be sufficient for access to a local PSTN egress point but for the end-to-end VoIP call through the backbone this value may be too low, causing too much packet drop. Since the PacketCable does not have any means of setting the Jitter Buffer Size per call, the provider has to make a compromise between PSTN quality and end-to-end voice quality.

The total absolute speech delay consists of many pieces:

Origination	
Coding/Decoding	17 msec
Cable/Network Delay	15 msec
Jitter Buffer Delay	10 msec
PSTN end-to-end Delay	30 msec
Termination	
Coding/Decoding	17 msec
Cable/Network Delay	15 msec
Jitter Buffer Delay	10 msec
	+
Absolute Speech Delay	114 msec

Using figure 1 to resolve the E-Model quality drop the drop can be found as 4.

Packet Loss

The amount of packet loss contributed can be partitioned as:

Router queuing	0.1%
Jitter buffer	1%
Cable Access (RF)	0.01%

Making 1.11% packet loss in origin and 1% packet loss in destination. Looking at the impact of 2.22% packet loss on figure 3, the e-model quality drop can be found as 7.66.

E-Model Result

When the G.711 coding (just sampling the voice with 8000 times a second) is being used, the base for voice quality would start from e-model score of 94.2. Calculating the drop due to delay (4) and packet drop (7.66) would result with an end-to-end quality of the 82.6 which is barely above the desired limit of 80.

CONCLUDING REMARKS

Even though at first glance it looks like that the desired of score 80 can be achieved, some points are worth mentioning:

- The calculation is for the worst case and for most of the geographical locations the score would be higher.
- The packet loss of 1% would only be seen on cases where severe jitter is observed.

- The overall experience with respect to PSTN landline services would be lower due to the fact that the introduced absolute delay is higher[1].

- When connected to high delay endpoints such as cellular phones, PBX's or international calls with long delays, the call quality may drop to an unacceptable level.

- The use of CODEC's that provide better bandwidth utilization will cause the voice quality to drop further.

- The use of a bigger Jitter Buffer to accommodate the connections to other end-points would cause voice quality to drop further.

- Any additional packet loss would cause the voice quality to drop further.

REFERENCES

[1] CCITT Recommendation G.711 (11/88), *Pulse code Modulation (PCM) of voice frequencies.*

[2] ITU-T Recommendation G.114 (05/00), *General Recommendations on the transmission quality for an entire international telephone connection.*

[3] ITU-T Recommendation G.107 (05/00), *The E-Model, A Computational Model for use in Transmission Planning.*

[4] CableLabs SP-RFIv1.1-I09-020830 (08/02), *Data-Over-Cable Service Interface Specifications Radio Frequency Interface Specification*

[5] ITU-T Recommendation G.113 (02/96), *Transmission impairments.*

[6] IETF Request For Comments RFC2475 (12/98), *An Architecture for Differentiated Services*

[7] IETF Request For Comments RFC2474 (12/98), *Definition of the Differentiated Services Field in the IPv4 and IPv6 Headers*

[8] Light Reading (7/01); The Internet Core Routing Test: Complete Results; *http://www.lightreading.com/document.asp?site=testing&doc_id=6411*

[1] Due to added delay most of the long distance calls will sound like high quality international calls.

BUSINESS AND PLEASURE: MIXED TRAFFIC ISSUES DRIVE NETWORK EVOLUTION

Robert L. Howald, Ph.D.
Motorola Broadband

Abstract

The term "triple play" originally was meant simply to convey the convergence of video, voice, and data in the network. It once encompassed everything that was important to know about network functionality. However, over the years the lines have blurred and the elements of the triple play have been further fine-tuned and splintered into a variety of different items. All three parts of the triple play are multi-faceted, and all contain important variables necessary for proper bandwidth and traffic management. Video has multiple, possibly interrelated, faces: analog, digital, narrowcast, on-demand, IP, and HDTV. Data, which once essentially meant DOCSIS 1.0, now includes the ability to support tiered data services that include best effort Internet traffic, like that offered via DOCSIS 1.0, as well as guaranteed and mission critical business services. Even simple residential Internet access is becoming a more complicated offering. With VoIP still in the wings, ideas such as online gaming have leap-frogged into play as part of the residential data mix. With data just as with video, possibilities abound that involve bandwidth and traffic management in both the access, backbone, and interconnecting points in the network. Successful deployment of tiered, prioritized, and guaranteed services include understanding aspects of the access network, including higher versions of DOCSIS, as well as non-DOCSIS solutions and technologies behind the HFC access network. Proper treatment of data services to and from the access network is a critical component of bandwidth management when considering architecture design options. Finally, voice circuits and IP voice also have a role in the redefining of the meaning of triple play.

This paper will analyze and characterize the traffic dynamics of the various service components above. Aggregation of these services in cases consistent with likely architectural scenarios will be discussed. Architecture and bandwidth conclusions will be drawn that align with the service and traffic mixes currently being offered. Finally, offerings such as gaming, security, and medical applications are some of the ideas among many potential services that have been mentioned recently. Their significance is magnified by the amount of highly interactive real-time voice and video needed to support them. The implications of such new service offerings will be discussed.

INTRODUCTION

At last year's NCTA show, a paper planted a stake in the quicksand [8], describing, as is self-described by the title, "A 10-yr Residential Bandwidth Demand Forecast and Implications for Delivery Networks." The presentation generated quite a few comments and questions, as any discussion related to predictions of bandwidth consumption might. In particular, the paper did not focus on perceived needs or the conventional "Field of Dreams" theory of bandwidth growth – that is, "if you build it, they will come." Instead, it was grounded in demand expectations based on market research, trend studies, and conversations with various individuals across the industry whom the information was shared with. The approach was to identify current and coming services that could be reasonably anticipated, evaluate and predict behavior, and aggregate the results. Obviously, the analysis was not done as an academic exercise, but as a tool that can be used to plan a business towards expected growth areas, and to engage operators in discussion that help them plan their networking needs.

As it stands today (now a third year into assessing the predictions), predicted behavior has deviated in a couple of areas, but none in earth-shattering ways. Only one service at this point is ramping more slowly than anticipated (VoIP). As a quick data point to encourage active minds, the study currently expects that, in 2010, there will be about 10x the bit rate demand on the forward path as there is in 2003, and about 25x in the return over the same period.

As valuable as this paper is to the company as a benchmark that is updated regularly, its original intent is only a piece of the puzzle needed to get a complete snapshot of the bandwidth and architectural evolution landscape. In particular, the results presented describe forecasts for the North American, residential market, and only the implications on the HFC access portion of the network. While there is certainly not a lot of extensive network infrastructure activity going on in the current slowdown, there is significant attention being given to services and equipment that have equivalent, if not more dynamic, impact on metro interconnect or backbone portions of the network. Examples include the growth in video-on-demand(VOD), the emphasis on supporting commercial services, and enhanced data aggregation and backhaul platforms. And, clearly, the paper described above, while concerning itself with evolution of the HFC plant aspects of service growth, implies impacts beyond HFC access. To complete the picture that enables the steadily predicted bandwidth growth requires a peek into what is going on outside the residential portion of the network, as well as at and behind the hubs that feed the distribution network. To do this, first we need to understand the relevant traffic engineering problems as well as access bandwidth problems for effective end-to-end system design.

This paper will introduce some of the concepts associated with the traffic engineering side of the problem. The problem at hand can actually be summarized quite easily. Never before has one network been asked to support so much content variety with such a wide range of quality of service (QoS) objectives. With this being the assumed case for growing cable operators, the following questions are being explored today:

1) What are the traffic implications of this multi-service, multiple goal situation?
2) What are the resulting architectural implications?

The fact that cable systems are able to encounter this type of problem at all and learn the important issues makes a strong statement for its competitive readiness in the larger picture of broadband providers. This paper will discuss scenarios that can be used to evaluate question one. There are as many answer to question two as there are opinions on optimal architectures. We will discuss common ones and general themes to be understood.

SOME MATHMATICAL TRAFFIC CONCEPTS

While widespread DOCSIS deployment brought data traffic management to the attention of the industry, the idea of understanding traffic characteristics and the effect on performance is not a new to cable. Archives at Motorola contain traffic studies aimed at understanding the response time of early settop IPPV request traffic, which used a basic ALOHA protocol. ALOHA is essentially a free-for-all that allows a user to send a message whenever ready and, basically, take their chances that no one else is doing so at the same time. If an acknowledgment is received prior to a time-out waiting for it, the user knows the message got through. The study goal was to understand how the settop loading and acknowledgement scheme effected the response time to a request from the user, determine the re-send likelihood, and understand the system breaking point. Implementation details were derived from this study. Through traffic modeling, the analysis was able to show that about 40% more settop returns could be accommodated at the HE equipment if the acknowledgments sent to the settop were modified in the way they were originally designed to be delivered. Clearly, equipment cost savings were directly obtained in this simple case.

As a second example, prior to full two-way activation of cable plants, Motorola had deployed an early SurfBoard® cable modem with a telephony return path. Traffic studies were commissioned to understand this drastic network asymmetry, the impact of PC hardware, and the effect of TCP/IP implementation on the PC. The analysis characterized how the telephony modems of the time (14.4 kbps and 28.8 kbps) compromised downloads that had much higher raw throughput capability. Performance of FTP transfer of large files was compared against symmetrical 10 Mbps and 100 Mbps point-to-point Ethernet to understand the user experience relative to, for example, the office environment. This data was use to support configuration guidelines for the product.

In the 1980's, Ethernet itself, standardized around a carrier-sense, collision detect multiple access (CSMA/CD MAC) protocol came under traffic analysis scrutiny. A widely referenced throughput analysis and testing study was performed as LAN technology began to explode during that period of time [3].

Telephone Network Simplicity

The purpose of traffic modeling is simple. By developing proper statistical models for data in the network, it is possible to predict pipe size requirements, bottlenecks, performance, and equipment requirements. Historically, there are two paradigms – telephone networks and data networks. The phone network is traffic engineered to minute decimal place (the "five nines") precision. The unit of Erlangs is used to describe voice traffic volume. The voice traffic arrivals are characterized statistically as a process with call arrivals that exhibit a Poisson characteristic. This information is used with the well-known Erlang formula to determine trunking capacity necessary to ensure that circuit availability can be guaranteed to the high level described above. Traffic engineering is possible with precision because of the well-understood nature of voice traffic with many years of historical precedent, and the single-service nature of that system at inception.

Data traffic, on the other hand, has not historically been heavily traffic engineered. Providing plenty of excess bandwidth has been the protection against performance degradation due to congestion, and there are still advocates of cheap bandwidth and less complexity as the way to continue. Others argue that, besides the inherent cost of higher performance equipment associated with under utilizing the network, flows are likely to encounter some bottleneck in an end-to-end system, particularly as the routes grow longer and more complex. Delivering repeatable QoS for high-performance services is not practical through pure bandwidth means in such cases.

We have mentioned the idea of Erlangs. Telephone system trunking curves – how many circuits must be deployed as a function of subscribers to assure a given blocking criteria – are available in many classic textbooks and papers. The results provide remarkably straightforward formulas for telephone network design – a formula that depends only on voice traffic offered (arrival of calls and duration) and the statistical assumption of a Poisson process for call arrivals. What statistical characterization can be used for other services, such as data? Are the answers as conveniently simple? Unfortunately, this answer is no.

Long-Range Dependence (LRD)

The finding that data traffic has a *self-similar* characteristic was one of the major traffic modeling discoveries to date for this relatively young discipline. Self-similarity – also called fractal or long-range dependent behavior – implies that, *regardless of time scale*, the traffic pattern has the same basic structure. When we say "traffic", we are talking about the baseband data volume and trends observed at the output of a CMTS or switch serving MPEG VOD streams, for example. This was an unusual finding, in that it indicates that there is correlation across much wider time scales than previously thought, and the assumption that smoothing occurs when observed over long periods was proven inaccurate. Another surprising way to envision this characteristic is to think in terms of our basic understanding that data traffic is bursty, which we usually associate with short time dependence in our minds. However, self-similar traffic indicates that long bursts separated by long time intervals are characteristic of the traffic as well. Intuitively, we would have expected the wider time scale to smooth out the peaks and valley around a mean.

The seminal paper showing self-similarity at work was based on an Ethernet analysis, but because of the astounding nature of the discovery, others were inspired to look closely at their own assumptions. Subsequent findings included self-similar properties of ATM traffic, metro area traffic (MAN), wide area network traffic (WAN), and also for multimedia traffic, such as compressed digital video streams and Web traffic.

The unearthing of self-similarity created a camp of network theorists that felt that the book on traffic theory now had to be re-written. Traditional models generally focused on Markovian behavior, which relies on limited memory of prior traffic – in other words, correlation is lost over time, and smoothing out occurs as the time scale is broadened. The stock market is a good example of this expected smoothing, although studying these curves are probably best avoided at this juncture. The impact of this correlation lasting over broad time periods has implications for policing, scheduling, congestion control, and statistical multiplexing gain.

The surprising finding naturally led researchers to search for the reasons for it. The cause of self-similar behavior was found to be associated with the fact that the distribution of transmission content is *heavily-tailed*. That is, the tails of the probability density function do not decay rapidly. What this means is that, rather than seeing the likelihood of the size of a transmission flow occurrence decreasing exponentially as the flow size increases, this drop-off in likelihood is not so drastic. There is a very wide variation in the size of packet flows that throws off traditional statistical models – large files, MP3's, JPEGs, database activity. In fact, it has been shown that such heavily-tailed characteristics are a *sufficient* condition for self-similarity. As important is to recognize what self-similarity is *not* caused by. The breadth of examples indicates that self-similarity is not associated with the delivery format generated to carry the information – i.e., it is not a protocol artifact.

Now, obviously, cable systems deploy equipment for carrying multimedia traffic. In particular, compressed video streams are on today's cable transport networks, with today's most relevant example in terms of equipment growth and network design being video-on-demand (VOD). Based on the above, the traffic characteristics will be the same whether the video delivery is MPEG over IP – such as GbE-based transport – or the other way around. And, of course, cable companies are interested in moving data around in the form of Internet traffic from CMTS's to ISP points of presence, and data from business services, both Internet directed or otherwise.

Summarizing, then, understanding the role of self-similar traffic patterns is valuable for the applications above in designing the HE to hub or hub-to-hub interconnects. As networks become more integrated, the value of understanding traffic increases as the aggregation pushes bit rates higher, making efficient use of resources yet more important. As movement of different types of traffic becomes integrated, there is the further need to ensure the QoS support for each. Providers must therefore understand the implications of traffic characteristics and the distribution of QoS needs of each.

<u>M/Pareto Model</u>

While we have explained and described a fundamental and surprising trait of many traffic types relevant to cable, making use of this model for statistical calculation requires fitting this knowledge into a distribution. The characteristic described has been shown to be a result of an aggregate of bursts of widely varying sizes. A model based on randomly arriving bursts with a heavily-tailed distribution is therefore called for. A Pareto distribution, commonly described in statistics texts, is combined with a Poisson arrival rate of overlapping bursts to create a mathematical realization of the situation. More specifically, data traffic is assumed to be bursts with a Poisson distribution and associated arrival rate, where each burst is of duration described by a Pareto distribution.

The M/Pareto model has several variables associated with it, including the Poisson arrival rate information. This portion of the model has been shown to be important to accurately curve fitting real traffic to it [2]. The essential "real" traffic property captured by varying the Poisson parameters is the amount of traffic being multiplexed onto a pipe for characterization. This approach is a valuable step to a traffic model comprised of an aggregate of multiple sources of independent information. The aggregation of traffic is not significant enough to use mathematical assumptions of Gaussian behavior driven by the central limit theorem. Models based on long-range dependence provide network designers with a tool for developing architectures and equipment requirements that support the aggregated traffic. This is important to capture, as issues associated with self-similarity drive network changes in queuing and congestion control mechanisms then Markov-based assumptions would imply.

<u>Gaussian Behavior</u>

What is occurring on the Internet and to a similar extent in breadth on HFC networks is the aggregation of more traffic and more traffic types from independent sources. It is not difficult to envision the challenge this growth entails; yet, this very growth and the evolution of integrated networks is potentially a blessing to the traffic modeler. The central limit theorem provides a fundamental statistical underpinning for what the nature of the traffic over time could evolve to – Gaussian behavior.

The central limit theorem is the basis for many natural phenomenons that exhibit Gaussian behavior, and, in the case of aggregated traffic, it becomes asymptotically so when many independent contributors - under some minor, but important, caveats - are aggregated. The convenience of this is that Gaussian statistics are very well-studied and understood, and if the traffic statistics can be assumed Gaussian, then many simplifications can occur and probabilities of occurrence characterized. Multiplexing gain can be predicted under Gaussian assumptions, and pipes designed efficiently for some pre-selected level of congestion avoidance. This can be used to support a desired set of policing, shaping, scheduling, and queuing mechanisms. While rapid traffic growth makes network evolution difficult, the bandwidth explosion, in general, is good for business, and the handling of traffic from a bandwidth boom potentially makes it more readily predictable.

SERVICE SET

DOCSIS

Since the wide acceptance and deployment of DOCSIS, all operators and vendors have interest in traffic characteristics of essentially this same basic system. As a result, there have been many articles on configuration of the CMTS and guidelines for a DOCSIS-based system setup. The paper described previously [8] suggests a doubling of DOCSIS return path traffic each year, a result that is a combination of take rates, modulation profiles, and usage of the medium by subscribers. This variable in the paper is dedicated to residential cable modems – i.e. Internet users at home. This doubling effect is corroborated by other, more general Internet traffic studies that suggest a “Moore’s Law” for data traffic [7]. This analysis notes that this trend has been pretty reliable except for a period in 1995-96 where there was a burst of greater growth attributed to simplified web browser breakthroughs that led to mass acceptance, and the subsequent changes made by online providers to graphically rich interfaces. Further traffic related information from trend studies indicate that access to broadband via DSL or cable modem results in a user increasing their time online by 50-100%, and that the bytes consumed per month increase 5x to 10x as well.

A very informative paper based on a project at CableLabs and also presented at last year’s conference [12] offered a first real comprehensive glimpse into DOCSIS traffic. Summarizing some of the key findings:

- Daily activity is a slow build throughout the day, with “busy hours” between 8 pm – 12 am (peak), and a subsequent rapid drop-off until beginning again at 5 am

- Traffic is seasonal – following school holidays and vacations

- Traffic asymmetry decreases from 3:1 to about 1.5:1 as familiarity and capabilities set in

- DOCSIS 1.1 enhancements to support voice traffic result in a 15% efficiency improvement over DOCSIS 1.0

The seasonal phenomenon represents the dominance of traffic by a younger generation of user. This particular phenomenon should become less pronounced over time as these kids become tomorrow’s adults, although a generally heavier level of usage by the academic community may linger.

DOCSIS 1.1 provides the ability to support VoIP traffic. It does so by supporting multiple classes of service (CoS), whereas DOCSIS 1.0 supports only one – best effort. DOCSIS 1.1 also allows packet fragmentation to ensure that latency-sensitive voice traffic is not bogged down behind large "best effort" data packets. From a traffic generating standpoint, DOCSIS 1.1 also implements pre-equalization at the CM side, a physical layer technique similar to pre-distortion, but for bits. This feature permits more practical use of the 16-QAM mode, which doubles the bit rate compared to a QPSK channel of the same symbol rate. In other words, the 160 ksps mode, which results in 320 kbps for QPSK, provides 640 kbps in 16-QAM mode. The result at the hub or HE is more bits-per-second pouring out of a CMTS spigot.

DOCSIS 2.0 also speaks foremost to raw throughput enhancements, It provides for a dual, selectable, medium access control, or MAC (S-CDMA or A-TDMA), and an enhanced modulation profile capable of 64-QAM at twice the previous maximum symbol rate. Raw capability is now about 30 Mbps. Built-in as well are enhanced interference mitigation techniques for narrowband and burst interference make use of the newest modes realistic, and use of lower return channels possible, creating more effective bandwidth. The CMTS spigot therefore just got wider, or the pipe became more fully utilized.

In summary, with DOCSIS we can expect rapid, raw, bits-per-second growth, more efficient bandwidth consumption in access, self-similarity in backhaul, and both best-effort and class-of-service (CoS) mapping.

Peer-to-Peer (P2P)

Although the proliferation of peer-to-peer communication still generally falls under a residential data (DOCSIS) discussion, this phenomenon is significant enough to warrant special mention. It is, of course, certainly the case that broadband access has, in fact, *enabled* P2P traffic to become as significant as it has, freeing downloaders from the limitations of dial-up speeds to deliver multi-megabit files. Peer-to-peer traffic – pioneered by music file sharing through Napster, but subsequently followed up by similar services (Kazaa, Morpheus) – raises a significant flag to operators of broadband networks who observe and/or police their traffic patterns.

The impact of heavy P2P traffic on cable modem systems offering no byte count or rate limits is twofold. The raw bits-per-second load sees a "bias" around which the web browsing peaks and valleys vary. The effect is to have a constant offset or mean value associated with the streaming file content much like any constant bit rate (CBR) application. The difference in the modern case is the high bit rates associated with the CBR-like traffic are rates that are considered high speed. Of course, with no rate capping or policies in place to limit this type of traffic, a very small number of users can essentially dominate the throughput of the link. Enough heavy P2P users or enough sharing of single return channels among users can therefore create a congestion scenario, as the demand for transmission time slots upstream outstrips supply. Simulations [12] support this effect. A single users acting as a source of MP3's, when placed among a dozen or so other users, managed to consume up to half of the return capacity over significant period of observed time. This creates a clear forward-looking argument for tiered service offering based upon either rate or volume limitations with the necessary prioritization and policing schemes to enforce the tiered structure.

To summarize, we can therefore add to our prior discussion of DOCSIS the need or objective of supporting tiered services.

Enterprise Traffic

The play for commercial services can be attacked primarily in two ways – DOCSIS-based and fiber-based. Which solution is determined by service needs at the business, which basically boils down to the size of the enterprise. The key features that DOCSIS 1.1 provides that make it a reasonable solution for a subset of the business market, in addition to cost, are the support of voice, enhancements that offer multiple service classes, enhanced security, and, finally, a more realistic opportunity to achieve 10 Mbps type of performance due to physical layer improvements that add pre-equalization. Of course, 10 Mbps has a nice ring to it in light of comparison to 10Base-T LAN environments. Finally, DOCSIS 2.0 encompasses the key features of DOCSIS 1.1, but also offers the 3x increased capacity return due to symbol rate and modulation improvements. In addition, the protocol advancements inherent in A-TDMA signaling and S-CDMA are designed to expose more return bandwidth to the operator for high-speed services that previously had been unusable for this type of traffic.

Based on the above, we can summarize by saying that business services over DOCSIS 1.1 represents for the operator a need for QoS tiers, and service level guarantees, similar to previously mentioned DOCSIS needs for peer-to-peer traffic.

Fiber-based solutions can deliver higher levels of service, corresponding to larger businesses or business campuses. Of course, the target here is to provide cost-competitive voice and data services with the same level of service experience to the end customer – data rates, security, reliability. While residential data traffic has grown rapidly as previously described, business data traffic has taken a more modest trajectory, and business voice traffic is essentially flat. Thus, while the per-subscriber business needs are higher, the growth in this data sector is more gradual, meaning the pipes assigned to support a fiber-to-the-business application may have longer legs than expected. This is somewhat intuitive, in that, while residential users find new and bandwidth consuming ways to exchange and download content, the shift in business transaction content has not been this dramatic. In terms of security, virtual private networks (VPNs) translate to both security and bandwidth guarantee issues. In terms of reliability, resiliency and 50 msec recovery characterize common “carrier class” attributes expected by business customers.

Fiber-based systems that segment spatially or via wavelength – the only non-DOCSIS, HFC-centric approaches today, have no aggregation issues in the access network. As the data hits the first aggregation point in a hub, it is at this point that it may have to be managed, prioritized, and scheduled alongside other traffic requiring bandwidth. If both business voice and data exist, an architecture that guarantees bandwidth for voice and queues data packets may be necessary. Architecture decisions are made with these types of network crossroads in mind – the service mix at these points could consist of VOD content, CMTS activity that could include voice and tiered data, business voice and data, even broadcast digital video content.

In summary, then, overall higher symmetrical bandwidth and quality of service guarantees are important to this market, and traffic growth to plan for may be less dynamic.

Networked Gaming

The average age of a “gamer” is on the rise. The Gen-X and Gen-Y demographic that grew up with the phenomenon of Playstation, Nintendo, and now X-Box are now growing up themselves – if perhaps in age only. They are bringing their bad habits with them, including their passion for gaming. At last year’s Western Show, not much was well-attended. But, one of the best attended sessions dealt with the gaming phenomenon and the impact on broadband.

Intuitively, what would we expect gaming needs to be? Certainly it is real-time interactivity, the magnitude of which needs quantification relative to the well understood baseline needs of voice traffic. Latency is of key importance, as thumb actions must be translated into game states and communicated rapidly to other players. As important yet is probably jitter variation among peers, so that the game can be fair to everyone, and no one has a built-in edge. Studies have shown that delays that are not noticeable to voice traffic users are quite noticeable to gamers, and, indeed, can affect game outcome.

Again, intuitively, the nature of gaming activity would not be expected to have the statistical look of web browsing or streaming media. Recent work has analyzed traffic distribution of the popular "Quake" application [4], observing both packet arrivals and sizes for clients and of the game server. The results of this study suggested that an Extreme distribution is a good fit for packet arrival of both servers and clients in most cases, as well as packet sizes of server traffic, where the server is the point from which game states are updated and broadcast to the clients.

The parameters of the distribution that define the exact shape of the broader family are functions of processing speed distributed among the servers and clients – another effect that would seem intuitive if there is no network bottleneck. Of course, making sure this is *not* the case is what our job is all about. Some observations suggest that a split statistical model – part deterministic and part exponential – is a better fit for client packet arrival in some cases. Client packet sizes were found to be deterministic. Future research is ongoing.

Today, gaming traffic is a minor contributor to overall volume. The number of gamers is relatively small, although the certainty of this growing is about a sure a bet is there is in predicting network usage. Secondly, however, gaming transmission volume is small, as the games merely transmit state information that is used by the software at the ends to actually render the complex video images. Migrating game activity to virtual reality based image transport could change the transmission volumes dramatically.

<u>Video-on-Demand</u>

Data growth has some historical precedent, including recent trends with cable modem users. These trends offer anecdotal evidence supporting the bandwidth growth assumptions predicted in [8]. Video-on-demand has not been characterized as carefully. However, that VOD has accelerated rapidly in the last couple of years is not news, with total VOD revenues increasing nearly 10x between 2000 and 2002 [14], and the expectation it will double again by the end of next year. It is also not news that the bandwidth needs of video content as a service to one subscriber greatly exceed that of a data service to that subscriber. VOD services are well down the path of some of deploying the widest-pipe and most cost effective technologies, including Gigabit Ethernet (GbE) and WDM.

The choice of technologies above are the same as those discussed to support the bandwidth growth in residential data. Because of this, carrying them intermingled with one another seems like a logical step for added efficiencies in the network. Important questions to consider are those associated with the QoS needs of each, and, if necessary, how Ethernet-based transport is augmented with added robustness to ensure such needs are met. For example, over 200 MPEG movies are carried on a single GbE pipe. The case for redundancy and fail-over becomes more compelling as pipes widen and carry more traffic. Whereas data implementations today may have Layer 4 mechanisms providing loss and flow control, loss of unidirectional streams represent a non-trivial protection situation, not to mention a bad day to be near the customer service center.

In terms of traffic, VOD has some obvious diurnal dynamics. Of course, most network decisions rely on demand requests during peak usage hours on peak days. The dynamic range is quite wide, and makes for some tempting unused bandwidth in off-hours. Fortunately, these behaviors are predictable, and fortunately, as an example, peak VOD hours will not coincide with, for example, peak busy hours of enterprise data traffic. By the same token, however, peak VOD busy hours may roughly coincide with peak residential Internet usage hours.

In summary then, aggregated VOD streams present to us a rapidly growing, high QoS, wide bandwidth application, the needs of which today are essentially met via silo networks. The technology trends, however, point towards the same technology choices expected for data growth. VOD traffic varies in daily and weekly trends in predictable ways over time.

IP Video/Audio

Streaming content has received quite a bit of air time in the past few years, while during that time P2P traffic was really what caught a buzz about it and took off. What constitutes streaming traffic and P2P begins to blur, but, in general, the concept of streaming media conventionally applied to the idea of a content providing service streaming IP to a computer terminal or settop box connected to a computer (or even a TV). The reference bandwidth projection predicts that this type of traffic will grow to be about 18 times as large between now and 2010. It current contribution in term of bandwidth consumption and traffic engineering is negligible, although it has potentially large architectural impacts if the enabling technology to light this fuse is all-IP, all the time. This is a visionary decision or timetable – depending on your perspective – every operator must make for the future growth and service providing capabilities of their system.

From a traffic standpoint, we have discussed what P2P traffic does to a dynamic set of flows of residential Internet. The effect is to create a steady “bias.” In other words, the mean bits-per-second increases, and the traffic dynamics exist on top of this mean. For streaming media, this same effect would be the expectation when it becomes significant enough to matter, except that this would be a downstream phenomenon. As such it could be more buried in the noise depending on the asymmetry experienced in the network. Obviously, if the mean is very large in comparison to the peaks and valleys, the impact of peaks and valleys on efficient pipe usage is very minor. In other words, if the streaming media content (or VOD content for that matter) dwarfs data transport content along the same conduit, the traffic engineering pipe size problem is simplified, since the ups and downs will be relatively small.

Now, a significant difference between streaming media and most P2P traffic today is that the former is real-time content, deserving of QoS capabilities that do not require the same level of sophistication as moving MP3’s and JPEG’s around. Since the content still exhibits LRD regardless of whether data or streaming media, mechanisms at the ends of the pipe that enforce policies and switch packets need awareness into the proper traffic models of this LRD, so that queues can be build and implemented to avoid dropped packets and blocking and supply the QoS expected for real time content.

Summarizing, streaming multimedia represents content characterized in prior work as having self-similar behavior. It is a relatively high bandwidth consumer on a single user session basis with the QoS needs of other real-time media, but the total bandwidth usage is low and growth path very dependent on architectural and service choices going forward.

Digital TV or HD

Broadcast digital TV has much the same characteristics as VOD. There are two main differences. First, in many cases, linear supertrunking is used rather than digital transport as a low-cost alternative when high bandwidth digital backbone is not in place. Second, the service group size is very large, making redundancy of path and equipment quite important. High Definition streams, for the most part, represent to the network engineer, digital TV traffic on steroids. The bandwidth hungry nature of HDTV is a promising possibility for inspiring networking bandwidth upgrades.

Similar to streaming media – and in fact analogous except for the more standardized format of delivery – digital TV provides a steady flow of packets and a nice averaging of bandwidth behavior to a bit rate number that is a function of compression and statistical multiplexing of a large number of streams. If anything rides along the same channel, such as VOD, then VOD dynamics would be superimposed.

ARCHITECTING FOR MULTIPLE SERVICES

QoS Parameters

There is no universal definition of Quality of Service QoS), just as there is no universal definition of "carrier class." Nonetheless, QoS encompasses basically five parameters:

- Latency – End-to-end absolute delay
- Jitter – End-to-end variation in delay
- Loss – Dropped transmissions
- Throughput – Bits-per-second or Bandwidth
- Availability – Likelihood of the network being "up"

QoS has achieved buzzword status quite recently, and its footprint is all over standardization committees. What is essentially going on are efforts to bring to the data world something it has always lacked – guarantee-able QoS – but doing so on a "connectionless" network while keeping as much legacy frame and protocol structure intact as possible. The result is primarily profitable to the acronym maker (or marketing team). The effort is a logical outgrowth of the indisputable fact that Ethernet dominates the LAN. As the LAN aggregates to the MAN and WAN, the goal of leveraging the broad familiarity with Ethernet, its cost points, and the flexibility of features within Ethernet and IP as protocols have driven traditional Ethernet and IP network designers to innovative approaches to solving this classic QoS shortcoming in the hopes of scaling the local LAN to a broader market.

Not surprisingly, some techniques to enhance Ethernet resemble old ideas. For example, the use of the DiffServ protocol (differentiated services), when implemented over MPLS (Multi-protocol label switching) has the look and feel of ATM, but with a lot of different acronyms describing the details. DiffServ provides the ability to classify a packet with a forwarding class describing its priority on a per-hop basis. The latter fact actually limits the overall QoS strength of DiffServ on its own, but increases its practicality. The role of MPLS is to expedite the forwarding of packets through the network by creating label-switched paths via tags on the Ethernet frames, directing packets at Layer 2, rather than making route decisions through the network that create processing bottlenecks and subsequent latency and jitter problems. Thus, these schemes together offer prioritization of payload types, and create predefined and expedited paths through the network. This scenario has indisputable similarities with ATM.

So, why re-invent the wheel? The answer is simply because the dominance of Ethernet and flexibility of IP have made riding this wave a necessity in network design, to the extent that incrementally upgrading the technology to carry more than best-effort is more palatable than addressing major equipment and protocol overhauls and learning curves.

QoS – Who Needs It?

Let's list some of the services encountered or viewed as on the horizon. How do these compare as far as who needs what for QoS? Let's use a simple scale: High (H), Medium (M), or Low (L) need for the particular QoS parameter below. A qualitative summary of QoS needs is shown in Table 1 below. Certainly, there is plenty of room for debate (I adjusted this chart more than a dozen times), and some still require more learning and evaluation.

Table 1 – Services and QoS Need

	Latency	Jitter	Loss
Residential Data (DOCSIS Internet)	L	L	L
Residential Voice (VoIP)	H	H	L
Business Data (DOCSIS 1.1 or higher)	L	L	H
Business Data (fiber-based)	M	L	H
Business Voice (VoIP or T1/T3)	H	H	M
Business Video (videoconference)	H	M	L
MPEG or IP Video or VOD	M	H	M
HDTV Broadcast	M	H	M
IP Audio (Radio AOL)	M	M	M
Interactive Gaming	H	H	H

Clearly, we can recognize that some applications are real-time, while others are not. This fact primarily drives the latency and jitter QoS needs. Also, based on the nature of the service, it may be loss tolerant or not. In general, if the content itself is to be transduced for human senses, it is likely to be loss tolerant to some extent. Human senses are quite effective as filters. If the content is information for a computer to interpret and process, it is likely to be less loss tolerant.

What tools exist to assure the level of QoS desired is achieved? In the DOCSIS world, the use of DOCSIS 1.1 or higher, and a next generation CMTS [13] provide this capability. The CMTS is a key element between the access and transport network, acting as both a media converter at layer 1 and protocol delineation point for layers 2 and 3. DOCSIS 1.1 provides the class of service capabilities on the HFC side, while advanced layer three implementation such as per-flow queuing provide traffic management functionality on the network side. Thus, for the first three items in Table 1, DOCSIS 1.1 and next generation CMTS enable the providing of QoS mapping from access network to interconnect ports. For business voice and data based on fiber connectivity rather than DOCSIS 1.1, and segmented spatially or via wavelengths in the access network, QoS schemes must reside in the aggregation equipment and supported elsewhere in the architecture.

On the video transport side, such as VOD and HD transport, QoS is assured by the fact that these systems currently are essentially silo systems. Statistical multiplexing occurs as server content traverses switches, but the rules of engagement are simplified by the singular content and rules easily developed from this simplified, application-specific architecture. Should these services become part of an integrated triple-play transport network, the dynamics of the traffic situation could change significantly. For example, for a single IP pipe supporting video and data, there would be a heavy reliance on IP QoS schemes through some of the existing standardization efforts to assure the video QoS needs are met. The end-to-end capabilities are not of the "guaranteed" variety, and would certainly require some traffic engineering and modeling.

Streaming media has the same type of architectural implications in the network, with the difference being that DOCSIS supports the access portion of the network. Thus, mapping of QoS mechanisms from one side of the CMTS to the other – again using the CMTS both the due media and QoS transducing – is needed.

Architecture Technologies

Clearly, many services with many different needs are set to co-exist. Sound business practice means finding an efficient means to handle them by judicious choice of technologies, levels of integration, and a healthy concern for operational costs and scalability. The term "triple play" alone sounds like an abbreviated set of services, but the flavors within the triple play, as shown above, clearly make the problem more complex. The access network itself is constantly being re-thought for fresh ideas, such as data overlays, wireless interfaces, and intelligent processing. At the aggregation points in the hub and Headends, various technology choices exist, some of which are deployed already in silo networks as previously indicated in VOD cases. VOD represents a good reference example because of where it is in the cycle – basically just at or past the "knee" of one of those classic marketing hockey stick charts, depending on which analyst you ask. It is a service experiencing significant growth, and it is using technology on the move in both the server and multiplexing arena, as well as in the transport pipe. VOD transport has been migrating from DVB-ASI transport to lower cost, more-flexible, GbE links. And, again, use of GbE technology makes its way into the data world as well. VOD also has been a driving application for another key technological option - wave division multiplexing (WDM).

Gigabit Ethernet has some notable shortcomings as an all-inclusive answer for network design. Ethernet as well as IP over Ethernet – both designed for data – do not inherently offer the resiliency, availability and network management attributes that become important elements of a "carrier class" solution when so much content is riding on the success of a single link. IP over Ethernet was developed as a "best effort" technology, and most of the ongoing efforts today revolve around finding way to be better than best effort. The previously mentioned DiffServ and MPLS developments fall into this category, and there are others. Thinking in terms of end-to-end IP as an attractive end game, MPLS, in fact, can be viewed as a way to skirt the routing limitations of an all-IP network by avoiding the per-hop calculation of routes through the network. Not all developments aimed at traffic engineering and hardening data systems are completely new, however. TCP/IP itself is a kind of QoS feature, and type-of-service (ToS) header bits have been around since 1981. Limited capabilities of these features – invented still in a "data world" context – limit their power to meet the kind of diverse needs expected.

Another Ethernet issue is that it does not inherently support circuit-based voice. Technologies exist to create virtual circuits over Ethernet. Similarly, at layer 3, voice over IP (VoIP) has been developed technologically. Each today has cost penalties. In short, however, Ethernet, even GbE or 10 GbE, and even if we include wave division multiplexing (WDM), cannot go it alone. All-IP looks attractive from an interoperability and flexibility standpoint, but the jury is out on guaranteeing that all of the QoS needs can be met even with the traffic engineering tools brought to the table with MPLS and other tools designed to optimize packet transport. Actually, right now, use of enough wavelengths and 10 GbE would be essentially the oft-practiced lazy man's QoS – gobs of bandwidth assuring that nothing gets held up. For the price of this overcapacity, if well managed, there would be no congested routes, and no pipe traffic peaks requiring buffering and queuing delay. There are smarter ways to solve the problem rather than relying on this brute force approach.

WDM itself is, in general, a brute force capacity enhancing tool. It allows a single fiber to carry multiple data-bearing streams, such as GbE or 10 GbE, by using a different wavelength for transmission for each. However, WDM does not easily offer QoS consciousness. This is not a show-stopping issue - intelligent wavelength management exists as a relatively mature technology. Integrating wavelength management as part of system resource management has been an anticipated direction of the telecom sector for some time, and could see relevance in cable networks as well since the same premise drove that thinking – bandwidth growth and support for advanced services.

The classic shortcoming of QoS is precisely the reason for equipment based on Next Generation Sonet. There is no debating the QoS features of Sonet transport. The knock against Sonet had to do with its rigid structure that made it inefficient as a packet data transport system. Coarse bandwidth increments left excess unused bandwidth, driving up the effective cost of doing business by decreasing fiber usage efficiency.

However, precisely these data issues are addressed with Next Generation Sonet, all the while holding firm on the guaranteed, proven, mature resiliency and reliability that has no comparison today among alternative technologies. Furthermore, there is lots of existing Sonet-based infrastructure. What modern equipment does is build the data flexibility into formerly coarsely-grained all TDM-only platforms through virtual concatenation (VC) and link capacity adjustment (LCAS) which allows dynamic – i.e. supporting data – provisioning of VC carriage. Such platforms contain port interfaces natural to data handling, such as 100BaseF and GbE, as well as the traditional voice-related interfaces of a traditional Sonet platform. Other data-oriented features are included as the platforms evolve to meet the shift in traffic demand. Support of standard framing protocols for data is gathering momentum (Generic Framing Protocol or GFP), as well as standards-based packet classification and guarantee-able class-of-service mapping.

Next generation Sonet platform, then, offer the guaranteed resilience and inherent QoS parameters for both TDM and Ethernet services – the resiliency still unique to Sonet – and now offer the added granularity and flexibility to efficiently support data needs.

The networking world never has quite enough protocols, and one of the latest is aimed at providing efficiencies of packet-based transport natively, but having the resiliency characteristics of Sonet. Avoiding the *traditional* Sonet limitations was a key target of this development effort. A standards body has been formed, IEEE 802.17, that is developing the layer 2 protocol known as resilient packet ring (RPR). Not surprisingly, the basic frame structure of RPR was based on Ethernet, and adds to it MPLS and Class-of-Service header content, as well as other fields.

RPR is a ring protocol, with the objectives of optimally supporting all previously described traffic types, maximizing efficient use of counter-rotating ring bandwidth, and simplifying network provisioning. The critical importance of packet resiliency is recognized as a key focus, as the traffic mix and amount no longer fit into a "best effort" paradigm as was once the case. As a layer 2 technology, RPR can run on entrenched physical layers (i.e Sonet and Ethernet PHYs), which is important considering the amount of deployed infrastructure. The standard is still an emerging one, and the discussions break down into two camps – Cisco and others.

WHAT NEXT?

What to make of all of these services, technologies, and, in general, the many choices that face network designers today? Needless to say, there is no single magic answer, and recommending what works best for 2010 will simply create competition for someone's infamous "who needs more than 64k of memory" comment. The good news is that cable operators are in the drivers seat at the moment. The competition for service provider of choice is arranged in their favor. But it is unclear how long this will last given alternative solutions and the allure of residential broadband as something people will pay for, a few-and-far between uptick in a beaten down economy. To capitalize on a game that is cable's to lose, and since there is no one-size-fits-all solution at this juncture, what we can recommend is a five steps "keys to success" approach:

1) Have a planned roll out of services to provide or support. A comprehensive sample set is provided in this paper. *This first step is, as always, a business exercise.*

2) Develop a bandwidth forecast of your own, and comprehend the traffic and QoS aspects of each service. The given information and references provide guidance towards both. *This is a business **and** technical exercise.*

3) Understand the current and growth capabilities and limitations of any existing infrastructure in the context of 2). This especially includes the emerging standards and techniques suited to the evolving service mix, and the flexibility available where there is infrastructure to build. The previous section points the way towards much of the

relevant activity in this area. *This is a technical reading & research exercise.*

4) Develop a time-phased service integration and network evolution plan for the traffic mix (not necessarily the same as an integrated network) aligned with 1), 2), and 3); the hard work of the first three steps make this more straightforward than it sounds. *This* is *a business **and** technical exercise, and the one that decides whether you grow, maintain, or flounder and become exposed to competition.*

5) Know what end-to-end means to your network responsibilities, and modify any "silo" role & responsibility organizational definition to smoothly evolve across the network interfaces, features, standards, and mapping schemes. *This is a technical management exercise. Technical managers often used to be technical people and carry with them technical biases, so this one may be more difficult than it sounds.*

CONCLUSION

On the QoS front, one of the more compelling human stories in recent years occurred when a researcher in Antarctica found herself stranded with her team during an inaccessible time of year for rescue missions and in need of critical medical attention. Much like the space shuttle engineering techniques being developed and discussed herein.

And what of 2010, based on the forecast previously referenced? Architecturally, will the triple play be implemented in a unified network, with today's darling being an all-IP format all the way to the home? Will some services continue to ride over separate parallel networks optimized to the bandwidth and QoS required? According to the forecast referenced, 59% of the forward path digital traffic demand will be Internet access, 26% will be some form of VOD, 13% will be streaming audio or video content of the PC variety, and the remaining IP telephony. Do

Columbia tragedy this past winter, the video phone images and stories of the researcher and her team offered a glimpse into the risks of pushing the envelope. Fast-forwarding to the present, recent literature describes the wireless infrastructure in an Alabama hospital [9]. The article also describes the network services the hospital uses to transport mission-critical data. The concept of remote medicine involves supporting transmission of high-resolution images and data in real-time to doctors to serve live patients anywhere, such as remote locations, a third-world country, or on a battlefield. Obviously, administrative and legal obstacles abound anywhere medicine is involved and lawyers are prowling. But the advantages possible are access to expert advice in a timely manner, access to trusted medical advice internationally, access to observation of expert and advanced procedures by other doctors and students, and less dependence on getting out and about to receive care when seriously ill. Can we invent a higher QoS service need? The military designs and implements its own private networks for mission critical data, because there is absolutely no margin for error. But, it can afford to as well, while applications such as the above would rely on the kinds of QoS and traffic

you buy this perception of today? Would this aggregation mix lend itself to a particularly convenient model? Some of the constraints of the previously introduced central limit theorem, aside from a variety of independent sources, is that the independent distributions have finite variance, and that there not be a singularly dominant distribution. Thus, while the broad traffic mix implies central limit simplification, a single dominant service can disturb this convenience. Furthermore, a cornerstone characteristics of self-similarity – heavy tails – can also imply infinite variance, another central limit theorem killer. The jury is once again out, as research continues to classify traffic trends and distributions for network modeling and optimal architectural design. Similarly, networking technologies will simultaneously evolve, making putting a stake in the quicksand that much more perilous.

Contact Info:

Rob Howald
Director
Systems Engineering
Network Infrastructure Solutions
Motorola Broadband
rob.howald@motorola.com

References

1) Adas, Abdelnaser, *Traffic Models in Broadband Networks*, IEEE Communications Magazine, July, 1997.

2) Addie, Ronald, Moshe Zukerman, and Timothy Neame, *Broadband Traffic Modeling: Simple Solutions to Hard Problems*, IEEE Communications Magazine, August, 1998.

3) Boggs, David, Jeffrey Mogul, and Christopher Kent, *Measured Capacity of an Ethernet: Myths and Reality,* Digital Equipment Corp, Western Research Laboratory, Reseacrh Report 88/4, September 1988.

4) Borella, Michael, *Source Models of Network Game Traffic*, 3Com Corporation.

5) Brownlee, Nevil and KC Claffy, *Understanding Internet Traffic Streams: Dragonflies and Tortoises*, IEEE Communications Magazine, October, 2002.

6) Chu, Tam-Anh, *WAN Multiprotocol Traffic Management: Theory and Practice*, Communications Design Conference, September, 2002.

7) Coffman, K.G., and A.M. Odlyzko, Internet Growth: Is there a "Moore's Law for Data Traffic?", AT&T Labs Research, July, 2000.

8) Nash, Randy, *A 10-yr Residential Bandwidth Demand Forecast and Implications for Delivery Networks*, 2002 NCTA Technical Papers, May, 2002.

9) *Look Doc, No Wires*, Communications News, March 2003.

10) Robert, Jim, *Traffic Theory and the Internet*, IEEE Communications Magazine, January, 2001.

11) Sahinoglu, Zafer and Sirin Tekinay, *On Multimedia Networks: Self-Similar Traffic and Network Performance*, IEEE Communications Magazine, January, 1999.

12) Shaw, Terry, *Bandwidth Utilization on Cable-Based High Speed Data Networks*, 2002 NCTA Technical Papers, May, 2002.

13) White, Gerry, *Tiered Data Services to Enhance Customer Value and Revenue: Why are DOCSIS 1.1 and Advanced Queuing and Scheduling Required*, 2003 NCTA Technical Papers, June, 2003.

14) MultiChannel News, Broadband Data Book

CABLE & CE INDUSTRY COOPERATION ON UNIDIRECTIONAL DIGITAL CABLE RECEIVERS

Brian Markwalter, David Broberg
Consumer Electronics Association, Cable Television Laboratories

Abstract

As consumer electronics companies and the cable industry continue to work together to accelerate the deployment of high-definition digital television (HDTV), they have created a memorandum of understanding (MOU) that defines how cable systems will deliver services and essential elements needed for unidirectional digital cable-ready receivers to receive such services. The MOU relies on Society of Cable Telecommunications Engineers (SCTE) and Consumer Electronics Association (CEA) standards to provide the framework for interoperability. This paper describes the December 2002 MOU and focuses on the standards referenced therein that define requirements for cable systems and receivers.

This paper provides an overview of the agreement as a foundation for providing a more detailed looked at the self-certification program it requires. The paper describes the categories of tests prescribed by the agreement including Critical Tests, Non-critical Tests, and Network Harm Tests. Taken together, these test make up the Test Suite, jointly developed by CEA and CableLabs®. The Test Suite is derived from existing work by CableLabs as part of the OpenCable project. This paper describes in greater detail the foregoing testing methodology and the expected benefits to the industry.

OVERVIEW OF DECEMBER 2002 MOU

In December 2002, 14 television manufacturers and eight cable system operators signed a memorandum of understanding covering interoperability of unidirectional digital cable products and cable systems. The MOU culminated months of work, facilitated by the Consumer Electronics Association and the National Cable Telecommunications Association, to reach consensus on how best to achieve the mutual goal of retail availability of cable ready receivers while ensuring cable services are delivered as intended. The MOU deals with four impediments that prevented television manufacturers from being able to introduce cable ready TVs through the OpenCable process: (1) legal concerns with the available POD Host Interface License Agreement, (2) certainty that a large percentage of cable systems nationwide would follow specific digital transmission standards, (3) the lack of encoding rules for copy protection, and (4) a test or certification regime in keeping with the way televisions are typically measured for compliance.

MSOs rightfully sought to ensure that in reconciling these CE manufacturer concerns their own goals not be sacrificed. These goals being: (1) cable services are delivered consistently whether through a leased device or a retail device, (2) cable not be competitively disadvantaged with respect to other video distributors, (3) operators have freedom to develop and market new services, and (4) retail cable ready devices not harm the cable network or allow theft of service.

Elements of the MOU obviously deal with certain aspects of these goals, as evidenced by the inclusion of a new DFAST license agreement and encoding rules. Enough ink will be spent on these mostly legal matters elsewhere. This paper instead focuses on the standards that both parties have agreed to rely on for compatibility and the self-certification process for the retail devices.

STANDARDS THAT APPLY

The Core Standards

In the MOU, cable system operators commit that cable systems with an activated channel capacity of 750 MHz or greater shall comply with the following SCTE standards.

- SCTE 40 2001, as amended by DVS/535
- ANSI/SCTE 65 2002
- ANSI/SCTE 54 2002, as amended by DVS/435r4

And all digital cable systems shall comply with these standards.

- ANSI/SCTE 28 2001, as amended by DVS/519r2
- ANSI/SCTE 41 2001, as amended by DVS/301r4

The ‘‘as amended by’’ notation reflected the need to point to these standards that were at the time being revised in the SCTE DVS committee. A quick description of each standard and its status as of this writing follows.

SCTE 40 2001, titled Digital Cable Network Interface Standard, is in the final SCTE approval stages and should publish as SCTE 40 2003. SCTE 40 defines the key characteristics of what the cable system delivers to the television in terms RF, transport layer, and other services, such as emergency alerts and closed captioning.

ANSI/SCTE 65 2002, titled Service Information Delivered Out Of Band for Digital Cable Television, is unchanged since the MOU was signed. This standard defines Service Information tables providing the data necessary to tune and display the services offered by the operator. The term Out Of Band indicates that the SI tables are delivered by a possibly proprietary transport to the POD and then forwarded in a standardized fashion to the cable ready device (Host) through the Extended Channel.

ANSI/SCTE 54 2002, titled Digital Video Service Multiplex and Transport System Standard for Cable Television, is now SCTE 54 2003 after completing its revision process. This standard builds on MPEG-2 Transport Stream coding to define how cable systems construct multi-program Transport Streams.

ANSI/SCTE 28 2001, titled HOST-POD Interface Standard, is in the final editorial stages after completing its ballot and should publish as SCTE 28 2003. This standard defines just what its title suggests – clearly necessary for developing unidirectional digital cable products.

ANSI/SCTE 41 2001, titled POD Copy Protection System, is near the end of a major revision related to switching the copy protection system to reliance on X.509 certificates. This standard defines how the interface between the POD and HOST is protected from having to expose video content in the clear.

The first three standards above are an obligation for 750 MHz cable systems to deliver digital video by these standards. The HOST-POD Interface and its associated copy protection standard are an obligation of all cable systems, regardless of whether digital transmission is used. Similarly, digital cable products marketing under this MOU are obligated to tune digital channels in accordance with SCTE 40, navigate using SCTE 65, respond to emergency alerts per SCTE 54, and include a POD interface compliant with SCTE 28 and SCTE 41.

Other Standards

The MOU relies on other standards, particularly related to certain interfaces on leased set top boxes and retail digital cable products. Television manufacturers commit to providing DVI or HDMI interfaces on a phase-in and resolution basis and cable operators commit to providing IEEE 1394 and DVI interfaces on HD set top boxes on a phase-in basis. Cable operators expressed an interest in DVI (uncompressed video) as the preferred interface, hence the commitment by television manufacturers to support it. Television manufacturers needed support for a compressed video interface on set top boxes for recordability, explaining the inclusion of this interface on leased boxes.

The IEEE 1394 interface described in the MOU is actually defined by a pair of standards, ANSI/SCTE 26 2001 and CEA-931-A. SCTE 26, Home Digital Network Interface Specification with Copy Protection, builds on EIA-775-A and EIA-779, which in turn build on IEEE 1394, to completely define how this interface is used between a cable device and another CE product. CEA-931-A, Remote Control Command Pass-through Standard for Home Networking, adds the usability feature that a display device can pass-through remote control commands to the video source at the other end of the 1394 interface.

ADDITIONAL REQUIREMENTS

Harm Prevention Tests, those meant to protect the cable system and its ability to deliver services, are singled out as applying to all products under the MOU. A mutually agreed upon set of harm prevention requirements does not exist in the form of an SCTE or CEA standard. The MOU recognizes this deficit by pointing to EIA/CEA-818-D and DVS/538 as sources for these requirements.

EIA/CEA-818-D, Cable Compatibility Requirements, collects together requirements from other standards for application to digital cable systems and compatible receivers. Part I states minimum requirements for receiver-compatible digital cable TV systems, and Part II states minimum requirements for cable-compatible digital TV receivers. SCTE DVS/538r1, Uni-Directional Receiving Device Standard for Digital Cable (Input), is a proposal for standardization of receiver requirements intended to complement the transmission standards used by digital cable TV systems. Neither of these documents are referenced directly by the MOU, except as sources for harm prevention test items.

PROTOCOL IMPLEMENTATION CONFORMANCE STATEMENT

One of the tools that often is used in the process of verification of a complex product that follows a number of industry standards is the Protocol Implementation Conformance Statement (PICS). This document is a detailed collection of every one of the requirements from all the referenced standards. This document creates a traceability matrix and serves as the basis for any conformance statement of a manufacturer seeking certification.

Since the MOU and the proposed rules for a unidirectional cable receiving device were written, a team of engineers from several manufacturers, along with staff of CableLabs and CEA, have been working to complete this critical piece of documentation. In the first quarter of 2003, this team participated in meetings and conference calls totaling more than 120 hours and spent in excess of $10,000 on conference call services to this end. This concentrated effort shows how critical is the element of accurately documenting each testable requirement.

The PICS document contains over 600 unique requirements. In many cases each line item includes a direct quotation of a normative statement from the applicable industry standard, along with a chapter and verse reference location. In some cases, a requirement was stated without any citable industry standard to reference. In those cases each new requirement is added to an appendix at the end of the PICS.

This process of including requirements without an external reference does represent a departure from the usual process of developing a PICS. This departure from past CableLabs practice was necessary since the MOU relies solely on published SCTE and CEA standards and some mutually agreed requirements derived from other sources, including OpenCable, EIA/CEA-818-D, and DVS/538.

The PICS documentation also serves as the detailed breakdown showing which requirements relate to Critical Tests and Non-Critical Tests. The Critical Test items are further divided to show which apply to ‘‘Tune and Display’’ requirements and which remain as Harm to Network, Security, or other harm related tests. The purpose of this division is to show which requirements apply to the different type of products defined in the MOU and proposed rules.

The final purpose of the PICS documentation is to list the requirements that need to be tested in the Acceptance Test Plan. This completes the traceability so that every test may be traced back to one or more line item in the PICS, each of which can be traced back to a normative statement of a referenced industry standard.

ACCEPTANCE TEST PLAN

The Acceptance Test Plan (ATP) is another document that is included in the Joint Test Suite (JTS). This document details each of the unique test procedures that are used to verify the requirements stated in the PICS. The ATP gives instructions to the test technician who performs the test and it details the equipment settings, connections, and other test conditions. The ATP also defines the range of acceptable results and how the results should be documented.

There are three basic guidelines that were used in creating the tests within the ATP: (1) All of the tests are ‘‘black-box-tests’’ meaning that the tests are performed on a closed box, using only the available input and output interfaces; (2) The tests are not meant to limit the type of test or procedure that can be used to verify compliance, but are simply a record of an agreed upon group of tests that are applicable; (3) The test plan is not static or complete, further revisions are expected as additional tests are developed and new test equipment becomes available.

The ATP is divided to match the breakdown of the PICS into Critical and Non-critical, with the Critical tests further divided to show Harm prevention tests, security tests and tune and display tests. This breakdown is prescribed by the terms of the MOU.

Each test within the ATP may be used to verify one or more of the numbered requirements of the PICS. Each test identifies what is being tested, the test equipment to be used, and the instructions on the exact settings of the controls and instruments used. Connection diagrams and further explanations of the setup are provided so that all tests are readily repeatable.

A variety of tests are necessary to fully determine compliance with the standards. One group of tests can confirm a portion of the requirements using a POD simulation tool that can be programmed to provide many of the message types that are used on the POD interface. This tool logs the response from the unidirectional receiving device and analyzes the response to confirm compliance.

The cable side has proposed also to include interoperability tests which use a genuine POD on a live cable plant. This type of test has been found to be important in previous CableLabs testing since the simulation tools do not contain proprietary circuitry needed to work on a real cable plant. Without those circuits, the tool is not able to receive messages from a cable headend which is necessary to confirm the receiver is not interfering with headend communications according to the requirements of the Harm tests. Further, there are a variety of requirements associated with the proper reception of the OOB signals, which vary widely from plant to plant that are not testable using the simulation tools. Television manufacturers believe this type of interoperability testing is not part of the MOU's self-certification process and offered instead to work with cable on interoperability events.

The ATP also includes the forms that record the results of each test. Blank space is provided to record the measured results right next to the defined range of acceptable results. This documentation becomes part of the first prototype test suite results that are recorded at CableLabs.

SELF CERTIFICATION PROCESS

Certification is the process of verifying compliance with the required standards necessary to earn the right to use the digital certificates necessary to operate on a cable plant. Without these digital security certificates, the product would not be recognized by the cable system. When the digital cable receiving product is first plugged into the Point of Deployment card (POD), a digital authentication process ensues. Each device verifies the authenticity of the certificates held by the other device. If both sides agree, the interface is said to be authenticated. If this process fails, cable services are disabled.

‘‘Self-Certification” is the form of this certification process that is prescribed by the MOU and that relies upon the individual manufacturer’s own statements and documentation. While the exact details of the self-certification process are not fully defined nor agreed upon at the time this paper was prepared, the following basic principles are expected to be used:

1) The first prototypes of the unidirectional digital television product will be brought to CableLabs or an appropriately qualified third party testing facility where the Test Suite will be executed. Test events will be scheduled at CableLabs to reasonably accommodate the demand and will be coordinated to make best use of resources.

2) If the test results reveal any failures of the Critical Tests and the product is a unidirectional digital television product, then corrections must be applied and the product resubmitted to CableLabs for re-testing as many times as it takes to correct all the Critical Test failures. If the first prototype submitted is not a television, and has critical test failures, only the corrections to the Harm Prevention Test failures need be retested.

3) Once the manufacturer has successfully passed all Critical Tests and corrected all other test failures as needed, the passing test results are submitted to CableLabs along with the self certification documentation. This additional documentation includes the affirmative conformance statement and other details that have not been fully defined at this time.

4) Once the passing test results and the Self Certification Documentation has been

submitted, CableLabs authorizes the assigned Certificate Authority to begin issuing the X.509 certificates to the manufacturer for the model and range of products specified.

5) Subsequent products by the same manufacturer have no obligation to be tested at CableLabs, but need only the Self Certification Documentation to be authorized for digital certificates.

NEXT STEPS

At the time of this writing, work remained on the PICS and ATP; they are expected to be completed by the time this reaches print. There will also be some further negotiation and documentation needed to fully define the details of the Self Certification process.

Of course, the FCC must endorse the proposed rules as submitted with the MOU in order for this process to be activated. In the mean time some manufacturers are going ahead and making products designed to meet the full OpenCable requirements under the PHILA agreement while others are waiting to take advantage of the MOU process.

There also remains some risk that the FCC may not endorse the exact proposal as submitted. If that happens the MOU says the deal is off and everyone will have to reassess how to proceed.

REFERENCES

The MOU is available under the FCC Further Notice of Proposed Rulemaking, FCC 03-3, in Dockets CS No. 97-80 and PP No. 00-67.

SCTE standards are available at www.scte.org.

CEA/EIA standards are available at http://global.ihs.com/.

CARRIAGE OF MPEG-4 OVER MPEG-2 BASED SYSTEMS

Ardie Bahraini
Motorola Broadband Communications Sector

Abstract

The MPEG-4 specifications have provided substantial advances in many areas of multimedia technology. In MPEG-1 and MPEG-2 System Specifications referred only to overall architecture, multiplexing, demultiplexing and synchronization of elementary streams. The MPEG-4 specification goes beyond these areas to encompass content description, interactivity, and scene description to name a few. This paper only addresses the overall architecture, multiplexing and synchronization of MPEG-4 content when carried in a system that already supports MPEG-2 Transport Stream.

INTRODUCTION

MPEG-4 is the first digital audiovisual coding standard that expands beyond defining compression algorithms to address emerging computing and telecommunication worlds. MPEG-4 system specifications were intentionally developed to be transport agnostic, enabling MPEG-4 content to be carried over many different transport systems such as MPEG-2, IP, ATM, etc. In particular, MPEG has amended MPEG-2 system standard ISO/IEC 13881-1 to allow carriage of MPEG-4 content over MPEG-2 Transport and Program Streams and this amendment is included in the published 2000 edition of ISO/IEC 13818-1.

The MPEG-2 system standard (ISO/IEC 13881-1) provides two alternatives to carry MPEG-4 content over MPEG-2 Transport Stream (TS). The first scheme is straightforward, and provides the capability for carriage and signaling of individual MPEG-4 audiovisual Elementary Streams (ES) by employing the MPEG-2 system-layer parameters such as PCR, PTS and DTS. This scheme could be used in existing systems that already use MPEG-2 Phy and Transport layers and want to take advantage of the better compression schemes as well as the synthetic video coding tools offered by MPEG-4 part 2 (ISO/IEC 14496-2). MPEG is also extending the MPEG-4 video standard in its specification ISO/IEC 14496-10 (also known as JVT) which will provide significant compression advantage over both MPEG-2 and part 2 of MPEG-4. More details regarding this implementation will be provided in the first part of this paper.

The second alternative defined in the MPEG-2 system layer provides the capability for carriage of MPEG-4 scenes in addition to carriage of MPEG-4 audiovisual elementary Streams. Carriage of this type of content over MPEG-2 Transport Stream follows both MPEG-2 system standards as well as MPEG-4 systems (ISO/IEC 14496 –1) SL_packetize or FlexMux tools specification. This scheme can also be implemented within existing systems in order to provide the MPEG-4 object-based coding and scene composition capability in addition to the better compression that is provided by MPEG-4. The second part of this paper will explore this alternative in more detail.

The last section of this paper will briefly explore how these implementations could be used to enhance current cable systems by migrating from dual carriage of Analog and Digital to an all digital network in order to address future bandwidth requirements as well as providing additional services and features such as HD, VOD and home Gateway based on MPEG-4.

A HIGH LEVEL OVERVIEW OF MPEG-4 DELIVERY LAYERS

MPEG-4 predecessors, namely MPEG-1 and MPEG-2, were designed to address specific systems. For example, MPEG-2 developed Transport Stream (TS) and Program Stream (PS) systems were targeted solutions toward TV Broadcasting and Local Retrieval of content respectively. Hence, the MPEG-2 system was specifically designed to optimize the transport of targeted data and delivery systems by integrating the Sync and Link layers. As a consequence, the MPEG-2 System is not efficient and cannot easily be ported to other mediums and delivery systems without substantial overhead.

MPEG-4, on the other hand, from the beginning was designed to be flexible and independent of under-layer technology such as the delivery system or link layer in order to be adaptable to different delivery systems such as TV broadcasting, IP and ATM systems. To address this separation and be adaptable by different systems, MPEG-4 defined three abstract layers namely: *Compression Layer, Sync Layer and Delivery Layer* as depicted in **Figure-1**. The Compression Layer specifies the encoding and decoding of audio-visual Elementary Streams and is specified by references [2] and [3]. The Sync Layer manages Elementary Streams, their presentation and synchronization information as well as fragmentation and random access information. MPEG-4 Sync Layer syntax is specified by ISO/IEC 14496-1 [4]. The delivery layer specifies the transparent access to other layers independent of delivery technology.

As depicted by **Figure-2**, one could further divide the delivery layer into two sub-layers namely, DMIF (Delivery Multimedia Integration Framework) layer that is specified by MPEG-4 ISO/IEC 14496-6 [*3*], and TransMux Layer that is not specified by MPEG-4 intentionally and is left to the transport technology such as IP, ATM or MPEG-2 to just name a few.

As shown in **Figure-1**, the abstract layer demarcation between Compression Layer and Sync Layer is referred to as ESI (Elementary Stream Interface) and the abstract layer demarcation between Sync Layer and Delivery Layer is referred to as DMIF.

In the following sections, a brief summary of terms related to the MPEG-2 System that are used throughout this paper is offered for those readers who are not familiar with MPEG-2 Sync and Link Layers terminology. Then, the carriage of MPEG-4 elementary streams over MPEG-2 Transport Stream are described as depicted in **Figure-3**. This implementation takes advantage of the MPEG-4 Compression Layer without using other features and functionality provided by MPEG-4. Next, system requirements and architectures are presented for those systems that not only attempt to use MPEG-4 compression, but also intend to use other features of MPEG-4 such as Link Layer and TransMux. This implementation is depicted in **Figure-4**. Finally, in the last section of this paper, the impact of MPEG-4 in the existing Cable TV system is briefly discussed; although a detailed discussion and analysis is beyond the scope of this paper.

A SUMMARY OF THE MPEG-2 TRANSPORT STREAM (TS)

The MPEG-2 system specification ISO/IEC 13818-1, defines two schemes for multiplexing Elementary Streams into a serial bit stream namely, Transport Stream (TS) and Program Stream(PS). The Transport Stream scheme is widely used in transmission of Audiovisual content in CATV today and the Program Stream is used mostly for storage media. In this section only Transport Stream hierarchy is examined since TS is the protocol that is applicable to CATV as noted above.

Each Elementary Stream (ES) contains coded video, coded audio or other data associated with a single program. These streams are separately packetized and formatted into a structure defined by MPEG-2 as Packetized Elementary Stream (PES) as depicted in **Figure 5**. As shown in this figure, each ES could expand into several PES packets. Each PES packet is identified by a stream_id in the packet header. The Stream_id that is associated with each PES packet is defined by MPEG-2 ISO/IEC 13818-1 System specification and identifies the type of steam that is contained in each PES. Each PES is then sliced into Transport Stream Packets that are 188 bytes that include a 4-byte header as shown in Figure 5. The MPEG-2 system specification defines each TS Packet to be identified by a field in the header known as Packet Identifier or PID that is 13 bits. Thus, the payload of each TS packet could contain up to 184 bytes since there are four bytes allocated for the header that include PID in addition to other fields. Each PID is associated with an ES of a service; therefore one program may have one video PID and several Audio PIDs.

Every MPEG-2 Transport Stream multiplex carries a set of tables known as Program Specific Information (PSI) tables. These tables contain information about services, which are present in the multiplex. PSI data includes the following tables: Program Association Table (PAT), Program Map Table (PMT), Conditional Access Table (CAT), and Network Information Table (NIT). Two tables that are relevant to our discussions are PAT and PMT tables. In a compliant MPEG-2 multiplex, there must be only one Program Association Table (PAT) that contains the list of services associated in the multiplex. The PAT in a multiplex associates each service number with a specific PMT PID in the same multiplex. PMT, in turn contains the list of PIDs associated with each service and other associated data. One field of interest to our discussion is the stream_type that is used to associate each component of a service identified by a PID with the type of elementary stream or payload carried within that PID.

CARRIAGE OF MPEG-4 ES VIA THE MPEG-2 TRANSPORT STREAM

This section discusses the encapsulation of MPEG-4 ISO/IEC 14496 audio-visual elementary stream in an MPEG-2 Transport Stream. As noted previously, MPEG has amended the MPEG-2 system standard ISO/IEC 13881-1 [1] to allow carriage of MPEG-4 content over MPEG-2 Transport and Program Streams. This amendment is included in the published 2000 edition of ISO/IEC 13818-1 *Ref*. [1]. According to this amendment, for the carriage of individual MPEG-4 elementary streams, only system tools from MPEG-2 [1] are used. This topology is depicted in**.** As shown in this Figure, MPEG-2 Link Layer and Sync Layer are used instead of MPEG-4 Link and Sync layers. Hence, elementary streams encoded according to MPEG-4 are carried in PES packets as PES_packet_data_types with no specific alignment. In another words, from a system point of view, encoded MPEG-4 audio-visual elementary streams are treated the same as MPEG-2 elementary streams. For example, elementary stream synchronization is accomplished according to MPEG-2 through decoding the PCR in the adaptation layer and the same time base is used to synchronize all the components of a service. This is contrary to MPEG-4 in which each component of a program could be synchronized to a different time base through the OCR.

In addition, Stream_Id values for video and audio elementary stream within the PES header have been defined by this amendment to indicate that PES payload contains MPEG-4 audio-visual elementary streams. As noted in the previous section, Stream_id is encoded in the PES header to indicate to the decoder what compression method is used. The new updated values for Stream_id could be found in table 2-18 of the 2000 edition of ISO/IEC 13818-1 *Ref* [1]. Furthermore, the 2000 edition of ISO/IEC 13818-1 amendment defines stream types that should be encoded in the PMT when MPEG-4 compression is used to encode MPEG-4 audio, and video elementary streams. As noted in the previous section, stream type is used to associate compression used for each component of a service to the PID that is pointed to by PMT table. This information is carried in the second loop of PMT.

New descriptors such as MPEG-4_video_descriptor() and MPEG-4_audio_descriptor() are defined by the 2000 edition of ISO/IEC 13818-1 for defining coding parameters of associated elementary stream. It is worth mentioning that these descriptors do not apply to the MPEG-4 elementary stream if MPEG-4 Link and Sync layers are used. These descriptors are carried in the second loop of PMT and flag to the decoder which level and profile of MPEG-4 compression was used to compress associated Elementary Streams. These descriptors can be found in the 2000 edition of ISO/IEC 13818-1, section 2.6.36 and 2.6.38.

CARRIAGE OF MPEG-4 SCENE VIA THE MPEG-2 TRANSPORT STREAM

This section discusses the encapsulation of MPEG-4 content which may consist of but not limited to: audio-visual, IPMP, OCI streams, Object Descriptor (OD), Scene Descriptor such as BIFS in the MPEG-2 Transport Stream. As specified by the 2000 edition of ISO/IEC 13818-1, these streams are carried in the SL_Packetized stream but use of FlexMux is optional since MPEG-2 offers multiplexing tools. MPEG has amended the MPEG-2 system standard ISO/IEC 13881-1 to allow encapsulation of MPEG-4 both SL_Packetized stream and FlexMux streams in PES packets as well as specifying additional descriptors and other relative fields such as stream_type and stream_id to aid the client side in distinguishing between MPEG-4 content and MPEG-2. Hence, one could use other functionality of MPEG-4 in the existing CATV beyond the improved compression that is provided by MPEG-4. This topology could be represented by Figure 4 and in more detail in Figure 6**.** In the balance of this section the features provided by the latest MPEG amendment for carriage of SL_Packetized and FlexMux in MPEG-2 Transport Streams are presented. First, a brief summary of Sync Layer and FlexMux features and functionality offered by MPEG-4 are provided as they relate to the discussion.

As depicted in Figure 6, the Sync Layer is located between the compression layer and the delivery layer. The Sync Layer provides a flexible set of tools that allow incorporating time base information, fragmentation of access units, and continuity information into data packets. The resulting packetization stream from the Sync Layer is referred to as the SL_packet stream.

The layer below the Sync layer is called the delivery layer that includes the FlexMux. The input to the FlexMux is the SL_packetized stream from the SL as shown in Figure 6. FlexMux is an efficient and simple multiplexing tool defined by MPEG-4 designed for low delay and low bit-rate streams. FlexMux was designed with low overhead since a presentation could have a large number of elementary streams.

Details regarding the Sync layer and FlexMux are beyond the scope of this paper and can be found in ISO/IEC 14496-1 system *Ref*[1].

Figure 6 illustrates the data flow from the compression layer to the Sync layer and then to the Delivery Layer. As depicted in this Figure, the Delivery Layer can be sub-divided into two sub-layers namely, DMIF that is specified by MPEG-4 and the TransMux layer that is not specified by MPEG-4 since different delivery layers already specify this layer. These layers may be any of: RTP, ATM or MPEG-2 just to name few. The following discussion focuses only on MPEG-2 TransMux.

The inputs to MPEG-2 TransMux are either SL_Packetized or FlexMux streams. The responsibility of this layer is to map these streams into the PES and TS structure defined by MPEG-2 with the following constraint: SL_Packetized streams are mapped into single PES Stream such that only one SL_packet constitutes the payload of one PES packet but in the case of TransMux, an integer number of FlexMux packets can be mapped into the payload of PES packets. MPEG-2 defines different a stream_id for PES packets with the payload of SL_packet versus PES packets carrying TransMux streams. The corresponding stream_id should be encoded in the PES packet header.

Furthermore, if SL_packets contain MPEG-4 Object Clock Reference (OCR) then PES packets carrying these SL_packets should have PTS in the PES header and a similar requirement applies to PES packets carrying FlexMux. In case of FlexMux encoded FCR, the timebase in the FlexMux packets should be carried in the corresponding PES header, if present.

Also, certain MPEG-4 streams such as Object descriptor and Scene Descriptor tables can be carried in the MPEG-2 section format. These data are normally static and are used for random access similar to the way in which PSI tables are used in the MPEG-2 system.

Newly defined stream_types are established for the PID streams carrying MPEG-4 content. These stream_types through PMT indicate to the client that the bit stream identified by PID in the PMT is a PES packet containing either SL_packets or multiple FlexMux packets, or that the bit stream contains MPEG-2 section data carrying OD or BIFS commands. Further, a set of descriptors for carriage of MPEG-4 [*Ref*] in MPEG-2

[*Ref*] has been defined. These descriptors provide more information about the stream and are included in the PMT. A list of these descriptors can be found in section 2.6 of MPEG-2 [*Ref*].

WHY MPEG-4

In the early days of cable television all systems focused on delivering clear analog broadcast channels to viewers in fringe areas beyond the reach of broadcast transmitters. As subscriber demand for more selection increased, additional channels were added to the analog line-up. MSOs soon discovered new sources of revenue from reservation PPV, IPPV and premium subscription channels, adding even more channels for these services. It soon became evident that traditional analog CATV system architecture could not support the growing service demand. MSOs were running out of bandwidth and had to upgrade their systems to modern Hybrid-Fiber Coax (HFC) and migrate services to MPEG-2 compressed, digital video transport in order to carry more channels and provide vastly improved picture quality. At the same time ever-increasing competition from the Direct to Home [DTH] satellite networks forced MSOs to seek service differentiation. MSO's began and continue today to offer enhanced features such as HD, VOD, Network PVR, VOIP, Streaming Media and Targeted advertising just to name a few *Ref* [9]. These advanced features pose new bandwidth on both the upstream and downstream HFC plant segments as explained in *Ref* [9]. Thus, once again, bandwidth resources in a typical HFC network are becoming scarce. In order to support these advanced features an MSO can choose to increase the physical capacity of the plant through a costly conversion to an all Fiber network such as FTTH. Or the MSO can optimize the existing plant bandwidth by employing advanced compression techniques such those offered by MPEG-4. Typically, MPEG-4 can provide two to three times better compression than MPEG-2 thereby lowering the effective bit rate without compromising picture quality. Thus a system migration to an MPEG-4 system can accommodate a three-fold bandwidth increase without a costly fiber overbuilds. Given the current investments in MPEG-2, it is not reasonable to retire all this relatively new equipment from service. But through a gradual transition to MPEG-4, as discussed above, it is possible to carry MPEG-4 over an MPEG-2 transport stream and use current or next generation set top boxes that could support this dual mode during a transition period similar to the transition that took place to migrate from analog to digital systems.

CONCLUSION

MPEG-4 is a promising, emerging technology that has been gaining momentum in the CATV industry. This paper has presented different mechanisms to carry MPEG-4 over MPEG-2 Transport Streams in current CATV systems that use MPEG-2 Transport Streams as the transport layer. The first scheme addressed the carriage of MPEG-4 Elementary Streams over MPEG-2 Transport Streams. This scheme enables a CATV system to take advantage of the improved compression offered by MPEG-4. The second part of this paper discussed the carriage of MPEG-4 content over MPEG-2. This scheme enables an MSO to offer the advanced features enabled by MPEG-4 in addition to the improved compression.

Other advanced features such as HD, VOIP, and various Streaming Applications can be supported by migrating to MPEG-4 and freeing the additional bandwidth needed by these applications. It is clear that MPEG-4 is proving to be an industry standard solution for increased efficiency in plant utilization, as well as a much-needed platform for the launch of diverse, evolving interactive services and overall feature enhancements.

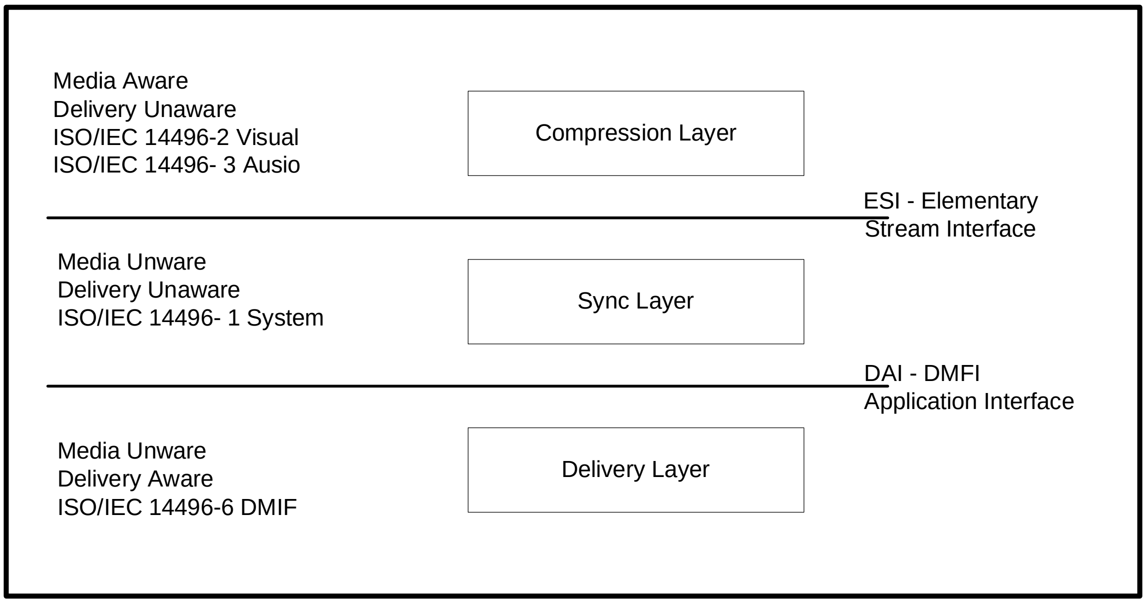


Figure 1 ISO/IEC 14496 SystemArchitecture

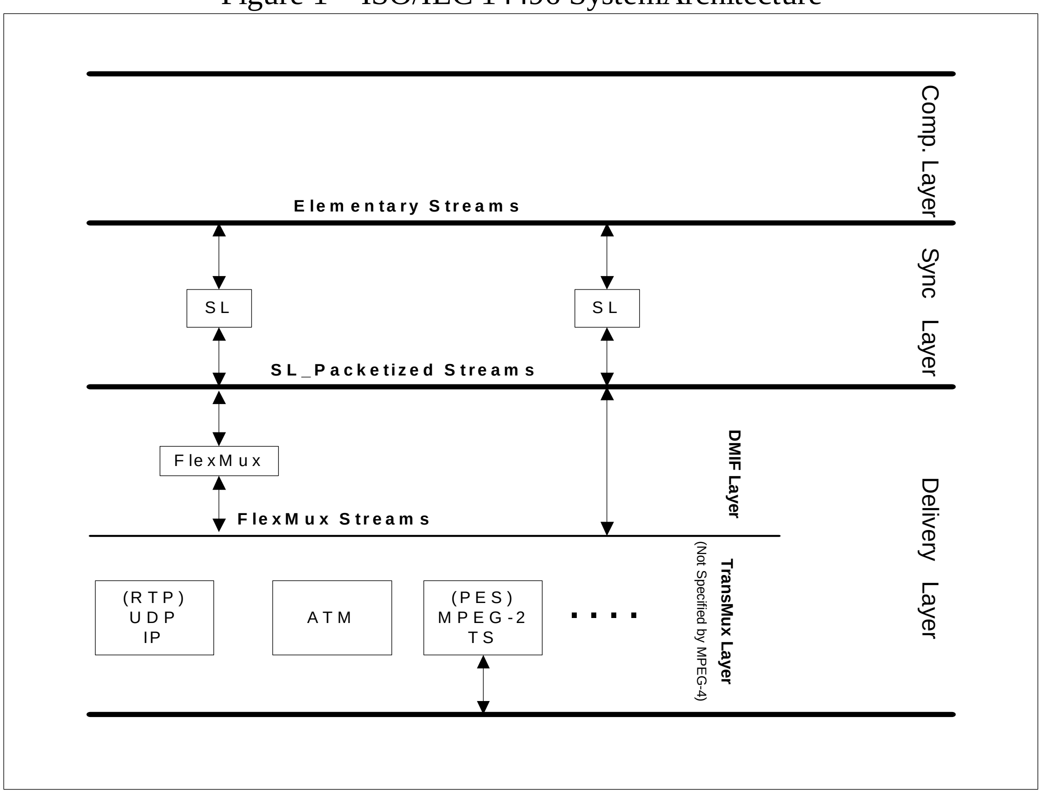


Figure 2 Detailed ISO/IEC 14496 System Architecture

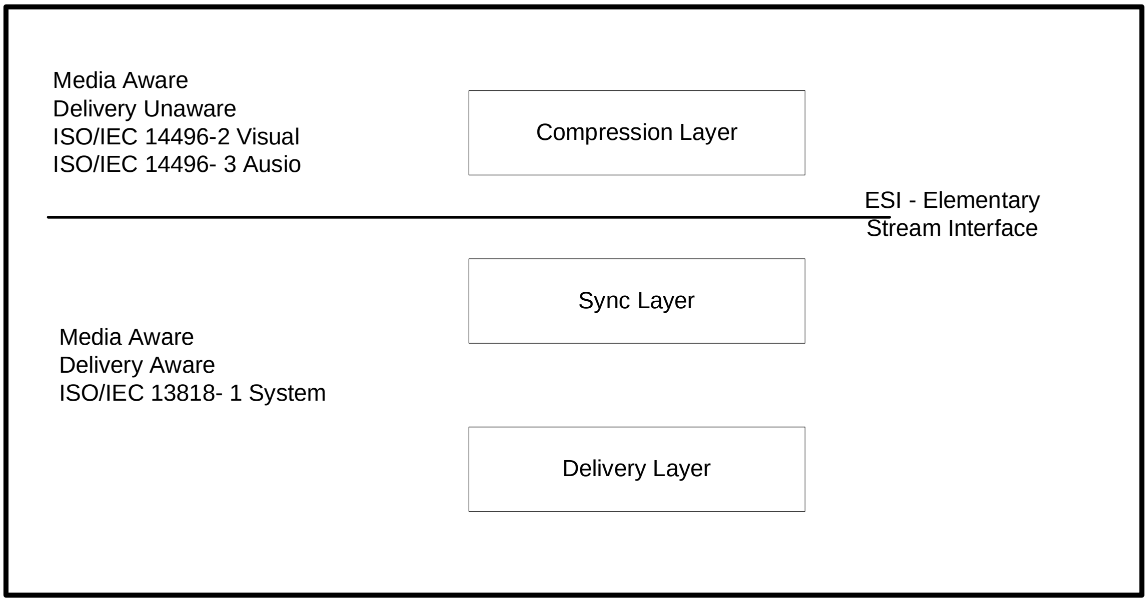


Figure 3 Carriage of MPEG-4 Elementary Stream Via MPEG-2 TS

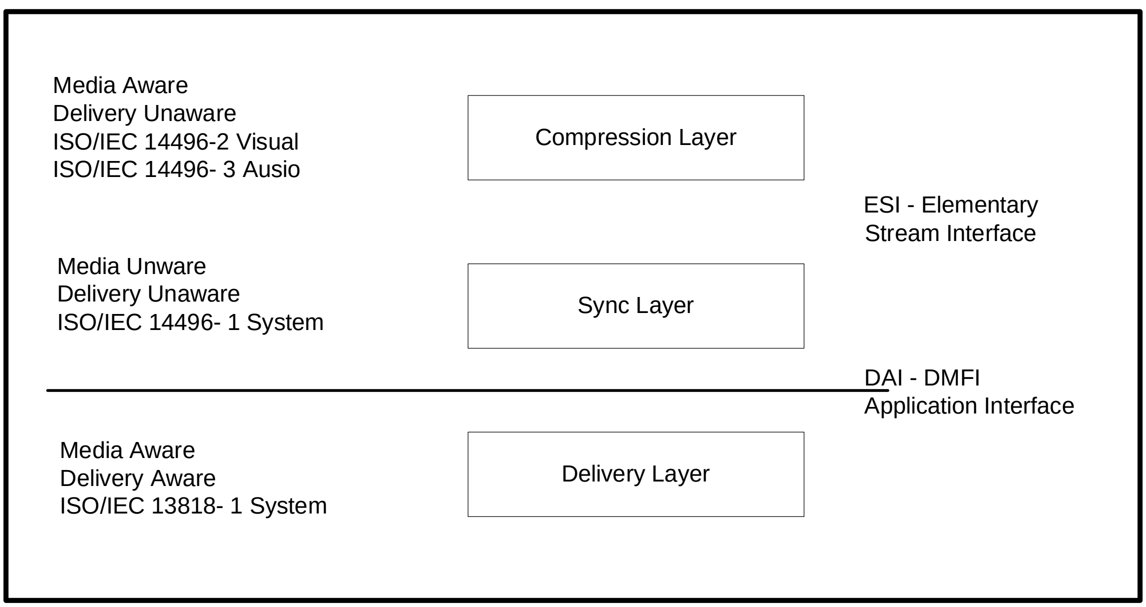


Figure 4 Carriage of MPEG-4 Content Via MPEG-2 TS

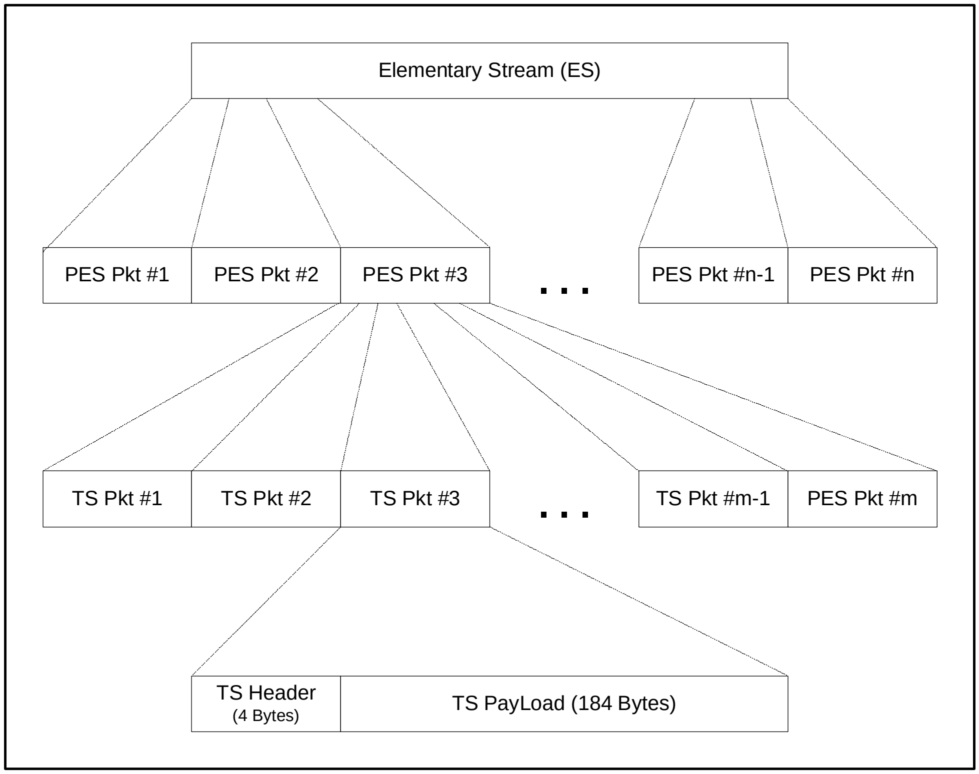


Figure 5 MPEG-2 Link Layer Hierarchy

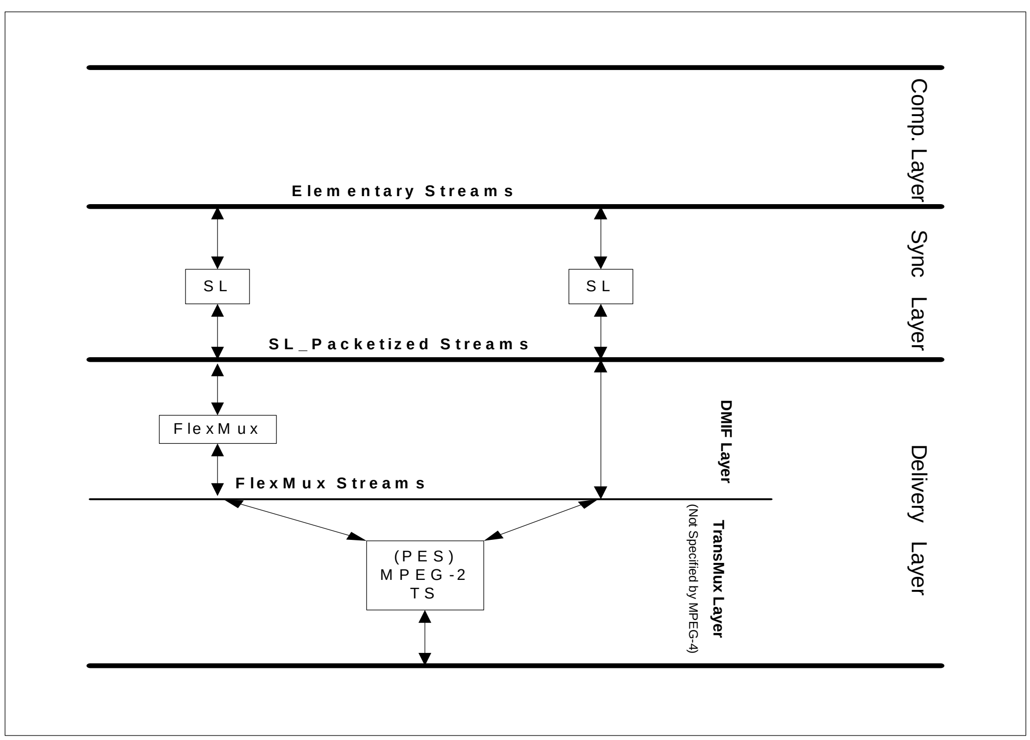


Figure 6 ISO/IEC 14496 System Architecture

REFRENCES

[1] ISO/IEC 13818-1, Information technology – Generic coding of moving pictures and associated audio information: Systems

[2] ISO/IEC 14496-1, Information technology – Coding of audio-visual objects Part 1: System

[3] ISO/IEC 14496-6, Information technology – Coding of audio-visual objects Part 6: Delivery Multimedia Integration Framework (DMIF)

[3] ISO/IEC 14496-1, Information technology – Coding of audio-visual objects Part 2: Visual

[5] ISO/IEC 14496-1, Information technology – Coding of audio-visual objects Part 3: Audio

[6] ISO/IEC JTC1/SC29/WG11 N4668, March 2002, Overview of the MPEG-4 Standard

[7] C. Herpel, A. Eleftheriadis, MPEG-4 Systems: Elemnentary Management

[8] G. Franceschini, The Delivery Layer in MPEG-4

[9] Use of MPEG-4 in Broadband Streaming Applications, Mukta L. Kar, Ph.D. and Nasser Syed, Ph.D. ,Sam Narasimhan, Ph.D. and Ajay Luthra, Ph.D.; NCTA Paper, Year

Author contact:

Ardie Bahraini
Sr. Director, System Engineering
Motorola
Broadband Communications Sector
(678) 957-3233
ardie@motorola.com

CONTROLLING AN INFINITE NUMBER OF CHANNELS

Doug Makofka
Motorola, Inc. Broadband Communications Sector

Abstract

A consumer sizes a typical cable system based upon the number of channels that can be accessed. Cable systems that are capable of delivering everything 'on demand' must be sized differently. The size of an on-demand system is related to the amount of media (movies, switched broadcast, live broadcast, etc.) in the on-line library. Systems capable of delivering any content 'on demand' present the experience of having an infinite number of channels.

Channel-oriented delivery systems and media-oriented delivery systems have significant differences. Media-oriented delivery systems need to offer generic, high-level delivery and transport functionality to the service and application control subsystems that manage the media. This is in contrast to channel-based delivery systems that only need to assign a service to a channel.

This paper presents an architectural and functional introduction to media-oriented delivery systems, including the ramifications to access control, bandwidth management, network management, and media transformation subsystems.

ON-DEMAND SYSTEMS STRUCTURE

Systems that are capable of supporting a large number of on-demand services are structured differently then systems built primarily to deliver broadcast services. On-demand services are transactional by nature. Some group of equipment in the network must actively process the request for content from the customer. This is in contrast to the broadcast service, which in most cases can be represented by a fixed channel map, and does not require any processing from the network in order for the tune operation to occur. A simple 'channel up' button push suffices. OnDemand systems require content caches of some type. This is in contrast to broadcast-oriented systems which simply act as a conduit between a service originator and the client devices.

On-demand services and the content/media associated with on-demand services can be managed separately from the delivery of the service/content. For this reason, OnDemand system management cleaves nicely into two pieces: Service-Application-Content support, and Media Delivery support. **Figure 1** shows the high-level structure of the OnDemand System, which reflects this division. This paper will give an overview of each subsystem, then focus on the Delivery Network, and how it provides resources to deliver 'infinite channels'.

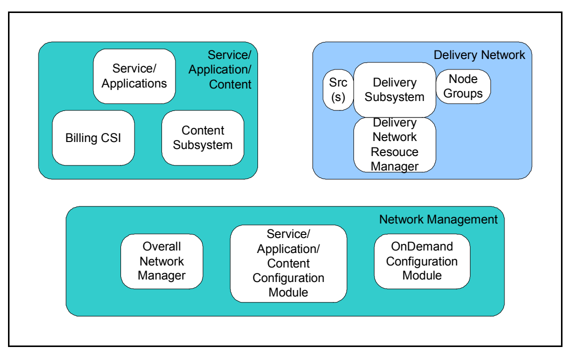


Figure 1 High-Level OnDemand System

Service/Application/Content (SAC) Subsystem

The Service/Application/Content (SAC) Subsystem has the following primary functions:

1. Determining how content is presented to the customer – host the service, access control;
2. Manage the content – including 'pitch-catch functions, and content distribution;
3. Primary billing interface;
4. Supports the automated provisioning and management of OnDemand services.

The SAC subsystem is where the majority of the added value of the current VOD systems resides. The Pegasus Interactive Services Architecture (ISA) standard is one definition of a SAC subsystem.

Access control is completely in the context of the SAC subsystem. This subsystem decides whether or not content/media will be sent to a client. The Delivery Network assumes that access control has already been applied. If it receives a request to deliver content it will. Note that encryption in the OnDemand system is client-based, not content-based. This is in contrast to broadcast systems where encryption is applied to content, and access rights are given to clients.

Delivery Network (DN) Subsystem

The Delivery Network is responsible for the delivery of content to the consumer/client device as directed by the SAC subsystem. The primary functions of the DN subsystem are:

1. Effectively manage DN resources (bandwidth, encryption, transcoding, insertion, QoS, etc);
2. Provide an open, high-level interface for requesting DN resources;
3. Manage and control devices from different vendors so that they cooperate seamlessly in the delivery of on-demand content;
4. Supports the automation of provisioning, configuration, and management of the DN.

In current VOD environments, the VOD systems are either the actual or *de facto* managers of the DN subsystem. This as been driven by the non-routable networking (ASI) between the VOD servers and the DN equipment (modulators, upconverters). With the advent of GIGE transport of content/media, the delivery of content is not necessarily determined by the output of the server. This is a strong motivation for removing the DN network management from the content environment.

Nework Management (NM) Subsystem

The Network Management (NM) subsystem, as it relates to On-Demand subsystems has the following functions:

1. Maintain a consistent view of system topology;
2. Maintain IP-level device configuration (DCHP device records, for example);
3. Maintain higher-level functional configuration for the Delivery Network;
4. Maintain and distribute higher-level resource to content source mappings between the Service-Application-Content subsystem and the Delivery Network subsystem.

The Network Management subsystem can 'see' across the two functional subsystems. It coordinates system provisioning and management activities. Even though the NM subsystem does not play an active role in the delivery of OnDemand services, it is needed to make large-scale OnDemand systems practical. Systems capable of delivering 'everything on-demand' will be quite large, and require automatic provisioning and management – to the extent that equipment can be placed in a rack, automatically determine its configuration and functional responsibilities when it powers up, without direct operator intervention. This will be discussed further in the *Automatic Configuration and Provisioning* section.

DELIVERY NETWORK (DN) STRUCTURE

The OnDemand system structure is based around the User-to-Network concept used in the DSM-CC User-to-Network Session protocol. It is natural to use this protocol as the basis of requesting and delivering resources from the DN; however, other protocols can also be accommodated – such as RTSP.

The Delivery Network (DN) offers 'resources' to the Service-Application-Content (SAC) subsystem. The SAC subsystem requests these resources from the DN as part of a session setup operation. The session is used to carry content/media associated with a service to a specific customer's client device. Examples of DN resources include bandwidth, multiplexing, encryption, transcoding, and rate shaping. A major goal of the DN subsystem is to provide these resources is a well-defined manner so that the SAC subsystem does not need to understand how these resources are mapped to specific devices in the network, or how to control those devices. This provides a basic decoupling of the DN and SAC subsystems' architecture.

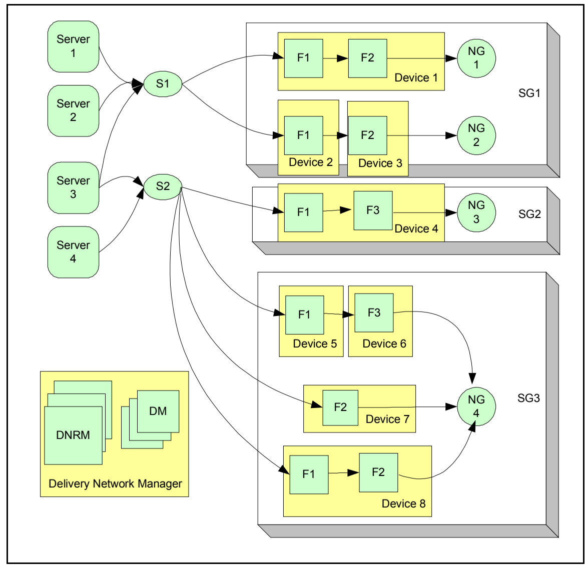


Figure 2 Delivery Network Structure

Figure 2 gives a high-level look a the structure of the DN. The arrows in the diagram show media paths. Control paths are omitted. The definitions of the elements in **Figure 2** are as follows:

Server1…n These are content/media servers. They are outside the DN, but are shown here as the connection points to the DN. The connections are likely GIGE, but ASI and other connections are also accommodated;

S1,S2 These are *Sources*, or SRC, which are the actual connection points into the DN. In the case of GIGE media transports, a Source is equivalent to an IP subnet. In the case of a ubiquitously switched system, there is only one Source;

Device 1...n These are the actual devices or products that are going to perform functions in the DN;

F1…n These are the *Functions* performed by the Devices (mux, modulate, encrypt, etc.). Each Function is described by a *Function Block*, or *F-Block* for short. F-Blocks map directly to the resources offered by the DN to the SAC subsystem;

NG1…n These are *Node Groups* – defined by the structure of the combining networks on the output of the Inband RF modulators. A Node Group represents the set of 'outputs' that can be seen by a particular client device. It also is a primary topological grouping of client devices;

SG1…n These are Service Groups. A Service Group is a collection of Node Groups. The motivating idea behind Service Groups is that they collect Node Groups that have the same Resources (F-Blocks). This helps to decouple NG topology from the SAC subsystem;

F-Block Chain Each path from a SRC to a NG is called an F-Block Chain. For example, there is an F-Block chain from SRC S1 through Device1, to Node Group NG1. Resources F1 and F2 are available on this chain;

DNRM The Delivery Network Resource Manager (DNRM) is the part of the Delivery Network Manager that handles the initial requests for resources from the Service\Application\Content (SAC) subsystem, allocates the resources to specific F-Blocks, then commands the Device Manager (DM) associated with the F-Blocks to do work;

DM The Device Manager (DM) translates standard F-Block behavior requests into device-specific commands.

Every device or product used to process media in the DN must be described as one or more F-Blocks. For example, referring to **Figure 2**, if F1 is a modulator, and F2 is an upconverter, then Device1 is a product that takes its input, modulates it, upconverts it, then outputs the processed RF signal. From the standpoint of the Domain Network Resource Manager (DNRM), it doesn't matter what Device1 is. Either the device supports standard F-Block commands directly, or there is a Device Manager (DM) that translates the standard modulator, and upconverter F-block commands into device-specific commands. The following is a partial list of potential F-blocks:

1. RF Modulator (QAM64, QAM256, etc.);
2. Upconverter;
3. Encryptor;
4. Inserter;
5. Multiplexer;
6. JitterBuffer/QoS shaper;
7. Transcoder;

<u>Delivery Network (DN) Transactions</u>

The DN subsystem provides resources used by the SAC subsystem to deliver services. There are many potential on-demand service types. Video On Demand (VOD) and its variants, Network Personal Video Recording (NPVR), Switched Broadcast, and others. The DN is unaware of these service types; however, all of these service types are accommodated by the DN using the same set of resources.

A User-Network model is used to structure the DN transactions. The SAC subsystem, and the customer's client devices are each Users. The DN is the Network. The primary transaction structure is one of the Users requesting sessions for media to be delivered by the DN. These sessions have resources associated with them. Bandwidth is the fundamental resource, but there are others – those defined by the F-Blocks supported by devices in the DN. The typical session setup trans action is shown in Figure 3.

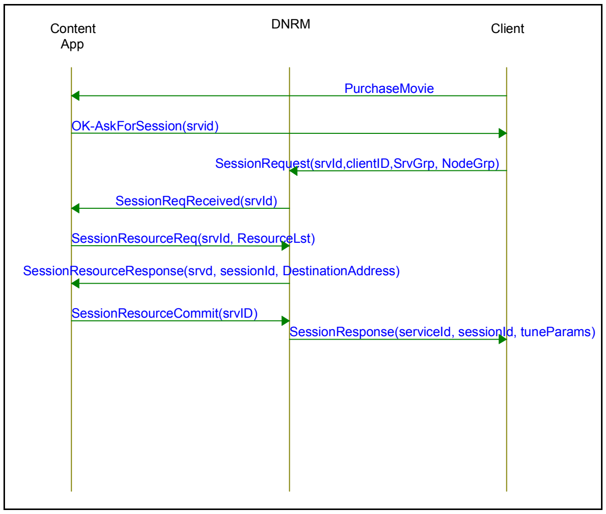


Figure 3 Session Setup Transaction

Notice that the session setup transaction is compatible with the DSM-CC U-N Session Setup Protocol. RTSP could also be accommodated by adding a few new functions. Another protocol variation is having the Content/App subsystem issue the SessionRequest command, or by just issuing a SessionResourceRequest indicating a new session is to be established. The DNRM could handle both scenarios from either protocol without any problems.

Managing the Delivery Network

Figure 4 shows the standard control flows in support of the DN. Notice that there are resource/session flows, and network management flows. Each are needed to support the OnDemand system. There are two levels of management in the DN:

1. F-Block or resource management
2. Device management.

Requests for resources are always at the F-block level. DN resources are allocated on the basis of F-Block allocation. Each F-Block type as an associated allocation scheme that allows the Delivery Network Resource Manager (DNRM) to know where unallocated resources reside in the DN. The DNRM handles the resource allocation based on the F-Block allocation scheme. The resource request is then translated to F-Block parameters which are passed to the Device Manager (DM) associated with the device that will provide the requested F-Block function.

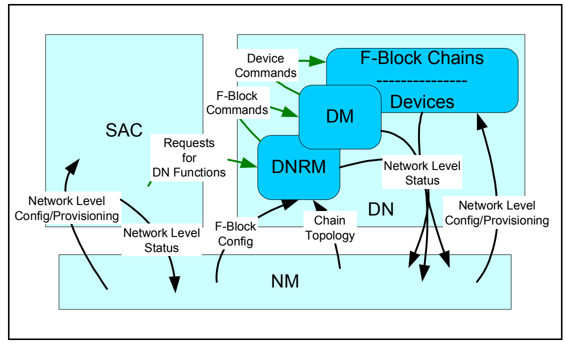


Figure 4 Management Control Flows

The Device Manager (DM) accepts the F-Block function request and issues the proper commands to the device. The DM may also request that the DNRM adjust the 'available resource pool' for the F-Block in questions. This allows device resource allocation to deviate from the standard F-Block allocation scheme.

Automatic Configuration and Provisioning

Delivering on-demand services to large subscriber populations requires networks that must support a large number of simultaneous 'channels'. These systems can be orders of magnitude bigger than broadcast-oriented system. In addition, there is are on-going control transactions between subsystems. The control transactions demand that each subsystem is 'in sync' with at least some portion of the overall topology and state of the OnDemand system. DN topology representation is an important part of automatic OnDemand system configuration.

OnDemand System Topology

The basis of DN topology is Network Topology. An example of Network Topology is given in Figure 5.

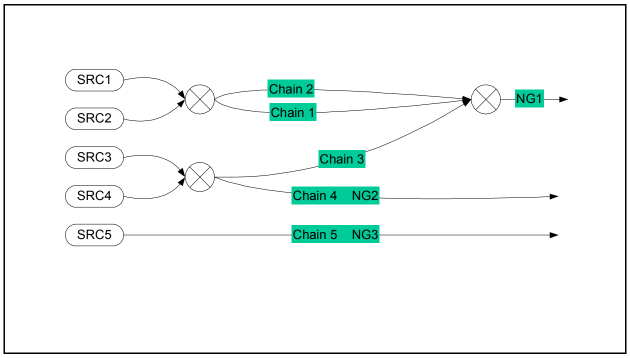


Figure 5 Network Topology

Network Topology is the framework networking that maps the DN sources into Chains, that reach Node Groups. In a GIGE-based system, Network Topology is a subset of the GIGE IP network configuration. This shows the close connection between the DN configuration and provisioning and the underlying network configuration – in this case IP.

Once Network Topology is defined, devices must be placed in the Chains between the Sources and Node Groups. This representation of topology is called Chain Topology. An example of Chain Topology is given in Figure 6. Chain Topology is the configuration information used by the Delivery Network Resource Manager (DNRM) to turn requests for resources into F-Block allocations.

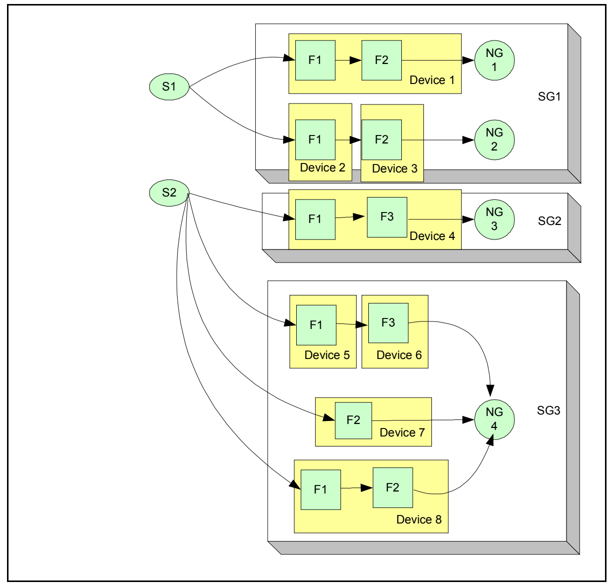


Figure 6 Chain Topology

The Service/Application/Content (SAC) subsystem may need some DN configuration information in order to efficiently distribute content to servers. Source Topology meets this need. Figure 7 shows the picture of the DN conveyed by Source Topology.

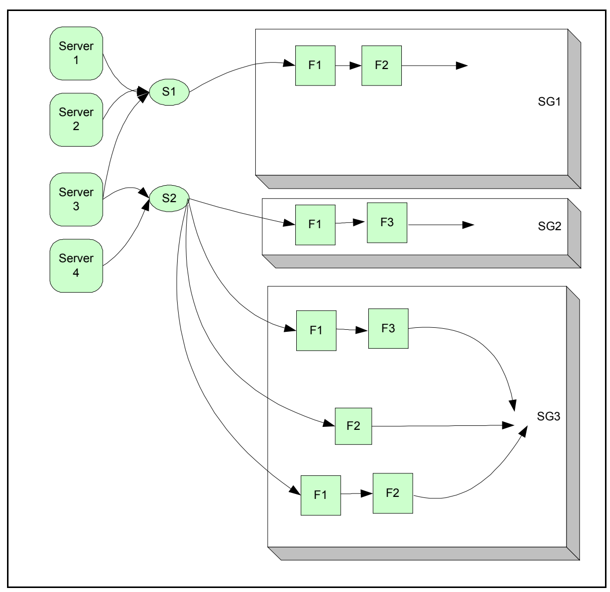


Figure 7 Source Topology

The SAC subsystem can use the DN without Source Topology. The content/media used by the services supported by the SAC will require specific resources of the DN. The

SAC could just request those resources without any notion of Source Topology; however, some optimizations in the SAC environment are possible with a knowledge of Source Topology. One example is that content/media requiring a resource not available in a specific Service Group could be hidden from customers in that service group. This would prevent requests for content it is impossible for the DN to deliver.

Delivery Network Configuration Module

The Delivery Network relies on topology, F-Block, and Device information. This information must also be used to drive or modify IP-level configuration information, such as DHCP records. This information cannot be hand-crafted. It must be formed from higher-level configuration operations. The Delivery Network Configuration Module (DNCM) automatically generates the topological and configuration information. Graphic interfaces are used to 'stage' the OnDemand system. The graphic interfaces allow manipulation of the topological diagrams shown in this paper, as well as detailed F-Block and device information.

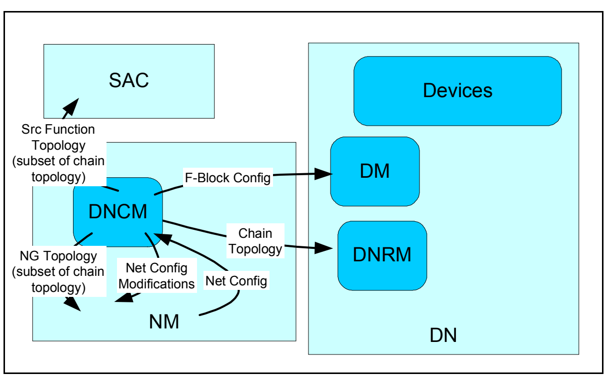


Figure 8 Delivery Network Configuration Module (DNCM) Environment

Figure 8 shows the environment in which the DNCM functions. It also shows the major information flows from between the DNCM and the other OnDemand System entities.

Among the DNCM's major operations are the following:

1. Load/Modify/Delete Network Layer Topology;
2. Load/Modify/Delete F-Block Defs;
3. Load/Modify/Delete Device Defs (each device participating in the DN will supply a device definition package);
4. Load/Modify/Delete NG Defs;
5. Configure Chain Topology;
6. Add /Remove Device to/from DN;
7. Enable/Disable Device in DN;
8. Associate Network Layer With Device Layer (effect any coordination between the subsystem managing DHCP);
9. Build Configuration Script;
10. Execute Configuration Script;
11. Configure DNRM and DMs (these managers need their own configuration based on system size, dedunancy strategy, etc.).

Ideally, an MSO can have individual systems configured off line by a knowledgeable systems engineer. These configurations can be sent to the locations where the DN exists. Technicians at the DN sites can then 'rack and stack' the devices needed in the DN. Using the DNCM, the technicians at the DN site can apply the configuration to the newly-installed devices without any additional configuration operations. Ongoing configuration and device changes can be handled by the local technicians at a high level. The DNCM will be able to d sanity checking on these changes. The DNCM can then make sure the changes are applied in a controlled, consistent fashion across the entire OnDemand System.

CONCLUSION

OnDemand systems are different the broadcast-oriented systems, hence, require a different structure and management strategy. Yet, it is possible to manage and control large OnDemand systems. Managing Delivery Networks is more than just assigning bandwidth. All Delivery Network functions must be available to the users of the Delivery Network via high-level functions that are well-defined, open, and allow competition among device vendors supplying products that provide Delivery Network Functions. Creating a Delivery Network management subsystem, that operates independently from the Service/Application/Content subsystems, will make delivering everything on-demand a technical and practical reality.

DELIVERING EVERYTHING EVERYWHERE IN THE HOME: WHOLE HOME NETWORKING

William Garrison
Thomas du Breuil
Motorola, Inc., Broadband Communications Sector

Abstract

This paper will describe the requirements for an in-home network. Specifically, it will address the data rate requirements for the various services and their delivery. Detailed descriptions of QoS requirements and the relationship between the Entertainment and Data devices and services will be given. Various wired and wireless networking technologies currently in the marketplace as well as new technologies that might serve this need will be covered Finally, a view of life in this home of the future will be explored.

INTRODUCTION

With the growth of broadband data services, many consumers have found it useful to install a data network. According to a 2002 Parks Associates report, 7.2 million homes now have a data LAN and this number will grow to 21.2 million in 2006. The primary application for these LANs is to share the broadband data access with multiple PCs in the home, but it is also used for printer sharing and file sharing.

Consumers are now buying PVRs and quickly realizing the benefits of video access via a hard drive. An obvious extension of this will be to access to this video content anywhere in the home. Wouldn't it be nice to view a program stored on your PVR downstairs on a TV upstairs? And consumers will want to add other media, such as music and photos, to this network as well as merge it with their data applications.

So why not just use the existing in-home data network for this new video application? Well, the reason is that requirements for an in-home data network are much different from an in-home entertainment network. An in-home entertainment network needs to support multiple entertainment streams (some at HDTV rates) with excellent QoS (Quality of Service). This network also must support other types of traffic, such as music, Internet, data and photo transport. Once this in-home network is in place, Voice-over-IP, i.e., telephony, and video telephony can be easily added. Cable operators have a unique opportunity in these in-home networks because they understand delivering audio and video best. But what is required to deliver these services?

REQUIREMENTS

Data Rates

A whole home network should be an infrastructure built to serve for a long time. Just as with AC power, you would never want to rewire your house just to add a new appliance – even if that appliance did not exist at the time you wired your house. Therefore both current and future needs must be considered when defining this network.

Relatively few homes currently have a HDTV display. Approximately 4.5% of all television households have a HDTV display now. However, 26% are expected to have at least one HDTV display by 2008. Currently, the average household has 2.7 TV sets and by

2008 some of those homes will have multiple HD displays. Table 1 outlines the typical services that may be expected in a fully networked home and their bandwidth requirements.

Application	Qty	Rate each Mbps	Total Rate Mbps
HDTV stream	1	19.4	19.4
SDTV stream	3	4.5	13.5
CD Stereo Audio	1	1.5	1.5
Multichannel Audio 5.1	1	4.5	4.5
DVD Audio, 6 channel	1	10	10
IP Data	2	1	2
IP Telephony	4	0.032	0.128
Total			**51**

Table 1. Home Network Bandwidth Requirements.

Quality of Service

Video requires a much higher QoS than data. Many networks provide reliable service by retransmitting a packet until it is successfully received. This is the correct approach to use for delivering data. However, video has a timeliness factor measured in milliseconds (or less!). If the video data is not delivered by the presentation time, it would be better to skip this packet and move on to the next.

In addition, a MPEG-2 TS (Transport Stream) has a jitter tolerance measured in nanoseconds. A common "solution" to the jitter problem is to use a large buffer at the receiver. This is demonstrated by most current PC streaming media players, where 5-10 seconds of video is buffered prior to playing. However, entertainment video is often interactive, so "solving" the jitter problem with a large buffer at the receiver will result in a system that seems "sluggish". Again, this is the typical experience with streaming media today on the PC where it takes several seconds to start playing or to resume play after pausing. And even with a large buffer, video glitches are common today in the streaming environment.

IEEE 802.11e is currently being developed as a standard QoS mechanism for wireless systems and promises to provide a QoS which meets entertainment video requirements. IEEE 802.11p exists for CAT-5 wired LANs, and HPNA 2.0 includes a prioritized QoS, but these schemes do not support entertainment level QoS. HomePlug 2.0 does not support entertainment level QoS. However, HomePlug AV (the next version of HomePlug) does plan to support entertainment level QoS.

Data and Entertainment

Data delivery is focused on accuracy and video is focused on timeliness. Video decoding is purposely designed to conceal errors while data transfers require perfection. Can both of these coexist on the same network? What tradeoffs need to be made between these?

HOME NETWORKING TECHNOLOGIES

Existing

If a home has an existing network, it is likely to be either a wired 10/100 Ethernet or a wireless 802.11b Ethernet. Unfortunately, neither of these is suitable for a whole home entertainment network. While 100 Mbps Ethernet is fast enough, it does not offer QoS. Plus, as a practical matter, few homes have Cat 5 cable running to all of the places where you would like to network. 802.11b offers neither the QoS nor the data rate required by an entertainment network. So, what else might be used for a whole home entertainment network?

Wired

Wired networking generally offers the highest data rates and the lowest device cost. A wired network could use dedicated wires (like Cat 5) or reuse existing wires (like phone line, power line or coax). Unfortunately, none of the currently available wired networks offer the bandwidth and QoS required by an MPEG-2 Transport Stream. HomePlug AV is the only proposed wired standard that promises to address this need, but the standard has yet to be defined and first products will not be available until Summer 2004. There are several proposed proprietary solutions for networking-over-coax that meet whole home networking requirements for Bandwidth and QoS, but none of these are adopted industry standards.

Wireless

Current wireless technology includes IEEE 802.11a, 802.11b and 802.11g. If an adequate QoS could be layered above it, IEEE 802.11b could theoretically support a standard definition video service with a stereo audio service. However, it certainly can not be the backbone of a home with the requirements of Table 1.

The data rate for 802.11a and 802.11g is adequate for most of the service set shown in Table 1, although they too will not handle the full service set. Why won't 802.11a or 802.11g handle 51 Mbps when it is advertised as a 54 Mbps standard? Because the effective payload rate is less than the advertised PHY rate. The advertised raw data rate does not subtract the MAC overhead and other inefficiencies. Table 2 shows the effective data rage for common networking technologies.

Home Networking Technology	Media	Raw Data Rate, Mbps	Approx. Effective Streaming Throughput, Mbps
100 Mbps Ethernet	Cat 5	100	90*
HPNA 2.0	Phone Line	10	6*
HomePlug 1.0	Power Line	14	6*
802.11b	2.4 GHz	11	5*
802.11a	5 GHz	54	20* - 34**
802.11g	2.4 GHz	54	13.5* - 34**
Magis Air5™	5 GHz	54	40**

Table 2. Existing Home Network Technologies

*ExtremeTech™ test results
**Theoretical limit

Also note these wireless standards do not include any provisions for Quality of Service, so in practice, they can not deliver a satisfactory media delivery experience without a large decoder buffer. 802.11e specifically addresses QoS through a prioritization scheme and may solve much of the QoS deficiency when it is approved. Meanwhile, proprietary solutions, such as Magis Network's Air5™ were designed specifically to meet the needs of video and audio distribution reliably.

Wired or Wireless?

Wireless networking is a must for portable devices. Every networked home will have portable devices and so every home will need a wireless network. So does a home already equipped with a wireless network also needs a wired network?

The likely answer is you will need both. Wireless is required for portable devices, but it may not reach all parts of the home, it may not be able to deliver enough throughput, and is subject to interference. This is acceptable for a portable device, but not for the backbone of a home entertainment system. In addition, portable devices are normally battery powered, which limits their processing power and hence the bandwidth they need from their connection to the home network. Wired devices generally have no such limitations and their bandwidth requirements will only grow with time.

EVOLUTION OF THE NETWORKED HOME

Where are we today? Today, many homes have RF distributed by coax and data networked by Ethernet or 802.11 wireless. In addition, more and more homes have an Entertainment Gateway that uses a hard drive to store various content that is received, usually referred to as a PVR in the current configuration.

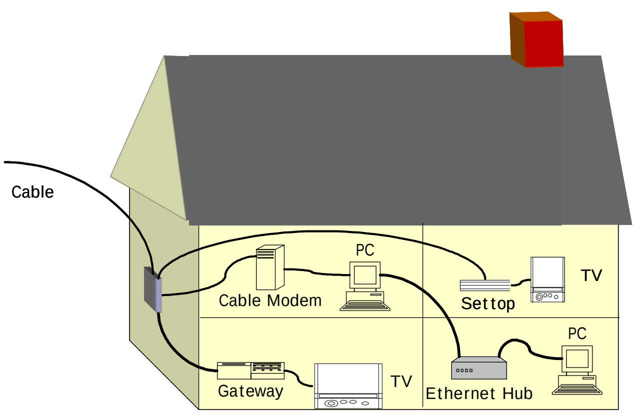


Figure 1 – The Current Home Network

There is a loose coupling between the RF and data worlds, in that the PC connects to a cable modem in order to connect to the Internet. However, for most purposes the two worlds of entertainment and data remain two separate worlds. Their closest linkage might be the DVD disk that can be played in either the entertainment center's player or the PC.

So, what does the networked home of today offer? As shown in Figure 1 in an Ethernet configuration, the consumer has these capabilities:

- Shared broadband for multiple PCs
- PC printer and file sharing
- Stand-alone PVR
- Digital Television and HDTV
- VOD and Impulse Pay-Per-View
- Audio sharing – MP3 to home entertainment center, digital audio to PC, etc.

AN INTEGRATED DATA NETWORKED HOME

The next step for home networks will be to add the entertainment devices to the data network. While this is less than ideal, it will add value to both the entertainment and data devices at very little cost.

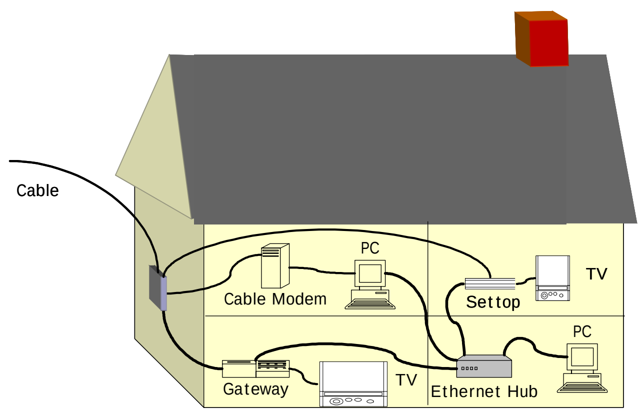


Figure 2 – The Integrated Data Home Network

What will the integrated data networked home of Figure 2 offer?

- Low bit rate video (< 1 Mbps) between PC and Gateway (with latency and some glitches)
- Archival storage – when your PVR disk is full, use unused capacity on your PC
- Remote access – move content in slower-than-real-time from one PVR to another for delayed remote viewing

- Pictures stored on the PC displayed on the TV

A FULLY NETWORKED HOME

In a fully networked home as depicted in Figure 3, the network backbone is robust enough to support any in-home application. Entertainment, data and voice applications are fully supported. Location does not matter. If your favorite program is recorded somewhere in the house, you can watch it anywhere in the house. If your favorite music is on any device in the house, you may listen to it on any device in the house. Format conversions are handled seamlessly.

What will the fully networked home offer?

- Quality video to/from the Gateway and PC, including high definition content
- Multiple high quality audio streams, including home theater
- Watch high definition TV on your PC even if you don't have a high definition TV
- Watch high definition TV on a standard definition TV via Gateway format down conversion
- IP telephony/video telephony

And what about your car? Why wouldn't you want to be able to listen to your favorite MP3s while on the road? Your car could automatically download the most recently played songs plus any ones you specifically designate every time you return home.

Note in Figure 3 the number of wires and devices goes down. This is because the best network is an invisible one. Communication can be via a coax network, wireless, Power Line or any combination of acceptable home entertainment networking .

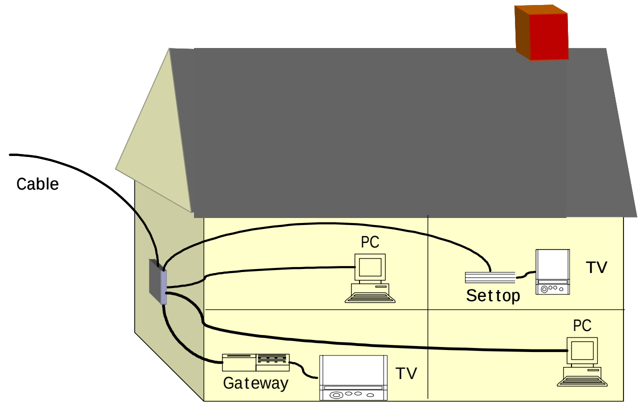


Figure 3 – Fully Networked Home

Where did the cable modem of Figure 2 go? Well, the home is sharing the cable modem that was already built into the Gateway. Future devices will have multiple network interfaces to make connecting as easy as possible for the consumer.

Whole Home PVR Scenario

You have just returned home and need some entertainment. So you plop down in the nearest chair and pick up the remote. Let's see, what is available? You want *Video, Recorded Programs, News*. Your home system knows that you like to get the latest news, and always records the most recent network news show for you. You don't know which device in the home recorded it (my PC? my settop?) and there really is no reason why you should care.

After you make your selection, the news starts. Well, after the first headlines, all you want is the sports. So, you fast-forward to the sports and see how your favorite team did. They blew the big play? You quickly go back to the menu and access the "Everything on Demand" system offered by my MSO, find the game, and Fast Forward to see that play. Yeah, they really blew it.

What Was Going On Behind The Scenes?

Your home devices have been autonomously recording content, based on your preferences. Some of your preferences were specifically enumerated when you set the system up, others were inferred by monitoring how you used the system. But when you plopped down, a content manager that was cognizant of every device in the network put it all together for you in one place.

After you made your selection, the first thing that happened is your current display device negotiated with the device that held the content. What is the best format to use? What is the best data rate? What QoS is available. As an example, presume the news was recorded in HD, but the in-home network is busy and only 5 Mbps is available with the QoS that you need. So, the network reserves 5 Mbps for this session and source device down-converts the news to a new data rate under 5 Mbps.

You start watching the news and decide to Fast Forward. The local device sends a message to the source, which starts the Fast Forward. Because the QoS minimizes the amount of buffering required at the display device, you see the news speed up within 200 milliseconds.

When you decide to go look at the big play, you are leaving you home network. Or are you? Your home system can record everything, so when you want something that is not available locally, you can fall back on your MSO to get the content. But the MSO might have known that many of their customers were going to look at that game, and so "pushed" the content to your home ahead of time. From your chair, it should not matter.

However, from the network's perspective, it does matter. For content from outside the home, the home network had to negotiate with a video server in the MSO's to select the content and setup the session. Playing that content now requires QoS all the way from the MSO's headend to your TV. Not a small challenge, because it spans multiple network domains.

Did the video come over a wired or wireless network? If the person flopped in the chair knows, then we have failed. The network needs to work seamlessly and invisibly.

CONCLUSION

The biggest remaining question is "When will all this happen?" The current home network isolates the entertainment and data networks. However, new products (such as Replay's and TiVo's latest generation devices) are starting to link the data and entertainment worlds. This is a start, and will likely grow over the next few years.

Full whole home entertainment quality networks are probably 3-4 years off. The devices required to build such a network will be available to early adopters at boutique prices early in 2004, but mass marketable whole home networks are probably still a few years out. Standards have to be established and production volumes must ramp up before price and ease of use meet mass market requirements. And there is a lot of software development required to make the network invisible and user friendly. The average consumer must be able to take a new device home and plug it in and find that it will simply work – like magic.

DOCSIS 1.1 – WHERE GAMING AND QUALITY OF SERVICE (QOS) INTERSECT

Kenneth Gould
Time Warner Cable

Abstract

Broadband on-line gaming is poised to be a key usage demand for residential high-speed data customers. With the recent releases of hugely multiplayer games such as Ultima Online and the availability of network enabled gaming consoles such as the Microsoft Xbox and Playstation 2, there are increasing opportunities for MSOs to cater to (and profit from) the demands of the broadband gamer.

VIDEO GAMING IS NOT A GAMBLE

Given the public launch of broadband-enabled gaming consoles in the last year, such as the Microsoft Xbox and Sony Playstation 2 consoles, considerable interest in cable modem service has been generated within the gaming community.

By January 7th, 2003, Microsoft announced that more than 250,000 subscribers had signed up for the Xbox Live service that was launched on November 15th, 2002 – this is twice as much as initial sales projections [1]. With 21.5 million Sony Playstation 2 consoles shipped to North America as of January 9th, 2003 [2], one can expect that quite a few owners will opt to purchase network adapters allowing for on-line game-play over a cable modem. To a lesser extent there is still demand for network access from Nintendo GameCube customers and the customers with the more aged Sega Dreamcast.

As the console gaming industry is a multi-billion dollar industry within North America [3], where are the opportunities for Multiple Service Operators (MSOs) to provide gaming services that provide added value to their customers and subsequently results in new revenue streams?

The most obvious possibility is to simply use gaming to attract new high-speed data customers to cable modem service. Every Xbox console is manufactured with an Ethernet port that is the sole interface for networked-based games. Xbox Live games are typically written for network play with the assumption that the bandwidth available will be less than 64Kbps upstream and downstream. It is relatively easy to provide a DOCSIS configuration file for a cable modem that limits its bandwidth consumption to 64Kbps. Likewise, the physical location of the Xbox relative to the physical location of a cable modem within the home is not a real problem given the availability of wireless Ethernet bridges and wireless-equipped cable modems. This opens up our potential pool of customers beyond households containing PCs.

The question is: Can MSOs offer this product without cannibalizing its existing high-speed data customer base? One stumbling block is that while it is easy to limit a cable modem service to 64Kbps, it is far more difficult to limit a service to only support console gaming. While the IANA list of well-known port numbers describes both TCP and UDP ports 3074 as being the “Xbox port”[4], our observations have shown that Xbox Live games use a wide variety of ports, of which 3074 is merely the most used. This greatly limits an MSO’s ability to create filters on a cable modem to allow Xbox traffic yet disable the customer’s ability to attach his PC to a "gaming cable modem" to surf the web or run peer-to-peer applications. Likewise, attempts to filter traffic based upon the MAC address of the console are fruitless given the end-user's ability to change the Xbox's MAC address at will. Similar behavior is seen from Sony Playstation 2s.

On a practical operations note, typically ISPs like to sign up customers with the minimum of paperwork. Customers are usually instructed to accept the ISP's "Terms and Conditions" electronically on a web page. This proves to be challenging for a new gaming-only customer to complete using only a gaming console.

It is worth pursuing the concept of attracting new customers from households which either only contains gaming consoles or which contain both consoles and PCs but have not yet opted for cable modem service, at a service tier whose bandwidth is less than the typical residential high-speed data tier. The MSO's market trials are still in their infancy, and there is not yet enough statistical data to determine whether offering lower-priced gaming tiers will cannibalize higher-priced PC-centric tiers, but anecdotal observations have so far indicated that downgrading very seldom occurs.

Co-Location Opportunities

Game publication is a multi-billion dollar revenue generator for large game publishers such as Electronic Arts [5], and as a result, these publishers spend a great deal of time and money to ensure that the servers on which the games are hosted are highly available, scalable to the number of customers playing, and well located within the network to provide low-latency gameplay. The Xbox gaming servers seem to provide a consistent "feel" to the gameplay as the servers for each title are managed by Microsoft. The game servers for PS2 games are not maintained by Sony, but are maintained by the individual publishers. Regardless of the model for server maintenance, the argument can be made that co-locating the gaming servers within an MSO's network can be a win-win situation for the publishers and the MSO. The MSO's customers experience even lower network latency which should make the customer and the publishers happy and the MSO benefits by keeping more gaming traffic on their network and off of the backbone.

Quality of Service Opportunities

What value-add could an MSO possibly bring to a gaming experience for which a gamer might actually pay? After all, console gaming over a cable modem works well today. One differentiator for MSOs is the ability to offer quality of service (QoS) guarantees. While console gaming works well today in a purely best effort data environment, MSOs will soon be offering many new services which will constrain how much bandwidth is available for best effort services. A perfect example is the offering of Voice-over-IP (VoIP) in a DOCSIS 1.1-enabled network. Assuming that the VoIP traffic is being transmitted using the DOCSIS Unsolicited Grant Service (UGS) traffic flows on the same upstream and downstream channels as the traditional Best Effort services, then for each phone call being made through a CMTS, there is obviously less bandwidth available for gaming.

In a bandwidth constrained environment, would gamers pay to possess a guaranteed amount of bandwidth and guaranteed latency dedicated to console traffic? Probably. One can argue that as new services are rolled out, MSOs will also be rolling out more efficient equipment (higher modulation profiles, DOCSIS 2.0, etc) that will offset any bandwidth constraints created by new services. The counter-argument is that this is unlikely given customers' penchant for consuming all bandwidth available to them, and even were it true, gamers may still be willing to pay a small fee just to achieve guaranteed low latency for their consoles. Gamers are constantly looking for an edge over their on-line opponents and are convinced that low latency gives them that edge.

QoS – Background

The DOCSIS 1.1 specifications created a foundation upon which products with quality of service requirements such as latency and bandwidth can be built. There are essentially two mechanisms for defining quality of service, "provisioned QoS" (pQoS) or "dynamic QoS" (dQoS). The parameters dictating the pQoS settings are pre-defined in the DOCSIS configuration file that the cable modem receives at the time that it boots. The DOCSIS configuration file would typically define a classifier that determines which packets are affected by the defined quality of service rules. The packets that meet the classifier's parameters make up a unidirectional stream of packets known as a "service flow". For example, since the majority of Xbox gaming traffic is transmitted to and from port 3047, a classifier can be defined which places all UDP or TCP packets transmitted to, or received on, port 3047 onto a particular service flow. That service flow has QoS parameters associated with it, such as a scheduling type (e.g. real time polling vs. best effort) and latency requirements (e.g. sub 150ms). All other traffic could default to a standard best effort service flow which would have a lower transmission scheduler priority.

Obviously, the DOCSIS 1.1 specifications only handle reserving and allocating bandwidth within the DOCSIS domain, specifically between cable modems (CMs) and the cable modem termination server (CMTS). End to end QoS can be setup with a combination of DOCSIS 1.1 and DiffServ or MPLS.

Dynamic QoS is typically used today in PacketCable-based voice over IP (VoIP) deployments. In this case bandwidth is reserved for voice calls "on the fly" between the cable modem and the CMTS only when a message is generated that indicates that a customer's phone has gone off-hook. An extension to the voice-centric PacketCable specifications is a promising possibility for future gaming services. The primary functions defined by the PacketCable VoIP specifications are QoS authorization and admission control, generation and capture of billing information, and security. These are all functions desirable in a QoS-aware gaming environment. CableLabs has been working on an extension of these specifications, known as PacketCable Multimedia [6], which expands the residential voice-centric specifications to be a general purpose platform for delivering many IP-based multimedia services that depend upon QoS. Note that while the PacketCable Multimedia framework is based upon the VoIP PacketCable specifications, the implementation of a VoIP PacketCable service is not a pre-requisite for PacketCable Multimedia-based gaming as gaming has no requirements for voice specific items such as wiretapping, PSTN interconnects, etc.

PacketCable Multimedia Architecture for Gaming

The easiest way to describe the PacketCable Multimedia Architecture is to provide a diagram of the architecture and discuss the functionality and interaction of each of the components as it related to providing QoS for gaming applications.

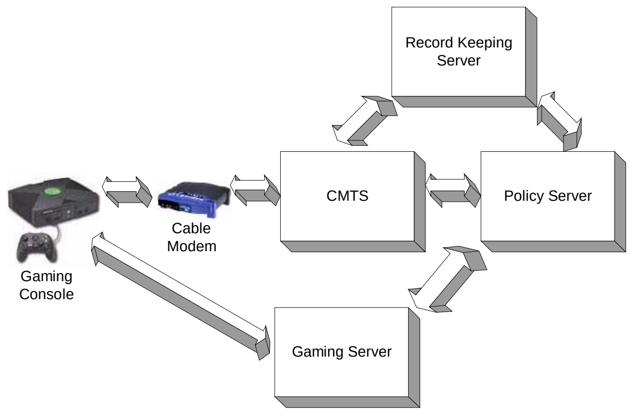


Figure 1 - PacketCable Multimedia Architecture for Gaming

Our assumption is that the gaming console has no concept of its QoS requirements nor of the PacketCable signaling like that available to a VoIP MTA (multimedia terminal adapter) to signal its desire for QoS reservations. Instead, a gaming console simply communicates with the gaming server as it does today. (e.g. Xbox's MechAssault game causes the Xbox to communicate with Xbox Live servers to set up a gaming session between players).

The CMTS is the gatekeeper (referred to as a Policy Enforcement Point or PEP) which determines whether the resources are available to reserve bandwidth between the cable modem and itself. Thus, the gaming server must communicate the console's bandwidth needs to the CMTS. It does so through an intermediary known as the Policy Server. As there could be many different applications all of which are contending for limited bandwidth resources, the Policy Server determines the relative priority of each request (based upon business rules) to determine which requests for QoS should actually be given to the CMTS. The Policy Server is also referred to as the Policy Decision Point (or "PDP").

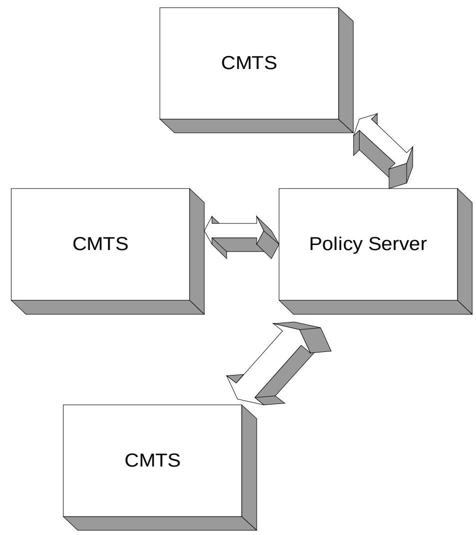


Figure 2 – A single policy server can serve multiple CMTSs.

Once instructed by the Policy Server of the gaming console's QoS requirements, the CMTS creates service flows for an individual cable modem's gaming traffic with the appropriate QoS characteristics. As an option, the PacketCable Multimedia architecture also takes into account the desire to track the actual usage of the QoS-based service flows for billing purposes. These billing records are gathered and maintained on the Record Keeping Server.

The gaming server and the policy server are expected to reside within the MSO network and are considered trusted devices. The gaming server must also take on the responsibility of authenticating the gaming console and assuring that the consoles are authorized to request gaming services.

Messaging Protocols

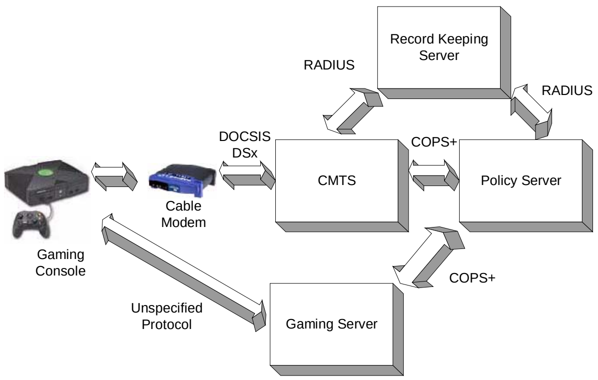


Figure 3 – Messaging Protocols

Obviously, there is also a messaging flow between the CMTS, Policy Server, Gaming Server, and gaming console which indicates the success or failure of the QoS provisioning. The messaging between the gaming console and the gaming server is outside of the PacketCable specifications. The messaging from the gaming server to the Policy Server and from the Policy Server to the CMTS is IETF's COPS based. Any event messaging sent from the Policy Server or CMTS to the optional Record Keeping Server is RADIUS based, and the CMTS/cable modem exchanges to establish QoS-based service flows is based on DOCSIS DSx messaging.

This has been a greatly simplified explanation of QoS allocation. Upstream and downstream service flows are handled by the CMTS in different manners. Upstream transmissions are made on a contentious, shared-access medium, where downstream traffic is handled by the CMTS as if it were a traditional IP router. The specifics of the QoS parameters that are associated with upstream and downstream service flows (these parameters are different) and the service flow scheduling types can be found in the VoIP-centric PacketCable 1.0 specifications. [7]

You will notice that there are a few things missing which simply fall outside of the PacketCable Multimedia domain, namely end-to-end network QoS setup including Policy Server to Policy Server communications. One can imagine that gamers desire low network latency on each network segment over which their gaming traffic travels. This can conceptually be handled by DiffServ or MPLS – most MSOs would argue that their network backbones are over-engineered and that the DOCSIS component is where the bandwidth is the most valuable resource.

Most relevant gaming servers have the ability to match gamers based upon their historical levels of quality of play (that is to say, based upon how good the player is at performing the game), and also based upon the latency of the gamer's network connections. Obviously, one goal of the PacketCable Multimedia framework can be to lower the latency of an individual gamer's network connection to the CMTS. The game servers would need to report the gamer's potential latency when matching up gamers rather than their pre-service-flow-setup latency. This could require some additional communication between the gaming server and the policy server and potentially inter-policy server or inter-gaming server communications.

The gaming consoles described above are referred to in the PacketCable Multimedia Architecture Framework as "legacy" clients as these consoles are unaware of the QoS capabilities and signaling necessary for the QoS negotiations within the framework.

A second type of client can have some PacketCable awareness built-in – when a network-based game is started, the client can request QoS. The console can now signal to the CMTS to add, change or delete bandwidth reservations, but the CMTS will only accept the reservations if the gaming server and policy server have authorized the console's reservation. This is very similar to the behavior of a VoIP MTA. This concept of building PacketCable awareness into a console or console game will probably not receive much enthusiasm for implementation by the game developers unless there is a considerable client base that could make use of it. For that reason, we anticipate that support for legacy clients must be well implemented first.

The third type of client is one which is totally PacketCable aware and does not depend upon a gaming server to setup its QoS. Instead the console is capable of transmitting its own bandwidth QoS requests to the CMTS along with authorization credentials. The CMTS passes the request onto the Policy Server which authenticates and authorizes the consoles request. The request message is then sent back to the CMTS which will then setup the appropriate service flows for that console.

The details of the PacketCable Multimedia signaling message structures, service flow scheduling types, service flow management, etc are outside of the scope of this document, but should be publicly available in the CableLab's PacketCable Multimedia Technical Report and Specifications by the time of the publication of this document.

Summary

Console gamers are able to use cable modem connections today with good results. As MSOs deploy new services that consume more of the limited bandwidth available between the cable modems and cable modem termination servers, the gamers' user experience could become less attractive. One method to enhance the user experience is to implement a PacketCable Multimedia Architecture that would enable QoS guarantees for gaming consoles without modification to those console or console games. This architecture would require enhancements to the gaming servers to make the server applications capable of interacting with the Policy Servers.

We have seen that there are a lot of customers playing games on our broadband networks, research shows that they are willing to spend vast quantities of money to do so, and they have voted with their wallets to use today's low-latency, high speed connections - it is up to the MSOs to cement the relationship by providing a service which is unobtainable from other providers.

References

[1]http://news.com.com/2110-1040-979556.html - "Xbox Live sign-ups beat forecasts", cnet.com
[2]http://news.com.com/2100-1040-980966.html - "PlayStation 2 shipments top 50 million", cnet.com
[3]http://www.idsa.com/IDSABooklet.pdf - "Essential Facts about the Computer and Video Game Industry", Interactive Digital Software Association (idsa.com)
[4]http://www.iana.org/assignments/port-numbers - "Port Numbers", Internet Assigned Numbers Authority (IANA.org)
[5]http://www.business2.com/articles/mag/0,1640,45482,00.html – "Could This Be the Next Disney?", Business 2.0 (business2.com)
[6]"PacketCable Multimedia Architecture Framework PKT-TR-MM-ARCH-V01", Cable Television Laboratories (cablelabs.com)
[7]"PacketCable™ Dynamic Quality-of-Service Specification PKT-SP-DQOS-I05-021127" Cable Television Laboratories (cablelabs.com)

DOCSIS™ TOOLS FOR TIERED DATA SERVICES

Doug Jones
YAS Broadband Ventures

Abstract

Given that tiered data service is a good economic idea (and a growing volume of data support this), how does the operator implement a solution? This paper discusses the technical tools available in DOCSIS for implementing both "Speed" and "Included Bytes" tiers.

INTRODUCTION

Business Case

Cable data system usage has been studied for several years now and a growing body of work is available that indicates tiering curbs extreme consumption behavior. On an untiered network, 80% of the total available bandwidth is consumed by only 12% of the subscribers. On a tiered network, 80% of the bandwidth is consumed by 25% of the subscribers, showing a more even distribution of consumption. Given that the majority (>70%) of High Speed Data (HSD) subscribers consume less than 2 GB (GigaBytes, where 1 GB = 1,000 MegaBytes) of data a month (combined upstream and downstream), curbing the extreme consumption of a few users will free up bandwidth for more "average" usage subscribers and the revenue they bring in.

That's about it for business motivation, the data are in and tiering makes economic sense. The remainder of this paper discusses technical methods to implement tiering on a DOCSIS network.

Types of Tiers

There are two types of tiers:

- Speed: Usually an instantaneous number measured in kilobits or megabits per second. This is how "fast" the CM is allowed to operate on the network. There can be separate speeds for the forward and return paths.

- Included Bytes: Usually measured over a period of time such as a month, this is the total amount of traffic through a CM. It is usually measured as an aggregate of both forward and return traffic, though separate tiers are possible for each direction.

DOCSIS provides a set of tools to implement both speed and Included Bytes tiers; however, the methods can differ between DOCSIS 1.0 and DOCSIS 1.1. Specifically the operator has more choice and arguably better options available to them with DOCSIS 1.1. But there are ways to get it done regardless of the version of DOCSIS deployed.

Overall System View

This paper discusses how tiers can be implemented on a cable data system. Collecting DOCSIS usage data is one part of the overall solution needed to implement tiers. A representation of the overall system is shown in Figure 1.

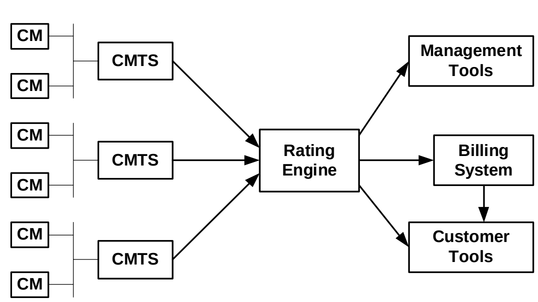


Figure 1

The items in Figure 1 to the right of the CMTS are not discussed in detail in this paper although they are important considerations for the back-office.

Usage data can be collected in either the CM or CMTS through methods described in this paper. A Rating Engine processes the data where business decisions are made to turn the raw usage information into a line item for the billing system. There are also tools to both allow the operator to manage the system and to allow customers to track their usage before the bill shows up at their door.

Collecting the usage data, while there are several methods available, is probably the most straightforward step of the entire process. Processing that data into billing information will be unique for each operator.

SPEED TIERS

Description

This type of tier defines the maximum speeds that a user will have over the DOCSIS connection. It is possible to define maximum speeds on both the upstream and downstream connection.

Example speed tiers are a user having speeds of 128 kbps on the return path and 1.5 Mbps on the forward path. The cable operator sets these numbers and it is possible to assign different speed tiers to different groups of subscribers.

The speeds are assigned to the Cable Modem (CM) through the CM configuration file, which is a list of instructions created by the cable operator and provided to the CM every time it boots. There are many parameters in the CM configuration file that the operator uses to define the data “service” provided to the user, but only a couple of the parameters are needed to create the speed tier.

Choosing a speed tier begins with the operators service activation system. When a subscriber requests HSD service, the operator generally offers a choice of several speed tiers to choose from. The service activation system communicates the speed tier information to the provisioning system where the corresponding configuration file is created and assigned to that subscribers’ CM. When the CM boots, it is provided that configuration file with the appropriate speed tier information as illustrated in Figure 2.

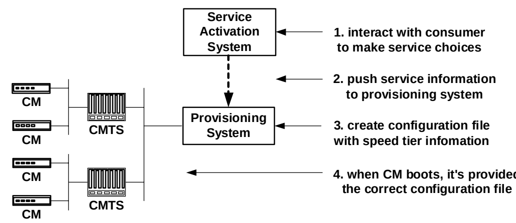


Figure 2

Speed Tiers: DOCSIS 1.0

In DOCSIS 1.0, the maximum speeds are not guarantees, rather the system will provide up to that speed if there is capacity available on the system. There are several reasons why the full speed may not be available, and primary among these is having too many users attempting to access the system at the same time. All networks are shared at some point and engineering enough bandwidth for peak usage can solve congestion.

In the DOCSIS 1.0 configuration file, the following two parameters are used to create speed tiers for the downstream and upstream paths:

- Maximum Downstream Rate Configuration Setting
- Maximum Upstream Rate Configuration Setting

These parameters are simply set to the desired speeds and the system enforces them to ensure the CM does not transmit at speeds higher than allowed by their tier.

Speed Tiers: DOCSIS 1.1

DOCSIS 1.1 supports many Quality of Service (QoS) parameters, the vast majority of which are not needed to implement speed tiers. While DOCSIS 1.1 QoS is complex, it is as simple as DOCSIS 1.0 to implement speed tiers.

In the DOCSIS 1.1 configuration file, the following two parameters are used to create speed tiers for the downstream and upstream paths:

- Downstream Maximum Sustained Traffic Rate
- Upstream Maximum Sustained Traffic Rate

The names of the parameters have changed to reflect that DOCSIS 1.1 offers a complete Quality of Service (QoS) package. These two parameters are part of that larger QoS package, however, they function exactly the same and cause the same effect as the DOCSIS 1.0 parameters.

INCLUDED BYTES TIERS

Description

Included Bytes tiers are sometimes referred to as consumption tiers. This type of tier counts how many Bytes of data are used by the CM over a period of time. An analogy is to the mobile phone industry that for example offers several “Included Minutes” tiers that include an allowed number of minute’s usage over one month. Similarly a fairly standard entry-level tier for HSD is including 2 GigaBytes (GB) of usage over one month. Data shows that the majority of HSD users consume less than 2 GB per month. For usage beyond the tier amount, the operator business policy implemented in the Rating Engine would determine the appropriate billing treatment for that subscriber.

The Included Bytes tier is generally a combination of both the upstream and the downstream usage as shown in Figure 3 below. An operator could choose to offer separate tiers for upstream and downstream usage.

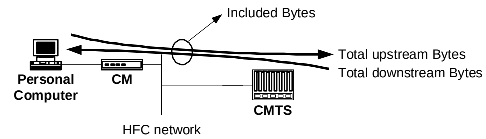


Figure 3

The amount of Bytes included in these tiers should come from the operators own investigation and business plan. Two GigaBytes is equal to 2,000 MegaBytes and is a reasonable amount of data for a subscriber just doing email and web surfing. Users that are heavy into peer-to-peer applications or that include large attachments with email or do a lot of file transfer may consume more that this.

Unlike Speed Tiers that are implemented using the CM configuration file through an interaction with the provisioning system, Included Bytes tiers are implemented by counting the number of Bytes of data that are sent and received through a particular CM.

Different methods are available for aggregating the Bytes of data through a CM depending if the system is DOCSIS 1.0 or DOCSIS 1.1. These methods are described in the following sections.

CM Byte Counters

While this method works with all DOCSIS versions, it is the only DOCSIS-defined method of gathering consumption information for DOCSIS 1.0 systems. A subsequent section describes enhancements available when using a DOCSIS 1.1 CMTS.

All DOCSIS CMs are required to implement Management Information Base

(MIB) objects that can be polled using the Simple Network Management Protocol (SNMP). Several of the required MIB objects include counters that track the number of upstream and downstream Bytes through that CM.

The operator can use an SNMP workstation, also known as a Network Management Station (NMS), to periodically poll each cable modem to collect downstream and upstream usage information as shown in Figure 4. Once polled for, the usage data is passed to the Rating Engine for analysis per operator business rules.

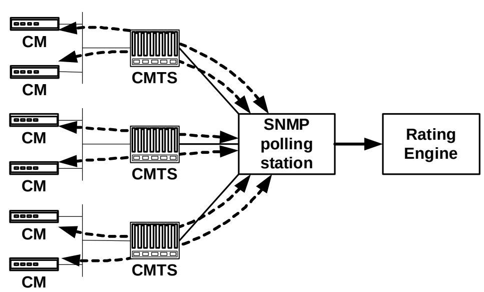


Figure 4 – Using SNMP to Poll CM counters

The time interval the NMS uses to poll all the CMs on the network is an issue to be considered for several reasons. As subscribers can power off their CMs, usage information may be lost from time to time. When the CMs are powered on, the MIB counters are not required to reset to zero (an implementation detail with MIBs, its just how they work). The NMS has to poll once just to get a baseline number from which to calculate further Byte usage.

In order to detect when a CM has been rebooted, there is a MIB object that contains the date/time of when the CM last rebooted. The operator can use this information to learn if the baseline number for this particular CM has changed.

While polling CM Byte counters is a simple and easy method supported by DOCSIS 1.0 to implement Included Bytes tiers, using CM counters may not be a highly reliable method due to the unpredictability of CMs being power cycled in the home. It will be hard to guarantee accurate counts, in fact, the operator can expect to undercount usage due to the issues listed above.

Another reason to carefully adjust the polling interval is the amount of traffic the SNMP polling of CMs places on the DOCSIS network. There can be thousands of CMs attached to a CMTS and polling too often can add appreciable traffic to the cable data network. Depending on the number of CMs on the network and the polling interval, the SNMP polling traffic can comprise up to 5% of the bandwidth of the cable data system. This is not a trivial number as this is bandwidth that could otherwise be charged for.

CMTS Byte Counters

DOCSIS 1.1 requires the CMTS to implement MIBs that count upstream and downstream Bytes on a per CM basis. Instead of polling all the CMs, the operator can now poll just the CMTS as shown in Figure 5. Note a DOCSIS 2.0 CMTS is required to have these same counters and this method is equally viable there.

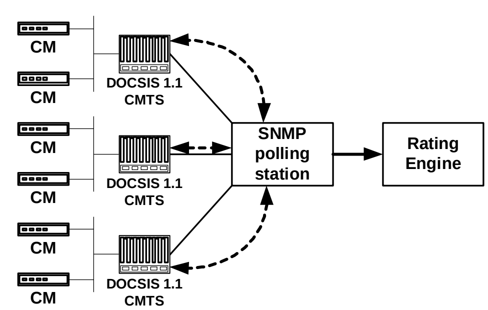


Figure 5 – Using SNMP to poll CMTS counters

This method still uses SNMP to poll the MIB counters at the CMTS, but since a CMTS is not supposed to be power cycled that often, the polling frequency can be greatly reduced to minimize the amount of SNMP traffic needed to collect the data. In fact the Byte counters required in the CMTS were designed to count very high specifically to allow the operator to poll the CMTS only once a month. As long as the CMTS is not power cycled, the counters will accurately count trillions of GigaBytes and it is highly unlikely a subscriber could consume that amount of data over a month. Using CMTS polling, subscribers can power cycle their CMs as often as they want and the CMTS will still keep accurate counts of their bandwidth consumption.

A complete rollout of DOCSIS 1.1 is not needed to take advantage of this easier , more reliable, and more accurate method to aggregate Byte count information. By only implementing a DOCSIS 1.1 CMTS and leaving the CMs at DOCSIS 1.0, the rest of the network, e.g., the back-office, does not need to be modified to support DOCSIS 1.1. Said another way, if only the CMTS is upgraded to DOCSIS 1.1 (all the CMs are 1.0), no changes are needed to the DOCSIS backoffice for provisioning DOCSIS 1.1 CMs. The already deployed DOCSIS 1.0 CMs will operate on the DOCSIS 1.1 CMTS with all the expected features available with DOCSIS 1.0.

3rd Party Counting System

Another option for measuring cable modem bandwidth consumption is to place a traffic counting device between the CMTS and the Metro IP aggregation network. This device is capable of counting the traffic into and out of an operator's DOCSIS network as shown in Figure 6.

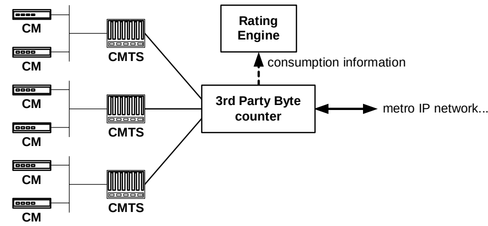


Figure 6 – 3rd Party Byte Counter

This solution does not depend on the version of DOCSIS deployed. In fact, this solution works with non-DOCSIS cable data systems too and so may be a consideration for operators that have both DOCSIS and proprietary data systems in the same metro area.

The 3rd party counting system can be approached in several ways. Some Ethernet switch equipment can aggregate traffic from several CMTSes into a single data stream as shown in Figure 7. This aggregation switch also takes on the additional processing task of Byte counting. On a periodic basis, consumption information is transferred from the switch to the rating engine.

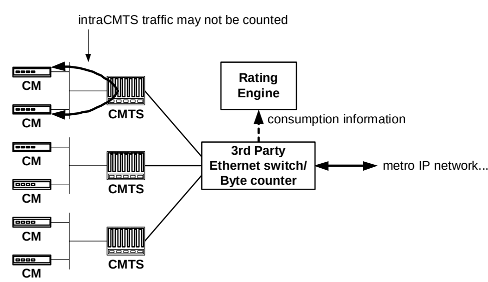


Figure 7

The configuration shown in Figure 7 is not capable of counting traffic that "stays at home" on a particular CMTS. That is, intraCMTS traffic from CM to CM on a single CMTS will not pass through the Byte counter as shown in Figure 7.

Another approach entails placing a traffic monitoring/traffic shaping device in a data path of already aggregated CMTS traffic as shown in Figure 8.

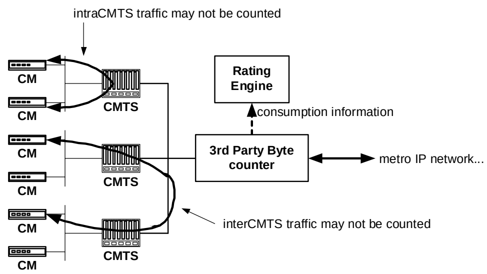


Figure 8

As shown in Figure 8, both intraCMTS data traffic and traffic between CMTSes may not be counted with this configuration.

Finally, using a 3rd Party counting system, in either configuration, has the potential to introduce a single-point of failure in the data network that could affect more than one CMTS worth of traffic. A system used for measurement purposes only, however, may not have this characteristic. It depends on the product.

SUMMARY

There are two types of data tiers, Speed and Included Bytes. Speed tiers are implemented through the CM configuration file. Included Bytes tiers are implemented by monitoring usage data from any of several sources, though some sources are more reliable than others. Tools exist in DOCSIS to implement both types of tiers.

DOCSIS 1.0 and DOCSIS 1.1 support very similar methods to implement speed tiers. However, DOCSIS 1.0 and DOCSIS 1.1 systems provide different methods to implement Included Bytes tiers. The DOCSIS community was more aware of the need for implementing data tiers in DOCSIS 1,1, therefore, that system has a more simple method to collect consumption data from the CMTS, whereas in DOCSIS 1.0 this information has to be collected from the CMs.

A key piece of equipment needed for the overall tiering system is the Rating Engine. DOCSIS only provides a technical means to implement tiers, whereas the Rating Engine is need to turn the raw data into billing information. The Rating Engine is not standardized in DOCSIS, rather, this functionality will be specific to each operator.

ACKNOWLEDGEMENTS

Jerry Bennington and Terry Shaw, both at CableLabs®, contributed to this paper.

AUTHOR CONTACT

Doug Jones
YAS Broadband Ventures
300 Brickstone Square
Andover, MA 01810
voice: +1 978.749.9999 x226
doug@yas.com

FAST ETHERNET IN THE LAST MILE

Oleh J. Sniezko
Aurora Networks, Inc.

Abstract

Broadband HFC network operators in North America are uniquely positioned to serve the increasing needs for telecommunications services among small and medium size businesses. This paper analyzes the market potential for providing bandwidth and other telecommunications services to this telecommunications market segment. Near-term opportunities for HFC network operators as alternative bandwidth providers and longer-term opportunities as Layer 2 and Layer 3 service providers are also discussed. It describes several technology solutions that can be deployed in the HFC plant to support transport of bandwidth-intensive applications and services. The paper also includes a discussion on the requirements such business applications will place on an operator's network.

Finally, the paper describes how a solution based on Ethernet transport can also be deployed to support delivery of high-bandwidth residential services to MDUs/MTUs in high-density urban environments.

INTRODUCTION

Broadband HFC Advantage

Broadband HFC network operators in North America have a distinct advantage over their competitors in providing telecommunications services to small and medium size businesses. With the right selection of the technology, high quality service can be ensured at lower incremental cost than the cost incurred by the competitors for delivering equivalent service quality.

In the U.S., HFC network footprints already cover approximately 80% of all SMBs. As of 1999, 1 in 5 SMBs already subscribed to cable TV at their business location, primarily for customer entertainment.[i] In most of these cases, HFC networks are already within the last few hundred feet from potential business customers. This translates into lower incremental fiber construction costs in fiber-to-the-business (FTTB) architecture.

With fiber to the business, high-speed connectivity at 100 Mbps or 1 Gbps can be provided cost-effectively today. This high-speed access can be throttled back or aggregated depending on individual customer needs. Moreover, fiber to the business offers sustained full throughput in contrast to such alternative offerings as DSL or cable modem.

Ethernet Advantages

Several data communications systems have been developed and implemented to serve internal and external telecommunications needs of enterprises. Among them, Ethernet has gained the broadest acceptance. The following numbers clearly support this assessment:

1. 80+% of all data packets begin and end their lives as Ethernet packets.
2. There are 250 million Ethernet ports deployed worldwide.

3. 90% of transported business data begins and ends as Ethernet on LANs.

Ethernet is well understood by small, medium and large enterprises. It is manageable by small businesses without a dedicated IT staff. Moreover, it is supported by the dominant standard and represents mature technology. The IEEE 802.3 standard on Ethernet was released in 1980. It also proved to be extremely flexible and future proof as new developments (Gigabit Ethernet, 10 GigE) do not obsolete previous implementations.

Beside the fact that Ethernet technology is characterized by low maintenance cost, it is also relatively inexpensive to implement due to economies of scale from large installed base. It has shown excellent price/performance trend: 10x the performance of the preceding generation at 3x the cost.[ii]

Alternatives

The enterprises use almost all technologies available today to interconnect their internal organizations located at different locations and to secure connectivity with Internet. The following data[iii] (Table 1 and Figure 1) provide the distribution of deployment of different connectivity technologies for data communication services.

	Medium Size Businesses	Extended Medium Size Businesses
Dial Up	83%	84%
T1/T3	41%	61%
ISDN	36%	58%
Switched 56 kbps	19%	35%
Frame Relay	15%	21%
Satellite	6%	3%
ATM	2%	1%
ADSL	1%	1%

Table 1: Type of Connectivity for Data Communication Services

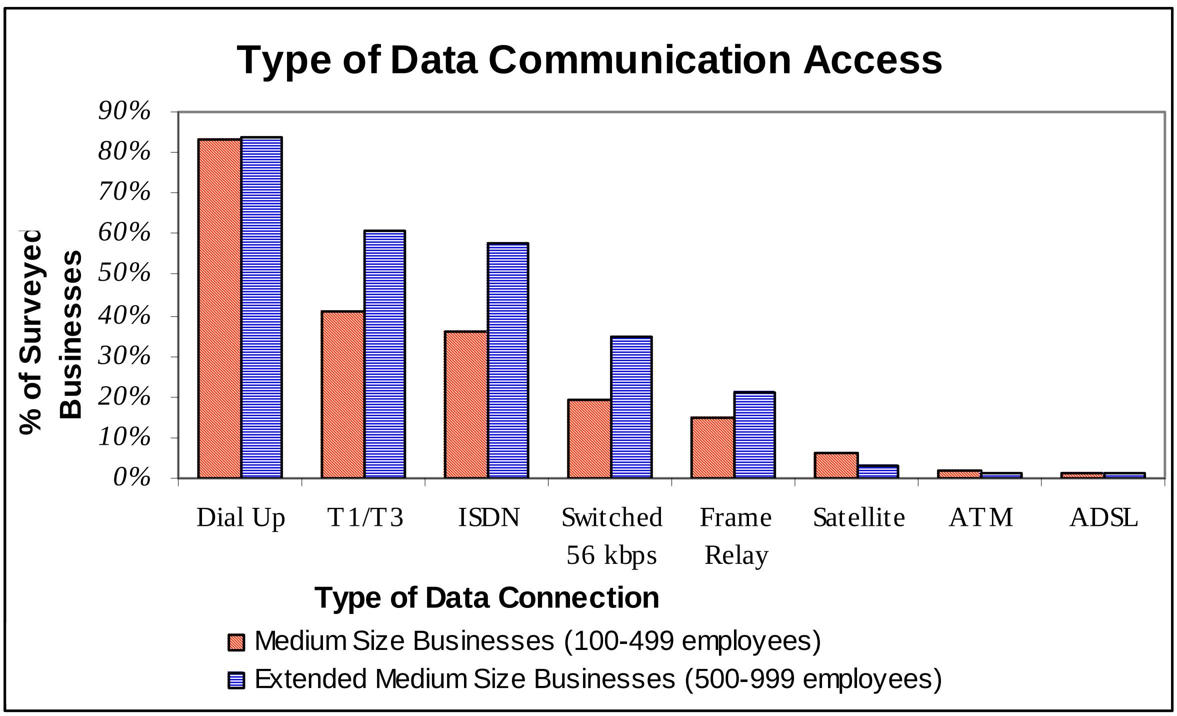


Figure 1: The Usage of Different Connectivity Technologies for Data Communications Services

<u>OPPORTUNITY</u>

<u>Market Definition</u>

There are many definitions of the small and medium size businesses. The following presents one of the accepted definitions with some characteristics of their communication needs and preferences:

1. Small size businesses:
   a. less than 50 employees,
   b. fastest growing segment,
   c. mostly ignored by local exchange carriers (LECs),
   d. highly receptive to competitive offerings,
   e. application and service needs: voice support for up to 24 POTS lines, Internet access (from 56 kbps to DSL speeds today) and IP management services,
   f. billing preferences: consolidated billing for voice and data (Internet) services.
2. Medium size businesses:
   a. up to 100 employees,
   b. multi-campus/branch offices, most often within a single metropolitan area,
   c. considered easiest target by many LECs,
   d. receptive to regional competitive offerings,
   e. application and service needs: voice support for digital PBX systems (requires fractional and full T-1 line provisioning), Internet access (high speed DSL today) and IP/WAN management services,
   f. billing preferences: consolidated billing for voice and data (Internet) services.
3. Large size businesses:
   a. in excess of 100 employees,
   b. multi-campus/branch offices,
   c. receptive to national/international offerings,
   d. application and service needs: voice and data networking (Intranet, private tandem switches, etc.), virtual private networking (today over frame relay and ATM networks).

<u>Market Statistics and Opportunity</u>

Based on the definitions presented above, the industry reports[iv] show that small and medium size businesses (SMBs) represent 95% of the entire U.S. business universe. They amount to 7.6 million entities approximately and this number is growing at approximately 2% annually. In 1998 alone, LAN penetration in SMBs grew by 10% to 36% while the percentage of LAN connected PCs grew by 24% to 13.4 million.

SMB spending on IT and telecom services is higher than $100 billion a year.

The costs and other characteristics of the alternative technologies for data connectivity today are presented below.

1. Low-speed T1 connections:
   a. typical installation times 30-45 days,
   b. installation costs range from $1,000 to $2,000,
   c. maximum symmetrical bandwidth of 1.5 Mbps,
   d. average monthly costs range from $400 to $1,200;
2. Low-speed, DSL connections:
   a. typical installation times 4-6 weeks,
   b. installation costs range from $200 to $300 plus CPE costs of $100-$400,
   c. 10+ DSL flavors results in complex pricing structure,

d. maximum symmetrical bandwidth up to 1.5 Mbps in various increments,
e. average monthly costs of $150 to $400 for symmetrical DSL service,
f. must be close to central office (CO);

3. Cable modem connections:
   a. typical installation times less than a week,
   b. installation costs range from $0 (cost to the customer, non-zero cost to the service providers) to $150,
   c. average monthly costs range from $40 to $80,
   d. shared bandwidth perceived as a disadvantage for business applications;
4. Data carrier connections:
   a. Installation costs range from $3,500 to $7,500,
   b. Dedicated symmetrical bandwidth offering: 100BaseT to 1 Gbps Ethernet,
   c. Average monthly costs range from $1,000 to $4,000,
   d. Limited availability depending on market.

The data carrier connections can be quite expensive even in the access network. The following table presents pricing structure for Worldcom Ethernet service offerings.[v]

	Service	**Customer Interfaces**	**Price/month**
Metro and WAN Private Line	Corporate MAN links	50 Mbps, 150 Mbps, 622 Mbps	Similar to ATM and frame relay
Dedicated Internet	Enterprise access to Internet	1 Mbps to 500 Mbps	$1,200 to $200,000 per circuit
Enterprise private line/VPN	LAN to LAN and corporate network connections	1 Mbps to 100 Mbps	$630 to $20,000 per circuit

Table 2: Worldcom Ethernet Service Profiles

<u>Market Requirements</u>

There are several basic requirements specified by most businesses and some specific requirements dependant on the business size and type. Almost the same requirements are defined by service providers. The basis requirements can be summarized in the following points:

1. Service scalability
   a. Bandwidth offering scalable from 1 Mbps to 1 Gbps
   b. Symmetrical bandwidth preferable
2. Deployment cost scalability:
   a. Equipment to provide services can be deployed based on the service demand (number of customers and bandwidth requirements)
   b. Cost scales with the SLA requirements (for example, equipment and route redundancy can be implemented on a as-needed basis)
3. Deployment simplicity
4. Future proofing
   a. Future protocols do not render equipment providing connectivity obsolete

b. Equipment upgrades are easy to accomplish at limited and demand driven locations
5. Acceptable service reliability and availability
6. Low maintenance costs

Some additional requirements from the following list may be critical dependant on business size and type:

1. Affordable and competitively priced
2. Security comparable to or better than provided by competitive technology and providers
3. Transparency to layer 2 and higher protocols:
   a. Allowing service providers (MSOs) and customers (SMBs) for leveraging the existing LAN/WAN hardware infrastructure
   b. Allowing for passing through Layer 2 and higher protocols, overlaying datacom protocols for robustness and security (e.g., CoS, QoS, IPSec)
4. Accessible to SMBs without a dedicated IT staff
5. Capable of supporting or transparent to critical business applications
   a. Transparent to voice, video and data applications
   b. Supporting:
      i. point-to-point LAN transport,
      ii. multipoint LAN interconnect and extension,
      iii. VLANs,
      iv. VPNs, and
      v. leased line replacement.
6. Low latency for latency-sensitive applications like VoIP and video streaming
7. Support for TDM (T1, E1, DS3) traffic and interfaces
8. Sustainable, non-degrading with distance performance (unlike various favors of DSL).

There are some requirements important to service providers:

1. ROI for equipment: in months
2. Clear network demarcation points
3. Compatible with existing HFC architecture and headend installation and equipment
4. Easy to install
   a. no additional active devices in the plant to install and manage
   b. capable of flexible connection topologies (point-to-point, ring, nested span, etc.)
5. Easy to manage
   a. SNMP compliant interfaces
   b. remote provisioning for new business customers
   c. upgradeable via software downloads.

TECHNOLGIES AVAILABLE TO BROADBAND HFC NETWORK OPERATORS[vi]

There are several data connectivity technologies that have been deployed in the past. This paper will concentrate on these that use Ethernet based interfaces. All of these technologies can be divided into two groups:

1. Technologies with network processing (intelligent) equipment distributed in the access plant.
2. Technologies with network processing equipment centralized in headends or main hubs.

Both groups require customer premise equipment. This equipment type and intelligence is mostly defined by medium type and internal LAN requirements. Both groups include several possible deployment scenarios and topologies.

Distributed Processing Equipment

Distributed equipment usually comprises IP switches/routers deployed in the field between headend (or CO) and the customer. Several architectures are being marketed by vendors:

1. FTTB/H EPON with:
   a. Switches and aggregation points in nodes and optical gateways on customer premises; or
   b. Switches at optical nodes and/or amplifiers and optical gateways on customer premises.
2. Hybrid fiber/copper pair architecture with:
   a. Switches and aggregation points at optical nodes,
   b. Switches in taps, and
   c. Cat6 copper pair physical layer mesh network between switches (Cat6 to homes).
3. Hybrid fiber/coax architecture with:
   a. Switches and aggregation points at optical nodes,
   b. Switches in taps,
   c. Coax drops to homes, and
   d. Coaxial physical layer with RF modulation and demodulation and up-and down- conversion at each switch location and at customer premises.

All the above technologies can be analyzed against the set of requirements presented previously. This detail analysis can be performed by the readers of this paper. Here are just few comments:

1. The history of IP and other higher layer protocols shows that the legacy equipment not always support and not always can be upgraded to support the new protocols and has to be replaced. This usually is not a problem when the legacy equipment is located at limited locations or on customer premises (demand-driven replacement) but may pose significant problems (cost, service disruption) if located in the access plant.
2. Addition of new switches and traffic aggregation points may lead to service disruptions for some architectures presented above. The service disruptions may affect data customers or any-service customers served by the HFC network.

<u>Centralized Processing Equipment</u>

Similarly to the distributed processing equipment solutions, the existing solutions for the centralized processing equipment offer several alternatives:

1. Ethernet over RF
   a. DOCSIS, a standard-based solution with:
      i. Steeply decreasing equipment prices
      ii. Improving performance (DOCSIS 2.0)
      iii. Limited total and per channel bandwidth (even with DOCSIS 2.0)
      iv. High maintenance cost of upstream HFC path
   b. Ethernet over RF based on proprietary solutions:
      i. Without bandwidth conversion or
      ii. With bandwidth up- and down-conversion.
2. Transparent Layer 1 pipe:
   a. Overlay systems with dedicated fiber or wavelength
   b. Integrated optical systems

The Ethernet over RF systems are comparable in performance. Proprietary systems may have advantage in delivering higher bandwidth, especially when combined with up- and down-conversion. However, the proprietary character of these solutions will most likely result in high equipment cost. Moreover, the coaxial shared medium is still perceived by many businesses as less reliable than dedicated fiber and copper media.

The transparent pipe alternative in its overlay configuration has been available for as long as IP switches and routers exist. Several established transport and IP equipment manufacturers and some start-up companies provide equipment for point-to-point, point-to-multipoint and ring topologies. The equipment has several standard interfaces ranging from T1/E1 emulated TDM circuits through 1 GigE interfaces. This type of equipment competes with ATM, frame relay and (recently) SONET solutions. As described above, thanks to the Ethernet proliferation, it has cost and other advantages over the competing technologies. This technology is suitable for larger business and is a simple extension of metro-market topology into access plant. It does not offer a significant advantage to HFC network operators as any operator with a capability of installing or with already installed fibers to POPs located in proximity of large businesses could implement this data communications technology. Moreover, this market (large businesses) has been successfully addressed by ILECs, CLECs and other data carrier companies.

The integrated solutions are usually a hybrid approach (at least as long as FTTH for residential services is not deployed). Usually it can be integrated to the node location. From the node to the business, it is delivered on a dedicated fiber. At least one of the vendors offering this technology has also capability for flexible increase in capacity to 1 GigE and above. The integrated technology has been enabled in the last several years with the introduction of digitized technology in upstream HFC links supporting the legacy RF two-way communication. The figure below shows an example of the integrated Ethernet solution.

The integrated solutions can be modified to provide a full, dedicated connectivity to larger businesses in an evolutionary and scalable manner. This is supported by a significant progress in passive component technology (colorless and WDM), especially in their capability to perform under harsh outside plant conditions.

APPLICATIONS AND REVENUE OPPORTUNITY

Near-Term Opportunities

Near-term opportunities for HFC network operators as alternative bandwidth providers can materialize in the following areas:

1. T1 replacement with T1 interfaces for both data and voice services:
   a. lower cost (cost of T1: $400-1200/month with $1-2K installation cost for a maximum of 1.544 Mbps bandwidth)
   b. shorter than 30-45 day waiting period for T1 installation and activation
   c. service to
      i. small/medium/large businesses
      ii. wireless backhaul to the PSTN for cell towers
      iii. virtually any customer of T1 services today
2. DSL replacement for data communication and Internet access
   a. dedicated, guaranteed bandwidth, no degradation with distance
   b. higher capacity than DSL at similar or lower cost (cost of DSL: $150-400/month for symmetric DSL for 128 kbps—2 Mbps; $60-200/month for asymmetric DSL for 192 Kbps—1.5 Mbps)
   c. shorter waiting period
   d. service to small and medium businesses
3. VoIP

a. integration of voice with existing data services for small and medium size businesses
b. transport of T1 over Ethernet to support legacy PBX applications with integrated or off-the-shelves interfaces (T1 and E1 emulators)
c. secondary voice lines for the residential market (MDUs)

These services can be provided to the following market segments:

1. SOHOs/SMBs
2. MDUs
3. Multi-campus businesses
4. Businesses with telecommuters
5. Institutions:
6. Schools
7. Universities
8. R&D facilities
9. Hospitals
10. Military and Government facilities
11. Sports facilities

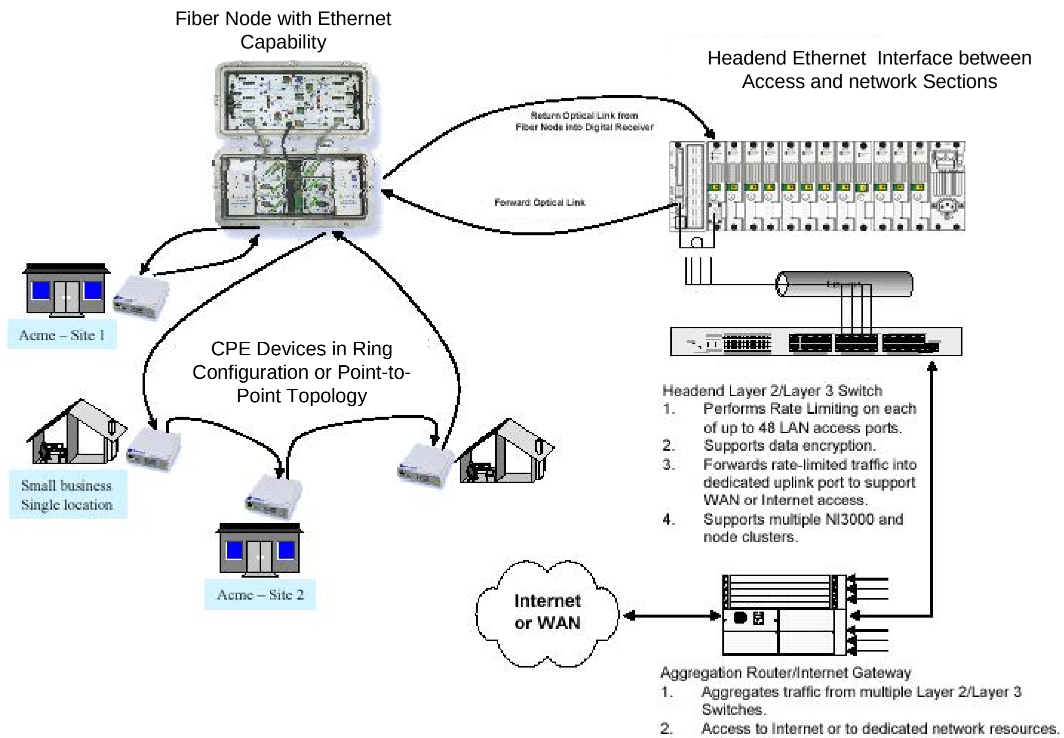


Figure 2: Integrated Ethernet over HFC Network – Example

<u>Long-Term Opportunities</u>

As the experience and familiarity with the networking technology grows, longer-term opportunities can be addressed and capitalized on. These include:

1. VLAN tunneling
2. Link aggregation
3. L2 and L3 Quality-of-Service (QoS) features
4. IP network services
5. Alternative high-speed ISP access
6. Alternative IP telephony services
7. Future services such as IP video streaming, etc.

Many MSO organizations have already expertise and staff to support these services to any businesses.

ETHERNET TO RESIDENCES

The integrated Ethernet approach can be easily extended to serve residences. This extension may happen selectively based on cost analysis. The residential Ethernet FTTH connections can take place initially in MDUs where the cost of CPEs and switches provisioning the service to individual suites can be shared among many tenants. This type of solution, besides minimizing network complexity, also allows freeing up valuable forward and return spectrum that can then be allocated to other services. Moreover, in most cases, it eliminates the need for costly coaxial cable rewiring in MDUs. With the development of lower cost CPEs with limited features, acting mostly as media converters, symmetrical Fast Ethernet (shared among several residences) can be also delivered to SFU residential areas. The figure below presents the evolution from FTTB to FTTH through intermediate step of deploying fiber to MDUs.

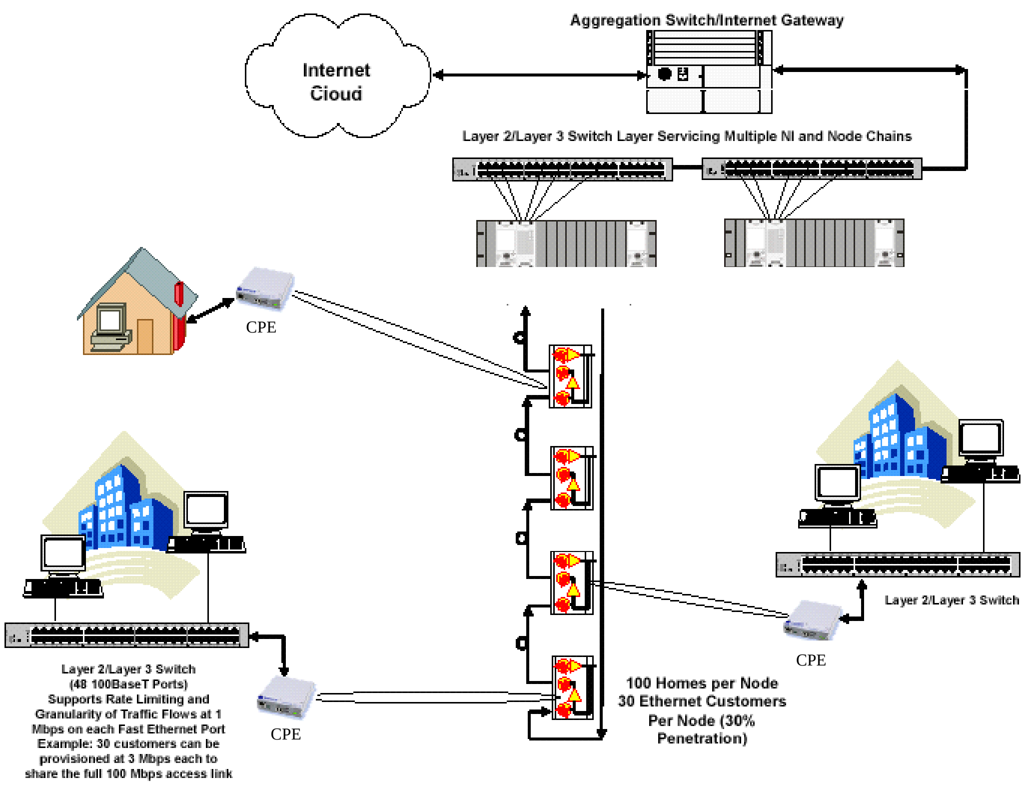


Figure 3: Evolution to FTTH Ethernet for Residential Areas

This evolution will be fueled by the trends[vii,viii] reported in the industry summary reports:

1. 400,000 FTTH subscribers worldwide implemented by the end of 2002,
2. 50,000 (or 22,500 by others) FTTH subscribers in North America
3. 50 FTTH served communities in the USA
4. Trends:
   a. 300,000 FTTH subscribers in 2003

800,000 FTTH subscribers in 2004 (1,400,000 by high projections for 2004)

SUMMARY

The demand for data communications and other telecommunications services from SMBs presents a lucrative opportunity for increased revenue. This increase can be achieved by leveraging MSOs' investment in broadband HFC networks.

Multiple choices of providing Ethernet to businesses are available to HFC network operators. These technologies and architectural choices should be evaluated against business and market (competition) requirements as well as against operator's objectives.

These solutions create an opportunity for HFC network operators to replace CLECs and ILECs as telecommunication service providers to SMBs. This market is dramatically underserved by the traditional telecommunication service providers.

Moreover, the technologies allow for competing with ILECs, CLECs and data carrier companies for large business market in a scalable and evolutionary manner.

FTTB applications can be readily and cost-effectively extended to the MDUs and MTUs today. Future price trends in the optical and digital technology may allow (and by some reports already allow) for cost-effective implementation of FTTH systems.

---

[i] AMI-Partners

[ii] Metro Ethernet Forum

[iii] AMI-Partners

[iv] AMI-Partners' 2002 U.S. Small Business Market Opportunity Assessment Report

[v] Lightreading, May 2002

[vi] Information from web-sites and published documents of the following vendors:

- Advent
- Aurora
- Cisco
- Harmonic
- Jedai
- Narad
- SwitchPoint
- Wave7Optics
- Xtend

[vii] J. Baumgartner, Fiber-to-the-home blazes an evolutionary path, CED, February 2003

[viii] R. Pease, Fiber-to-home proponents speak out against negative deployment myths, Lightwave, February 2003.

FEEDER FIBER INFRASTRUCTURE FOR THE SMALL TO MEDIUM BUSINESS DATA SERVICES

Donald Sorenson
Scientific-Atlanta, Inc.

Abstract

This paper investigates the business development and engineering advantages of utilizing ring topologies in last mile fiber-to-the-business applications. The use of traditional star or bus-star oriented network topologies become less than optimal when the realities associated with business services market dynamics and geographic circumstances are considered. It will be shown how the use of rings in the access network can reduce business development risks, is highly synergistic with existing fiber feeder plant, simplifies engineering and operations tasks over the life of the plant and can improve the MSO's service deployment velocity. All critical success factors for the MSO in the profitable deployment of fiber based business services.

INTRODUCTION

The access and transport fiber infrastructure upgrade investment made by MSO's in support of their residential broadband initiatives has positioned them well to become the major data services carriers in the growing 70B$ to 90B$ small to medium business data services market. ***In the U.S., MSO's collectively now have more access fiber passing small to medium businesses than any other competing communications entity.*** This paper assesses the advantages and potential problems associated with the use of fiber ring based access technologies which leverage fiber inventories to capitalize on this promising business development opportunity.

HFC feeder fiber is the key differentiating strategic element that is working in the favor of the cable service provider. For any carrier attempting to address the small business market space, the means for effective and efficient backhaul has been demonstrated to be one of the key barriers to market entry. Accordingly, MSO's must optimally leverage the use of existing dark fiber. Each fiber that passes pockets of businesses must be able to support as many subscribers as possible while providing meaningful service levels. Solutions must be able support multiple subscribers per feeder fiber (fiber-gain) while controlling risk, maintaining simplicity and providing low first-in cost.

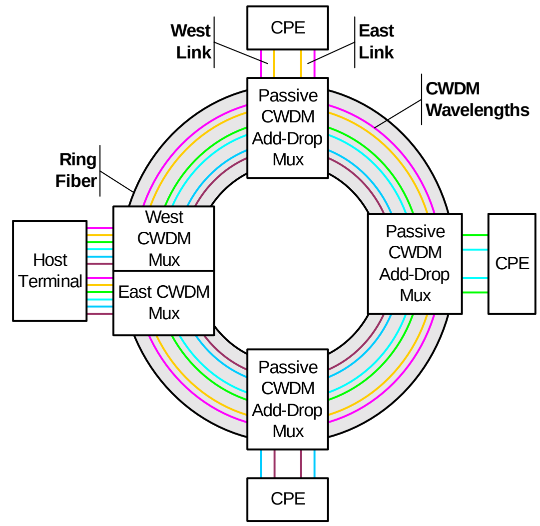
Many of the data oriented access solutions being circulated today are based upon classic bus or star local area network (LAN) topology principles that are implemented over either fiber or coax infrastructure. To achieve any degree of feeder fiber-gain these approaches generically require the use of either passive or active edge aggregation elements within a few thousand feet of the subscribers being served. Such architectures are quite suitable for applications where subscriber densities and service take rates are high and can be accurately predicted, e.g. urban or residential applications. Unfortunately for the MSO neither of these conditions typically exists when extending existing spare fiber for business building access.

There is a better way! This paper will explore the comparative strategic advantages of utilizing a fiber ring topology for accessing business buildings. Coupled with the emerging low cost ring based networking technologies such as Resilient Packet Ring[a] (RPR) and Multi-Protocol Label Switching (MPLS) Fast-Reroute[b] the MSO's can minimize deployment risks and while maintaining a high of degree of service flexibility. ***This paper will show how access fiber rings are highly adaptable to varied circumstances, reduce construction and operations demands and ease the impact on existing HFC feeder fiber.***

RING TECHNOLOGY OVERVIEW

Broadly, fiber ring technologies are based upon either optical add-drop (OAD) multiplexing or electronic add-drop (EAD) approaches. Each has its own advantages, however it will be shown that EAD approaches based upon emerging Ethernet technologies can greatly simplify deployment and operations demands while increasing the number of subscribers supported per ring.

Each station in a wave division multiplexed (WDM) OAD ring communicates to a host headend terminal over two pairs of dedicated optical wavelengths using a standard link aggregation protocol such as that found in Ethernet[c]. Each wavelength pair is routed in opposite directions around a diverse or collapsed path ring, as illustrated below.



WDM OAD Ring

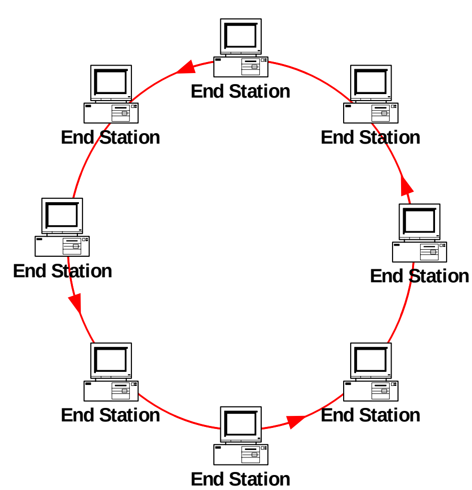
The number of stations which can be placed on an OAD ring is governed by a number of factors. For example, based upon the available wavelengths the maximum number of ring stations is one-half of the available wavelength count, e.g. a CWDM based ring can support a maximum of 9 stations per ring, assuming 2 wavelengths per station in each direction with a total of 18 wavelengths available. The combination of passive optical multiplexer and ring segment optical losses, e.g. water-peak attenuation, may also limit the number of ring stations based upon optical budget restrictions. The incremental insertion of new stations into a ring must be carefully planned to ensure that the optical performances of the new and existing ring stations are not adversely affected.

The OAD ring approach requires the MSO to accurately track to whom wavelengths have been allocated. If successful, hundreds or thousands of subscriber wavelengths, subscriber and headend-host switch ports and network service allocations will have to be accurately correlated and tracked on a regional basis over the life of the network. The problem is further complicated by the fact that network management automation tools do not typically extend to a manually provisioned optical layer. These challenges could become a serious liability threatening the MSO's ability to consistently provide a high degree of service reliability and integrity. The approach also presents a physical port scaling challenge at the host terminal, demanding two OAD multiplexers and optically interfaced switch ports per subscriber. ***A key advantage of the OAD approach is that the links between the host terminal and each CPE unit are dedicated and private; enabling the simple and effective physical layer security, privacy and individual station failure immunity associated with dedicated media connections while operating in a resilient shared media environment.***

The apparent ease by which the add-drop function is performed in an OAD solution leads one to anticipate a reduction in ring station and headend host switch complexity and expense. In a linear cascade of stations, used strictly for the purpose of providing best effort traffic over a single physical port, elegance in station design will indeed prevail since stations may be composed of simple physical layer fiber-to-twisted pair media converters. However the headend host switch must still bare the full brunt of being a carrier-class device. An additional complication arises from the fact that the headend host terminal(s) must economically scale on a per physical port basis for each port type required (e.g. 10baseT, 100bastT, 1000baseT, T1, T3 and etc.) as each subscriber is added, which can in turn drive up costs, headend space requirements and the need to track physical fiber or twisted pair port connections for each subscriber. Further, station complexity and cost increase dramatically with the inclusion of higher level functions such as support of OAD rings, multiple subscriber physical and logical ports, varied data types, managed services and QoS sensitive traffic; effectively becoming equivalent to the complexity and cost of a comparably featured EAD based ring station.

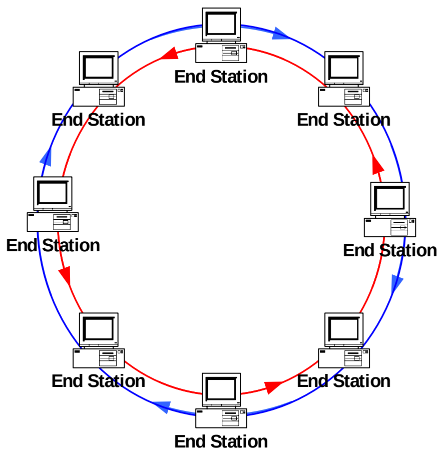
From a business development and management point of view these circumstances present a potential condition where OAD based solutions add unneeded equipment and operations expense. Over the life of the plant, headend and drop CWDM optical passives, headend host switch physical ports, headend O/E interfaces along with wavelength and switch port record keeping complications will greatly encumber the overall cost-performance of the solution. Fortunately, all of these costs can be minimized or eliminated through the alternative use of an EAD based approach.

All stations in an EAD based ring share a common set of fibers or wavelengths which greatly increases the number of stations a ring can support, for example RPR supports up to 255 stations per ring. Ring traffic is added to, dropped from or transits through each station in the ring. As previously mentioned, many of the data network architectures being derived for service provider applications, including rings, have roots in existing enterprise LAN technologies. The most common ring based LAN technology is Token Ring[d] (IEEE 802.5), which is based upon a single ring typically utilizing twisted pair copper media.



Single Ring Network

A more robust duel ring approach is used in Fiber Distributed Data Interface[e] (FDDI) and is implemented over either twisted pair copper or fiber media. The use of a dual ring gave FDDI a resiliency advantage not available in the simpler single ring approach.



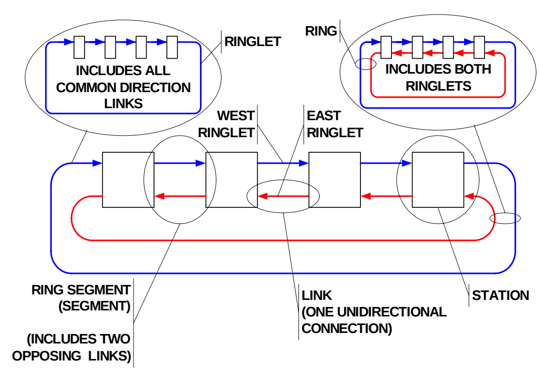
Dual Ring Network

In either case, Token Ring or FDDI, the use of rings within the enterprise has clearly not enjoyed the widespread success of star or bus-star technologies such as ATM and Ethernet. This has been principally due to higher costs and operational reliability issues associated with disruptions in ring continuity when workstations were inadvertently switched off, unplanned moves or interconnecting cabling failures[f]. However FDDI has been successful as a reliable high capacity link between core network elements such as servers and routers as well as between facilities in campus applications. In these cases the devices and interconnecting media are managed and stationary – thus they are far less vulnerable to inadvertent user manipulation.

Within the telecommunications domain, clearly the most widespread utilization of the ring topology is in the form of Synchronous Optical Network[g] (SONET). Like FDDI, SONET is based upon the principle of resiliency through the use of multiple rings, however by contrast SONET is not a data link layer technology and is incapable of performing packet switching functions on its own. SONET is the defacto standard for high reliability transport within the telecom industry, however its capabilities come at a price point that is sustainable for only the most demanding high revenue subscriber requirements. ***The reliability of SONET rings between fixed managed facilities has been extraordinary, virtually eliminating network outages due to fiber or equipment failures.***

EAD RING FUNCTIONALITY

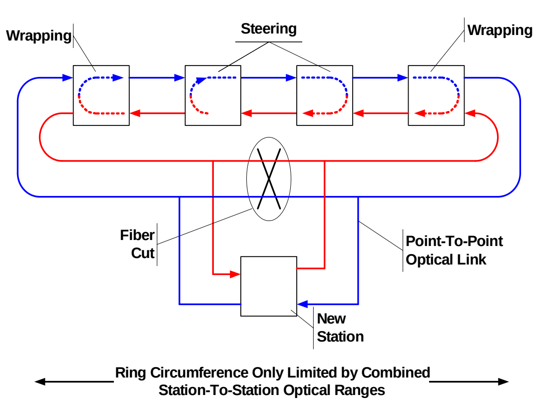
It will be useful to briefly review terminology and functionality associated with ring networks. The focus of this section will be on EAD rings while highlighting the key contrasts and similarities to OAD rings. A degree of commonality exists in the vocabulary used in the aforementioned ring standards, and for discussion purposes here the terminology adopted in the IEEE 802.17 RPR draft standard will suffice and is illustrated in the diagrams below.



Ring Terminology

The active elements on a ring are referred to as either stations or end-stations based upon the nature of the station's traffic. Stations that terminate all traffic are commonly referred to as end-stations. The term "station" will be used here since CPE or headend host devices rarely terminate all traffic. One station on the ring will serve as the ring's interface to the MSO's backbone network by aggregating non-local traffic into a few high-speed interfaces. Ring segments exist between stations and are composed of two opposing links. Most, though not all, access rings will be composed of two ringlets. Each ringlet-link is a half-duplex connection with directional signaling over an optical wavelength transported on a fiber or with the opposite ringlet-link in a single fiber WDM arrangement. Occasionally, due to circumstances associated with a lack of fiber, it may be desirable to operate the ring in an open single-ringlet mode where stations are simply daisy chained together, thereby forgoing many of the key carrier-oriented benefits of using ring-based equipment.

The counter rotating ringlets combined with a station's ability to reroute traffic in the event of a fiber cut facilitate the topology's well known resiliency protection capability. As illustrated below, stations can implement protection by either wrapping traffic at the stations adjacent to a ring element failure or by having all stations steer their traffic that was transiting the failure point away from the failure. Stations determine the condition of the ring by either continually circulating topology status information or by monitoring optical signal levels and are thus able to detect the failure and redirect traffic in less than 50ms.

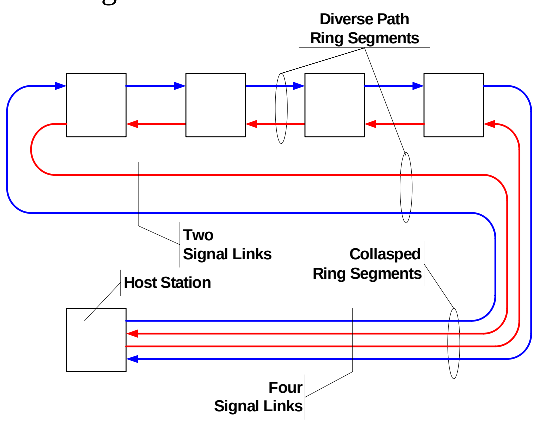


EAD Ring Protection

Typically fiber failures are accidental, but in the case where a new station is being added to an OAD or EAD ring a fiber cut is intentional and appropriate. Whether intentionally or accidentally cut the ring enters its protection mode preventing protected service failure to existing subscribers, for example a DS1 would not loose frame-lock. Once a new station is added to the ring all of the ring's stations automatically update their ring topology information and resume normal operation. It will be shown that the ability to insert a new station at any arbitrary location within a ring without interruption to protected services is a crucial benefit to using rings in last mile access applications.

A key attribute of an EAD ring is that each ringlet link is constrained to a simple point-to-point optical connection with no intervening passive devices. Each station regenerates the link for transmission to the next station allowing the link to operate over great distances with standards based optical devices. A signaling constraint is typically not placed on the length of a ring segment, and in turn the ring circumference is only limited by the combined lengths of each ring segment. In the last mile applications, station-to-station distances are typically less than 10km, virtually eliminating optical budget and dispersion constraints for both the initial ring deployment and future bandwidth enhancements. ***Decoupling the fiber plant from station optical budget constraints helps to both ease deployment concerns and insure that the plant will remain transparent to future upgrades.***

A fiber ring can be implemented around physically diverse path or within a single collapsed path as illustrated below. Two signaling links are required for each diverse path segment of an EAD ring, with each segment typically being supported by two fiber strands. Four signaling links are required for collapsed segments typically consuming four fiber strands.



Ring Segment Paths

As previously discussed, a segment's fiber strand usage can be reduced through the use of simple two wavelength WDM techniques utilizing optical combiners where existing fiber inventories are insufficient to support one link per fiber strand. Where WDM is used, each station's east and west optical interface's transmit wavelength must be appropriately coordinated, e.g. all west interfaces transmitting on 1310nm and all east interfaces transmitting on 1550nm.

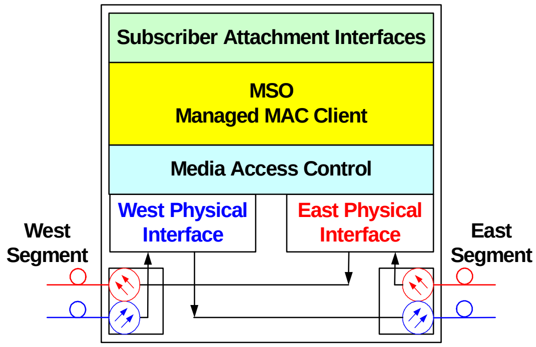
It is important to understand that a ring's segment's paths can diverge and merge multiple times along a common feeder fiber route and within one fiber feeder serving area. This attribute enables a great deal of planning flexibility while using existing fiber strand inventory still available within an HFC fiber feeder cable. ***Rather than focusing on the feeder's physical end points and spare fiber strands at each HFC node, a ring design proceeds by focusing on the feeder's many splice locations and a common set of spare fiber strands that are used in all of the feeder trunks and branches.***

ACCESS RING CHALLENGES

Though the use of ring topologies in access applications does address many of the difficulties encountered with alternative approaches, they can and will present challenges of their own. ***Under specific circumstances problems with regard to security, reliability and fiber utilization can arise in access rings.*** This section will point out these circumstances and explore methods that alleviate or eliminate their impact.

Since a ring is a shared media point-to-point topology, security, service theft and multi-point failures in either the ring fiber or stations must be accommodated. OAD ring applications must limit CPE access to those wavelengths specifically allocated. This is routinely achieved by placing the service drop's passive optical multiplexer in the outside plant (OSP) where it is secure and under the complete control of the MSO. Additionally, measures must also be taken to ensure that the ring's host headend aggregation switch provides suitable security measures to prevent successful substitution of a foreign station and/or the hacking of switch traffic or network control plane information. ***It is important to understand that, because OAD ring stations do not process neighboring station traffic, security and service theft issues are not eliminated, they are simply moved directly back to the ring's hosting switch.***

EAD ring technologies will pass a portion or all of a ring's traffic through each station on the ring. ***Mechanisms must be in place such that stations can be assured to be trusted entities; fortunately such mechanisms are native to the operation of EAD stations.*** An EAD station is typically composed of four functional layers as illustrated below.



EAD Station Function Block Diagram

The east and west physical layer interfaces provide optical linkages to neighboring ring stations. The media access control (MAC) entity manages add, drop and transit station traffic. The functional entity that is responsible for controlling all traffic presented to the subscriber attachment interfaces is the MAC client. The client is directly managed by the MSO's network management system. Traffic cannot be presented to subscriber interfaces until the client has been authenticated as a trusted network element and services have been setup by the management system; the station does not flood its switch ports to build its forwarding tables as found in a standard Ethernet environment. Station authentication and control processes are typically equipment-vendor specific, for example a system may restrict access to the network's control plane with an access control authentication technology such as RADIUS[h].

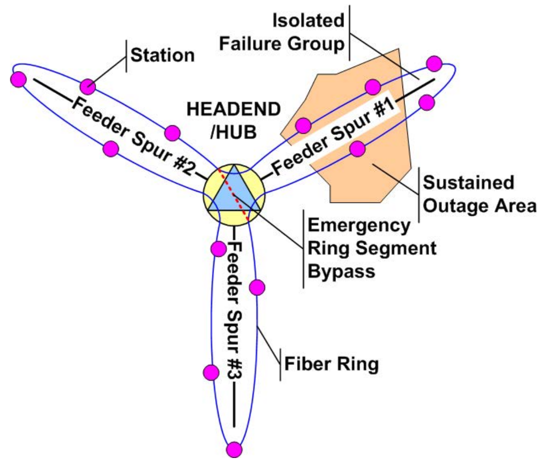
The conditions that can trigger service outages on a ring are dependent upon the networking technology used and the physical fiber path, collapsed or diverse. ***The behavior of fiber-cut induced service outage in collapsed path OAD or EAD ring is identical to that found in bus or bus-star topology; all stations downstream of a fiber cut experience an outage.*** Conversely a diverse path ring is typically resilient in the case of a fiber cut. However, multiple station failures can present a problem for EAD rings.

A multiple EAD station failure condition, though rare, may isolate or island subscriber traffic creating a service outage irrespective of how the ring fiber is deployed. The condition can be addressed simultaneously at both the local station and serving area levels. Multiple station outages can be caused by either a common widespread condition, such as sustained power outage, or a collection of simultaneous isolated station failures. To mitigate the possibility of multiple simultaneous isolated station failures in a carrier environment, experience has shown that each station should be a fixed-position managed device that is operated from a standby-power source. ***If a station standby-power source is not available, a local automatic optical bypass switch can be implemented at either the station or the building's service drop.*** Thus local station powering conditions are not likely to cause multi-station service outage conditions. A sustained utility power outage(s) within the ring's serving area that interrupts the operation of multiple stations is best addressed on a serving area basis.

Recovery from a sustained power outage may be possible through the use of optical bypass switches at each station. Historically in LAN environments this approach has not been successful due to the added cost and limitations in the number of sequential bypasses that can occur at once due to optical loss limitations. Given the relatively rare conditions under which multiple stations simultaneously fail, a strategic manual-bypass approach may be more appropriate.

Collapsed rings composed of fiber feeder segments can be equipped with sets of mechanical fiber strand splices at key feeder splice locations. Splice locations can then serve as a manual bypass point. In the event of a sustained local power outage, the effected sections of the ring can be bypassed at the nearest splice for the duration of the outage. Unfortunately, failed segments of diverse path rings cannot be so readily isolated, and measures must be taken at the hub/headend level.

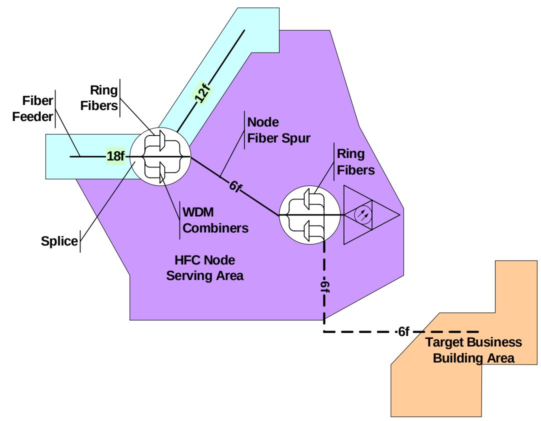
The business subscriber density within a headend or hub serving area is likely to result in a condition where a single EAD ring will route through a hub multiple times which facilitates an alternative bypass approach, as illustrated below.



Ring Spur Bypass

In this case an emergency manual or automatic optically-sensing bypass function is implemented at the hub for each ring-spur such that an effected spur is isolated for the duration of the outage. This restores the ring to full operation for the remaining subscribers on unaffected ring spurs.

When planning the deployment of collapsed fiber rings, circumstances will be encountered where the available spare fiber strands in a feeder cable will be insufficient to support a four fiber EAD ring segment. This condition is most likely to occur when attempting to extend ring fibers to a business location from a fiber feeder segment where fewer than four spare fibers are available. For example, a feeder spur to a lone HFC node typically contains 4 to 6 fiber strands with 1 to 2 strands held in reserve for advanced services; this is an insufficient amount of free fiber for the extension of a collapsed electronic add-drop ring using a conventional four fiber approach. In these cases ring fiber counts can be reduced by using either passive WDM or CWDM techniques, as illustrated below.



WDM Fiber Ring

In this case WDM combiners are implemented at the splice cases located at each end of the node's fiber spur. Implementing a WDM ring segment does impose additional optical requirements on the segment's stations. Each station's optical interfaces must transmit using different optical wavelengths and with sufficient power to overcome optical combiner insertion losses. For example, one station may transmit on 1310nm while the other may transmit on 1550nm. Other CWDM wavelengths may be selected as well. In any case, stations equipped with modular pluggable optical interfaces can easily accommodate this requirement while facilitating the use of higher power DFB transmitters for the support of ring segments up to 80km in length.

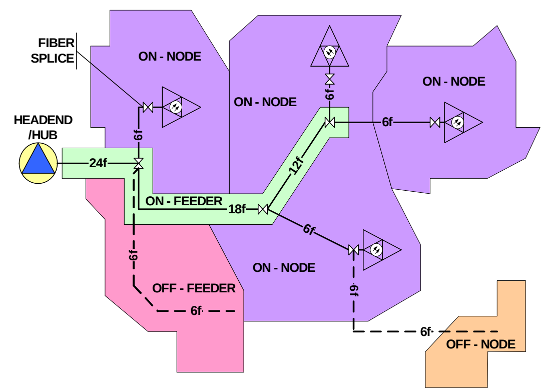
FIBER ACCESS CHALLENGES AND SOLUTIONS

Clearly the primary barrier to offering services via fiber is the cost to extend the fiber to the subscriber's building and subsequently to the subscriber's demark point. Fortunately in the majority of cases some portion of the MSO's fiber plant is within a few thousand feet of the building, however just being close usually is not sufficient. ***Effective strategies for both the targeted marketing of communications services and fiber build-out can greatly reduce risks and enhance business development success.***

Until recently the implementation of a fiber-based business service by an MSO has been based upon an incremental opportunity brought to the MSO by a service broker or by the subscriber directly. Historically, the service is typically implemented quickly and easily using a point-to-point approach. One or more feeder fiber strands are dedicated to the service at each of the subscriber attachment points to the MSO's backbone network. Unfortunately, due to limitations in available spare fiber media, this approach is not sustainable or scalable within a general business services deployment program. Alternatives which allow businesses to securely share common strands of fiber media are essential and available in both ring and non-ring topologies. ***The key attribute of a successful topology lies in its ability to flexibly adapt the diverse array of circumstances that are driven by market, geographical and operational circumstances.***

Physically, the business service market areas where the MSO is likely to enjoy the most success and velocity are those that are near existing fiber inventory and are typically underserved by legacy business service providers. Commonly these areas are suburban in nature with a relatively low density of business buildings and potential subscribers per street mile. Densities of 10 to 15 buildings and 15 to 20 businesses per street mile can be expected. This fact coupled with a low fiber service take rate will force the use of build-out strategies that can cost effectively light enough fiber to pass a large number of business buildings within a target area.

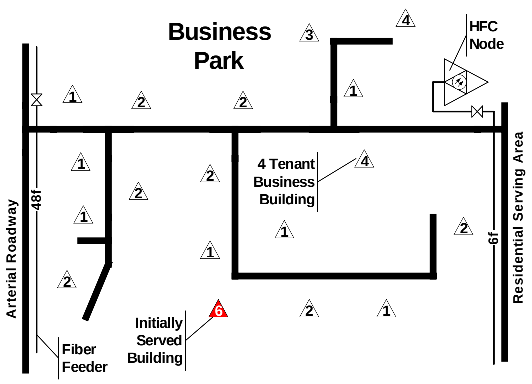
Determining how to extend fiber to business buildings is driven by where the buildings are located with respect to the existing fiber inventory, how many buildings are involved and how they are physically laid out. ***The good news is that the majority of the business buildings in an MSO's addressable market are within one or two thousand feet of the MSO's fiber plant.*** With respect to the fiber plant, business buildings can be broadly categorized as being on or off the fiber-feeder buffer or HFC node area as illustrated below.



Business Building Locations

Geographically, where businesses tend to cluster relative to the fiber is dependent upon regional circumstances as well as how the MSO chose to build the plant originally. In the majority of cases the primary fiber-feeder runs are located along arterial roadways which fortunately are also where most businesses are located. However larger businesses with more advanced communications needs may be more likely to cluster off of the feeder or node areas. Typically off-node clusters will present the greatest challenge in that the fewest possible spare fibers will be available to serve the cluster.

The number of business buildings in a cluster and their geographical layout can complicate how fiber will be extended, this is where selecting the right network topology can reduce complexity, cost and investment risk. The following example illustrates a combination circumstance where a small business complex is bracketed by arterial roadway on the left and a residential neighborhood on the right.



Business Park Example

A large distribution feeder cable runs along the arterial, while a small 6 fiber stub from another distribution branch reaches a lone fiber node. The business complex is composed of fewer than 40 businesses housed in a number of single and multi-tenant buildings. The building represented by the shaded triangle houses the initially targeted business subscriber. The simplest and lowest cost approach to extending fiber to the initial target location would be the allocation of two fiber-feeder strands that are extended directly to the target building from the nearest splice point. Such an approach is certainly low cost day-one, but quickly becomes a serious liability when the core business development objective is to sell services to other businesses in the area.

Ideally the task of extending fiber to the first business subscriber would cost-effectively position the MSO well for future sales within neighboring buildings without significant additional cost and without consuming additional feeder fiber. Unfortunately, circumstances are greatly complicated by the fact that the sale of business services rarely proceeds in an orderly fashion. Business opportunities arise in a nearly random sequence based upon existing subscriber service contracts and evolving subscriber needs and available services. Additionally, in each case the decision to serve an opportunity must be justified on its own return on investment (ROI) merits. Typically the ROI model for the first subscriber in an area will bare the brunt of the cost of extending fiber from the nearest point on the existing plant. In those cases where the physical route to the subscriber's building passes a number of single or multi-building clusters of business subscribers, as illustrated, the MSO may choose to augment the initial service extension, building out with additional fiber such that other businesses along the route can be easily provisioned for service as they are signed up over time; a common sense strategy, but one that increases the ROI risks and that has practical limits in terms of how many additional fibers can be extended and to where.

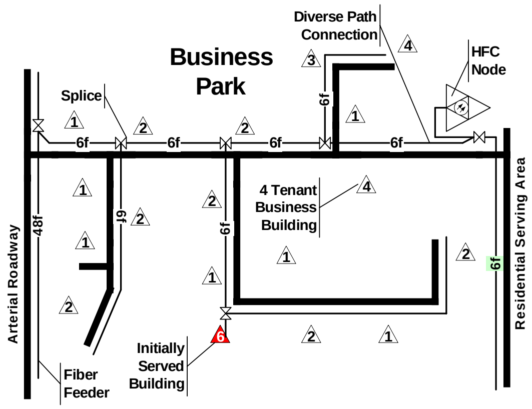
Once the decision has been made to extend fiber into a business building area it is very difficult to predict which of the remaining businesses could become customers and when. It will be to the MSO's advantage to pass fiber by as many buildings as budget will allow knowing that in a competitive environment service turn-up velocity will be a key to winning service contracts once they are up for renewal. Thus, once a service is sold service turn-up can be limited to installing a drop and terminal equipment. Both of which can proceed in a much more timely fashion than roadway construction.

Perhaps based upon prior knowledge and practice in these circumstances the shared fiber deployment approaches most likely to be initially explored are bus-star or star topologies such as PON and remotely switched Ethernet. In either case engineering challenges quickly mount, the following are just a few:

- How many buildings in the area are expected to house customers?
- How many fibers per roadway branch?
- Within budget can fiber be routed along each roadway in the complex?
- Where should fiber access points be placed?
- How and where should the fiber runs aggregate?
- Where is the fiber coming from, is the cluster too far from the host hub or headend?
- Is there sufficient remaining optical budget to reach all of the buildings?
- If an active star topology is used, where is the aggregating switch to be located and how will it be powered?
- Is there a solid record keeping system such that two, three or four years after bundles of fiber have been installed and wavelengths have been assigned will operations be able to quickly and accurately tap into the available fiber strands without disrupting existing services?
- Will service resiliency be required?
- Will existing customers tolerate the service outages that may be required to add new subscribers?

By contrast a ring based topology, particularly EAD approaches, offers several key advantages that greatly simplify or eliminate these issues. A ring approach eliminates all design variables associated with how many fiber strands and/or wavelengths to use and where. It is very likely that the same common set of strands will be used for the entire area. ***A ring also eliminates the need to know ahead of time where access to a fiber buffer will be needed. This is very important considering the fact that often business park developments are not fully mature and new buildings are being added periodically.*** Fiber rings are passive and EAD rings do not require field installed optical couplers or combiners. Additionally, all CPE optical connections are point-to-point greatly easing optical level concerns. Thus plant design considerations along with ranging considerations are greatly simplified. Since the same fibers are being used no matter where a subscriber is being added to a ring a simple color based fiber tracking methodology is all that need be followed, e.g. a four color coding scheme that identifies the east and west ring segments irrespective of location on the ring. EAD rings further simplify optical layer record-keeping by eliminating the need to track subscriber wavelength allocations. Finally, rings offer the added bonus of being able to offer resiliency in the form of diverse path fiber routing.

An example of how a ring topology can be used to address every business building in our business park example is illustrated below. Note the lack of centralized fiber aggregation points, outdoor active equipment and the ability to address all of the business park buildings with a simple six fiber buffer routed along each roadway. Fiber splice cases can be added arbitrarily for service drops as subscribers are added over time, deferring costs and simplifying attachments. Technicians need only know the east and west ring segment fiber pair color codes to successfully splice into the ring. However the wavelength allocations associated with an OAD ring must still be carefully tracked. If a diverse fiber path is desired for the feeder portion of this ring a diverse path connection can be made to the nearby HFC node’s fiber stub.



Fiber Ring Access Example

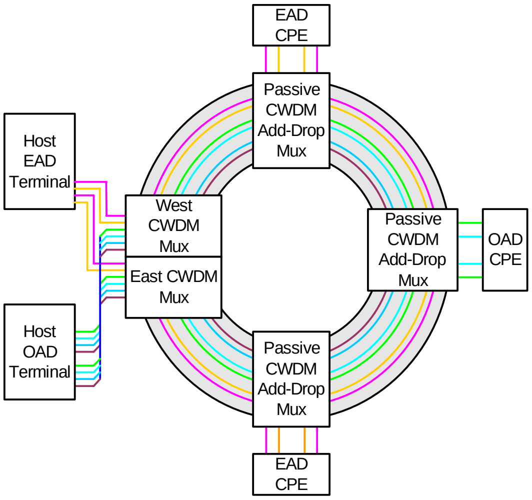
Engineering considerations aside, probably the most attractive aspect of using rings for business access is the overall reduction in deployment cost and business risk. ***Fewer components, simpler signaling and easier tracking lead to lower CapEx and OpEx costs, but the fact that a modest fixed amount of fiber strategically routed that can be accessed arbitrarily over time puts the MSO in a favorable position; one of being able to extend service quickly as opportunities arise, irrespective of take rates, building densities and geographical layouts.***

CONCLUSION

The development of business services programs will present MSO's with significant technical and managerial challenges that have little historical precedence. An integrated set of simple, clear and effective business development and technical strategies that leverage existing strengths and infrastructure will be critical to success. This requirement clearly must include fiber deployment strategies that maximize service delivery capabilities while minimizing complexity, cost and business development risks. To this end it has been shown why ring based access topologies should be seriously considered as a cornerstone to the MSO's business services program.

Two fundamental techniques for implementing fiber rings based upon optical or electrical add-drop functionality have been outlined and shown to offer common and unique properties that are largely beneficial, but can also present unique technical and operational challenges that must be effectively accommodated. Ring implementations based upon optical layer add-drop (WDM) techniques offer elegant physical layer means of extending private dedicated links to the subscriber premise that are immune to localized CPE failures. However these techniques encumber engineering, operations and management functions with difficult service extension and tracking challenges with little or no promise for an offsetting cost benefit. By contrast, ring implementations based upon electronic add-drop techniques greatly simplify deployment and management demands by supporting the automation of all service related management and tracking functions at the network management system level; enabling the normalization of how physical ring connections are made irrespective of where or when they are implemented.

A combination OAD/EAD strategy may offer a best of both worlds compromise, as illustrated below.



Combination EAD/OAD CWDM Ring

In this case two OAD ring wavelengths are allocated for EAD ring operation and are dropped into those business buildings where one or more small to intermediate business subscribers are located. The EAD ring's innate ability to aggregate at the CPE and the headend is highly scalable in this circumstance. Larger business subscribers who can afford access to a dedicated optical link can be hosted from the remaining OAD wavelengths. These subscribers are likely to be much fewer in number thus reducing scalability and plant management demands. This strategy allows the MSO to offer a enhanced differentiated high-end service to larger customers while at the same time easing plant scale and management issues.

REFERENCES:

[a] IEEE 802.17 Resilient Packet Ring Working Group, http://www.ieee802.org/17/index.htm

[b]Internet Engineering Task Force, Fast Reroute Extensions to RSVP-TE for LSP Tunnels, http://www.ietf.org/internet-drafts/draft-ietf-mpls-rsvp-lsp-fastreroute-02.txt

[c]IEEE 802.3 - 2002, "Carrier Sense Multiple Access with Collision Detection (CSMA/CD)
Access Method and Physical Layer Specifications", Section 4, Link Aggregation, http://standards.ieee.org/getieee802/portfolio.html

[d] IEEE 802.5 – 1998, "Token Ring Access Method and Physical Layer Specifications", http://standards.ieee.org/getieee802/portfolio.html

[e] ISO 9314-2:1989 Information processing systems -- Fibre Distributed Data Interface (FDDI) -- Part 2: Token Ring Media Access Control (MAC), http://www.iso.ch/iso/en/StandardsQueryFormHandler.StandardsQueryFormHandler

[f] "FDDI Technology and Applications", Authors: Sonu Mirchandani and Raman Khanna, Copyright 1993 by John Wile & Sons, Inc

[g] "SONET, Second Edition" Author: Walter J. Goralski, Copyright 2000 McGraw-Hill

[h] Internet Engineering Task Force , RADIUS (Remote Authentication Dial-In User Service), http://www.ietf.org/proceedings/96dec/charters/radius-charter.html

BIOGRAPHY

Donald Sorenson is Manager of Advanced Access Planning in the Transmission Network Systems sector of Scientific-Atlanta, Inc.. Mr. Sorenson is responsible for the engineering evaluation and planning of Scientific-Atlanta's next generation access network technologies with particular emphasis on technology migration strategies for HFC networks. Donald obtained his BSEE from South Dakota School of Mines and Technology in 1978 and has maintained a professional engineering focus throughout his career. He brings 25 years of product and technical business development experience to S-A and the telecommunications industry, of which the past twelve have been spent in the Cable Television and Telecom market segments.

email: donald.sorenson@sciatl.com

FLEXIBLE WHOLE-HOME NETWORKING STRATEGIES IN A MULTI-TV ENVIRONTMENT

Carlton J. Sparrell
Ucentric Systems

Abstract

Intense competition between cable and satellite is becoming the driving force for whole-home digital services such as multi-TV PVR. As service providers deploy customer premise equipment that enables new applications to reach all corners of the digital home, new approaches are required to manage this hardware to ensure ease of installation, quality of experience, and flexibility for different customers' needs.

This paper introduces a powerful architecture for managing distributed hardware resources in the networked home, and describes how a centralized resource manager and a flexible QoS enabled IP LAN can be used with advanced set-top boxes, low-cost clients, and other consumer electronic devices to deliver advanced whole-home digital services today while allowing service providers to maintain control of the consumer experience.

THE NEW DIGITAL HOME

Connected devices targeting the consumer are gaining momentum. Personal computer networking has taken off as broadband connections have driven data gateway sales and Wi-Fi has simplified installs. At present, some 8 million U.S. households have a network to enable Internet sharing[1].

At the same time, new TV-focused consumer applications are driving the need for whole-home connectivity. Personal Video Recording (PVR) technology has become a requirement of advanced set-top boxes as service providers seek to increase revenue and reduce churn. While many providers are taking delivery of single-TV PVR units in 2003, consumers have demonstrated their desire for a multi-TV PVR solution with 43% of PVR households having two or more units, and 74% of them saying they want PVR on all TVs in the home[2].

Adding one PVR box to each television is not only costly, but is confusing and frustrating. Viewers must deal with different recording libraries on every TV, and also coordinate passwords, recording schedules and settings. A less costly and more convenient solution is to use in-home networking with an advanced set-top box providing recording and storage. Additional televisions can access that same content through the use of low-cost media clients delivering content on demand over the home network. Instead of a PVR device on each TV at a cost of $350 to $500, an add-on media client can bring PVR to an additional TV at less than $100 retail.

While today's advanced set tops are one candidate to provide this functionality, CES 2003 saw the commitment of the consumer electronics world to the concept of the media center and media gateway. Panasonic, Philips, Pioneer, Samsung, Sharp, Sony, Thomson, and Yamaha all introduced connected entertainment devices adding networking and hard drives to devices such as PVR, DVD, and CD appliances.

COMPETITIVE FORCES

The trend of network-enabled products is the result of three powerful industry forces converging on the home:

- Cable operator competition with satellite is pushing both to look for ways of extending digital tier services to every TV in the house;
- Cable has recognized the importance of the CE channel with agreements that allow CE manufacturers to build digital cable ready devices; and
- The PC, the set-top and the CE media center suppliers are all trying to be the control center of the home.

For cable operators to prevent disintermediation of their services, they need to embrace an in-home networking platform that installs easily, supplies the quality of service their subscribers are accustomed to, provides extensibility to match evolving customer needs, and delivers a predictable, intuitive user experience.

NETWORKING OPTIONS

As service providers move to embrace whole-home strategies, there are a number of options for the delivery of media throughout the home. The most common options are:

- Analog distribution;
- Legacy MPEG distribution; or
- IP networking

Analog Distribution - One approach to distributing media around the home is modulating video and audio onto the existing coax network in the home. For example, a single advanced set-top box in the living room might provide a PVR session to a TV in the family room by blending the video and applications graphics into a single NTSC signal, and modulating that signal onto an unused or notched-out channel while an IR receiver in the family room communicates received key-presses back to the set-top on a separate channel.

Analog distribution allows a set-top or media gateway solution to distribute content throughout the house without a digital network. Unfortunately, this solution degrades digital cable to analog quality while creating privacy concerns by broadcasting every session "in the clear" to every TV in the house. The analog approach also does not scale well as hardware for every television needs to be added to the advanced set top.

Legacy MPEG Distribution - Another approach to distributing media around the home is to adapt the advanced set-top box to broadcast PVR sessions digitally using HFC modulation standards such as in-band MPEG encapsulation with out-of-band backchannels.

The one advantage of applying the HFC solution to home media distribution is that it allows existing low-end set-top boxes to be repurposed as terminals for additional sets within the home. The repurposed equipment can then utilize the in-band and out-of-band interfaces for communicating with an advanced set-top box media server which becomes a mini-head end.

Drawbacks of this approach include a) the high cost of head-end networking components, and b) the asymmetric communications channel. While the overall cost of this approach is partially mitigated by the reusability of legacy boxes, the high cost of the low-volume, HFC networking chips means that new equipment cannot follow the same low-cost volume curves of designs using commodity off-the-shelf silicon.

The HFC networking chips also present a problem in that while they allow the home server to act as a mini head-end they recreate the inherently asymmetric communications channel supported by the legacy equipment. The lack of a high-speed back-channel in the legacy equipment eliminates the possibility of pooling networked resources such as tuners and storage. This approach is also not very suitable for IP traffic and hence cuts off an attractive and important array of emerging consumer services and product.

IP Networking - The IP protocol is the most broadly deployed networking solution worldwide and is available in several forms appropriate for whole-home media distribution including wireless and over-coax.

The significant advantage of standards-based IP networking, regardless of the physical layer, is that it leverages commodity hardware available at low-cost. IP networking provides the flexibility of supporting both bursty and streaming high-speed connections over a symmetric channel. With appropriate quality of service and resource management, IP networks support the widest range of applications and efficiently handle video (including trick-play), digital audio, and data networking.

For service providers wishing to pursue the consumer applications and associated revenue opportunities enabled by IP networking, while supporting legacy equipment during the migration, a hybrid approach is possible with intelligent resource management of both legacy and IP-based network resources.

THE DISTRIBUTED HOME

As service providers adopt whole-home networking approaches to digital content, the challenges include ease of installation, quality of service, extensibility, and interoperability of equipment from different vendors. Add to this the desire to have a flexible platform adaptable to different subscribers' needs and competitive market requirements driving time-to-market.

The solution is to deploy network-ready equipment with software components capable of auto-detection and auto-configuration of all devices and resources distributed on a quality of service enabled home network.

Distributed Resources

The Ucentric approach is to employ a centralized resource manager capable of discovering, allocating, and controlling distributed resources. Consider the hardware resources common to the whole-home experience:

Video Tuners – Not long ago, video tuner requirements were straightforward: One TV, one tuner. With PVR, and now whole-home PVR, the right mix of tuners changes depending on the household. Different households will require a different number of tuners depending on the number of simultaneous live-pause and pre-recording sessions needed for their viewing habits.

Service providers must recognize that a whole-home CPE network is not one size fits all. What is required is a means for adding tuners incrementally based on the customer, and allowing the number to change as the customers' needs evolve.

Conditional Access Modules - As with tuners, the conditional access equation of the past was on a per-TV basis. With a whole-home solution, the number of conditional access modules required is more a function of a family's premium content tier and viewing habits. While the ideal is to have every tuner CA-enabled, the economics of CA-licensing or POD cost suggest that in the short term there may be premium content homes where not all tuners are CA-equipped.

Persistent Storage - The central component to time-shifted media is the hard drive. The problem with storage economics is that it is expensive for a service provider to deploy, and consumers will always want more. The solution to this problem is to leverage traditional and CE channels to allow a consumer to add more storage to their network as their needs evolve. Examples of add-on storage include 1394 drives connected to the primary ASTB or drives located in other CPE on the IP network. A resource management solution needs to be able to discover, configure and control persistent storage located anywhere on the network.

AV Encoders - MPEG-2 is the default AV encoding technology used today for time-shifting and storing analog content. As consumer oriented products evolve, this will migrate to support for more efficient encoding techniques such as MPEG-4 and H.26x. Whole-home PVR solutions today require an MPEG-2 encoder per analog tuner, but in this evolving world, it is important to plan resource management solutions to be advanced codec-ready, and be able to distinguish different encoder capabilities on mixed resource networks.

AV Decoders - As with the AV encoders, there is a migration underway towards multi-codec capable decoders. As these products are deployed, a resource management solution must be able to recognize their capabilities.

Bandwidth - Control of bandwidth is critical to maintaining quality of service in IP networks. A whole-home network may contain a single subnet, such as an 802.11g-over-coax backbone, or it may contain multiple subnets including a wireless network and local IEEE1394 buses. Disk bandwidth management is also critical for time shifting multiple high-def and standard-def streams. A successful resource management solution must be able to reserve and maintain bandwidth throughout the system for all viewing sessions even during trickplay.

Outputs - Not all output interfaces are the same. A resource management solution must be able to distinguish between active high-def and standard-def ports, and track which interfaces are configured to support DRM protected content, and which are not.

Other Resources - In addition to these common AV resources, other devices are converging on the home network. POTS lines provide necessary connections for voicemail and caller-ID information. Broadband connections provide a data link to the Internet. PDAs and cell phones offer alternative GUI opportunities. A resource management solution must remain flexible enough to support new devices and remain extensible to future unforeseen consumer applications.

Centralized Resource Management

Each of the above resources shares the common element that they all operate on streams of media or data content. Some elements, such as tuners, are stream producers. Other elements, such as decoders and their associated display outputs, are stream consumers. There are also elements that connect producers and consumers. Any of these resources can be chained to form a media pipeline to provide a service, such as live-pause TV.

Standards such as UPnP and HAVi have provided mechanisms for networked devices to broadcast their availability and allow negotiation between devices for control of each other's resources. While this ad hoc negotiating technique works for some lean-forward activities like downloading a video from a camcorder for digital editing, it lacks the level of sophistication needed to provide a single point of control to reserve resources in the future, string together chains of distributed resources, and resolve conflicts in a manner essential for the lean-back experience of whole-home media on demand.

The Ucentric centralized resource manager is capable of discovering resources as they are added to the home network and providing a single point of arbitration for the reservation of these resources among authorized applications. Around each resource is implemented a unified API allowing applications to request and control the resources they need.

HOW IT WORKS

Devices deployed within the Ucentric environment provide a service wrapper around each resource. This service wrapper provides a standard API for stream operations associated with that device. For example, an MPEG-2 decoder provides a streaming media interface supporting the Ucentric Media Protocol (UMP). This API allows other components, such as a disk-stream media server, to deliver trick-play enabled MPEG-2 content to the decoder regardless of whether the two resources are in the same box or in different rooms of the house. When a new device is connected to the network, the device reports its available resources and their interfaces using Simple Service Discovery Protocol. Applications communicate to each other and the resource manager using a networked XML interface.

Each network is required to have at least one resource manager-capable device. In the typical home this is a single advanced set-top box, but in homes that contain more than one ASTB, a negotiation protocol is used to determine which resource manager is active. The active resource manager maintains a table of all available devices, their stream resources and resource interfaces. The resource manager also maintains reservation information allowing resources to be assigned in the present or future.

Applications requiring resources communicate these requests to the resource manager. These requests take the form of a pipeline graph connecting resource requirements. For example, when a new video output is activated on the system, a new session application is instantiated and associated with that display. This session reserves the resources necessary to provide a graphical user interface and pre-recorded MPEG-2 content to that display. If that session later needs additional resources, such as tuner for watching live-pause content, an additional request will be made to the resource manager at that time.

Other applications, such as the EPG, may request resources for a future event. For example, when a user requests that a show be recorded at a future time and date, the application requests the associated resources (e.g. tuner, MPEG-2 encoder, disk capacity) for the time window required.

Example

Consider a typical evening at home with the whole-home enabled network described in Figure 1.

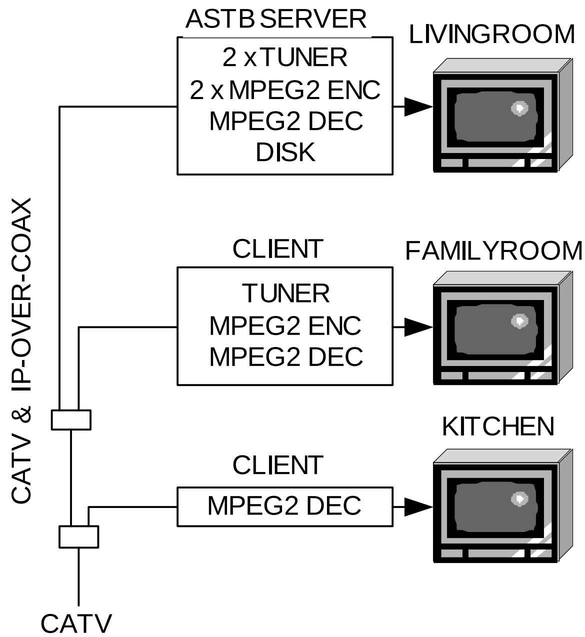


Figure 1 –Typical whole-home network

This diagram illustrates a three-TV household. The living room TV is connected to an advanced set-top box configured to be a media server. The media server contains a number of resources including two analog/digital tuners, two MPEG-2 encoders, and one MPEG-2 decoder. The media server also includes a hard disk with associated media services capable of sending and receiving media streams. The family room TV is connected to a media client with a tuner, MPEG-2 encoder, and MPEG-2 decoder. A third TV in the kitchen is attached to a client with an MPEG-2 decoder. The advanced set-top box and the media clients communicate over a 2.4GHz IP network sharing the same coax as the in-house CATV.

One morning, Dad programs the system to record a hockey game at 8:00pm on digital channel 150 using the EPG on the TV in the kitchen. The EPG scheduling application requests a reservation of an audio-video pipeline with the pipeline in Figure 2:

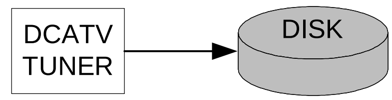


Figure 2 – Digital record pipeline

The requested pipeline includes a digital tuner. If channel 150 had been a premium station, the scheduling application would have made additional requirements on this tuner, such as associated Conditional Access or POD module. The requested pipeline also indicates the required disk bandwidth and storage capacity needed to record the program.

The resource manager searches the resource database for resources that match the request. The network contains one disk and three tuners. All three tuners have the same capabilities, differing only by their location. The resource manager uses a least-cost algorithm to construct a pipeline choosing a tuner in the server to avoid using network resources.

The resource manager checks for disk space both when the user schedules the recording and shortly before the recording begins. If no disk space is available when the user schedules the event, the resource manager checks to see if any “delete-able” files are available for deletion. If all the files on a full disk are marked as “do not delete”, the user will be alerted.

Upon successful reservation of the required resources, the reservation is stored in the resource manager reservation table for use when considering future reservation requests. A successful reservation is communicated back to the application with a reservation identification.

At 7:30 that evening, the kids want to watch a show in the family room. The show they want to watch is on analog channel 32. When they select this program from the EPG, the application calls the resource manager to request resources. The pipeline in Figure 3 is requested by the application:

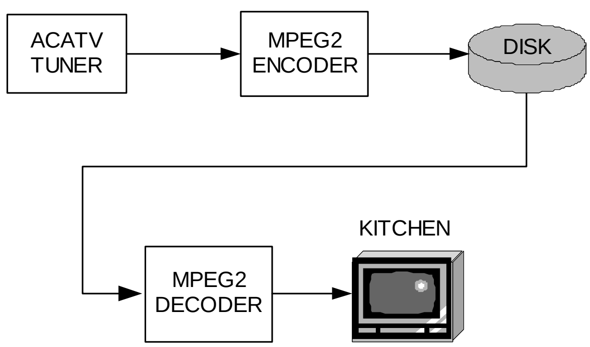


Figure 3 – Requested live-pause pipeline

There are two unassigned tuners on the network, one in the server in the living room and one in the client in the family room. While the tuner in the family room is local to the TV session requesting the tuner, the stream will be written to the disk in the living room for live-pause. The least-cost algorithm leads the resource manager to assign the tuner/encoder pair in the living room to the pipeline, saving a transfer of the encoded data twice across the network. This method preserves more network bandwidth for other uses including best-effort data transfers between PCs sharing the network.

Once the resource manager has successfully mapped the requested pipeline to actual network resources, the instantiated pipeline is returned to the application and the resources are marked as reserved (in this case indefinitely). Note that the resource manager has added one component to the pipeline (Figure 4).

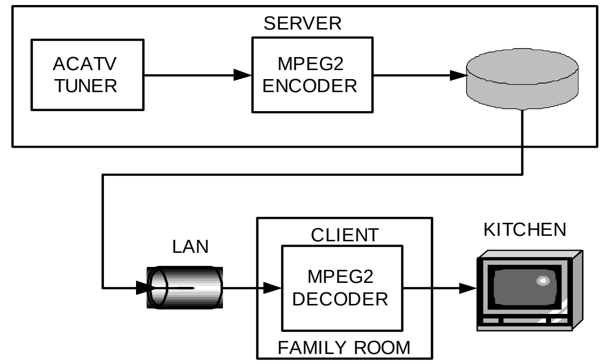


Figure 4 – Granted live-pause pipeline #1

The LAN connection is required to connect the components in the server to the components in the client. The LAN is a managed resource with guaranteed quality of service. Bandwidth allocation is controlled by the resource manager. The resource manager assigns the bandwidth requested to send one MPEG-2 stream.

At 7:45pm, Mom wants to watch a program in the kitchen. The resource manager asks for a pipeline identical to that in Figure 3. In this case, the only tuner remaining on the network is the tuner in the family room. The resource manager completes the graph in Figure 5:

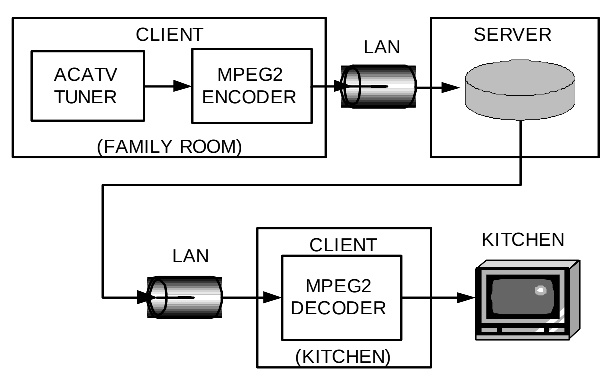


Figure 5 – Granted live-pause pipeline #2

Two network components need to be added to the graph, and twice the bandwidth reserved on the network.

At 7:50pm, the system prepares to record the hockey game, verifying that disk space is available, and removing “delete-able” content or alerting the viewers as needed.

At 8:00pm the recording of the hockey game commences.

At 8:05pm, Dad sits down in the living room to watch a program. He chooses not to watch the game, but to look through the video library, selecting a James Bond movie recorded earlier that week. The system now makes an updated request for resources (Figure 6):

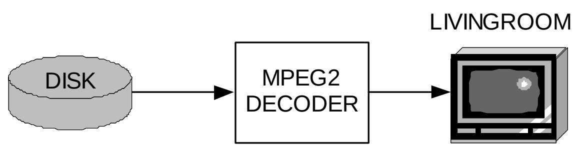


Figure 6 – Play from disk pipeline

The resource manager is able to construct this graph out of resources available in the server in the living room.

Pipelines are torn down when they are no longer needed. For example, the recording resources in Figure 2 are freed when the scheduled recording of the hockey game completes. Typically, a minimal video playback pipeline (**Figure 6**) is maintained to allow every television to have instant access to at least pre-recorded content.

FlexMedia LAN™

The distributed resource model and centralized resource manager discussed here has been implemented in the Ucentric FlexMedia LAN™ software solution. The Ucentric resource manager is capable of managing multiple LAN technologies simultaneously, including hybrid topologies such as wired/wireless or legacy MPEG and IP over coax.

FlexMedia LAN™ also provides control of bandwidth allocation for multiple streams and applications through the Ucentric bandwidth broker. This bandwidth broker provides QoS mechanisms including priority services, dynamic stream management and bandwidth allocation.

The strength of FlexMedia LAN™ stream management over the network is combined with the Ucentric distributed resource manager to pool all tuner, storage and network resources - allowing the number of tuners to be provisioned flexibly in relationship to actual household need. Additional supported services include expandable storage located anywhere on the network providing convenient installation options.

CONCLUSIONS

Cable operator competition with satellite is pushing operators to look for ways of extending digital tier services to every TV in the house. This competition has in part lead cable operators to recognize the importance of the CE channel as a means for making every corner of the home “cable ready”. With consumer applications evolving, and PC, set-top and CE media centers all trying to be the control center of the home, cable operators can leverage their quality experience with new CE relationships to deliver a compelling whole-home experience. To accomplish this, service providers need to adopt a flexible network approach to maintain control of the consumer experience.

Successful deployment of new consumer applications requires the flexibility of IP networks. The heterogeneous nature of these networks together with the need for quality of service guarantees requires intelligent resource management. Whole-home applications require a whole-home approach to deployed resources, and this requirement has led Ucentric Systems to adopt a centralized resource manager in the FlexMedia LAN product offering. This approach provides the most flexible, predictable, quality user experience available today.

REFERENCES

1. Parks Associates, "Drivers for Home Networks and Residential Gateways," 2002
2. "The PVR Monitor III" from C Cubed, October 2002

CONTACT INFORMATION

Carlton J. Sparrell
Networking Systems Architect
Ucentric Systems
2 Clock Tower Place, Suite 550
Maynard, MA 01754
carltonj@ucentric.com

IMPLEMENTING AND VERIFYING OFF-AIR DTV CARRIAGE CONTRACTS IN CABLE HEADENDS

Nandhu Nandhakumar, Jian Shen, and Gomer Thomas
Triveni Digital, Inc

Abstract

Cable-carriage of off-air DTV broadcast streams may involve the selection and transformation of different components of the stream. Carriage agreements may specify constraints on these processes. This paper lists some of the more common technical aspects of carriage agreements and describes how they can be implemented conveniently and with low operational cost. Systems that can be used to implement the agreement can also be used to monitor and verify compliance with such carriage agreements. An integrated solution is described to meet these needs.

1. INTRODUCTION

A typical ATSC terrestrial DTV broadcast stream contains one or more video programs, audio programs, and/or data programs. It also contains the Program and System Information Protocol (PSIP) metadata used by DTV receivers for tuning, electronic program guides (EPG), and other functions[1],[2]. Because a terrestrial broadcast stream has lower bit rate than a QAM-modulated cable signal, and because cable operators may sometimes carry only a selected subset of services that appear in an off-air broadcast stream, off-air DTV broadcast streams are typically re-multiplexed in the cable headend. In the re-multiplexing, the cable operators often make various changes to the original signal, such as filtering out unwanted programs, elementary streams, data packets and/or optional metadata. In addition, the cable operators may transcode input video and audio streams to reduce the bit rate in order to maximize the usage of cable bandwidth.

Cable operators face two problems when performing such re-multiplexing:

1. The modifications to the original signal may be governed by various regulatory requirements and industry wide agreements, as well as by carriage contracts between the terrestrial broadcasters and cable operators. Moreover, it is possible that a cable operator may have different carriage agreements with different broadcasters. Thus, it may be necessary to apply different policies to different signals in the same cable headend multiplex.
2. There is increasing desire to retain PSIP metadata in off-air broadcasts when passing them through a cable plant, in order to accommodate "cable-ready" consumer DTV receivers. However, most of the MPEG-2 multiplexers used in cable headends are not designed to handle the PSIP metadata.

This paper describes a headend system architecture that can be used by cable operators for processing PSIP and other metadata while simultaneously enforcing and verifying carriage agreements between terrestrial broadcasters and cable operators. Such a system can work in conjunction with one or more MPEG-2 transport stream multiplexers to produce output streams that are fully compliant with carriage agreements and the various metadata standards including the ATSC PSIP standard, and the ANSI/SCTE 65 2002 standard[3].

2. Technical Components of Carriage agreements

Carriage agreements negotiated between the local broadcaster and the cable operator may contain many technical elements. Some of the more common ones are described below. Later we present ways in which these elements can be implemented and monitored.

1. Specific video and audio services in the broadcast stream that will be carried may be specified. This can range from all services being carried to only a single (primary service) being carried. The services to be carried can be specified in the agreement by the MPEG-2 program number, stream packet identifiers (PIDs), and/or virtual channel number used by the broadcaster.
2. In the case where multiple services are to be carried, the agreement may specify if the broadcaster can allow services to disappear and appear, or if there are restrictions – e.g., if the services must be present at all times, perhaps in the form of a low bit rate "barker" channel.
3. The agreement may specify whether the operator can reduce the bit rate of, or modify in any way, the packetized elementary stream (PES) structure of the services which usually results in increased compression/reduced image quality. This operation is usually implemented by exercising the transcoding/rate-muxing features of a multiplexer, and sometimes by decoding the signal to uncompressed baseband and then re-encoding the signal.
4. If bit rate reduction is to be allowed on any service, the agreement may specify upper and lower limits of the incoming and outgoing bit rates.
5. An MPEG-2 program in the incoming stream may contain multiple elementary streams, e.g. one video, two audio (one of which could be a secondary audio program/SAP). The program may also contain one or more data streams that are associated with the A/V program. The agreement may allow only specific elementary streams to be carried, e.g., only the main audio track and no data even if it is program-related.
6. The stream emitted by a local broadcaster may contain services that have no A/V content at all and only carry data. Carriage of data–only services may be (dis)allowed. These services may be specified similar to the manner in which A/V services are specified. Additionally, specific descriptors in MPEG-2 Program Specific Information (PSI) and descriptions in ATSC DST's may be used to identify and refer to specific content types.
7. The agreement may disallow carriage of streams that contain encrypted or proprietary content that cannot be received by all cable customers on a non-subscription basis.
8. The carriage agreement may stipulate the amount of in-band EPG information (in the form of ATSC A/65 compliant PSIP data in the incoming stream) that will be carried. The minimum amount is likely to be as specified in the February 2000 NCTA-CEA PSIP agreement [4],[5].
9. However, the agreement may exceed the minimum requirement described in the NCTA-CEA PSIP agreement, and can specify the following parameters per service carried:
    - Number of hours/days of program event titles
    - Number of hours/days of program descriptions

- Various EPG table cycle times or bit rates

10. The agreement can specify branding as signaled by in-band EPG data. Specifically, whether one-part of two-part virtual channel numbering will be used and what the specific number will be. The short name and description of the virtual channel can also be specified.
11. The off-air signal typically has content advisory information, and closed-captioning signaling information carried in the EIT's. The cable operator may additionally require that these be signaled in the MPEG-2 PSI tables, or may agree to copy these descriptors (during multiplexing) in the headend from the EIT's to the PSI tables – if legacy digital STB's require this information to be in the PSI tables.
12. If the cable operator can support updates to out of band EPG data - either standards-based or proprietary in format – the agreement could specify if the in-band EPG data will feed the out-of-band EPG data. This The latter is used to drive EPG's in operator provided STB's, and in future POD-host equipped retail, digital-cable-ready TV's and STB's.

3. IMPLEMENTATION OF AGREEMENTS

DTV services arrive at the cable head-end over satellite and terrestrial links, as well as via other means (over an ATM network, for example). Multiple transport streams originating from different sources are typically multiplexed into a higher bandwidth transport stream, and are then modulated by a QAM modulator before being sent out to customers via cable. Available multiplexers can select audio/video services from each incoming stream and create a new, output transport stream with only the desired services. Depending on their source, some incoming transport streams may contain PSIP data while others may not. For example, transport streams received from off-air terrestrial broadcasts will typically contain PSIP, while streams originating from cable networks (offered as premium, hence scrambled services) may not include PSIP.

Available multiplexers can implement some of the elements of carriage agreements listed in the previous section. However, currently available multiplexers are designed to handle the grooming and bit-rate modifications of the audio and video services only. Metadata such as the ATSC PSIP tables are typically ignored by present-day multiplexers, and are either blocked from passage, or are erroneously passed through – without the needed modification to reflect the program and A/V PID grooming that the multiplexer implements. Also, the multiplexers are meant to be used in a relatively static configuration, and are neither easily easily, automatically nor dynamically reconfigurable to handle some of the cases described in the previous section.

A more general approach for implementing the various elements of cable carriage agreements described above is to use a separate metadata processing system that can process and interpret the in-band metadata in the off-air stream to (1) transform and generate new metadata, and (2) trigger pre-programmed control actions that instruct the multiplexer to implement needed service selection and A/V bit rate modification actions. Figure 1 illustrates the metadata processing system in the cable head-end environment.

The metadata processing system monitors each incoming streams in real-time and provides detailed information about the contents of the stream, including the existence of various services, such as data services, that are not easily discovered by a traditional multiplexer. To protect the investments already made by the cable operators, the system should work with existing MPEG-2 multiplexers.

The metadata processing system can either get the relevant metadata (i.e., PSIP/SI/PSI) directly from the input transport streams, or can have them passed to it from the multiplexers. The latter can be achieved by a private or standards-based protocol to exchange desired information between the multiplexer and the metadata processing system. By decoding the MPEG-2 PSI tables and the PSIP tables, the metadata processing system can identify every component in the transport stream and tell whether it is a video, audio, data or PSIP/PSI packet. Thus, a cable operator does not need to continuously and manually monitor and analyze the transport stream content to effect a multiplexer control action. The system can automatically identify the PID streams that show up in the transport stream, determine the actions such as re-mapping, blocking or passing through each PID stream according to the carriage policy parameters, and send appropriate control commands to the multiplexers.

When multiple streams containing PSIP data are multiplexed to a single transport stream, the system needs to process the different incoming PSIP streams and create new in-band PSIP data for the output stream. The system decodes the original PSIP data, obtains their semantic content, translates the data, and merges the metadata at the table level. The tables can be played out as MPEG-2 transport packets either by the metadata processing system, or the system can send the tables to a multiplexer for it to play out MPEG-2 packets at table-specific rates. Note that not all multiplexers can support download and playout of externally specified tables. The table playout rates should take into consideration the bandwidth limitations indicated in the carriage agreement. The output PSIP data stream from the metadata system is multiplexed into the output transport streams along with audio and video and other elementary streams. This forms the in-band PSIP data required mainly by cable-ready DTV receivers for tuning to in-the-clear streams.

In addition to handling in-band PSIP, the metadata processing system also generates an OOB SI stream. The aggregated SI data contains the information for all the “in-the-clear” virtual channels (VCs) in the cable system, as well as any scrambled services that the cable provider chooses to include for the purposes of discovery by POD-enabled cable-ready DTV receivers. For incoming streams that contain PSIP data, the system optionally extracts the EIT and ETT data and converts them to the aggregated SI format described in SCTE 65 2002 [3], sometimes also referred to as SCTE DVS 234. For incoming streams that do not contain PSIP, the system allows manual or programmatic input of the VC information so that it can be included in the OOB virtual channel map.

In addition, the metadata processing system may also export information to the cable operator’s proprietary program guide service to perform real-time update of the service information. Typically, the database used by the cable operator for EPG service is days or weeks old. When a program, such as a sports game, runs over time, the EPG information following the overrun program event is out-of-date. If the incoming stream contains updated PSIP information, this information can be used to update the cable guide.

4. OPERATIONAL ISSUES

The operation of the metadata processing system consists of specifying the following parameters during the initial setup phase:

1. Mapping of specific input programs/PIDs to specific output PID's, or blockage of specific input PIDs/MPEG-2 programs. This operation should be performed only once. A tight integration between the metadata processing system and the multiplexer control system will ensure consistency between PID grooming and metadata grooming.
2. Mapping of an incoming virtual channel's major-minor channel number to an outgoing/cable channel number. Mapping may specify either a two-part or single-part channel. In addition, the channel name can also be changed if desired.
3. The number of hours of EPG data to be passed through.
4. The frequency of optional PSIP tables being played out, mainly EITs and ETTs, as defined by the actual table interval time or limitation on the PSIP bandwidth.
5. Enable/disable the copying of various descriptors from the on-air event's EIT to the PMT. For example, the content advisory descriptor that may be only in the EIT should be copied to the PMT.
6. Allow/disallow transcoding or bandwidth reduction by the multiplexer for a specified transport stream, program, or elementary stream. If allowed, then specify the maximum/mimimum bit rate for that program/elementary stream.

The control information can be entered with the help of a simple user interface and the data can be saved in a format that is easy to view and edit, e.g., using XML. This XML file may capture an agreement that may be in force at multiple locations, e.g., a nationwide agreement between a broadcast network and an MSO. Instead of duplicating the manual configuration process at every location, the metadata processing system should be able to export and import the XML file (corresponding to a specific agreement) for easy implementation of the contract elsewhere.

5. MONITORING & VERIFICATION

A low cost approach for monitoring and verifying compliance with various carriage agreements within a headend is of high value to a local system's operations. To reduce cable headend operations burden and minimize the equipment cost, it is desirable to have the monitoring device be the same as the metadata processing system described above, so that all aspects related to carriage contract implementation can also be verified and monitored continually by a single system. Thus, the system should provide convenient user interfaces for contract specification, metadata transformation & generation, multiplexer configuration, stream monitoring/visualization, and configuring error-based alarming.

5.1 Error Monitoring

In addition to monitoring compliance with carriage agreements, cable operators may need to monitor input streams to prevent errors from being propagated to their customers. Typical errors that may be present in a DTV broadcast stream include:

1. Missing, invalid, or infrequent PSIP and PSI tables. Because of these metadata errors, DTV receivers may not be able to tune channels or block unwanted programs properly. Furthermore, the on-screen program guide may be missing or incorrect.

2. Missing video or audio elementary streams.
3. PCR jitter, which can cause incorrect synchronization of the audio and video signals, sometimes resulting in a “lip sync” problem.
4. Video and audio buffers overflow or underflow, which can cause ragged, stuttering video and audio.

Depending on the type of errors found, the cable operators may have different options to deal with the erroneous conditions in the transport stream. Some errors may be fixed or filtered out by the metadata processing system and multiplexers. For example, PSI and PSIP tables are normally regenerated by the metadata processing system. Therefore, certain syntax errors, inconsistent data, and incorrect table time intervals may be automatically corrected. Other type of errors may be corrected with some level of manual intervention. In addition, some errors may be filtered out if they are associated with data that are not critical. For the errors that cannot be easily fixed, the cable operators may need to contact the original broadcasters so that the problems can be promptly addressed. Since transport stream errors can also be introduced in cable headends, comprehensive stream monitoring (at various points in the path of a stream from ingest to emission) can be used to detect the occurrence and cause of errors – such as equipment failure or operator error during provisioning, reconfiguration, maintenance, etc.

Besides error monitoring, another important reason for cable operators to monitor input streams is to know the exact content and bandwidth usage of an input stream so that they can optimize bandwidth usage in their backbone and QAM feeds to the home. For example, when the program lineup changes in an input stream and leaves significant unused bandwidth, the cable operator can detect and use the available extra bandwidth to offer other services.

5.2 Contract Verification

A monitoring system at the headend that continually checks for compliance with carriage agreements is of high value to the local systems operations. All of the technical components of carriage agreements described in section 2 above can be verified automatically as described in table 1.

Different stream monitoring strategies may be applied in cable headends. One possible approach is to check the input or output streams periodically and manually using a stream analyzer. However, this approach only shows a snap shot of the stream and requires tedious routine inspections that result in high operational cost. An alternative approach is to monitor multiple streams simultaneously and use the system to verify carriage contract compliance automatically. This approach is the most cost-effective. The operator is only required to enter the contract items that are to be monitored in a standard template using a simple user interface, he can repurpose the contract specification file used for implementing the contract. To meet the needs of automatic contract verification for multiple off-air feeds, the monitoring system should provide the following features:

1. Allow multiple inputs and performs simultaneous analysis on different streams in real-time.
2. Provide remote user interface for displaying and analyzing data.
3. Provide an alarming function to inform the cable operators or the original broadcasters once any errors are detected in the stream.

4. Provide a recording function that is automatically triggered by defined error conditions in a transport stream for further analysis.

6. SUMMARY

Whether the cable operator implements must-carry type of contracts or private carriage agreements, a large number of technical parameters must be addressed in implementing the agreement. The technical issues will only continue to become more numerous and complex as penetration of DTV increases and service providers and content providers seek to innovate and expand offerings. In the near future the deployment of retail, cable-ready DTV receivers based on OpenCable and related standards, and the launch of interactive services will create these additional technical requirements.

Current digital cable headends are equipped to deal with only a limited number of these technical issues. Equipment deployed today can only groom and rate-shape audio and video programs, and they also require static channel line-ups. In-band PSIP metadata is ignored and cannot be correctly transformed. The ability to handle this new type of stream element, as well as the ability to handle frequent and periodic changes to incoming stream composition, requires new types of stream processing, monitoring and operational support systems.

Triveni Digital has developed and deployed the StreamBridge system to meet a critical need in this area. StreamBridge is designed to be a low cost, highly integrated system to facilitate the bridging of off-air streams to digital cable systems. This system is designed to analyze the composition of incoming streams, comprehend carriage contracts specified in convenient XML format, instruct multiplexers in the headend to implement appropriate grooming and rate-shaping actions, and monitor/verify the resulting streams for carriage contract compliance. The system is scalable and a single unit can support a large number of multiplexers / transport streams. The StreamBridge system provides remote user interfaces, and can interface with multiplexers that are not co-located. The system software can be easily upgraded to evolve with changing standards, operational needs, and other headend equipment changes.

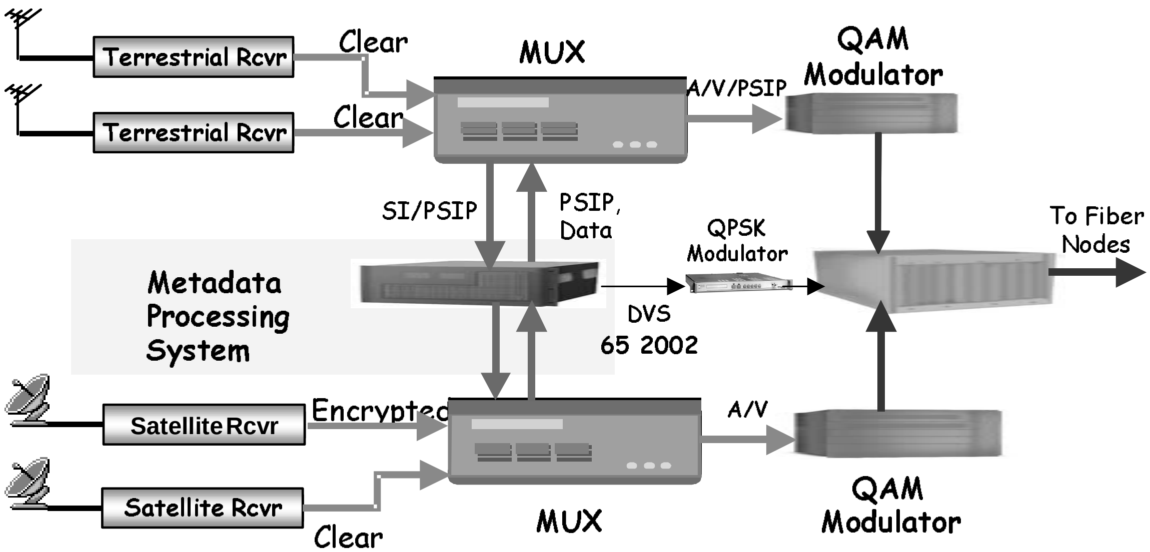


Figure 1: Metadata Processing System in Cable Head-End Environment.

Table 1. Common technical contract items and related monitoring parameters

Contract Item	Monitoring Parameters
No bit rate reduction	Monitor and compare the bit rate of specified elementary stream.
No video trans-coding	Periodically compare the input stream with the corresponding portion of the output stream for changes in bit rate.
Meet minimum bit rate requirement	Monitor the bit rate of specified elementary stream and compare it with a specified value.
Carry specified program or all programs	Discover the specified programs in the input stream. Identify the same programs in the output based on program number, PMT PID or virtual channel mapping information.
Carry specified elementary stream or all streams of a program	Discover the specified elementary streams in the input stream. Verify that the specified elementary stream is present in the output stream based on PID mapping information.
Carry specified data program	Monitor that the data program is present in the output stream.
Carry required PSIP data	• Verify the presence of the virtual channel in the virtual channel table. • Verify the number of EITs present in the transport stream is the same as specified. • Verify the presence or absence of ETTs as specified. • Perform cross table analysis to check the data among different PSIP tables and between PSIP and PSI tables are consistent after re-multiplexing.
Verify channel branding	Check if the correct virtual channel name, major and minor channel numbers are present in the output VCT.
Verify the bit rate of PSIP data.	Identify the PSIP data that are related to the particular input broadcast streams. Calculate the bit rate of all PSIP data that belong to the input stream and compare it to a specified value.
Carry required descriptors	Analyze the tables that carry the descriptor and verify that the specified descriptors are present and correct.

REFERENCES

[1] Program and System Information Protocol for Terrestrial Broadcast and Cable (Revision A) and Amendments 1A, 2 and 3, ATSC Document A/65A, 31 May 2000.

[2] ITU-T Rec. H. 222.0 | ISO/IEC 13818-1:1994, Information Technology — Coding of Moving Pictures and Associated Audio — Part 1: Systems.

[3] Service Information Delivered Out-of-Band for Digital Cable Television, ANSI/SCTE 65 2002.

[4] NCTA-CEA PSIP Agreement, February 2000

[5] Memorandum of Understanding Among Cable MSOs and Consumer Electronics Manufacturers, 12 December 2002

CONTACT INFORMATION:
N. Nandhakumar,
Triveni Digital, Inc.,
40 Washington Road,
Princeton Junction, NJ 08550 USA

INGEST & METADATA PARTITIONING: *REQUIREMENTS FOR TELEVISION ON DEMAND™*

Robert G. Scheffler, Chief Architect
Broadbus Technologies, Inc.

ABSTRACT

On demand video services, such as today's Video on Demand (VOD), Subscription Video on Demand (SVOD), and the fast-approaching Television on Demand™ (TOD®) are enhancing the consumer television experience and creating new, exciting revenue opportunities and increased cash flow for cable operators and content owners alike. However, the technical requirements to support these services are becoming more demanding and complex. In VOD, cable operators are seeing solid buy-in rates, repeat purchase patterns, and concurrency rates of 3%-10% with limited marketing and promotional support. With recent trials of SVOD and an increased number of popular titles, concurrency rates have 'smoothed' the peak usage rates throughout the week to numbers that often approach 10%-20%. However, with Television on Demand (TOD) services, consumers will have considerably more programming choices including movies, subscription-based content, and the most popular broadcast content. It is anticipated that concurrency rates of TOD may steadily climb to levels that approach 30%-65% -- rates that mirror the total concurrent U.S. television viewing audience as measured by rating services such as Nielson.

Increased service usage, additional content, and new business models are challenging MSOs to conduct unprecedented network architecture preparation and planning. In addition, decisions related to asset distribution, content propagation, network loading, metadata and rules issues need to be addressed to make Television on Demand a commercial reality.

This paper will address the issues and requirements associated with server ingest of broadcast content and content propagation. It will also discuss the architectural implications for the VOD server and propose a new class of server to support TOD requirements. The paper will also discuss how TOD content is managed through the creation and distribution of enhanced metadata formats in an environment that is controlled by studios, distributors, and cable operators.

New video server architectures and rules-based content control and propagation systems become integral contributors to the success of future on-demand services.

VOD/TOD CONTENT INGEST

The issue of the ingest of broadcast television content is one that will become more and more important for advanced video services such as Television on Demand to become a reality. As more content is made available and concurrency rates increase, architectural decisions will have to be made to support these increased demands on the network. A new architecture comprised of higher density VOD/TOD servers with the capability to ingest broadcast television will be required to support ever increasing content libraries and stream counts. However it is important to look at the evolution of VOD architectures to understand how those requirements will change in the future.

VOD in the Past

In the early days of VOD, movies were distributed on tapes. These tapes were shipped to each site that required a specific movie title. Using an encoding rate of 3.375 Mbps and an average movie length of 100 minutes, the total size of each movie was roughly 2.4 GBytes. A typical installation might contain a library of under 100 movies and was capable of streaming to less than 1,000 subscribers simultaneously.

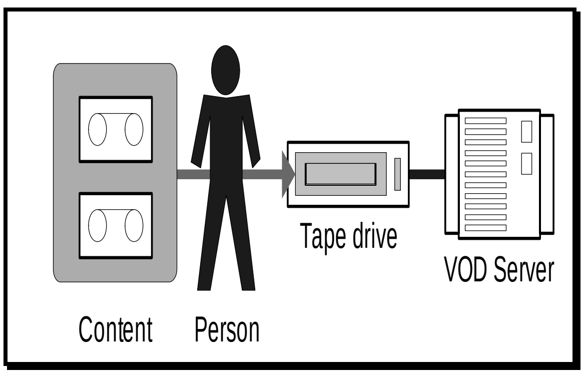


Figure 2-1 Content Ingest for VOD in the past

In early VOD deployments, metadata or other business rules weren't typically supplied with the content. The operators themselves were responsible for deciding what rules applied to particular content and for entering the appropriate rules into the VOD server or control system. This relatively simple model meant that most of the attention was focused on the billing interfaces, set top box (STB) client, and head-end control. With low stream counts, movie titles could be loaded during off-peak hours when the VOD server had more processing capacity to focus on the ingest functions. This was very labor intensive with a single operator feeding tapes and entering rules to instruct the STB guide software about the pricing and availability of new titles (see Figure 2-1). Keeping up with content ingest was quite manageable for the operator and the conventional VOD server.

VOD Today

As an industry, VOD has matured beyond the simplistic example described above. VOD installations now enable 1,000 to 3,000 customers to access a library of 150 to 300 movies. As a result, shipping tapes to VOD enabled head-ends has proved to be a logistical challenge and has evolved to a newer model called pitch-and-catch, where content is distributed by private broadcast to remote stations and syndication partners via satellite (see Figure 2-2). With increased library sizes, increased stream counts and more diverse suppliers sending data, the distribution and propagation of content has shown itself to be quite a challenge. Content can still arrive on tapes and is caught by catchers along with trailers, posters, and rules that are required to put it all together.

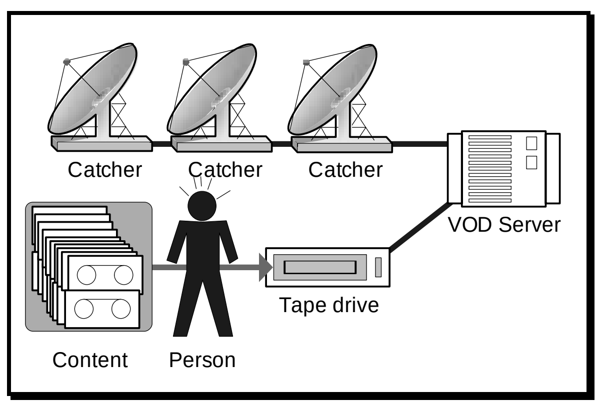


Figure 2-2 Content Ingest for VOD today

Content aggregation companies have risen to the challenge by offering services to edit, adjust, and compile these diverse formats and metadata into a nice bundled package to be pitched and caught. However, a fundamental problem is that while quite adept at low-volume streaming, conventional VOD servers usually lack in their ability to simultaneously ingest large quantities of content. The situation multiplies itself as we add streams, services, storage, and begin to distribute more hardware throughout the network.

Combining SVOD with VOD

Subscription VOD (SVOD) increases the existing VOD content library by adding 50-100 movies and other content and making them available to an increased number of subscribers. Even with a limited amount of content offered, trials of SVOD to date have resulted in increased concurrency rates that may be as high as 10%-20% or 3,000 to 5,000 streams in a typical system.

These concurrency rates place tremendous demands on the streaming capacity of the network. Also as stream counts increase, so does the problem of content ingest. To increase the stream count, additional streaming servers are required. These additional servers need access to the library of ingested content. If a given piece of content is to be made available to every customer on the network, the content needs to be either locally stored or remotely accessible. One way to make the content accessible is to add an ingest server or propagation server at the point where the content is caught or loaded from tape. This ingest server could then locally store the content, making it available to the rest of the servers. Alternatively the ingest server could be used to propagate or distribute the content to the streaming servers, whether local or remote (see Figure 2-3). Remembering that the streaming servers are primarily intended for streaming, there is a fixed amount of bandwidth available for large amounts of content propagation. To now handle the ingest of a significant amount of content, a conventional VOD server will typically lose some, or much of its streaming performance.

Today's VOD server systems adequately accommodate the demands of low-concurrency VOD/SVOD deployments. However, adding the task of ingesting numerous channels of broadcast content to conventional VOD servers creates a massive hardware and software infrastructure that takes up a lot of space, consumes a lot of power, and is inherently less reliable.

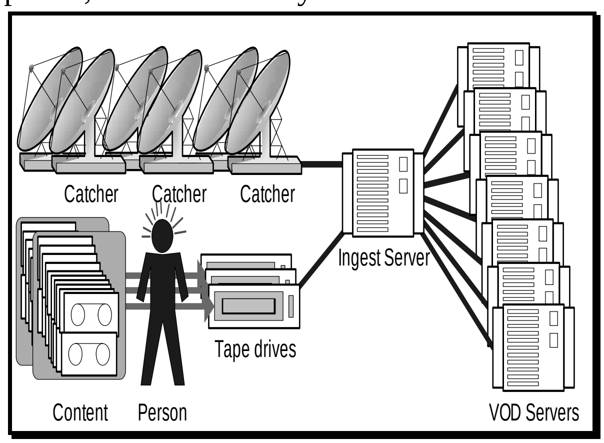


Figure 2-3 Content Ingest for VOD in the future

Television on Demand using Conventional VOD Servers

Now let's look at an example where we expand the VOD/SVOD service offerings to include Television on Demand (TOD). TOD enables cable operators to provide on demand delivery of live or pre-recorded broadcast television services as well as the movie and subscription-based content that VOD/SVOD offers. TOD is especially attractive to television content owners because it allows the viewing and sale of older programming that is out of syndication. TOD enables the consumer to have PVR functionality during broadcast television viewing without requiring a hard-drive in the STB. At a minimum a TOD system should be capable of storing 1,000 movies for VOD/SVOD customers, plus 10,000 hours of captured broadcast television.

With Television on Demand, ingestion, propagation, and streaming of content needs to occur such that the customer still feels like they are watching broadcast television. In addition to the plethora of content, trailers, posters, and rules that VOD/SVOD requires, there is now a real-time requirement for low latency content ingestion. Current VOD/SVOD systems, complete with catchers, tape drives, and content ingest propagation now have to support the ingestion of broadcast television feeds (see Figure 2-4). The path of the broadcast feed to the broadcast ingest server to the ingest and propagation server to the VOD server and then to the customers is an operation that will take many seconds and must occur at the same time as the propagation of VOD content to the VOD servers.

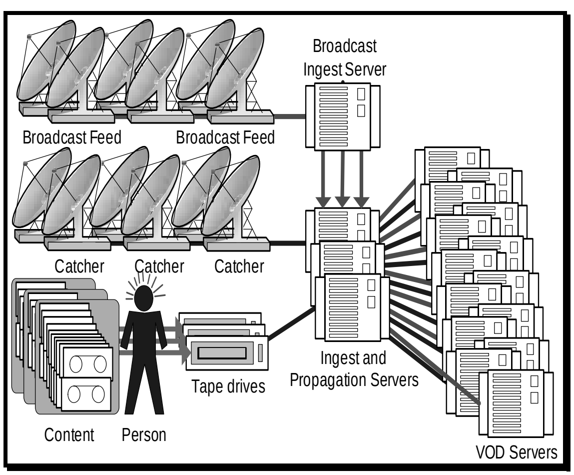


Figure 2-4 Content Ingest for TOD using VOD servers

Consumer concurrency rates for TOD will require a much higher stream count than the current growth projections for VOD offerings. When VOD, SVOD, and TOD are combined a typical system may require 20,000 to 40,000 simultaneous streams. For example, using conventional VOD servers capable of 500 streams each would require 80 servers to satisfy the stream requirement. However, as more conventional VOD servers are added, the problem of propagating the content to all the servers increases exponentially and creates the need for more ingest servers to propagate the content so that eventually there is a hierarchy of ingest servers to streaming servers. A conventional VOD server is designed for streaming to customers, not for moving, propagating and ingesting television content. Therefore today's VOD severs are not the optimum solution for this compelling, new application.

Deploying TOD with TOD servers

The critical issues that must be addressed to adequately support TOD are content ingest and stream count. A new class of TOD server is required that can ingest dozens of channels of broadcast television while simultaneously redistributing thousands of streams with zero-latency. The associated delays can be removed by running the broadcast feeds for ingest directly into a TOD server where they can be directly streamed to customers without requiring an external hierarchy of propagation servers. This solves the content ingest and propagation problem presented by TOD. However, a hierarchical approach to storage is also required for off-line VOD/SVOD/TOD content access. What is needed is a distributed storage strategy with shared local storage as well as shared remote storage that decouples the streaming functions from the storage functions. By decoupling these functions, stream-count and storage-size can be scaled independently while storage can be placed in the network where it can be used in the most cost-effective way. A master head-end containing a pooled storage library would allow a group of servers to access lesser-used programs without requiring local copies. By using this distributed storage architecture, each type of content can actually be moved and positioned in the network for the perfect balance between hardware and transport costs. As the needs of the network change, the placement of system components can change as well.

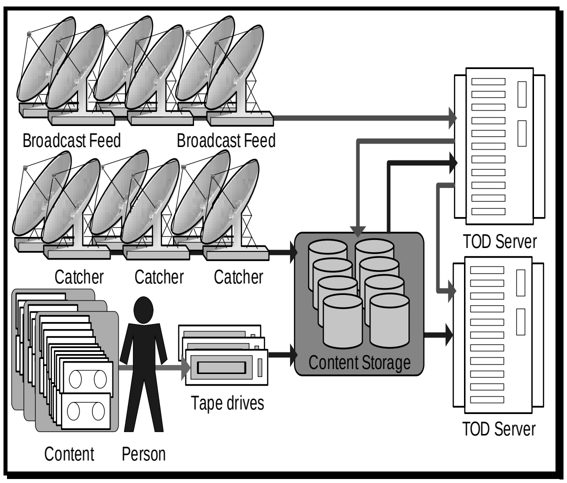


Figure 2-5 Content Ingest for TOD using TOD servers

A flexible architecture that can handle low-latency live-ingest as well as pitcher-catcher and tape based distribution models would be ideal for cost-effectively supporting TOD applications. The capability to decouple streaming from storage, while being able to distribute the storage anywhere in the network, would also significantly improve the economics of TOD. With streaming positioned in one place and storage distributed throughout the network the new architecture will scale to support even the most demanding TOD applications (see Figure 2-5). The future of VOD, SVOD, and TOD are dependent on a new architecture where scale can be controlled and each environment can be tailored for specific applications with unique requirements.

Summary of Content, Streams and Ingest

As needs grow and new business models are introduced, the capacity and scaling of VOD streaming servers are being tested. Content libraries are increasing and greater concurrency is leading to higher and higher stream counts (see Figure 2-6). With the introduction of TOD, the added ability to ingest broadcast television with low-latency is transitioning from an interesting feature into an absolute requirement.

Conventional VOD servers are being taxed to the limit with only a modest library change rate per month. As content libraries grow, to prevent libraries from becoming ‘stale’ with old content, an increased demand is being placed on off-line ingest. Even now, conventional VOD servers are reaching their limits in being able to keep up with SVOD and VOD applications. Regardless of how much streaming requirements increase as TOD begins to proliferate, the cable operator will be forced to add additional servers just to handle ingest tasks. Even then, the resulting system will not adequately address the problem of broadcast ingest to streaming latency. The clearly superior solution is to use a new class of specialized TOD server capable of ingesting and directly streaming with no perceivable delay.

Application	Movie Library	Library Change	Real-Time TV Ingest	Concurrency Rate	Stream Count
VOD	150-300	15/month	0 streams	5%-10%	1,000-3,000
SVOD/VOD	200-400	40/month	0 streams	10%-20%	3,000-6,000
TOD/SVOD VOD	1,000	100/month	100 streams	30%-65%	20,000-40,000

Figure 2-6 System Capacities for VOD, SVOD, and TOD

METADATA AND CONTROL

Rules are needed

The business of broadcast television today is very complex. The participants are numerous -- content owners, content aggregators, content distributors, broadcast and cable networks, MSOs -- and the relationships between the players are dynamic. What keeps content flowing from creators to consumers is the execution and enforcement of detailed contracts. These contracts determine the rules of "how", "when", and "by whom" content may be viewed. Whether it's a re-run episode of "Friends" that airs in syndication on TBS or a live broadcast of the New York Knicks on ESPN 2, there are specific contract-based rules that govern the manner in which content is handled. Therefore, it should be no surprise that a system of contract-based rules will continue to govern (and perhaps with greater emphasis) in a business that combines broadcast television content with on-demand content.

When VOD was initially deployed, the rules were relatively simple. MSOs would license a window of time when a movie would be made available to its subscribers. During the licensing window, the movie would be placed on the VOD Server and be available to subscribers. After the window was over, the movie would be deleted from the server. A set of rules, or metadata, capturing the pre-negotiated License Window Start and End Times would be read and enforced by the VOD server.

As the industry moves towards SVOD and ultimately TOD, the same set of complex rules and attributes must be applied to each piece of content. Examples of additional rules for handling television content could include:

- Specific days of week when content is available
- One or more timeslots during the day
- Time range that the program is available on a particular day
- Specific commercials that must be carried with the program
- Trick-mode rules and attributes (specific speeds, enabled/disabled functions)
- Specific customer groups by demographic or geographic regions

Rules should be entered and applied as early in the process as possible. There are rules from many levels. Examples include:

- Content owner or studio
- Studio distribution arm
- Content aggregator
- Television network
- Local television station

- Cable MSO
- Cable local unit

Some of the rules apply to VOD, some to SVOD, and some only to TOD. The key is that there are many rules that can come from any number of places. While it can seem daunting, it is quite easy to create and manage these rules.

Partitioning Metadata

The Video-on-Demand Content Specification as published by CableLabs has become the de-facto standard of how metadata is created and how it can incorporate many of the rules necessary to describe how on-demand content is to be handled. Initially written to support VOD (movies), it has been expanded to support SVOD. Moving forward, it is likely that the specification will need to be expanded to support all forms of on-demand content, including broadcast television.

Some metadata rules pertain to the specific content itself, while others apply to how that content is distributed and sold. One piece of content from a studio can be sent to many cable systems across the country. If the studio had to regenerate the content metadata each time, it would become a painful process that nobody would want to use. However, if the content specific metadata were attached or imbedded in the content itself, and the distribution specific metadata was separate, then the same content with metadata attached could be sent to many locations, with a different version of the distribution metadata. Thus, the content metadata and the rules-specific metadata has been partitioned.

1. Content Metadata

Content metadata includes program specific things such as a unique identifier, title, rating, description, time, actors, directors and crew, category, trailer file names, poster file names, etc. This type of metadata does not change, no matter who, what, when, or where it is distributed. This metadata could clearly be embedded in the actual content file and would stay with the file no matter where it goes.

2. Rules-specific Metadata

The rules-specific metadata starts at the content creation studio. The studio decides if there are any specific restrictions on the distribution and sale of this content and passes those rules along to the content distributors. For example, there may be a requirement to restrict a specific category of commercial - A “Friends” episode may require Coke commercials, but not Pepsi. From there, the studio distribution arm may require more specific rules. “Friends” may be allowed from Monday through Friday anytime, but not Thursday from 8-9 pm, to prevent intruding on first-run episodes. Further downstream, the television network may decide to allow viewing anytime on Tuesday and Wednesday because those are non-peak days. The local television station may want to restrict viewing from 10-11pm during the local news hour.

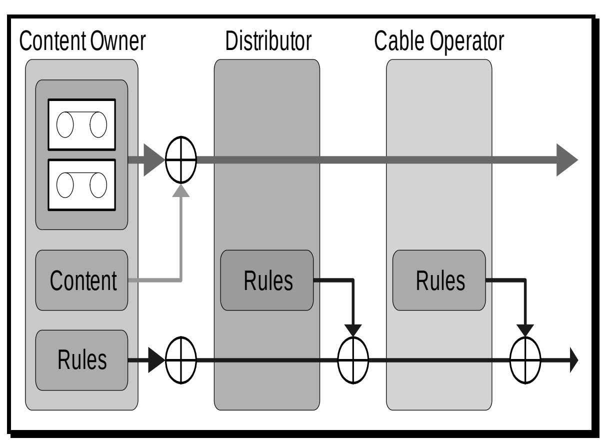


Figure 3-1 Rules-specific Metadata Flow

At each step along the way, the rules can become more restricted, but cannot be less restricted. In this manner, the content rules become more and more defined as they propagate downstream to the network operator and eventually the consumer (see Figure 3-1). Each system along the path is responsible for obeying the rules imposed upstream, and can expect each system downstream to obey the rules it passes on. When they reach the cable system, the TOD menu or EPG is built using these rules for the content received. By using this approach, the menus for the STB can be automatically and dynamically constructed.

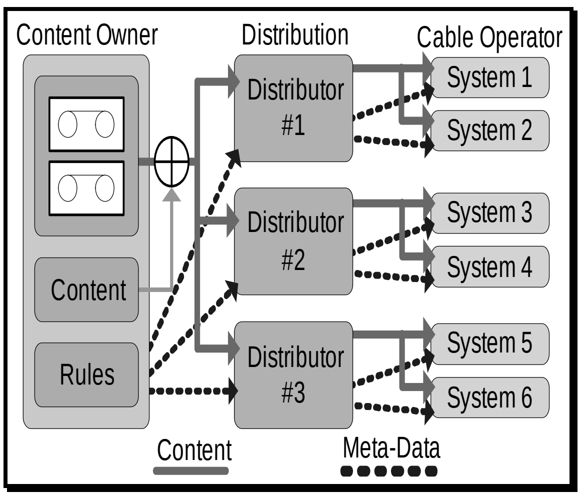


Figure 3-2 Metadata Flow to Multiple Downstream Paths

At each step in the process, there can be multiple downstream paths (see Figure 3-2) to both multiple distributors and cable systems. For example, the studio could sell a "Seinfeld" episode to the WB for certain nights in a specific week, and TBS on other nights. From each step facing down, the metadata can fragment, meaning there is a one-to-many relationship at each step of the way. This is important because at each level, a seller can sell to multiple customers. However, it would be inconvenient to have to re-record and re-master content each time it was sold. An improved solution would be to ship the exact same content to each downstream customer, but each would be supplied with unique rules-specific metadata which can be changed or updated at any time without requiring the entire piece of content to be resent.

Creating Metadata

With the two distinct types of metadata, appropriate software will be required to author and control its creation. A key ingredient is a unique identifier used to tie the asset together with both forms of metadata.

1. Content Metadata

The content specific metadata is created at the earliest possible point in the production and distribution chain. The best place for this is at the studio or encoding provider. In cases where the content is broadcast television, the content metadata could originate from the television network, or other production company supplying the network feed.

2. Rules-specific metadata

The rules-specific metadata can be created and adjusted at any point in the production and distribution chain, but would typically be originated at the same point the content is generated. For live television events, the rules could and should precede the actual content transmission. By sending the rules ahead first, the STB EPG can be populated, or other similar guide related decisions can be made.

Propagating Metadata

Both forms of metadata need to be sent along the same path as the actual content.

When any piece of content is sold or distributed downstream, the content metadata is included with the actual content along with an edited copy of the rules-specific metadata. Every copy of downstream content could have a unique set of rules-specific metadata, but the content metadata would stay the same. This allows each downstream provider to receive different rules, and allows them to be changed at a later time. When the rules change only the rules-specific metadata need be resent, not the content metadata or the entire program content. With this approach, any distributor in the chain can revise and update their rules-specific metadata as necessary.

Enforcing Rules-specific metadata

1. Asset Distribution

To make this system viable, each video server or file server along the asset distribution path must receive and obey rules encoded in the metadata. Typically in the role of asset distribution, all that is required is to pass-on the rules given to us. At any point in the path, the rules can be edited to become more restricted, but never less restricted. As assets are moved downstream to the cable plant, appropriate TOD software will pick-up the rules-specific metadata. The TOD software will use this rules-based data to build the availability matrix of programs, and associate a local time-slot for the consumer. The TOD server software is then responsible for ensuring that the studio/distribution/network rules and permissions are obeyed.

2. Content Propagation

When propagating content throughout the cable system, there can exist specific rules related to perishable content, or content that has a limited availability window. When this type of rule is implemented, it is important that the system remove such content and make the storage and streaming space available as quickly as possible. Another situation where the propagation of content is important is when a known high-concurrency program arrives and needs to be propagated to many places in a large network to facilitate the expected high demand.

CONCLUSION

In this paper, we have examined how conventional VOD servers are limited in their ability to ingest content and support the increasing stream requirements of TOD. There is a considerable impact in the output stream count as a VOD server is asked to ingest more content. With most existing systems, there is a non-linear loss of streaming capability while ingesting content. Specifically, many output streams may be lost for each single stream ingested. As the number of titles increases in VOD libraries the problem becomes more and more apparent. To reduce the impact on a VOD server, ingest of new content can occur after-hours. However this is just a temporary solution and won't scale as ingestion requirements continue to increase. With the upcoming everything on demand revolution, including Television on Demand, the ingest limitation of existing VOD server architectures becomes catastrophic. The more bandwidth consumed by ingest, the less bandwidth is available for streaming functions. Therefore more servers are required to keep the same stream count. As more servers are added, ingest and propagation becomes more and more complex. Elaborate ingest servers with content propagation services are a short-term solution but problematic longer term as unacceptable latencies are introduced to the distribution of broadcast television.

A new breed of servers designed specifically for Television on Demand is required. These servers need to handle over 100 streams of live ingest while simultaneously redistributing the ingested content to over 20,000 output streams. The server must not suffer any performance degradation in output streams while ingesting live or non-live content. The latency through such systems must be low enough to enable live television with trick-mode functionality similar to that of DVD. The streaming elements and the storage elements must be separately scalable and movable within the network.

With the plethora of ingested content from VOD, SVOD, and TOD, new means for authoring and propagating metadata must be implemented. In addition to content metadata, a new class of rules-based metadata will be required to protect revenue streams by allowing a rules-based distribution and STB presentation of content. The metadata must be partitioned and carried separately from the actual content to allow updating as well as customization depending on the MSO and region that the content is destined for.

A new breed of specialized, high performance TOD server with low-latency and live content ingest capabilities, plus a new metadata methodology, is a requirement to realize the potential of Television on Demand for cable operators.

About the Author

Robert Scheffler is Chief Architect at Broadbus Technologies, a provider of next-generation server systems that enable cable operators to effectively scale and migrate their networks from Video on Demand to Television on Demand™ – (TOD®)

Robert can be reached at:

(978) 264.7900
Robert.Scheffler@broadbus.com
http://www.broadbus.com

LEGAL ISSUES IN A TRUSTED DOMAIN[a1]

by David St. John-Larkin[a2]
edited by Jud Cary[a3]
Cable Television Laboratories, Inc.

Abstract

A Trusted Domain generally provides for the delivery, retention and utilization of copyrighted content within a secure residential network. This paper attempts to identify and to address some of the key legal issues of copyright law that are presented in a Trusted Domain in the abstract sense. An in-depth discussion of technical and business issues raised by Trusted Domains is beyond the scope of this paper.

Contrary to those commentators who criticize trusted systems as parochial or limiting,[1] the thesis of this paper is that the Trusted Domain can (1) preserve, if not increase, current copyright law privileges enjoyed by consumers, (2) assure content owners of a secure network and (3) provide distributors a new product offering. Ultimately, the Trusted Domain may serve as a model for the next generation of content-related services that preserves the expectations of consumers, and protects the rights of copyright owners alike.

I. INTRODUCTION

The digital world of the 21st Century is no different than Alice's Wonderland.[2] In both cyberspace and the fictional land at the end of a rabbit's hole, there exist communities that do not rely upon the scientific laws of nature or real-world social and legal norms. Just as an invisible cat makes sense in a world of talking playing cards, a "worm" that unobtrusively embeds itself within vulnerable computers to monitor suspicious activity makes sense.[3] In other words, the laws governing both worlds are entirely self-imposed. In Alice's world the Queen of Hearts (presumably) sets forth the law, in the digital world "*code is law*."[4]

One key attribute of "code," including copy control software and digital rights management systems (DRMs), that underlies our digital world is that it is mutable. Code is not bound to follow rigid structural or architectural guidelines; rather, code is flexible and can adapt to new or changing circumstances. A second important attribute of code is that it can facilitate fast distribution of perfectly replicated information. These attributes have led some to proclaim the vision of a technological utopia, a modern Enlightenment where individuals share information, knowledge, and culture at the press of a button or pulse of light.[5] Some commentators, however, offer that code leads to a "dystopia [where] digital technology is the handmaiden of copyright infringement" and the death of copyright law.[6] The fear expressed by these commentators is that digital technology will supplant copyright law, and that owners of digital content will use code to "undermin[e] the utilitarian balance of copyright [law] and threaten free expression."[7] While not entirely unfounded, these fears are reactionary. It is certainly true that code or digital technology could be used to usurp the general provisions of copyright law. Conversely, code could strictly enforce copyright law and restrict traditional fair use privileges that most consumers in the digital world now assume as a right.

As originally suggested by Mark Stefik, the concept of trusted systems offers a model for code to exercise complete control of digital content.[8] The protection of digital content from unlawful distribution, especially in the post-Napster age of peer-to-peer networking (e.g., KaZaA), is an important reason to implement trusted systems. However, the existence of a trusted system does not of itself eradicate the privileges bestowed by copyright law. This paper discusses a type of trusted system, called the Trusted Domain, that can preserve, if not increase, copyright law privileges enjoyed by consumers while concurrently assuring content owners of a secure network. Because trusted systems rely upon code, the Trusted Domain can flexibly incorporate and closely model copyright law, as well as appurtenant copyright privileges such as fair use. Moreover, there is ample reason why the Trusted Domain should be crafted to model real world copyright law. Simply stated, Americans love fair use–fair use privileges are marketable goods that increase the value of content to the consumer.[9]

This paper discusses the Trusted Domain as applied in the context of a home network consisting of a plurality of multimedia components (see Illustration A., below). Part II sets forth the general architecture of the Trusted Domain, and describes the possible range of specific characteristics a Trusted Domain may implement. Part III explains how the Trusted Domain affirms, rather than annihilates, certain copyright law principles, including the first sale doctrine and fair use, and may even be used to preserve or enhance certain privacy protections. This paper concludes by submitting that the libertarian Trusted Domain protects digital content and ensures the continuation of copyright privileges that are consistent with the expectations of both content owners and consumers alike.

II. ARCHITECTURE OF THE TRUSTED DOMAIN

Conceptually, trusted systems consist of a set of protocols or rules[10] that govern the use, management and protection of copyrighted material. Physically and logically the Trusted Domain is embodied in a network architecture[11] that can include various rules or functions related to the use of content in the Trusted Domain, including the backup,

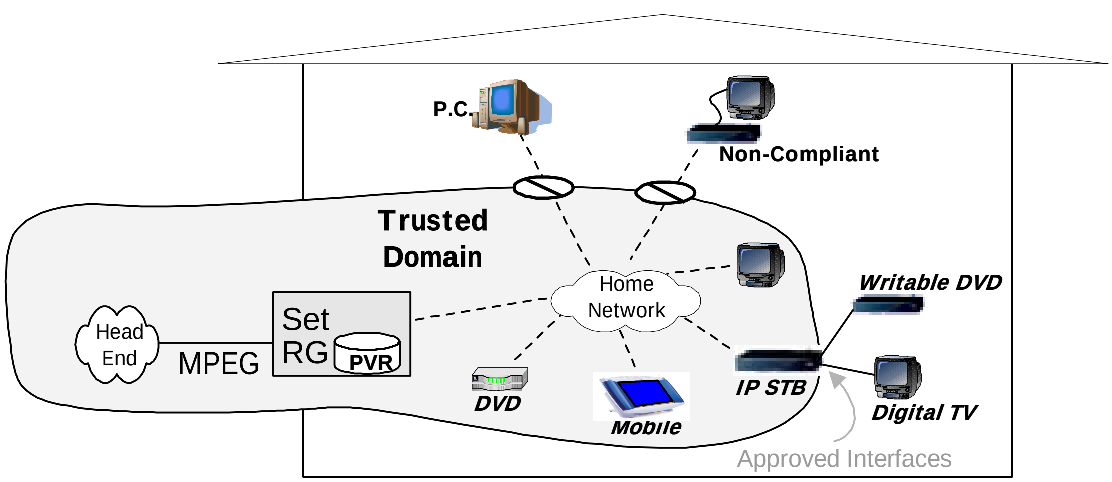


Illustration A.

conversion, distribution, playback, recording, storage and transport of copyrighted material. The most important attribute of a trusted system, implied by its name, is that it must identify Trusted and Non-Trusted Devices.[12] A Trusted Device is an application or electronic device capable of identifying itself and implementing the rules of the Trusted Domain. Likewise, a Non-Trusted Device is any application or electronic device that does not identify itself or cannot implement the rules of the Trusted Domain.

Any set of Trusted Devices or Non-Trusted Devices may be combined into a Trusted Domain.[13] The purpose of a Trusted Domain is to enforce rules applicable to individual Trusted and Non-Trusted Devices. Importantly, a Trusted Domain must establish, manage and enforce rules for each device connected within the domain. In other words, the set of Trusted Devices that comprise the Trusted Domain must establish trust within the network and maintain a secure means of managing the input and output of content within the Trusted Domain.[14]

A. Usage Rules

Content usage rules imposed upon the Trusted Domain can generally be divided into two categories: distribution (or transport) rules and content rules. Distribution rules enable the Trusted Domain to verify that content is transferred only to devices implementing the requisite security safeguards, e.g., transfer to other Trusted Devices. Content rules enable the Trusted Domain to implement requisite control over the content that is utilized by a Trusted Device. The content usage rules of a Trusted Domain are entirely self-imposed. Because they generally rely upon code, the content usage rules imposed upon Trusted Devices are customizable and may be as restrictive or unrestrictive as necessary. As discussed in the following subsections, there are two fundamental paradigms for asserting content usage rules within a Trusted Domain: the libertarian Trusted Domain and a rule-based Trusted Domain.

B. The Libertarian Paradigm

One embodiment of the Trusted Domain implements only a single, simple, content usage rule: within a Trusted Domain there are no content or distribution rules, apart from the requirement that content only be distributed to other Trusted Devices. This open or libertarian paradigm builds upon the assumption (set forth above) that a Trusted Domain is able to identify and regulate the connection of Trusted and Non-Trusted Devices to the network. Subject to initial access and authentication of Trusted Devices, in this simplified form, a Trusted Domain would eliminate the need for complicated copy control and content encoding rules. Within the secure network of a Trusted Domain, a person would then be free to use and distribute content without restriction: any content, any time, anywhere within the Trusted Domain.

For example, a person could distribute a movie purchased for the Trusted Domain to all Trusted Devices capable of video-playback that are within the Trusted Domain.[15] Instead of restricting playback of the movie to a single DVD player, a person could simultaneously transfer or play the movie on other Trusted Devices such as a personal video recorder (PVR), LCD projector or the digital television on the front of your refrigerator. Furthermore, use of a movie within the libertarian Trusted Domain would also be free of any copy control restrictions (e.g., copy-once, copy-never, view-only, view-once, etc.). Instead, the movie could be freely consumed and used within the Trusted Domain. So long as the content remains within the Trusted Domain, the content can be utilized without restriction.

For simplicity, and for the purpose of raising and discussing general legal topics, this paper focuses on this "libertarian" version of a Trusted Domain–once within the Trusted Domain, content is generally available any time, anywhere within the Trusted Domain. Of course, a wide variety of distribution and content rules, and every combination thereof, could be imposed on a Trusted Domain system. And, different technical or business considerations may influence a particular desired Trusted Domain. However for academic discussion, those issues are outside the scope of this paper. Section III of this paper therefore proceeds to address the legal significance of this libertarian model in greater detail.

C. The Rule-Based Paradigm

An alternative to the libertarian Trusted Domain is a rule-based Trusted Domain that implements one or more rules to control or to regulate the use and distribution of content within the Trusted Domain. In contrast to the libertarian paradigm, the rule-based paradigm establishes a set of rules to regulate any or all of the activity within the Trusted Domain. Various functions could be made subject to such rules, including backup, conversion, distribution, playback, recording, storage and transport of content. Various rule-based paradigms already exist in the digital domain. For example, content "Encoding Rules" are required when using the Digital Transmission Copy Protection (DTCP) system (e.g., on a 1394 digital connector). Copy protection schemes that exist in physical media can also be honored or modeled in a Trusted Domain; for example, CSS protection on a DVD, or the various copy protection methods applicable to CDs could be enforced in the rule-based Trusted Domain.

Although existing copy protection rules can be modeled in the Trusted Domain, it is arguable that content owners might be more willing to "soften" such rules within the Trusted Domain because they know that the Trusted Domain network is secure, and is limited to the Trusted Devices on the Trusted Domain. This reasoning may especially hold true in a Trusted Domain limited to the home environment where the number of Trusted Devices is relatively small, and the audience is limited. In other words, copy protection rules that apply to a particular piece of content outside the Trusted Domain may differ from the copy protection rules that are applied to the same piece of content within the Trusted Domain. The rule-based Trusted Domain paradigm offers a wide variety of options to content owners, distributors and consumers. As expected, the technical and business issues are also more complex. For simplicity, this paper focuses on the libertarian Trusted Domain noted above. However, many of the core legal issues remain the same.

III. COPYRIGHT LAW AND THE TRUSTED DOMAIN

The libertarian Trusted Domain paradigm (or even a rule-based Trusted Domain with fairly lax copy protection rules) has the potential to preserve in the digital domain two fundamental copyright law principles: the protection of copyrighted content, and the preservation of fair use privileges. Additionally, the distribution of copyrighted content within this paradigm comports with the first sale doctrine by allowing the consumer to freely distribute content to other

Trusted Devices. The Trusted Domain also may preserve, or even enhance, certain privacy expectations. The following subsections discuss the legal implications of the libertarian paradigm for the protection, use and distribution of content within the Trusted Domain.

A. The Trusted Domain as a Compliment to the Law

The protection of copyrighted content is typically accomplished via *ex ante* or *ex post* enforcement measures. Generally speaking, technical prophylactic measures protect content *ex ante*, whereas legal enforcement measures protect content *ex post*.[16] Technical prophylactic measures include the use of encryption, third-party verification, device and user identification, self-healing software and digital certificates (that may be embedded in silicon). Legal enforcement measures include the use of contract law, copyright law, and the anti-piracy (anti-circumvention) provisions of the Digital Millennium Copyright Act. The Trusted Domain, and trusted systems generally, are best classified as a technical prophylactic measures:

> Trusted systems . . . achieve what copyright law achieves. But [trusted systems] can achieve [copyright protection] *without the law doing the restricting*. [Trusted systems] present a much more fine-grained control over access to and use of protected material than law permits, and it can do so without the aid of the law.[17]

The general distinction between *ex ante* and *ex post* copyright protection, however, does not suggest that code and law are substitutes. *Ex ante* enforcement must be responsive to the immediacy of potential copyright infringement. In a world where data can be instantaneously replicated and transmitted, legal protection is much too slow. On the other hand, technical measures gain legitimacy through the law and the law is much better equipped to sanction people who try to infringe upon copyrights. The use of technical and legal measures to protect content therefore establishes a complementary or symbiotic relationship. Some commentators downplay the differences underlying this relationship and suggest that code and law are substitutes in their protective ability.[18]

The Trusted Domain, as an *ex ante* copyright protection mechanism, is a necessary and unique compliment to the legal protections afforded by the Copyright Act of 1976 (Copyright Act) and, as amended, by the Digital Millennium Copyright Act (DMCA).[19] The Trusted Domain implements a flexible, but still robust and secure, transport layer that rides on top of a network layer (e.g., a hybrid fiber-coax cable plant).[20] If history is our guide, however, it is apparent that technological safeguards will “probably not be 100 percent effective.”[21] The Trusted Domain, by implementing multiple renewable copy protection mechanisms (enumerated above, e.g., digital certificates), implements a corrective means of quickly resolving potential security holes.[22] Because the circumvention of technological copyright protection measures implicates the reproduction right,[23] Congress passed the DMCA as a complementary *ex post* legal enforcement regime.[24] Section 1201 of the DMCA prohibits the manufacture and distribution of devices (and the rendering of services) for the purpose of circumventing technological measures that protect against unauthorized access to works.[25] So, Section 1201 addresses the conduct of circumventing a technological measure that protects access.[26] Congress passed this *ex post* enforcement measure because it recognized the urgency and importance of protecting digital content: once digital content is copied, it is very easy to duplicate and distribute.[27] The effect is that Section 1201 publicly discourages the circumvention of copy protection measures through the threat of an *ex post* application of copyright law.

Another complementary *ex post* copyright enforcement measure is provided by contract law. Generally speaking, contract provisions governing aspects of copyrighted works are enforceable.[28] There is, however, disagreement among courts as to the scope of "specific contractual provisions that would otherwise be enforceable under state law."[29] An expansive interpretation[30] of Judge Easterbrook's opinion affirming "shrink-wrap" licenses in *ProCD, Inc. v. Zeidenberg* highlights this disagreement, and distinguishes state contract rights from the exclusive rights in the federal copyright regime:

> Rights "equivalent to copyright" are rights established by law–rights that restrict the options of persons who are strangers to the author. . . . A copyright is a right against the world. Contracts, by contrast, generally affect only their parties; strangers may do as they please, so contracts do not create "exclusive rights."[31]

Thus, bilateral contracts, contracts that exist between two parties, "may be enforced."[32] As it pertains to preventing the circumvention of trusted systems, contract law thus provides the Trusted Domain with another means of enforcing copyright protection measures beyond technical safeguards. Establishing contracts that define the boundaries of permissible behavior within the Trusted Domain provide yet another tool to safeguard content and reinforce *ex ante* technical content protection measures. However, as explained below, contract restrictions in the digital world may encroach upon traditional first-sale concepts and thus may diminish the value of content in the Trusted Domain without adding any more protection to the content than is already incurred by the use of other *ex ante* and *ex post* copy protection measures.

<u>B. The Trusted Domain and Benefits of the Fair Use Doctrine</u>

The libertarian Trusted Domain may preserve, if not expand, the fair use privileges enjoyed by consumers in the analog world and respond to the difficulty of post-sale fair use valuation problems that were historically left unaccounted for by market pricing mechanisms. The Copyright Act grants copyright owners six exclusive rights, generally enumerated as: adaptation (derivative works), distribution, display, performance, reproduction and convergence (digital performance and transmission rights).[33] Fair use is a defense that can be asserted where there is infringement of one of these six exclusive rights.[34] The doctrine of fair use is highly contentious and was at one time labeled "the most troublesome doctrine in the whole law of copyright."[35] At the heart of the fair use doctrine is an ongoing debate about whether the doctrine is itself dependent and restricted by technology and subject to economic constraints imposed by the market forces. This debate is stereotypically between copyright owners, who regard the fair use doctrine as an artifact of the analog or print world that should slowly recede with time, and consumers, who view fair use as an immutable right that is necessary for promulgating one of the Copyright Act's purposes to convey copyrighted content back into the public domain.[36] Copyright owners, in this generalized sense, assert that fair use only applies where the "transactions costs associated with clearing rights sometimes exceeded the value of the proposed use."[37] Consumers, alternatively, would claim that fair use is core to the principle establishing copyright laws in the first place–i.e., to benefit the public–and is "not merely a matter of economics" nor of technology.[38]

The rule-based Trusted Domain paradigm, as a means of regulating or controlling the specific use and distribution of content, may perpetuate the same quandary presented in this fair use debate. Rule-based usage rules permit the copyright owner to price the use of content on a *pro rata* basis.[39] Accordingly, the hypothetical copyright owners would say that the increased technological capability to control use piecemeal does not run contrary to the fair use doctrine:

> Fair use, [the copyright owners] argue, defined rights in an area where it was not possible to meter or charge for use. In that context, fair use set a default rule that parties could always contract around. The default rule was that use was free.
>
> But as the limits of what it is possible to meter and charge for changes, the scope of fair use changes as well. If it becomes possible to license every aspect of use, then no aspect of use would have the protections of fair use. Fair use, under this conception, was just the space where it was too expensive to meter use.[40]

Alternatively, the hypothetical consumers would state that the fair use doctrine is "inherent in the copyright – required whether technology makes it possible to take it away or not."[41] As presented below, the libertarian Trusted Domain paradigm not only recognizes these divergent positions, but presents a model much better suited to reconcile them.

The libertarian Trusted Domain paradigm fundamentally allows unrestricted use of content within the Trusted Domain. A book could be paraphrased within an electronic document, a movie clip embedded within a home-movie, or a song transformed instantaneously to play on multiple devices simultaneously. Assuming the actual use otherwise satisfies the other parameters of fair use,[42] the possibilities are endless. Another premise of the Trusted Domain, that all use within the Trusted Domain will not be metered or charged, assures the consumer that their fair uses continue unencumbered by *pro rata* licensing fees and protects that individual's personal content. This model, however, is also structured to protect content by assuring copyright owners that it remains solely within the network of Trusted Devices. Moreover, the libertarian Trusted Domain paradigm recognizes the legal importance and the monetary value of fair use by allowing copyright owners to set initial distribution prices at the convenient point-of-sale entry to the Trusted Domain and thereby capture the marginal costs of fair uses that were previously considered a market failure (and thus allowed free of charge).[43] Certainly, Americans love fair use. Instead of punishing or restricting fair use, the libertarian model markets fair use and creates new business models.[44] Ultimately, then, the libertarian Trusted Domain paradigm allows consumers to continue to enjoy their fair use privileges while providing content owners a convenient mechanism to set a price for fair use in a secure environment.[45]

C. The Trusted Domain and Preservation of the First Sale Doctrine

The libertarian Trusted Domain paradigm provides copyright protection measures that do not preclude application of the first sale doctrine. The first sale doctrine relates to the distribution right and is a limitation that prohibits a copyright owner from exercising control over the distribution of a tangible copyrighted work past the first-sale. In other words, a copyright owner may attach conditions on the first-sale of a copyrighted work (e.g., payment of a specific price) but may not thereafter condition resale or further distribution of the tangible copyrighted work upon any criteria. The first sale doctrine is a default rule "origin[ating] in the common law aversion to limiting the alienation of personal property" and policies opposing restraints of trade.[46] Codified in Section 109 of the Copyright Act, the first sale doctrine heralds back to *Bobbs-Merrill Co. v. Strauss*[47] in which the U.S. Supreme Court "construed the exclusive right to [distribute] . . . as applicable only to the initial sale, so that absent an appropriate contractual provision, there could be no restriction on re-sales."[48]

In the days of the *Bobbs-Merrill Co. v. Strauss* case (1908), the first sale doctrine was also practical to implement with respect to the media of the day – books, newspapers, etc.[49] Historically speaking, it was difficult or impossible to monitor further sale or distribution of such copyrighted works, or to collect compensation for such. However, with the advent of code, further distribution of copyrighted works can be easily tracked, monitored and regulated subject to technological controls. And, in some cases, such information can prevent further distribution or use of a copyrighted work, e.g., digital content marked "view-only" would also prevent any further distribution.[50] The distribution of digital content, however, can be fundamentally different than the distribution of books or other analog media. Where distribution of the digital content itself necessarily requires creation of a copy prior to distribution, "[S]ection 109 does not apply to [the] digital transmission of works."[51]

The libertarian Trusted Domain paradigm somewhat restores the historical and distribution-specific conception of the first sale doctrine, at least in spirit. Whereas a rule-based Trusted Domain may attach conditions that restrict the distribution of content to certain Trusted Devices, the libertarian model allows distribution to all devices within the Trusted Domain. Notably, to truly comply with the first sale doctrine, "distribution" in this sense would technically need to be a "move." That is, in the operation of transferring content, the storage place of the original content would need to be deleted or rendered unusable. [52]

Enabling the first sale doctrine through the libertarian Trusted Domain allows consumers to make use of digital content no different than how analog content, or books (with respect to further distribution, not copying), are utilized in the real world. Moreover, with the addition of a few simple content rules, consumers could distribute digital content to other Trusted Domains implementing the same management paradigm. Thus, the entrance to the libertarian Trusted Domain acts as the point-of-sale to provide the bargained-for uses that the first sale doctrine originally enabled under earlier technological constraints.

C. The Trusted Domain and Privacy

The advent of trusted systems prompted many commentators to reexamine the role of privacy norms in the digital world.[53] One early commentator suggested that "the freedom to read, listen, and view selected materials anonymously should be considered a right protected by the First Amendment . . . ."[54] The commentator also argues that the civil and criminal enforcement provisions of the pre-DMCA legislation may prove susceptible to constitutional challenge.[55] Trusted systems were seen as a form of "private legislation" that could potentially disrupt the balance between preservation of a copyright owner's exclusive rights and enrichment of the public domain.[56] Trusted systems, it was argued, could potentially marginalize, if not entirely eviscerate, copyright law.[57]

Contrary to these and other dire predictions forecasting the end of copyright law, more recent commentators noted the practical benefits that may arise by allowing trusted systems to manage consumer information. For example, automated information that covers the "provenance . . . and conditions of sale or license" may "substantially reduce . . . transaction costs."[58] Consumers and copyright owners also may benefit by a system that assures the authenticity and integrity of digital content delivered to the home.[59] Finally, it is now technically recognized that consumer-specific information can be anonymized. Anonymizing or aggregating an individual's preferences with the preferences of other people allows copyright owners and distributors to lower transaction costs, ensure the authenticity and integrity digital transmissions, while also directing sufficiently targeted information to consumers (e.g., targeted advertising).

The libertarian Trusted Domain may preserve, or even enhance, certain expected privacy norms. It is recognized that some *de minimus* form of metering must be established at the point of entry into the Trusted Domain in order to enable proper billing and payment.[60] However, once inside the libertarian Trusted Domain, no further metering is required; content may be used anytime and anywhere within the Trusted Domain. This is not to say, however, that consumers may not want more monitoring within the Trusted Domain. It is foreseeable, that given the option, many consumers may wish to monitor and store information to help backup and restore digital works or facilitate interactive services.

As such, we submit that the libertarian Trusted Domain may actually preserve certain expectations of privacy, now known in the analog world, in the digital domain.

IV. CONCLUSION

As this paper sets forth, the libertarian Trusted Domain paradigm protects digital content and recognizes the value of preserving copyright privileges that are consistent with the expectations of both copyright owners and consumers alike. The apparent benefits accruing from implementation of the libertarian Trusted Domain paradigm are numerous.

Consumers receive a convenient and standardized media platform that minimizes confusion about how to use content. This platform securely and transparently protects content within the Trusted Domain and preserves, if not expands, content usage expectations.

Content providers may also benefit from considerably more protection and security for the distribution of high-value digital content. The unrestricted nature of the libertarian Trusted Domain in particular increases the value of content, and allows content providers and distributors to create flexible new business models to capture this value.

Likewise, consumer electronics manufacturers may benefit by a network that offers new market opportunities for devices and standardized interfaces for compatibility.

Finally, the Trusted Domain offers distributors a unique competitive network architecture for packaging and delivering content into the residential home.

In summary, the libertarian Trusted Domain can be used to affirm copyright law principles, including fair use privileges, establish a digital media platform that creates value to consumers, content owners, device manufacturers, and distributors.

[a1] The opinions expressed in this paper are that of the author and editor individually, and do not represent the opinions of the companies or industry in which they are employed.

[a2] Third-Year Law Student and Production Editor, Journal on Telecommunications and High Technology Law, University of Colorado School of Law. B.S., Environmental Studies & Social Policy, Northwestern University, 1997; B.S., Computer Science, *magna cum laude*, Michigan Technological University, 2002.

[a3] Assistant General Counsel, CableLabs. B.S., Applied Math & Computer Science, University of Colorado, 1986; M.E. Engineering Management, University of Colorado, 1993; J.D. *cum laude*, Santa Clara University, 1993.

[1] LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 136-139 (1999); Yocchai Benkler, *Free as the Air to Common Use: First Amendment Constraints on Enclosure of the Public Domain*, 74 N.Y.U. L. REV. 354, 418 (1999); Julie E. Cohen, *A Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace*, 28 CONN. L. REV. 981, 981-82 (1996); Mark Gimbel, *Some Thoughts on the Implications of Trusted Systems for Intellectual Property Law*, 50 STAN. L. REV. 1671, 1671-72 (1998).

[2] ROBERT LOUIS STEVENSON, ALICE'S ADVENTURES IN WONDERLAND (1865). In some cases, there is no distinction between the two worlds. *See e.g.*, *Alice in Wonderland – An Interactive Adventure!*, *at* http://www.ruthannzaroff.com/wonderland (last visited Mar. 9, 2003).

[3] LESSIG, *supra* note 1, at 17 (Lawrence Lessig offers the surveillance "worm" story as a means of showing how cyberspace is both like and unlike real space; the forced entry of a non-invasive worm may be so unlike the real world such as not to raise Constitutional eyebrows.).

[4] *Id*. at 7 (emphasis in the original).

[5] Mark Stefik, *Letting Loose the Light*, *in* INTERNET DREAMS: ARCHETYPES, MYTHS AND METAPHORS 220-21 (Mark Stefik ed., 1996).

[6] Gimbel, *supra* note 1, at 1671; John Perry Barlow, *Selling Wine Without Bottles: The Economy of Mind on the Global Net* (Dec. 1993) ("Copyright is dead."), *at* http://www.eff.org/pub/Publications/John_Perry_Barlow/HTML/idea_economy_article.html.

[7] Gimbel, *supra* note 1, at 1671-72.

[8] *See* Stefik, *supra* note 5, at 220. The concept of trusted systems is discussed in Part II of this paper.

[9] To realize the value of maintaining fair use privileges, we need look no further than the rising popularity of personal video recorders (PVRs) such as those offered by Tivo or SonicBlue (ReplayTV) that allow television shows to be digitally recorded for later-viewing ("time-shifting").

[10] *See* Mark Stefik, *Shifting the Possible: How Trusted Systems and Digital Property Rights Challenge Us To Rethink Digital Publishing*, 12 BERKELEY TECH. L.J. 137, 139 (1997).

[11] The network architecture of a trusted system is comprised individually or as a union of hardware and software.

[12] *See* Mark Stefik, *Trusted Systems*, SCIENTIFIC AMERICAN 78, 79 (Mar. 1997) ("[T]rusted computers have the capability to recognize another trusted system, to execute usage rights and to render works so that they either cannot he [sic] copied exactly or else carry with them a signature of their origin."). Discussion of the various technological tools used to verify trust are beyond the scope of this Article. Notably, the technological measures that verify trust must be renewable, revocable and robust. These characteristics generally require the implementation of self-healing (automatically upgradeable) software and silicon or embedded encryption safeguards.

[13] In computer science terminology, a Trusted Domain would most likely consist of a closed parallel network of devices communicating via an encrypted challenge-response system.

[14] *See* Stefik, *supra* note 5, at 229.

[15] As noted earlier, business models are outside the scope of this paper; but, query what the appropriate price for this type of a purchase for use within a Trusted Domain would be relative to other "traditional" purchases of the same movie, e.g., DVD movie rental, VOD.

[16] Citing a former research assistant's paper that attempts to discern the most efficient protections in cyberspace, Lawrence Lessig suggests that the real-world analog to the *ex ante* and *ex post* distinction is the difference between a fence and the law. LESSIG, *supra* note 1, at 122 (citing Harold Smith Reeves, *Property in Cyberspace*, 63 U. CHI. L. REV. 761 (1996)).

[17] LESSIG, *supra* note 1, at 129.

[18] For example, while Lawrence Lessig recognizes the possibility of this symbiotic relationship, he advocates that code currently displaces law:

> What copyright seeks to do using the threat of law and the push of norms, trusted systems do through the code. Copyright orders others to respect the rights of the copyright holder before using his property. Trusted systems give access only if rights are respected in the first place. The controls needed to regulate this access are built into the systems, and no users (except hackers) have a choice about whether to obey these controls. The code displaces law by codifying the rules, making them more efficient than they were just as rules. Trusted systems in this scheme are an alternative for protecting intellectual property rights – a privatized alternative to law. They need not be exclusive; there is no reason not to use both law and trusted systems. Nevertheless, the code is in effect doing what the law used to do. It implements the law's protection, through code, far more effectively than the law did.

Id. at 130.

[19] Copyright Act of 1976, 17 U.S.C. § 106 (1976); Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1998).

[20] *See* Kevin Werbach, *A Layered Model for Internet Policy*, 1 J. Telecomms. & High Tech. L. 37 (2002) (offering a layered-model approach to understanding vertically-related communications platforms); *see also*, Douglas C. Sicker & Joshua L. Mindel, *Refinements of a Layered Model for Telecommunications Policy*, 1 J. Telecomms. & High Tech. L. 69 (2002) (suggesting that a layered model must incorporate the

technological characteristics of individual telecommunications platforms).

[21] U.S. COPYRIGHT OFFICE, *A Report of the Register of Copyrights Pursuant to §104 of the Digital Millennium Copyright Act*, 98 (Aug. 2001), *available at* http://www.loc.gov/copyright/reports/studies/dmca/dmca_study.html.

[22] The cost of developing sufficient copy-control mechanisms is not to be underestimated. Certainly the recovery of high fixed development costs would be difficult to recover, in the short term, if charged only to a single company's consumers. The economic rationale for development of robust security measures is better realized when internalized by a standard-setting organization.

[23] The Copyright Act provides that the copyright holder, or his or her agent, has the exclusive right to "reproduce the copyrighted work in copies or phonorecords. . . ." 17 U.S.C. § 106 (a)(1).

[24] U.S. COPYRIGHT OFFICE, *supra* note 21 at 98. Notably, Congress did not "prohibit the conduct of circumventing . . . [all] copy control measures," as this conduct may be permitted when claimed as a fair use defense, such as for library archival work. However, "fair use and other copyright exceptions are not defenses to gaining unauthorized access to a copyrighted work." *Id.* at 11-12. The analogy is clear: quoting a lawfully acquired book may be a fair use, quoting a book stolen out of a safe is not.

[25] *Id.* at 10.

[26] *Id.*

[27] STAFF OF HOUSE COMMITTEE ON THE JUDICIARY, 105TH CONG., SECTION-BY-SECTION ANALYSIS OF H.R. 2281 AS PASSED BY THE UNITED STATES HOUSE OF REPRESENTATIVES ON AUGUST 4, 1998, 2 (Comm. Print 1998) (Serial No. 6). The Senate Judiciary Committee explained that:

> Due to the ease with which digital works can be copied and distributed worldwide virtually instantaneously, copyright owners will hesitate to make their works readily available on the Internet without reasonable assurance that they will be protected against massive piracy. Legislation implementing the treaties provides this protection and creates the legal platform for launching the global digital on-line marketplace for copyrighted works.

S. REP. NO. 105-190, at 8 (1998).

[28] Selby v. New Line Cinema Corp., 96 F. Supp. 2d 1053, 1059 (C.D. Cal. 2000) (courts have found, generally, that breach of contract claims are not preempted by § 301 of the Copyright Act).

[29] U.S. COPYRIGHT OFFICE, *supra* note 21 at 162.

[30] An alternative interpretation of Judge Easterbrook's *ProCD* opinion is that it limits "shrink-wrap" or "click-through" license terms past the first sale, only where there is full disclosure between the contracting parties.

[31] 86 F.3d 1447, 1454 (7th Cir. 1996) (citation omitted).

[32] *Id.*

[33] 17 U.S.C. § 106 (1996).

[34] The fair use doctrine was first articulated in *Folsom v. Marsh*, where Justice Story enumerated factors to decide issues of fair use:

> [W]e must . . . look to the nature and objects of the selections made, the quantity and value of the materials used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects, of the original work. . . ."

9 F. Cas. 342 (C.C.D. Mass. 1841) (No. 4901). A full examination of the history and development of the fair use doctrine is

complex and beyond the scope of this paper. It is sufficient, herein, to note that Congress codified this judicial doctrine in § 107 of the 1976 Copyright Act, which sets forth four factors that are determinative of fair use:

> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
> (2) the nature of the copyrighted work;
> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
> (4) the effect of the use upon the potential market for or value of the copyrighted work.

17 U.S.C. § 107.

[35] Dellar v. Samuel Goldwyn, Inc., 104 F.2d 661, 662 (2d Cir. 1939) (per curiam).

[36] U.S. CONST. art I, § 8, cl. 8 ("[T]o promote the Progress of Science and useful Arts, by securing for *limited* Times, to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.") (emphasis added). Another conception of the fair use doctrine was as a "proxy for a copyright owner's implied consent." Tom W. Bell, *Fair Use v. Fared Use: The Impact of Automated Rights Management on Copyright's Fair Use Doctrine*, 76 N.C. L. REV. 557, 581 (1998) (citing MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.05 13-151 (disapproving of the notion) [hereinafter NIMMER]). Tom Bell notes that the proxy conception of fair use fell into disfavor because it did "not explain why fair use protects parody and other uses of copyrighted material that owners find disagreeable." *Id*. at 582.

[37] MARSHALL LEAFFER, UNDERSTANDING COPYRIGHT LAW § 10.17[A] (1999).

[38] *Id.*

[39] A proponent of rule-based usage controls, Jane Ginsburg states:

> As we move to an access-based world of distribution of copyrighted works, a copyright system that neglected access controls would make copyright illusory, and in the long run it would disserve consumers. Access controls make it possible for authors to offer end-users a variety of distinctly-priced options for enjoyment of copyrighted works. Were delivery of works not secured, novel forms of distribution would be discouraged, and end-users would continue to be charged for all uses, whatever the level in fact of their consumption.

Jane C. Ginsburg, *From Having Copies to Experiencing Works: the Development of an Access Right in U.S. Copyright Law*, Columbia Law School, (Public Law Working Paper No. 8, 2000).

[40] LESSIG, *supra* note 1, at 136 (citations omitted).

[41] *Id.*

[42] *See infra* note 34.

[43] Bell, *supra* note 36, at 581-84. Tom Bell notes that technology is more effective than fair use in responding to market failure. *Id.* Where transaction costs preclude value-maximizing uses of copyrighted content, "automated rights management radically reduces the transaction costs of licensing access to copyrighted works . . . it responds to market failure." *Id*. at 583.

[44] Responding to Lawrence Lessig's contention that a "small [ ] number of large companies" necessarily force consumers to choose restrictive usage architectures, David G. Post poses the rhetorical question:

> [I]f there are diverse architectures of privacy, of identity, and of content

> protection laid before the public, why is it so obvious that we will end up choosing the one(s) that deny us those things that Lessig (and I) think are so important?

David G. Post, *What Larry Doesn't Get: Code, Law, and Liberty in Cyberspace*, 52 Stan. L. Rev. 1439, 1454 (2000) (citing LESSIG, *supra* note 1, at 130.).

[45] A "property rights [regime, allows] the producer of intellectual property [to] charge more than marginal cost, and thus cover the total cost of producing and disseminating the works." J. Frank H. Easterbrook, *Technological Innovation & Legal Tradition: Enduring Principles for Changing Times*, 4 Tex. Rev. L. & Pol. 103, 105 (1999); *see also*, Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 Harv. L. Rev. 1089 (1972) (presenting a framework for determining whether property, liability or inalienability rules should protect an entitlement).

[46] NIMMER, *supra* note 36 at § 8.12[A], fn 5 (*citing* Sebastian Int'l, Inc. v. Consumer Contacts (PTY) Ltd., 847 F.2d 1093, 1096 (3d Cir. 1988)).

[47] 210 U.S. 339 (1908).

[48] NIMMER, *supra* note 36, at § 8.12[B][1].

[49] Some commentators have called for the expansion of Section 109 to include digital works. U.S. COPYRIGHT OFFICE, *supra* note 21, at 40. The Office recognizes that, "*[p]hysical* copies of works in a digital format . . . are subject to [S]ection 109 in the same way as *physical* copies of works in analog form." *Id*. at 78 (emphasis added). However, this is *not* to say that digital distribution of the work itself—which generally includes copying, not just distribution of the physical medium—is intended to be covered under the ambit of Section 109.

[50] However, "[t]he first sale doctrine does not guarantee the existence of a secondary market [(called an aftermarket)] or a certain price for copies of copyrighted works." U.S. COPYRIGHT OFFICE, *supra* note 21 at 74 (responding to arguments that CSS technology limits the resale or aftermarkets for used DVDs). Furthermore, "[t]o the extent that there is a concern that region coding may limit the number of purchasers outside North America who are willing to buy [region-encoded] DVDs . . . that concern has nothing to do with § 1201 [of the DMCA]." *Id.* at 74, n. 263.

[51] *Id*. at 80. The U.S. Copyright Office explains why the reproduction right takes precedence over the distribution right, saying:

> The Supreme Court [in the *Bobbs-Merrill* decision] drew a sharp distinction between the two rights, creating an exception to the vending (i.e., distribution) right *only to the extent that it didn't interfere with the production right*.

Id. (citation omitted) (emphasis added). In fact, the *Bobbs-Merrill* decision explains this purpose, noting that "a grant of control to the copyright owner over resales would not further [the] main purpose of protecting the reproduction right." *Id.* at 20-21 (citing 210 U.S. 339 (1908) ("[T]he main purpose [of the copyright statutes is] to secure the right of multiplying copies of the work . . . .").

[52] It is noted that the DTCP encoding rules allow for such "move" operations.

[53] *See e.g.*, Cohen, *supra* note 1, at 982 ("[T]he new copyright management technologies force us to examine anew the sources and extent of that freedom."); Niva Elkin-Koren, *Copyrights in Cyberspace–Rights Without Laws?*, 73 CHI-KENT L. REV. 1155 (1998).

[54] Julie E. Cohen, *Some Reflections On Copyright Management Systems And Laws Designed To Protect Them*, 12 BERKELEY

TECH. L.J. 161, 185 (1997) (citing Cohen, *supra* note 1, at 1003-30.).

[55] *Id*.

[56] *Id.* at 163 ("[Copyright management systems] may enable both pervasive monitoring of individual reading activity and comprehensive 'private legislation' designed to augment–and possibly alter beyond recognition–the default rules that define and delimit copyright owners' rights."); Elkin-Koren, *supra* note 53, at 1186 (The "technological ability to restrict access to information . . . provide information suppliers with an inherent advantage over users . . . .").

[57] Echoing John Perry Barlow's claim that "copyright is dead," *see infra* note 6, Julie Cohen suggested that:

> Copyright owners . . . envision using [self-enforcing digital] contracts to secure–and redefine–their 'informational rights.' Within this vision of private ordering and technological self-help, contract law rather than copyright law is paramount. Limits on information ownership set by the public law of copyright are conceived as optional restrictions that can be avoided using appropriate contractual language.

Julie E. Cohen, *Copyright and the Jurisprudence of Self-Help*, 13 BERKELEY TECH. L.J. 1089, 1090 (1998).

[58] Jane C. Ginsburg, *Essay–How Copyright Got a Bad Name for Itself*, 26 COLUM.-VLA J.L. & ARTS 61, 70 (2002) (suggesting that § 1202 of the DMCA reduces transaction costs and increases transaction reliability).

[59] *Id*.

[60] Trusted Domains generally use some form of authentication to discern Trusted and Non-Trusted Devices. And, it is recognized that such authentication methods need to be sensitive to First Amendment privacy concerns. General technical techniques exist to ensure a base level of privacy, including the method of anonymizing or aggregating data to ensure anonymous use. w

MANAGING PEER-TO-PEER TRAFFIC – BEYOND DOCSIS 1.1

John R. Pickens, PhD
Greg Hutterer
Allot Communications, Inc.

Abstract

Managing traffic rates and assuring that networking resources are fairly offered and consumed is perhaps the penultimate requirement and opportunity for HFC cable networks. Failures translate rapidly to customer dissatisfaction and lost revenue. Current mechanisms such as DOCSIS™ 1.1 and PacketCable™ DQOS provide strong traffic management functionality for select specified applications (e.g. VOIP). However the rapid demand for new applications and services (e.g. peer-to-peer), coupled with the long cycle-time of specification, implementation, testing, and deployment is rapidly bringing networks to their knees. New mechanisms such as in-line flow classification and application signature detection enable operators to quickly understand and adapt to new application paradigms (e.g. peer-to-peer), and fulfill rapidly changing subscriber demand. New tools and interfaces are needed to accelerate service revenue and enhance customer satisfaction.

BACKGROUND

Few network designers anticipated the exponentially growing traffic levels and ever-increasing, almost viral, portfolios of applications, architectures and protocols seen in broadband networks today. In particular, the original application traffic assumptions driving design of HFC cable networks have been vastly exceeded. The end-user cable environment is rife with subscriber-popular, bandwidth-hungry applications, each one vying for its share of the available HFC first-mile bandwidth.

Peer-to-Peer (P2P) file sharing in particular has emerged as a highly popular IP technology, primarily among home users. The rapid rise of protocols such as KaZaA, Morpheus, and Gnutella allow virtually every computer to become a server, freely sharing enormous audio and video files at will among users of an uncontrolled global community.

Community	Users
FastTrack	4,464,221
iMesh	1,421,256
eDonkey	582,030
Overnet	315,592
DirectConnect	151,898
Blubster	93,883
Gnutella	83,439

Figure 1 - P2P Communities
(**www.slyck.com**)

The largest P2P file sharing communities have millions of users – The FastTrack community has nearly 5 million users (FastTrack uses KaZaA Lite, Grokster, KaZaA, and iMesh clients). No ISP or MSO can fail to notice the impact of Peer-to-Peer file sharing.

P2P file sharing applications can consume the majority of the total bandwidth, even with only a few subscribers active, making this issue a major concern for cable providers.

The following issues are increasingly being noted by operators:

- P2P generates high traffic loads - delaying more urgent traffic of other subscribers and negatively affecting customer satisfaction.

- P2P traffic consumes even higher upstream bandwidth resources. As knowledge of each new subscriber node is propogated through the network, more and more peer nodes request uploads, thus exponentially increasing upstream traffic.

- P2P increases operator costs by forcing WAN capacity upgrades and HFC capacity upgrades through reducing CMTS subscriber density and splitting fiber nodes.

An opportunity exists for MSOs to manage the bandwidth crisis imposed by new application paradigms like P2P file sharing and also reap new service revenue by offering strong and flexible Quality of Service functions to subscribers. Not only the traditional applications (web, email) and the operator applications (VOIP) can be serviced, but new application paradigms such as peer to peer (from KaZaA to Grid computing), broadcast streaming (from MPEG video to internet radio), and two-way voice and video conferencing (from PacketCable voice to Voice/Video Instant Messaging) can be serviced as well. Also the operator can detect new applications as they are activated by subscribers, and can implement policies with respect to traffic prioritization or even traffic-blocking.

APPLICATION PARADIGMS

Original Broadband Application Paradigm

When broadband networks were first conceived, designed, and deployed, the assumed applications model was occasional web browsing and electronic mail, with infrequent activation of bandwidth hungry applications like file transfer. This model was primarily downstream assymetrical with intermittently active users.

This assumed applications model has evolved to include new planned subscription services such as symmetrical-bandwidth cable telephony. Both deployed networks and DOCSIS [1,2] and PacketCable [3,4] protocol standards have been carefully architected to support these sorts of carefully designed applications.

This initial set of applications and services paradigms is server centric – web servers, email servers, and VOIP soft-switch servers. The server model is strong in its ability to manage the scaling of services, centralize provisioning, and support trusted authentication and authorization schemes.

But, as users, and especially developers, become aware of the high-bandwidth, low-latency, and always-on attributes of broadband, new application paradigms rapidly emerge. Subscribers begin adopting the applications. Because of the bandwidth-intensive traffic profiles of these new applications, the existing best-effort QoS mechanisms fall apart.

New Broadband Application Paradigms

"The media is the message" – Marshall McLuen, 1967 ***[5]****.*

Client-server is not the only (or even best) model for distributed processing. As the speed of the communications link rises, latency drops, and peer-to-peer reachability becomes pervasive, other distributed processing models become possible.

Peer to peer (P2P) has achieved recent notoriety, especially for its seemingly unconstrained penchant for bandwidth consumption. P2P is notable in that there is often no central point of control (although there are some hybrid P2P with super server architectures for scalability). The list of P2P applications is long, and new applications are coming online every day.

Is the appearance of P2P a surprise? It should not be. P2P is a member of a class of new distributed applications model types, each of which are enabled in the new world of always-on, high-bandwidth networking.

There exist at least six distributed computing models, each of which is enabled by Broadband **[6]**:

- Ad hoc distributed computing model – no specific architecture constraints. The applications developer has a networking API at his disposal and is free to generate any partitioning of functions using any protocol.

- Remote Procedure Call model – the application uses a procedure call interface. The procedure function name and all calling parameters are shipped to the remote node, and the application waits until a procedure return is invoked with the return result.

- Remote Evaluation model – fragments of applications are moved to the remote system on which the data is contained. The application works on local data without incurring any cross-network latency. The remote environment may require call-outs or data-sharing facilities to access data from the invoking environment.

- Remote compute cluster model – all file, database, and computational resources are collocated on a remote high speed low latency network. What flows between the client and the compute cluster is keyboard/mouse input and screen output (bitmapped buffers or 2D/3D graphics operations).

- Memory mapped model – the network is viewed as a logical extension of paged memory, and paged-memory working set algorithms are used to maximize locality of data.

- Distributed object model – applications are structured as communicating objects. Objects are migrated by the network operating system to maximize throughput and minimize latency.

Some of the paradigms can be very traffic intensive. The compute cluster model can generate an average of almost a megabit per second of downstream bandwidth for display updates; the memory mapped model can utilize the entire available bandwidth of the Broadband-enabled virtual bus.

A recent set of initiatives called Grid computing **[7]** encompasses several of the models outlined above. The notable shift is that the network itself is becoming the interconnection bus of a massively parallel distributed virtual computer. As the bus speed increases, and as the latency drops, applications are being developed that utilize the bus bandwidth and connectivity matrix.

P2P is not an aberration – it is an innate reflection of the speed and latency of always-on broadband networking, and is the first of many bandwidth-consumptive distributed applications that will be seen by MSOs.

What is needed are tools and mechanisms to classify and enforce traffic in a fair manner, both to the MSO provider and to the revenue generating subscriber, and which can quickly accommodate new distributed applications and application paradigms.

PACKET AND FLOW CLASSIFICATION

Packet Classification

The foundation of traffic enforcement is a function called Packet Classification. Packet Classification inspects incoming packets and hands off the packets to the QoS enforcement function (priority queueing, congestion management) of the packet processing engine[1].

Typically packet classification functions inspect source and destination addresses, port numbers, and priority fields. Some of the features that operators expect from packet classification include:

- Fair and policy-driven oversubscription management
- Monitoring and accounting
- Class of service management
- QoS on an application specific basis
- P2P awareness
- QoS on a subscriber or tiered service level basis.
- Usage based billing
- Denial of service protection

In order for packet classification services to be implemented, however, a set of filtering parameters – often called flows - must be learned and populated into tables used by the in-line packet classification function.

Methods for Learning Flows

There are two existing methods for learning flows – QoS-Smart Application Signaling and In-Line Flow Classification. HFC Cable standards currently are defining interfaces for the first method. This paper identifies requirements and interfaces for the second method. Both are ultimately required and both are compatible with DOCSIS in-line packet classification methods.

QoS-Smart Application Signaling

Judging by the existing technical work within HFC cable (DOCSIS 1.1 **[1]**, DOCSIS 2.0 **[2],** PacketCable **[3]**, and PacketCable DQOS **[4]**), there seems to be consensus that flow classification (learning flows) is a required element of the total QoS solution. The current solution focus is based on the paradigm of QoS-smart applications (smart with respect to QoS control). In this paradigm the application learns flows via some unspecified internal protocol mechanism. The application client and/or its trusted server then explicitly signals QoS control and authorization elements through standardized signaling interfaces to in-line networking nodes (CMTS edge router) or cascaded policy servers[2].

[1] "Packet classification" is used synonomously with "QoS enforcement" for the remainder of this paper.

[2] The PacketCable Multimedia project is currently underway and is defining an expanded QoS control model for multiple QoS-smart applications. The specifications are not yet public.

Figure 2 shows how the QoS-smart model is used within PacketCable[3] VOIP telephony. MGCP **[8]** and SDP **[9]** protocols communicate QoS control information at the application layer. The VOIP application server (Call Agent) uses PacketCable DQOS signaling to communicate flow classification parameters to the CMTS. The CMTS utilizes DOCSIS 1.1 to communicate flow classification parameters to the CM.

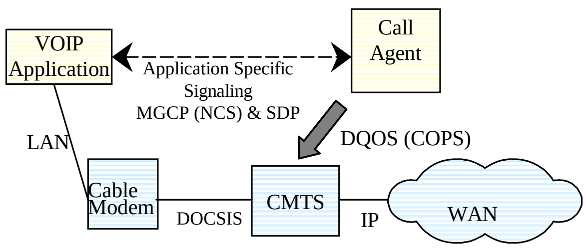


Figure 2 - PacketCable DQOS

The QoS-smart model works well for those applications that are developed in conformance to the model. Specific targeted services such as PacketCable VOIP telephony are beneficiaries of the QoS-smart style of learning flows.

However, for other new or legacy applications the QoS-smart model incurs long lead times between the specification, implementation, testing and deployment phases. Other delay inducing factors include:

(1) Today's applications development environment consists of many independent vested interests and non-collaborative standardization authorities. A major personal computer operating system vendor, for example, removed RSVP signaling and its associated APIs needed for application development from its latest generation windowing platform in 2002. Thus new applications on this platform have no formal QoS-smart signaling interface even available.

(2) In the multi-vendor application developer environment a variety of implementation platforms need to be harmonized – versions of Windows for clients, embedded systems with various RTOSs, Linux and Windows platforms for servers.

(3) Even given the presence of QoS signaling interfaces, the application developers must choose to utilize such interfaces.

From the subscriber's perspective new applications are appearing at an accelerating rate, and they are easy to access and install. The subscriber wishes to utilize QoS services for favorite applications (a revenue opportunity for operators), but since QoS-intelligence is lacking within the applications the subscriber (and any nearby neighbor sharing a common DOCSIS MAC domain) is destined for an unsatisfying experience.

In-Line Flow Classification

The second approach for flow classification (learning flows) is in-line classification. In this model all control-plane traffic is dynamically inspected and flows are dynamically learned.

The in-line classification engine is primed with external definitions of control-plane signaling mechanisms, applications, users, networks, and application signatures. The engine is also primed with policy definitions reflecting both operator and subscriber QoS policy attributes which govern applications traffic management in the packet classification phase.

[3] PacketCable Dynamic QoS (DQOS) is simplified for purpose of understanding. Other elements such as Record Keeping Server exist in the full architecture.

Figure 3 shows the in-line flow classification model. All packets received from any port are inspected by the Flow Classification function (FC) which tracks and maintains current state for the control flows of active application instances. The Flow Classification function creates parameters in a flow description table. Typically the parameters consist of IP addresses, UDP port numbers, protocol Ids and various traffic handling policies for the flow.

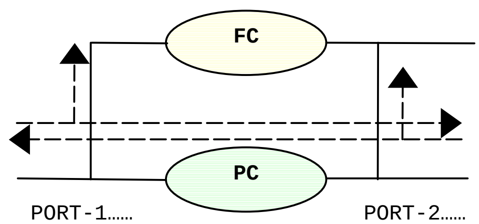


Figure 3 - In-line flow & packet classification

The Packet Classification (PC) function inspects all packets in both directions and categorizes each packet into a specific flow based upon the parameters stored in the flow description table. The packets are placed into queues and traffic policies are implemented according to the flow definitions in the flow description table (priority, rate-limit, packet discard).

This model is similar to the model for managing virus signatures in virus checking systems today. The difference is that the classification functions and classification definitions are maintained within the network, rather than on the end-subscriber's PC or workstation.

The advantage of the in-line classification method is that new applications can be rapidly identified and given packet classification and QoS enforcement functionality. Rapid distribution of new classification parameters and subsequent configuration by operators and subscribers quickly implements policies for new applications. A reasonable goal is to reduce the time for supplying QoS functionality (and generate revenue) for new applications from months (or years) to weeks (or days or hours).

Flow Classification Mechanisms

Flow classification[4] requires a number of processing-intensive mechanisms:

1. Portless flows – the flow is simply defined as a combination of IP addresses, and perhaps also a protocol identifier. No real learning is required other than detection of active packet flow with timeouts.

2. Fixed port mapping flows – the flow is simply defined as IP addresses and port numbers. No real learning is required other than detection of active packet flow to/from fixe addresses/ports with timeouts for inactivity.

3. TCP with well-known-ports – TCP has a standard application specific method for establishing flows [ref: IANA assigned number authority]. Inspection of TCP packets and looking for the session establishment commands (TCP SYN packets) can identify specific flows. The well known port number identifies the application.

4. Out of band control protocols – many application protocols use out of band methods to communicate IP addresses and port numbers. All

[4] The term "stateful classification" is sometimes used and is equivalent to flow classification as used in this paper.

packets of the out of band control protocol are monitored, and IP addresses and port numbers are extracted from the control signaling. H.323, MGCP, and SIP are three examples of out of band control protocols.

5. Protocols that use random port numbers for session setup. Many applications such as some P2P file transfer applications use random not-well-known port numbers. These are sometimes called "port hopping" applications. Applications signatures must be detected for these protocols (see below).

6. Masquerading protocols – applications masquerade as existing well-known applications protocols such as HTTP (web) and FTP (file transfer). In some cases these are well-intentioned methods to pass through firewalls without having to impose on firewall managers. For example, some implementations of voice/video instant messenger services can be carried as HTTP traffic both to utilize HTTP security and to achieve firewall traversal. In other cases they are pernicious ways of fooling the network into thinking this is a friendly application (e.g. KaZaA), and obtain preferential bandwidth and access rights. Application signatures must be utilized to ferret out any of these sorts of flows.

For several of the application types above, application signatures are required. Application signatures are defined as application-specific protocol elements embedded deep within packets, e.g. HTTP fields, which contain application specific values. There is no standard for definition of application signatures, and in some cases multiple fields and Boolean expressions must be computed before a specific application signature is matched.

An in-line flow classification function must inspect all combinations of the above application types in control streams in order to unambiguously differentiate application flows. Once the flow is identified, then specific flows and policies can be defined for use by the in-line packet classification function.

NEW TOOLS

The in-line flow classifier is a new architectural element. Since it inspects all packets in order to classify flows, it also performs packet classification functions in order to enforce administrator and subscriber defined policies. We call this element the Classification and Enforcement Engine (C&EE). See **[10, 11]** for a specific example of a C&EE and its application in managing peer to peer traffic.

The C&EE is typically external to existing network nodes (e.g. CMTS and CM), although it could be implemented internal to a CMTS or CM if the platform has enough packet processing horsepower and is not limited by inflexible ASIC functionality.

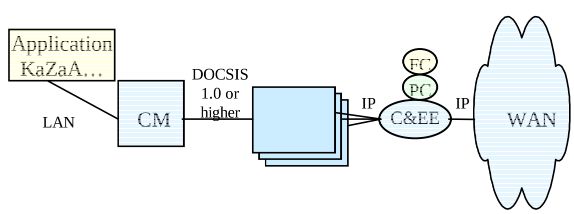


Figure 4 - C&EE with CMTS/CM

The C&EE provides both flow classification (FC) and packet classification

(PC) functions. Policies can be defined and enforced for all applications, including P2P file transfer, even in DOCSIS 1.0 systems. New applications can be detected and configured by the system administrator. Database updates from the C&EE supplier can update the knowledge base of application and protocol types with quick turnaround times.

End-to-End System Architecture Evolution

Since the C&EE, CMTS, and CM are all capable of providing packet classification (PC) functions (both DOCSIS 1.1 and 2.0), a future opportunity exists to integrate all networking elements into a strong end-to-end QoS management architecture. In the integrated architecture the flow classification and packet classification function of the C&EE is combined with the packet classification functions of the CMTS and CM via standard signaling interfaces. Administrator and Subscriber defined traffic policies can then be concurrently applied to both QoS-smart applications and QoS-unaware applications.

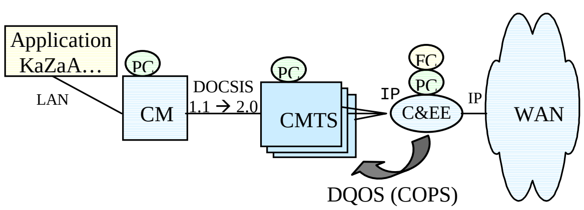


Figure 5 - C&EE and DQOS

Figure 5 shows a configuration for the integrated end-to-end system. The C&EE aggregates traffic from multiple CMTS platforms (and their downstream CMs and applications). The C&EE performs flow classification (FC) and utilizes the DQOS signaling interface to communicate relevant parameters and policies to the packet classification (PC) function of the CMTS. The CMTS utilizes the DOCSIS 1.1 (and above) signaling interface to communicate relevant parameters to the packet classification function of the CM. No application server is required in order to support or add new applications. This integration is achieved using existing PacketCable DQOS signaling interfaces and parameter definitions.

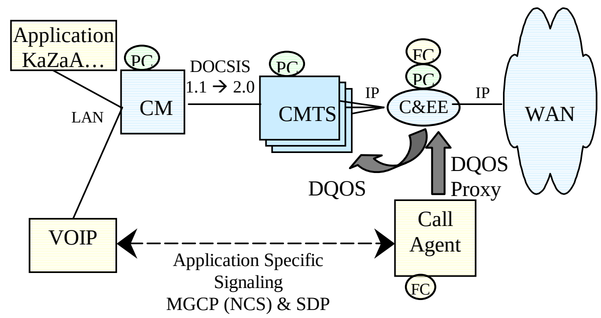


Figure 6 – C&EE with QoS-smart Server

Figure 6 shows concurrent operation of C&EE in-line flow classification with QoS-Smart application servers flow classification (e.g. PacketCable Call Agent). In this configuration the C&EE node proxies between the downstream CMTS nodes and the Policy Server (e.g. PacketCable Telephony Call Agent). In this configuration both QoS-smart flow learning (e.g. for PacketCable Telephony) and in-line flow learning can coexist.

Some protocol and parameter extensions are likely required in order to maximize the functionality of the full end-to-end QoS architecture[5]. For example, the total available bandwidth is divided between the flows known by the QoS smart application (PacketCable Telephony) and the C&EE node. Also, given the wider variety of

[5] The current proposal is based upon the PacketCable DQOS framework. Once PacketCable Multimedia is published appropriate modifications can be introduced to the proposal.

application types supported, some new traffic policy types may need to be defined.

CONCLUSION

A new approach is proposed for dealing with the explosive growth of new applications, new application types, and rising subscriber expectations for fair and configurable quality of service in DOCSIS HFC networks. The approach features a new element which can be incrementally added to the end-to-end architecture. This element is called the Classification and Enforcement Engine (C&EE) and provides in-line flow classification functions for both existing and new types of application traffic. The C&EE is compatible with the current QoS-Smart model for flow learning in PacketCable, DOCSIS, and emerging PacketCable Multimedia standards.

The C&EE can be implemented in DOCSIS 1.0 systems, and, with appropriate future extensions, can utilize the PacketCable DQOS signaling interface to fully utilize the packet classification functions contained within the CMTS and CM for DOCSIS 1.1 (and higher) systems.

Using new tools like the C&EE the operators and subscribers will have the ability to manage and control bandwidth utilization on an application by application basis. More importantly, operators will both retain subscribers and reap new revenue for managed bandwidth services.

REFERENCES

[1] CableLabs, DOCSIS™ 1.1, "Data-Over-Cable Service Interface Specifications, Radio Frequency Interface Specification - SP-RFIv1.1-I09-020830", 8/30/2002.

[2] CableLabs, DOCSIS™ 2.0, "Data-Over-Cable Service Interface Specifications, Radio Frequency Interface Specification SP-RFIv2.0-I03-021218", 12/18/2002.

[3] CableLabs, **"**PacketCable™ 1.0 Architecture Framework Technical Report - PKT-TR-ARCH-V01-991201", 12/1/1999.

[4] CableLabs, "PacketCable™ Dynamic Quality-of-Service Specification - PKT-SP-DQOS-I05-021127", 11/27/2002.

[5] Marshall McLuhan - The Medium is the Massage. New York: Bantam Books, 1967.

[6] Pickens, "Transport protocol revolution", in SNA and TCP/IP Enterprise Networking, Manning Publications, 1998.

[7] Ian Foster and Adriana Iamnitchi, "On Death, Taxes, and the Convergence of Peer-to-Peer and Grid Computing", http://iptps03.cs.berkeley.edu/final-papers/death_taxes.pdf.

[8] CableLabs, "PacketCable™ Network-Based Call Signaling Protocol Specification - PKT-SP-EC-MGCP-I06-021127", 11/27/2002.

[9] IETF, RFC 2327 "SDP: Session Description Protocol", April 1998.

[10] Allot Communications Inc., "NetEnforcer Data Sheet", Allot Communciations, Inc., www.allot.com.

[11] Allot Communications, Inc., "KaZaA v2 and other new P2P Applications Technical Note", www.allot.com.

John Pickens is a Technical Advisor to Allot Communications, Inc. and can be contacted at jpickens@ieee.org.

Greg Hutterer is the Director of Service Provider Sales at Allot Communications, Inc. and can be contacted at ghutterer@allot.com or (952) 944-3100.

MATHEMATICAL MODEL OF INTERACTIVE PROGRAMMING GUIDE

Yakov Kamen
Entelel Corporation

Abstract

The classic interactive programming guide (IPG) was designed over 20 years ago using a grid data-presentation model. This design was perfectly suitable for a small number of homogeneous video channels and a short (few - hour-long) schedule. Today's IPG must manage over 300 heterogeneous video, PPV, VOD, and music channels in a two week schedule. It also has to manage time shifting (PVR) capabilities. The classic grid-based IPG was never designed to handle these tasks, and has to be significantly modified to reflect this new reality. The big question is how to modify the IPG so that it wins consumers' minds and solves the new problems? A mathematical model of the IPG is necessary to make the right decision.

This article describes the first mathematical model of the IPG based on the cognitive information theory. Different popular IPG solutions are analyzed and compared based on the proposed model.

IPG COMPONENTS AND STRUCTURE

TV Event Descriptions

The Interactive programming guide (or IPG) allows the viewer to view and manipulate TV schedule data directly on the TV screen. The schedule data can be described as a structured set $\{E\}$ of TV event metadata or events. Each event E consists of a channel ID that defines the event's channel, starting time, event length, event name, and event description. Formally an event E is defined as a structure:

$$E = \langle C_{ID}, T_S, \Delta T, E_N, E_D \rangle, \qquad (1)$$

where

- E - is an event description,
- C_{ID} - is a channel ID, that may include the channel name, number, ID, etc.
- T_S - is an event's start time (time stamp)
- ΔT - is an event's length (in minutes)
- E_N - event's name (can be empty)
- E_D - event's description (can be empty or can be a complex structure of different multi-resolution description representations)

Note, that according to definition (1), when the same TV program is shown on different channels or at different times, it is considered two different TV events.

TV Channels

There are two types of channels in the IPG: regular (TV) channels and special "on-demand" (OD) channels. In the case of TV channels, all events are linearly ordered by time and can not intersect in the time domain. In the case of OD channels, events are not linearly ordered by time. In this article we consider all channels to be TV channels.

Major Components of IPG

Each IPG represents schedule data differently, but there are common rules that affect the guide's logical structure. For example, descriptions of events that have already passed are never shown on the screen. The conventional IPG that follows existing rules consists of the following components (Fig.1): sorting and searching control component, date and time, event listing (which consists of a subset of TV events described by their names), time and channel IDs, and the description of the highlighted event or dynamically updating help information [Kam01].

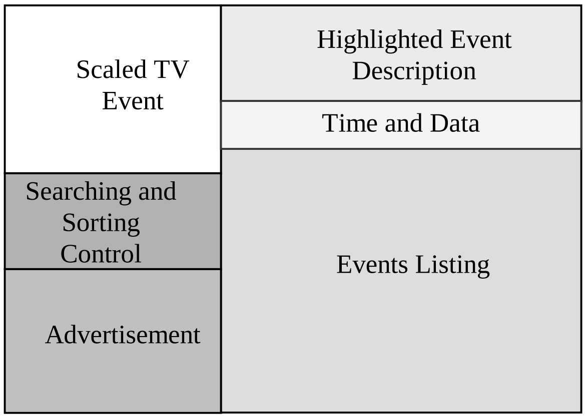


Fig1. Typical Components of an IPG

The Event listing is the most important component of the IPG, and therefore we will concentrate on its modeling and optimization. The "surfability" of the schedule data is the second most important component of the guide.

IPG OPTIMIZATION CRITERIA

Criteria specification

It is very difficult to define numerical IPG optimization criteria.

First of all, there are different formidable traditions of event listing presentation in different countries.

Second, depending on subscription package and location, different users have access to different channel packages, and, as a result, have different needs in an optimized IPG. For instance a user that watches 12 public broadcast channels would be satisfied with any IPG. However a user that is subscribed to 300+ channels and actively uses different recording devices (PVR, DVD, VCR) is significantly more sensitive to the IPG's efficiency.

Third, the description language significantly affects IPG event listing design, because a hieroglyphic language demands a different data presentation esthetic than alphabetic languages.

Fourth, different viewers have different TV-watching habits. The same person usually has different behavior patterns on working days, weekends, on vacation, and on holidays. These patterns are not stable. They tend to change over the years depending on health, family, and living conditions.

Fifth, optimization criteria must be "computable" and verifiable. This means that a criteria like "create an event listing such that all users will be happy" would not satisfy the goal of this work.

Sixth, the IPG user interface is, after all, a work of art. This means that the best and most useful solution may not be the most practical or ergonomic.

With all of the considerations above, the proposed mathematical model is based on a synthetic criteria C that consists of two separate criteria C_1 and C_2. The first criteria C_1, called "maximum listing information" criteria, estimates event listing information value. The second criteria C_2, called "minimum energy surfing" criteria, estimates the effort a user has to exert to find a TV program he would like to watch or record.

First we define criteria C_1 :

Formally, each event name listing (event listing) is a projection P of a subset of events on the screen. Each projected event in the listing is represented as

$$P(E) = < C_{ID}, T_S, \Delta T, P(E_N) >, \qquad (2)$$

where

$P(E_N)$ - is a projection of the event's name.

The criteria C_1 is defined as the maximization criteria comparing event listings by total information projected on the screen:

$$C_1 = \max_P \sum_1^{N(P)} P(E_i), \qquad (3)$$

where

$P(E_i)$ - is a projection of the i-event to the screen;

$N(P)$ - is the number of event names on the screen projected by P.

Now we define criteria C_2 .

The criteria C_2 is a minimization criteria that compares different IPGs by the average energy the user has to spend to go from an event E_0 to an event E_1. In this article we define an energy unit as a single key press of the remote controller. As a result the criteria C_2 allows us to find an IPG that requires a minimal average number of key presses to go from one arbitrary event name to another. To formalize criteria C_2 we define a distance function R in the event space such that $R(E_i, E_j)$, or the number of remote controller key presses needed to move the focus from event name E_i to the name E_j, is minimal. Lets assume that A(.) is an averaging operator as it has been defined in [Kam94]. The average distance between all pairs of events we will call "the IPG surfing diameter" or just IPG diameter. It is a good measure of IPG surfing energy. Formally, Criteria C_2 is an IPG diameter minimization criteria described as

$$C_2 = \min_R \underset{(i,j)}{A}(R(E_i, E_j)), \qquad (4)$$

where

$A(x_1,\ldots x_n)$ - is an average function between $x_1,\ldots x_n$;

$R(x,y)$ - is the distance between objects x and y.

In the proposed mathematical model all IPG solutions are measured by the criteria C_1 and C_2.

ASSUMTIONS AND CONSTRAINTS

Any mathematical model is based on a set of basic assumptions and constraints that allow one to make non-trivial general conclusions.

Described below are the major assumptions and constraints of our IPG mathematical model.

- Homogeneous Event Value. All TV events in the schedule have the same priority value for all users. In the real world this assumption is not correct.
- Transmission Continuity. The current model assumes that all channels are always transmitted without interruption 24 hours per day, 7 days per week.
- Channel Structure. In the current model we assume that all available channels are TV channels where the events are linearly ordered. OD-channels do not exist in this model.
- Channel Distribution. A real user has his own list of "informative" channels and "non-informative", "noisy" or "garbage" channels. In the model below we assume that all channels are equally informative.
- Event Independence. Information located inside two arbitrary event descriptions is independent, i.e. for every two events E_1 and E_2 information I located in the pair of events is equal to the sum of information located in each event: $I(E_1 + E_2) = I(E_1) + I(E_2)$.
- Channel Independence. Information located inside two arbitrary channels is independent, i.e. for every two channels c_1 and c_2 information I located in the pair of channels is equal to the sum of information located in each channel: $I(c_1 + c_2) = I(c_1) + I(c_2)$.
- Semantic Equivalence. All descriptions that consist of the same number of symbols have equal amounts of information
- Event Information Equivalence. All event names viewed at the same time are equally important for users and consist of an equal amount of information

EVENT LISTING INFORMATION

Event Listing Modeling

There are numerous event listing models. The event listing information criteria C_1 measures the quantity of the information not its quality. From C_1 point of view all event names viewed at the same time are equally important for a user and consist of an equal amount of information I_0. For simplicity we set $I_0 = 1$. Each event name is projected on the screen into the event listing's fixed-sized "cell". If the cell is smaller than the event name, the event name is truncated and it looses some amount of information. The same video content has a different value to the user depending on whether it is already in progress, starting now, or will be playing in the future, because the starting time matters. Obviously, an event has the maximum value for the user if it is starting now. However it is a fairly rare case: usually events have already started (currently playing event) or will start in the future (future event). Both playing and future events have less information for the user than "starting now" events.

A major assumption of this mathematical model is that the total information value of the event listing is a sum of the information values of all projected events (2), and each projected listing event can be completely described as follows:

$$I_{P(E)} = I_{\mathrm{E}}(\mathrm{s_E}, \mathrm{s_C}) \cdot G(T_0, \Delta T, T_c) \cdot I_0,$$
$$I_{\mathrm{E}}(\mathrm{s_E}, \mathrm{s_C}) \in [0,1]; G(T_0, \Delta T, T_c) \in [0,1], \qquad (5)$$
$$I_0 = 1,$$

where

$I_E(s_E, s_C)$ - is a name value function that describes the amount of information that has been "left" in the original name after the projection P;

$G(T_0, \Delta T, T_c)$ -is a time value function that describes the information value that the current event has compared to the information value it would have if the event started immediately;

s_E -is the size of the event name;

s_C -is the size of the cell.

Most Popular Event Listing Models

Three examples below describe the most popular event listing organization schemes.

Example1. Grid based event listing. Grid data representation is the most popular listing design approach in the US. It consists of a set of time-proportional rectangular cells that are used to show event names. A simple example of a grid listing is shown in Fig. 2. An abstract description of the grid listing page is shown on Fig.3.

Fig.2 Grid based listing page

<table>
<tr><th></th><th>T^1</th><th>T^2</th><th>T^3</th><th>T^4</th></tr>
<tr><td>C_1</td><td colspan="2">$E_k(C_1)$</td><td>$E_{k+1}(C_1)$</td><td>$E_{k+2}(C_1)$</td></tr>
<tr><td>C_2</td><td>$E_k(C_2)$</td><td>$E_{k+1}(C_2)$</td><td colspan="2">$E_{k+2}(C_2)$</td></tr>
<tr><td>C_3</td><td>$E_k(C_3)$</td><td>$E_{k+1}(C_3)$</td><td>$E_{k+2}(C_3)$</td><td>$E_{k+2}(C_3)$</td></tr>
<tr><td>C_4</td><td colspan="4">$E_k(C_4)$</td></tr>
<tr><td>C_5</td><td colspan="2">$E_k(C_5)$</td><td colspan="2">$E_{k+1}(C_5)$</td></tr>
<tr><td>C_6</td><td colspan="3">$E_k(C_6)$</td><td>$E_{k+1}(C_6)$</td></tr>
<tr><td>C_7</td><td colspan="2">$E_k(C_7)$</td><td colspan="2">$E_{k+1}(C_7)$</td></tr>
</table>

Fig. 3 Formal model of the grid listing

In the formal model (Fig.3) $C_1,\ldots, C_7$ are sequential channel IDs; $T^1,\ldots, T^4$ are standard time intervals (usually 30 minutes). $E_k(C_i)$ is the name of the event that is being transmitted on channel "i" during the time interval T^1. $E_{k+l}(C_i)$ is the name of the event that will be transmitted on channel "i" after the event with the name $E_k(C_i)$ at the "l"-step.

Example2. Link-list based event listing. The link-list listing shows the maximum number of events of the currently highlighted channel on the same screen. The Link-list solution is very useful for digital video recording (DVR or PVR) enabled systems. In this example we defines two link-list schemes based on wide and narrow cells. The first scheme (Fig. 4) is using wide cells to present event listings, the second scheme (Fig.5) is using narrow cells to present event names.

	T^1	C_3 page $_1$
C_1	$E_k(C_1)$	$E_{k+1}(C_3)$
C_2	$E_k(C_2)$	$E_{k+2}(C_3)$
C_3	$E_k(C_3)$	$E_{k+3}(C_3)$
C_4	$E_k(C_4)$	$E_{k+4}(C_3)$
C_5	$E_k(C_5)$	$E_{k+5}(C_3)$
C_6	$E_k(C_6)$	$E_{k+6}(C_3)$
C_7	$E_k(C_7)$	$E_{k+7}(C_3)$

Fig. 4 Link-list based event listing (wide cells)

	T^1	C_9 page $_1$	T^1	
C_1	$E_k(C_1)$	$E_{k+1}(C_9)$	$E_k(C_1)$	C_8
C_2	$E_k(C_2)$	$E_{k+2}(C_9)$	$E_k(C_2)$	C_9
C_3	$E_k(C_3)$	$E_{k+3}(C_9)$	$E_k(C_3)$	C_{10}
C_4	$E_k(C_4)$	$E_{k+4}(C_9)$	$E_k(C_4)$	C_{11}
C_5	$E_k(C_5)$	$E_{k+5}(C_9)$	$E_k(C_5)$	C_{12}
C_6	$E_k(C_6)$	$E_{k+6}(C_9)$	$E_k(C_6)$	C_{13}
C_7	$E_k(C_7)$	$E_{k+7}(C_9)$	$E_k(C_7)$	C_{14}

Fig.5 Link-list based event listing (narrow cells)

Example 3. Event Matrix Listing. This type of data representation is popular in some European countries (Fig.6).

	C_1		C_2
$t(E_k(C_1))$	$E_k(C_1)$	$t(E_k(C_2))$	$E_k(C_2)$
$t(E_{k+1}(C_1))$	$E_{k+1}(C_1)$	$t(E_{k+1}(C_2))$	$E_{k+1}(C_2)$
$t(E_{k+2}(C_1))$	$E_{k+2}(C_1)$	$t(E_{k+2}(C_2))$	$E_{k+2}(C_2)$
$t(E_{k+3}(C_1))$	$E_{k+3}(C_1)$	$t(E_{k+3}(C_2))$	$E_{k+3}(C_2)$
$t(E_{k+4}(C_1))$	$E_{k+4}(C_1)$	$t(E_{k+4}(C_2))$	$E_{k+4}(C_2)$
$t(E_{k+5}(C_1))$	$E_{k+5}(C_1)$	$t(E_{k+5}(C_2))$	$E_{k+5}(C_2)$
$t(E_{k+6}(C_1))$	$E_{k+6}(C_1)$	$t(E_{k+6}(C_2))$	$E_{k+6}(C_2)$

Fig. 6 Formal model of the matrix listing

In the figure above C_1 and C_2 are two sequential channel IDs; $E_k(C_i)$ is the name of the event that is being transmitted on channel "*i*" during the time interval T^1. $E_{k+m}(C_i)$ is the name of the event that is transmitted on channel "*i*" at the "*m*"-step; $t(E_k\ (C_i))$ is the starting time of the event "*k*" on the channel C_i.

How would one decide which name listing representation is more informative? This can be done by comparing information presented on the "average" page of the listing using formula (5) and its realization described below.

Name Value

At first glance it is beneficial to show as many event names on the same listing as possible. However, screen space is always limited and the visible part of the name inevitably shrinks when new "cells" are added to the screen.

The name value function (name value) estimates the amount of information left in an event name of size s_E after its projection into an event listing cell of size s_C.

Below we will define the name value as a monotonic function of two variables

$$I_E(s_E, s_C) = \begin{cases} 1, & \frac{s_C}{s_E} \geq 1 \\ 0 < f(\frac{s_C}{s_E}) < 1, & a < \frac{s_C}{s_E} < 1, \\ 0 & \frac{s_C}{s_E} < a \end{cases} \qquad (6)$$

where

f(x) -is a monotonically increasing function;

a -is a threshold parameter $a \in [0,1]$ that defines the average loss of information that transforms data into noise.

For simplicity we will linearly approximate *f(x)* as:

$$f(d) = \frac{1}{1-a}d - \frac{a}{1-a}, \quad d = \frac{s_C}{s_E}, \qquad (7)$$

As a result, formula (6) will look like the following (Fig.7):

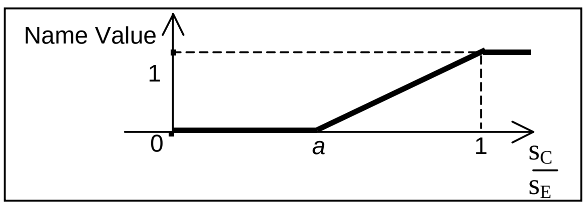


Fig.7 Name value function approximation

According to experimental data analysis a = 0.6 is a good approximation of the name threshold value for English-language based US TV schedule.

Time Value

As mentioned above, the information value of an event name in the event listing depends on the event's starting time. In the chosen model, the time value for future events monotonically decreases over time. The time value of the currently playing event is a monotonically growing function and as a result, the longer the event has been playing, the less value it has to the viewer.

The exact functional tie between the event starting time and its time value depends on many parameters, including subjective characteristics of the user, event's genre, structure, etc. Schematically the typical function may look like the following (Fig. 8)

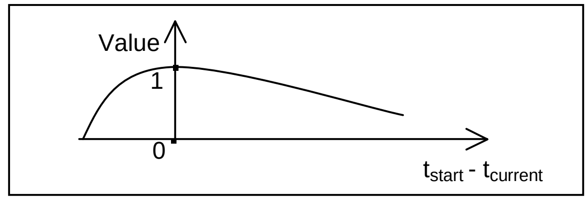


Fig.8 Time value

For simplicity of the model we assume that future events' time value decreases with a constant speed. Accepting this assumption, $G(T_0, \Delta T, T_c)$ can be approximated with the formula (8) below:

$$G(T_0, \Delta T, T_c) = \begin{cases} \left(\frac{T_c - T_0}{\Delta T}\right)^m \\ b^k, & k = \frac{T_c - T_0}{\Delta T} \end{cases} \qquad (8)$$

where

m -is a parameter that defines time value degradation speed for currently playing events;

b -is a parameter that defines time value degradation speed for future events.

In practice, both information degradation speed parameters *m* and *b* would vary per user, time of day, or type of equipment (PVR). Based on experiments we found that m belongs to the interval [0.2, 1] and *b* belongs to the interval [0.7, 0.9]. Assuming that the currently airing event's start time is distributed uniformly we approximate (8) with:

$$G(T_0, \Delta T, T_c) = \begin{cases} 0.5^m \\ b^k, & k = \mathrm{int}(\frac{T_c - T_0}{\Delta T}) \end{cases}, \qquad (9)$$

where

int(*x*) -is the integer part of *x*.

Comparison of Models

With a few additional assumptions formulas (5)- (9) allow us to compare different models of event listings. In this article we compare models described in the examples above: "Grid", "Link-list wide", "Link-list narrow", "Matrix wide", and "Matrix narrow" . We will compute the average listing information values for the four cell sizes: 8, 10, 12, and 14. For comparison we assume that the standard time interval T is 30 minutes, and the number of visible cells is 28 (7x4). In the model comparison process we used empirical data collected from real US TV schedules (English language). This data includes a tabulated average name value function for each cell size, called a cell power table. A fragment of the cell power table is presented in Table 1.

Table1. Cell power table (fragment)

	cell size (symbols)				
	8	9	10	11	12
Name value	0.24	0.30	0.38	0.41	0.50

It also includes a tabulated histogram of event duration distribution (see Table2).

Table2. Event duration histogram

Event duration (min)	Percent (%)
1-30	56.8
31-60	20.7
61-120	10.4
>120	12.1

Final comparison results are presented in Table 3 below

Table3. Models Comparison

Listing type	cell size (symbols)			
	8	10	12	16
Grid	3.98	5.07	5.84	6.83
Link-list wide	4.09	5.64	6.40	7.58
Link-list narrow	2.29	3.96	6.49	9.17
Matrix wide	4.03	5.37	6.15	6.89
Matrix narrow	0.80	1.92	4.12	7.13

Table3 shows that there is no event listing model that is "the best" for all cell sizes.

GUIDE SURFING

Surfing Control

The minimum energy surfing criteria C_2 would benefit IPG solutions that use a lot of special keys that "short cut" the most popular step sequences. But the idea of improving the surfing experience by adding special keys does not work. First, screen space and remote controller buttons are limited. Second, it is impossible to convince a user to learn an "F-16 cockpit" style remote controller to surf TV in the dark. To make the minimal energy surfing criteria meaningful, we assume that all designed models must use the same minimal set of surfing keys: up, down, left, right, select, and ten digits 0-9.

Channel and Time Distance

Without limitations we would consider that the distance $R(E_i, E_j)$ between two arbitrary events E_i and E_j, used in formula (4), is a "manhattan" metric in the channel/time coordinate space. In other words,

$$R(E_i, E_j) = R^C(E_i, E_j) + R^T(E_i, E_j), \qquad (10)$$

where

$R^C(E_i, E_j)$ -is the distance between the channels of events E_i and E_j ;

$R^T(E_i, E_j)$ -is the distance between the times of events E_i and E_j.

In formula (10) $R^C(.)$ is called a channel distance and $R^T(.)$ is called a time distance.

Users' Tasks and Models

Users surf the IPG to solve three main tasks:

Task A. Find something to watch now.

Task B. Find something to watch soon.

Task C. Find something to record or to watch in the future.

In the case of task A, time distance is equal to zero and only the channel distance has to be estimated. When the user knows the channel number, the optimal surfing solution is dialing that channel number. In this case, the IPGs diameter is equal to the average number of digits in a channel number.

When the desired channel number is not known but the channel name is, task A is to tune to the channel based on its name with the minimal number of key presses. Assuming that channels are uniformly distributed, and that the name listing is the only surfing solution, the expected IPG diameter is approximated with the following formula:

$$\underset{(i,j)}{A}(R(E_i,E_j)) = \underset{(i,j)}{A}^{C}(R(E_i,E_j)) \approx \frac{0.5K}{N} + 0.5N, \qquad (11)$$

where

K -is the total number of channels;

N -is the number of channels visible on the event listing.

Function (11) achieves its minimum when $N = \sqrt{K}$ and it is equal to N. When N is close to $\sqrt{K}$, simple event listing scrolling is the optimal surfing method in the channel domain. However when the difference is large enough, there are additional opportunities to minimize the channel diameter by implementing multi-resolution data representation modules. The simplest idea of multi-resolution channel list representation is the idea of a "channel matrix" (Fig.7, courtesy iSurfTV Corporation).

Fig.7 Channel Matrix example

The channel matrix module uses screen space to show the maximum number (L) channel IDs (in visual or textual format) on the screen ($L>N$). The channel matrix (Fig. 7) has almost no information about the playing event names. This means that the user has to surf inside the matrix page to check several channels before he will make his decision to switch to a channel. The channel diameter of an IPG that includes a channel matrix module can be approximated with the formula:

$$\underset{(i,j)}{A}^{C}(R(E_i,E_j)) \approx \frac{0.5K}{L} + 0.5L, \qquad (12)$$

where

L -is the number of channel IDs in the channel matrix ($L>N$).

Note that diameter (12) is smaller than diameter (11) only if L is closer to $\sqrt{K}$ than N, i.e. if $N < L < 2\sqrt{K} - N$. This means that the matrix module would improve the surfing experience only in the case of a large number of playing TV channels.

Another solution is to create a new module that stores channel IDs alphabetically in the "notebook" style. Using the "optimal" notebook module channel diameter can be decreased to:

$$\underset{(i,j)}{A}^{C}(R(E_i,E_j)) \approx \sqrt{K} + 2 \qquad (13)$$

Let us now analyze tasks B and C. Both of them require the user to surf in the time domain. Below we will compute the time diameter in both tasks B and C. As with (11) we will approximate the time diameter of the linear event listing, when surfing in the time domain, with the following formula:

$$\underset{(i,j)}{A}^{T}(R(E_i,E_j)) \approx \frac{0.5K}{N} + 0.5N, \qquad (14)$$

where

K -is the total length of the schedule in hours (K varies from 1 to 720 hours);

N -is the average period of time visible at the event listing (N varies from 0.75 to 6 hours).

Formula (14) allows us to formalize the concept of "playing soon" events as events that would start playing in the time interval when the linear time surfing is the most optimal solution. Formally this time interval is defined as the interval $[T_C; T_C+N^2+N]$, where T_C is the current time.
Based on the definition of "playing soon" events we will estimate the time dimension in task B based on formula (14).
The optimal solution of task C is based on a multi-resolution time representation. In this model we will analyze three competing implementations.
The first implementation is a homogeneous grid that positions days on the first dimension and time intervals on the second dimension. The time diameter in this approach could be approximated with the formula:

$$A^T_{(i,j)}(R(E_i,E_j)) \approx \frac{K}{48}+\frac{12}{N}+1 \qquad (15)$$

The second implementation is a set of two screens: a day listing screen, and a screen with 12 one-hour intervals. The time diameter is approximated by:

$$A^T_{(i,j)}(R(E_i,E_j)) \approx \frac{K}{336}+13 \qquad (16)$$

The third implementation is a set of three layers: day, part of the day (morning, day time, prime time, evening, etc.), and one hour time intervals inside each part of the day. The time diameter in this implementation is fractionally smaller than diameter (16).
Comparing formula (14), (15) and (16) we can conclude that solution (14) is preferable for a very short schedule (less than 2 days), solution (15) is preferable when the schedule fluctuates between 2 and 10 days and solution (16) is preferable when the schedule is longer than 10 days.

FUTURE DEVELOPMENT

In this article we presented the first mathematical model of an interactive programming guide. A lot of assumptions and constrains make the usability of this model limited in practice. Therefore many of these constrains can be waved without serious complications. A model's complication would be compensated by its improved practicality. The maximal event listing information and minimal energy surfing criteria also can be generalized and improved. For instance, a channel's probability of being watched can be added to the event listing information criteria. In the minimal energy surfing criteria we can replace the "key press", as a measurement unit, with time.
Several important questions had not been discussed in this article. For example a model's robustness to the parameter's variation is extremely important for practical implementation.

REFERENCES

[Kam 01] Yakov Kamen, "Components and Advanced Functionalities of the Next Generation Interactive Programming Guide," Proc. "The 2nd Annual Digital TV Application Software Environment (DASE) Symposium 2001:End-to-End Data Services, Interoperability & Applications", Gaithersburg, June19-20, 2001, p.307-316

[Kam 94] Yakov Kamen and Sergei Ovchinnikov "Meaningful Means on Ordered Sets", Proc. "Fifth International Conference IPMU", Paris, July 4-8, 1994, p.1009-1013

METHODS TO INCREASE BANDWIDTH UTILIZATION IN DOCSIS 2.0 SYSTEMS

Daniel Howard, Laura Hall, Keith Brawner, and Hans Hsu
Broadcom Corporation
Nick Hamilton-Piercy, Reynold Ramroop, and Sheng Liu
Rogers Cablesystems

Abstract

Several methods to increase bandwidth utilization in DOCSIS 2.0 systems are presented, and results of a field trial that demonstrated several of the methods are reported. On the downstream where the majority of packets transported are large packets, the use of wider bandwidth RF channels of 12 or more MHz is discussed, where statistical multiplexing gains improve the channel capacity above and beyond the factor by which the bandwidth is increased. Similarly, going to higher order modulation on the downstream (1024 QAM) also increases the downstream capacity, and in particular can provide MSO's the ability to get more HDTV channels per 6 MHz of downstream bandwidth or 25% more data capacity in DOCSIS downstreams.

On the upstream where small data and voice packets can be the majority of packets transported, bandwidth saving techniques such as dynamic header suppression and synchronous operation reduces the packet overhead that can account for a significant fraction of the minislots required to transport the packet. For medium and large packets on the upstream, the use of higher order modulation up to 256 QAM can be used to expand the capacity, and when combined with the ingress and impulse robustness features of DOCSIS 2.0 systems, 256 QAM upstreams can be supported on today's cable plants, even in bands which previously could not support more than QPSK. These facts are born out by the results of a field trial in which both 1024 QAM on the downstream and 256 QAM on the upstream were demonstrated, the latter in the presence of ingress noise on the cable upstream. The conclusion is that many cable plants are already capable of supporting higher orders of QAM modulation, and thus capacity increases of up to 33% above that already provided by DOCSIS 2.0 are possible.

INTRODUCTION

It is generally recognized that bandwidth needs of users on cable plants will continue to increase as new applications and higher quality versions of existing applications emerge. Applications that increase bandwidth needs on the downstream of cable plants include high definition TV (HDTV), video-on-demand (VOD), voice over IP (VoIP), and higher speed data service. The growth in data rate requirements over time is still considered exponential, doubling every year at least, however the trend can perhaps be better appreciated via applications that are known to be increasing the current data rate requirements for cable operators in particular. For example, activities such as more frequent and larger file transfers via email and web browsing due to imbedded high quality photographs, audio, and video will drive both upstream and downstream data rate requirements upward in the near future. In particular, the transfer of video files over broadband Internet connections is already part of current product lines for personal video recorder devices such as RePlay TV's current offering [1], where the capability to share movies over the Internet with friends and family is touted.

On the upstream, home based servers and peer to peer networking applications a la Napster (and its current look-alikes) will continue to drive bandwidth needs upward. And upstream bandwidth must be provided in a more robust manner, since RF interference is frequently present. When RF spectrum below 25 MHz must be used for additional upstream data channels, the modulation technology must be robust to ingress, impulse, and higher thermal noise conditions.

The upstream capacity has recently been addressed by DOCSIS 2.0 technology, which provides significant increases in upstream capacity and robustness on cable plants. Compared to the DOCSIS 1.1 maximum rate of 16 QAM operation at 2.56 Megasymbols/sec, DOCSIS 2.0 technology enables at least 64 QAM at 5.12 Megasymbols/sec, an improvement of three times in raw capacity. DOCSIS 2.0 provides this improvement by providing a new modulation technique, SCDMA, which is inherently robust to impulse noise, by increasing the robustness of TDMA via byte interleaving and increased FEC, and by the proprietary schemes most vendors provide for cancellation of ingress.

DOCSIS 2.0 does not address capacity increases in the downstream, however, and if data rate requirements are doubling every year, then in two years the three-fold increase in upstream capacity provided by DOCSIS will be used up and additional improvements in bandwidth may be needed.

Hence, there is still a need to consider techniques that increase capacity in both the upstream and the downstream on cable plants. In this paper, several techniques for increasing the capacity on cable plants are presented, and field tests of some of these techniques are also presented as proof of their viability. The techniques include higher order modulation on both the upstream and downstream, synchronous operation on the upstream, and dynamic payload header suppression on the upstream.

METHODS AND BENEFITS OF INCREASING CAPACITY ON THE DOWNSTREAM

Wider Channels

On the downstream of cable plants, there are two main techniques that can be used to increase the raw capacity of data services: increasing the total channel width, and increasing the order of modulation. It should also be noted that increased video compression via MPEG4 is another way to get more channels in the same RF bandwidth, however in this paper we focus on the media access control (MAC) and physical (PHY) layers.

Increasing the downstream channel width can be done in two manners: first, the symbol rate can be increased. Since cable downstreams are channelized on 6 MHz spacing in North America and 8 MHz spacing in Europe, the most efficient manner to use for increasing the symbol rate of DOCSIS downstream signaling would be in integer multiples of 6 MHz (or 8 MHz in Europe). Hence, the first method of increasing the downstream channel capacity is to increase the downstream symbol rate from about 5 Megabaud to 10 Megabaud and the subsequent downstream RF bandwidth to 12 MHz, or twice the current 6 MHz RF bandwidth.

Note that doubling the channel width does not necessarily increase the spectral efficiency of the transmissions since the alpha factor used in symbol shaping may remain constant. Even though the guard band in between the two individual channels is removed, more guard band on the edges of the signal spectrum is required in terms of Hz for the same value of alpha when a larger symbol rate is used. Since doubling the channel width does not increase the spectral efficiency, the main benefit of this technique lies in the additional statistical multiplexing gain that comes from wider channel widths. Essentially, leftover capacity in each of the individual channels from gaps in scheduling transmissions can be combined and used for additional transmissions in the single, wider channel. Further, latency can be reduced by exploiting earlier opportunities to transmit in the combined channel instead of waiting for opportunities in a single, smaller channel. Estimates of statistical multiplexing gain vary from 10% to 40% [2], depending on the traffic, size of the original channel, and the number of channels being multiplexed.

One of the issues associated with using larger channel widths on the downstream is that legacy modems cannot use the larger channels, being limited to conventional 6 MHz channels. Thus, the second method of increasing the statistical multiplexing gain on DOCSIS downstreams is to combine multiple downstream channels logically so that a modem can receive on multiple 6 MHz channels simultaneously. This method allows future modems to access larger channels, and headend schedulers to use leftover capacity in the individual channels more effectively, but at the same time permits legacy modems to continue to use the individual 6 MHz channels. A version of this technique was described in a recent NCTA/SCTE paper by ATT [3], where 40% gains in channel utilization from statistical multiplexing were shown for combinations of four downstream channels.

Higher Order Modulation

Wider downstream channels will clearly provide some level of increased performance on the downstream, however traffic variation and the ratio of new to legacy modems cause statistical multiplexing gains to be variable and difficult to predict for multiple and/or wider downstream channels. Another technique, which gives clear and predictable gain in the downstream channel, is the use of higher order modulation, for example 512 QAM or even 1024 QAM. In the latter case, the spectral efficiency is increased from 8 bits/symbol (256 QAM) to 10 bit/symbol, an increase of 25%. Further, by increasing the channel capacity directly, there will also result a statistical multiplexing gain in that channel. Figure 1 depicts the received constellation of a 1024 QAM downstream signal.

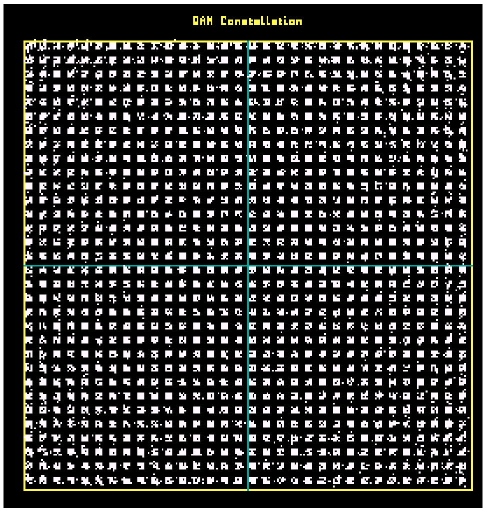


Figure 1. 1024 QAM DS Signal

But higher order modulation on the downstream has been criticized in the past based on the difficulty in reliably operating even 256 QAM on the downstream [4]. In particular, the cresting of CTB and CSO on the downstream and higher SNR requirements are often quoted as factors that could limit or prevent the successful operation of higher order modulation on the downstream.

For the SNR issues (which also arise in using double- and quadruple-wide downstream channels), it may be noted that in current cable downstreams, the transmit power of digital channels is backed off from the level used by analog channels, by up to 10 dB. The reason is that analog TV requires an SNR of up to 46 dB for high quality video, while a practical digital TV receiver requires only about 30 dB of SNR. In going to higher order modulation such as 1024 QAM, another 6 dB of SNR would likely be required, which is still 10 dB below analog levels. But if the noise floor of the plant was insufficient, clearly when all analog channels are replaced by digital carriers, the resulting digital carriers can be transmitted at higher levels due to the laser power margin freed up by removing the analog channels. And if only a subset of analog channels were boosted in order to support higher order modulation, the effect on overall plant balancing would be minimal.

The CTB and CSO issues can be dealt with using many of the techniques described in [4], examples of which include additional interleaving and/or coding, better equalization, and also offsetting higher order QAM frequencies to avoid the strongest CTB and CSO ‘tones’.

The fact that modern transmitter and receiver technology has mitigated many of these issues is born out by a field test of higher order QAM on the downstream, described in a subsequent section below. Further tests are planned.

Benefits of Downstream Improvements

While 25% improvement from higher order QAM and 10-40% from statistical multiplexing may not sound like drastic improvements in downstream capacity, consider the benefits for an application that is currently in the headlines for cable operators: HDTV. Currently in a 6 MHz downstream channel using 256 QAM, cable operators can deliver 2 High Definition (HD) channels without degradation, 3 HD channels using statistical multiplexing and allowing a slight degradation in quality. If the channel width were doubled to 12 MHz, then a 40% statistical multiplexing gain would mean that 5 channels could be delivered in 12 MHz with no degradation, and a 24 MHz wide downstream channel could deliver 11 HD channels.

But now consider increasing the order of QAM to 1024 on the downstream. The additional 25% raw capacity plus an additional statistical multiplexing gain leads to 15-16 HD channels in 24 MHz of RF bandwidth. This translates to 4 HD channels per 6 MHz of RF bandwidth with no degradation, or double the current number of channels. And this doubling would also apply roughly to DOCSIS data downstreams, where users could double current download speeds as an effective counter to competition, or as a means of attracting small to medium businesses to cable modem service.

Note that as mentioned earlier, additional compression technologies such as MPEG AVC (also termed MPEG4 part 10 or ITU H.264) will further improve the bandwidth utilization of digital video. High definition video using AVC is projected to use between 3 and 7 Mbps, depending on the content type, with sports programming being one of the more difficult types. By comparison, current MPEG-2 HD transmissions typically use approximately 18 Mbps for the video stream. Higher data rates can be used with AVC for even better quality, and lower date rates can be used where some degradation is acceptable and for content that is relatively easy to compress. As a result, in general approximately 2.5 to 3 times as many AVC HD video streams can fit into the data rate previously occupied by MPEG-2 HD. Combining this with the previous example of 1024 QAM/quad channels, this would translate into 10-12 HD channels per 6 MHz of RF bandwidth.

The notion of HD video on demand becomes quite viable under such scenarios.

METHODS FOR INCREASING UPSTREAM CAPACITY

Higher Order QAM

Higher order QAM can also be used on cable upstreams, which means greater than 64 QAM TDMA or 128 QAM/TCM SCDMA can be transmitted. If for example, 256 QAM TDMA is used on the upstream, up to 33% additional capacity is provided by using 8 bits per symbol instead of 6. And this additional capacity can be provided in a completely compatible manner with existing, legacy cable modems, since the burst nature of upstream transmissions means that higher order QAM transmissions can be mixed with lower order QAM in the same manner as DOCSIS 2.0 transmissions are mixed with DOCSIS 1.x transmissions.

But the upstream must be robust to ingress, impulse, and thermal noise conditions. As it turns out, most CMTS vendors have included some form of proprietary ingress cancellation processing in their designs, and the addition of SCDMA and TCM to the DOCSIS 2.0 specification significantly improves the robustness to impulse noise as well as providing several dB more robustness to thermal noise.

Consequently, higher order QAM turns out to be quite viable even on today's upstreams. In the next section, field tests of higher order QAM on the upstream, including 256 QAM TDMA, are presented, where the higher order QAM was operated reliably even in the presence of three ingress signals.

More Efficient Small Packets

The burst nature of upstream transmissions leads to variation in the benefit of higher order QAM on the upstream, however. Since longer preambles are typically required when transmitting higher order QAM on the upstream, and the preambles of small packets can be a significant portion of overall packet duration on the upstream, the benefits of higher order QAM on small packets can be less than 33%. Note that on medium and large packets, which account for the majority of bandwidth consumed on upstreams without VOIP service, the preamble is such a small fraction of the packet duration that 256 QAM provide a 33% improvement in bandwidth utilization. But especially for small packets using the conventional TDMA approach, the improvement can be less than 10%.

Since transmitting VOIP packets using compressed voice will significantly increase the number of small packets on the upstream, methods of improving the efficiency of small data and voice packets will be required. Several methods are available as extensions to DOCSIS 2.0. First, using synchronous SCDMA, instead of the TDMA currently in use, which is quasi-synchronous, permits reduction of the preamble of small packets without degradation in robustness. The synchronous mode is required to maintain code orthogonality in SCDMA mode, but has the added benefit of reducing the preamble overhead on small packets. For example, a 20-30% reduction in packet size can be obtained for highly compressed voice packets using synchronous transport, depending on the specific burst profile that is in use.

But the packet payload itself can also be reduced. A technique known as dynamic payload header suppression (DPHS), which extends the current fixed payload header suppression scheme of DOCSIS in a simple manner, can be used to reduce the header of small packets to the point where the packet duration is about a third of the original duration for small data packets such as TCP ACKs. This translates into three times the bandwidth utilization of small packets, and when synchronous operation is added, small packets can be up to 4 times more efficient. Unlike schemes that only address TCP ACK packets, DPHS also applies to other data packets and to voice packets. On a data-only network, where only 12% of the bandwidth is consumed by small TCP ACK packets, DPHS can provide about 12% overall network bandwidth utilization improvement, while a TCP ACK-only technique would only provide about 8% capacity improvement. On a network with say 50% compressed voice and 50% data traffic, the results are 16% improvement for DPHS and 4% for ACK-only techniques, while on an all-voice network, DPHS can provide up to 25% improvement while ACK-only techniques provide no improvement.

Thus by combining higher order QAM with techniques to address small voice and data packets such as synchronous CDMA and DPHS, the overall bandwidth utilization on the upstream can be increased by up to 33% regardless of packet size. Synchronous mode of operation and DPHS have no impact on ingress robustness, and both can be applied to SCDMA mode in order to be robust to impulse noise. The fact that higher order QAM on the upstream is robust to ingress is born out by field tests described in the next section.

HIGH-ORDER QAM FIELD TESTS

1024 QAM Downstream Test

A test of 1024 QAM was performed on a live cable plant in Rogers Cablesystems that was well-maintained, but nonetheless had measurable levels of CTB and CSO. The test setup is shown in Figure 2 at the end of this paper. The 1024 QAM transmitter was located in the headend while the cable modem 1024 QAM receiver was located in a van and was connected to the cable plant in a residential location. There were 3 active components (amplifiers) between the fiber node and the CM 1024 QAM receiver. No degradation to existing adjacent 256 QAM carriers resulted from the test.

First, a check of 256 QAM operation was made using transmit power levels that were identical to those used by current digital transmitters, and no errors in transmitted packets were observed. Next, since the SNR appeared to be sufficient for 1024 QAM, the modulation was increased to 1024 QAM using the same power level. Errors were detected, and thus the transmit power level was increased by 6 dB, and the system rebalanced, with a resulting SNR of about 36 dB. The majority of errors disappeared, however occasional errors were seen at random times which thus could not be ascribed to CTB and/or CSO cresting as described in [4] since they were not periodic. Possible causes include hardware and software issues in the prototype system used.

Further tests of higher order downstream modulation are in process and may be reported at the NCTA National Show.

256 QAM Upstream Test

The same prototype system was used to test high order QAM on the upstream. In this case, a range of upstream frequencies was made available for testing, some of which included up to 3 ingressors. The range was small however, hence lower symbol rates had to be used in the test in order to compare ingress free operation to operation in the presence of ingress. First, 64 QAM operation (the current maximum TDMA mode for DOCSIS 2.0) was validated in RF spectrum that was free of ingress using an 800 kHz wide signal. Next, the signal was moved so that three ingressors were present and using ingress cancellation processing, 64 QAM operated reliably with less than 0.01% packet error rate (PER).

Next, higher order QAM was tested, both with and without ingress present. 128 and 256 QAM TDMA were seen to operate reliably with less than 0.1% PER when no ingress was present. The 256 QAM signal was then moved to a frequency where 3 ingressors were present and with the ingress cancellor disabled. The PER rose to 96%, however when the ingress cancellor was engaged, the PER dropped to less than 1% PER.

CONCLUSIONS

The deployment of high definition TV will challenge cable operators to find new ways to expand their downstream bandwidth. The techniques presented here have the potential to double the number of HD channels, and when combined with emerging video compression schemes, can provide up to three times the number of HDTV channels in a 6 MHz RF downstream channel.

Wider channels can and have been implemented in current silicon for cable technology. The fact that error-free operation could be achieved at 36 dB SNR on real cable plants confirms the viability of 1024 QAM as a downstream modulation technique as well.

On the upstream, higher order modulation, when combined with techniques such as dynamic payload header suppression, provides robust and reliable data return service to residential customers that provides up to 33% improvement in bandwidth utilization, even on plants with ingress present on the upstream. In particular, 256 QAM on the upstream was shown to be ready for deployment in today's cable plants.

REFERENCES

[1] See www.replaytv.com
[2] "System-level capacity and QoS in DOCSIS1.1 upstream," C. C. Lee and Jérôme Bertorelle, SCTE Emerging Technologies Conference, January, 2002
[3] , "FastChannel: A higher-speed cable data service," B. N. Desai, Z. Jiang, N. K. Shankaranarayanan, D. Shur, A. Smiljanic, T. J. Totland, J. van der Merwe and S. L. Woodward, SCTE Emerging Technologies Conference, January, 2002
[4] "Statistical properties of composite distortions in HFC systems and their effects on digital channels," R. D. Katznelson, NCTA Cable 2002, May 2002

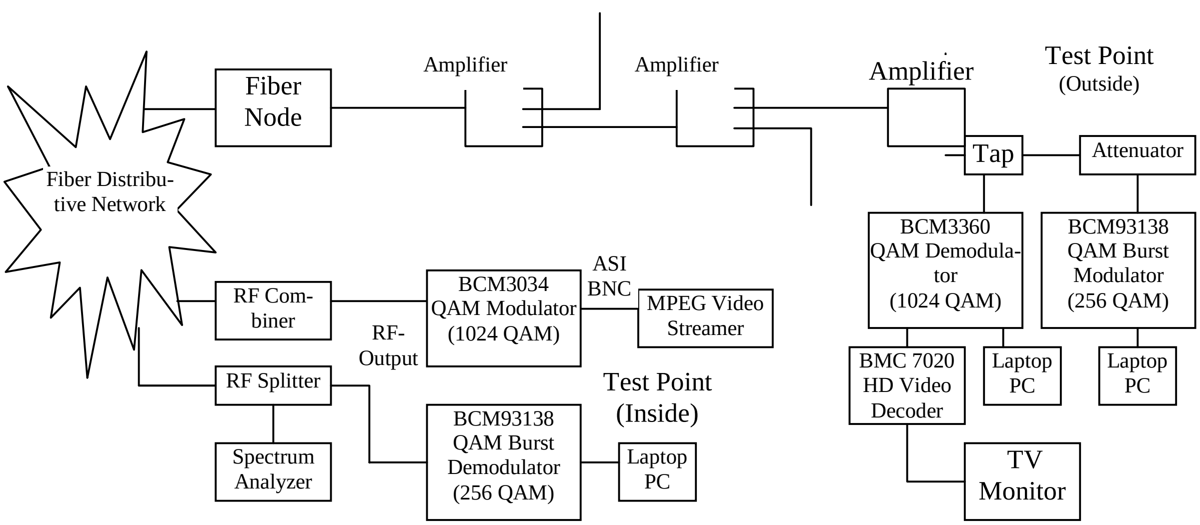


Figure 2. Field Test Setup

MODELING THE SCALING PROPERTIES OF VIDEO ON DEMAND ACCESS NETWORKS: SIMULATED TRAFFIC AND WORKLOAD ANALYSIS

Junseok Hwang, Srinivasan Nallasivan

ABSTRACT

Several different approaches –Pay per View, Near Video on Demand (NVOD) and Video on Demand (VOD) have been used to deliver video services to customers over cable networks, and a variety of network architectures have been proposed for VOD. This paper will model the performance characteristics of different VOD architectures and pay special attention to their scaling properties. To observe fundamental video stream traffic characteristics and the scalability of servers and the transmission infrastructure, we propose to perform simulation experiments for various VOD architectures to reveal which bottlenecks were the most serious. Different VOD architectures assume different locations or types of bottlenecks. Sensitivity analysis will be conducted by changing the values of various inputs (including technical ones such as headend locations, content distribution and streaming mechanisms). Simulation is done using different load balancing scenarios such as server load, round robin and the scalability issues are discussed by using server caching at the local hubs. Failure mode recovery analysis is also conducted as one of the scenarios to study the fault tolerance issues in VOD networks.

INTRODUCTION

Internet broadcasting and streaming contents has recently been attracting a great deal of attention, despite their inadequate content quality. The demand for such services is projected to continue to increase in the near future, and streaming contents are expected to play a major role among applications for the next-generation Internet. On the other hand, digital broadcasting and compressed audio/video such as DVD (MPEG2) and Video on demand are considered high-quality services and have become increasingly popular at home. Distribution of On-demand digital content is one of the major issues that need to be solved with a tradeoff between bandwidth and customer satisfaction. Cable operators have been constantly battling around to provide the best services to their subscribers by cost efficient, scalable and suitable technology to support these high bandwidth applications. This paper will discuss the scalability issues of various architectures by various simulation experiments considering load balancing on the head end and suitable server caching at the distribution end.

TECHNOLOGY OVERVIEW

Architecture Studies On Performance And Scalability

There are three main common architectures that are being used to locate the video servers and edge devices. They can be a fully distributed, fully centralized and partially centralized. One of the key issues that cable operators are still working through is where to locate the servers in their networks. They can choose to locate the bulk of their servers in the headed or in the hubs (which are closer to homes). The architecture must be capable of providing a good scalable, flexible model and it should also support high efficient bandwidth applications at a low operational cost per stream.

Fully distributed architecture involves installation of servers at the hub. This approach greatly helps to reduce the transportation cost, but since there is a duplication of content in the hubs, it increases the storage cost and it becomes difficult to manage the network thereby increasing the operational cost. Distributed architecture involves distribution of streaming transport in QAM/RF channels which is highly bandwidth demanding [1] and requires efficient use of bandwidth at HFC network .

Fully centralized architecture involves use of a one-server farm and other edge devices at the head-end. It overcomes the drawbacks of the distributed architecture by providing a low storage and operational cost, but it requires high bandwidth transportation between headend and the hubs. This architecture is not suitable for larger distances exceeding 25 kilometers [2].

Partially centralized architecture involves using the video servers at the headend and edge devices at the hubs. This architecture overcomes the drawback of both centralized and distributed architecture and it can be effectively utilized by using a Gigabit Ethernet backbone thereby providing long distance transport, increased carrying capacity and providing a flexible architecture.
Most cable operators have not yet decided on one scheme or the other, with many using different schemes in different markets.

Components Of Video On Demand

a.Vod Servers

VOD servers host large volumes of digital content supporting MPEG compression format. VOD servers encapsulate individual MPEG streams as single program transport mechanisms (SPTS) or into multiple program transport streams (MPTS). These SPTS or MPTS are mapped into ASI/ATM/Gbe/packet ring. IP-based servers are used as storage servers as IP takes advantage over other servers in using MPEG-2 over IP, thereby reducing the cost significantly in the system. Gigabit Ethernet interface helps in increasing the throughput per rack unit and the no of streams that can be encapsulated can be calculated as available bit rate to the total bit rate supported by Gigabit Ethernet. [3]

b.Edge Qam Nodes

QAM devices convert VOD server output (MPEG-2) to coax channels (6-8 MHz). It initally receives the MPEG-2 video and it re-stamps the packet that were delayed by the jitter and re-routes the packet to appropriate destinations. Its main role is in fixing up the jitter introduced due to the encapsulation in the network.[2]

c.Transport Network

The video content is distributed from headends to hubs. The transport architecture may be ranging from ATM or IP over Giga bit ethernet or IP cloud. The resilient packet can also be used to provide redundancy over the circuit in transmitting the video traffic incase of a ring failure or central /remote node headend failure.

d.Setup-Top Boxes

Set-top receivers at the customers premise acts as client nodes to VOD servers and terminate QAM signals to extract incoming VOD streams.

RELATED WORK ON DESIGN CONSIDERATION FOR SCALABILITY ISSUES

Scheduling Disk Issues

Real-time constraints make traditional disk scheduling algorithm, such as first come

first serve, short seek time first, and scan, inappropriate for VOD. Studies on VOD networks [6] suggest that two scheduling algorithms can be used for real time scheduling.

The best-known algorithm for real-time scheduling of tasks with deadlines is the *earliest deadline first algorithm (EDF)*. The media block with the earliest deadline is fetched first. The disadvantage of this algorithm is excessive seeks and poor utilization of the server's resource. [6]

Under round-based algorithms, a server serves all streams in units of round. During each round, the server retrieves a certain number of blocks for each stream. Since MPEG-2 results in variable-bit-rate compressed streams, the number of blocks that must be retrieved for each client in each round will vary according to the compression ratio achieved for each block.

A simple scheme that retrieves the same number of blocks for each stream (generally referred to as a round robin algorithm) is inefficient since the maximum playback rate among all streams will dictate the number of blocks to read. This results in streams with smaller playback rates retrieving more data blocks than needed in each round. This may overflow some clients' buffer as well as decrease the capacity of the server. Consequently, more clients can be accommodated by reducing the number of data blocks retrieved per service round for streams with lower playback rate.

The Placement Scheme decides the cluster size and stores video files across all clusters and verifies that a proposed placement scheme meets the placement requirements. It is an important factor for load balancing on servers. David Du [7] suggested that the placement scheme has to satisfy the following two requirements.

It is necessary to include video data from each video file in a cluster. This is because that the types of video files requested by retrieval processes are unpredictable. Sub requests within a service cycle may read video data from any video files available in the server.

The continuity of data block for each video file should be maintained between clusters. All data blocks should be stored within one cluster and their corresponding next data blocks should also be stored in a cluster range. If this method is not followed, after serving current sub requests, the following requests in next service cycle will read data blocks not from a cluster range**.**

Load Balancing

Whenever a load balancer receives a packet from the client machine, it must choose an appropriate server to handle the request. Load balancer will use the policy to determine as which server is appropriate.

Popular load balancing methods include:

Random: The load balancer chooses one of the candidate servers at random.

Round-Robin: Round robin policy is a method of managing server congestion by distributing connection loads across multiple servers. The load balancer cycles through the list of candidate servers depending on the selection weight specified. It can be classified as

Server Load: The load balancer chooses the candidate server with the lowest CPU load among all the servers

No of connections: The load balancer checks the server for the number of connections. When a new request is made, the load balancer checks and makes the connection with the server with least number of connections.

SIMULATION MODEL AND ANALYSIS

This section explains the various simulation models and the assumptions done for conducting our study.

Requirement Assumptions

In order to model the VOD architecture properly,we use the following requirement assumptions in our study.

- High-speed connection exist betweent the headend and the hubs which helps to avoid delay thereby reducing the packet loss rate and minimum latency in the network

- Connection oriented transfer is necessary for the timely arrival of packets and to provide high level of QoS.
- The data stored in the servers must be efficiently managed. This aspect deals with disk scheduling, data placement schemes and cluster management. Servers with local buffers are necessary for transfer of data without delay and jitters.
- ATM backbone is used for simulation analysis between the headend and the edge devices. The other alternatives could be 10 Gigabit Ethernet or resilient packet ring [3]
- The end-to-end delay measured in the network must be minimum for efficient transport of high bandwidth streaming applications
- Jitter is one of the factors affecting the quality of service.A smaller jitter provides with a high quality picture to the customer
- The utilization in the network is proportional to the number of connection that the server supports.The utilization must be initally low in order to support multiple connections in the future therby increasing the scalability demands of the network.

Model Description

The goal of this simulation is to evaluate the end-to end delays of a VOD network using load balancing and server caching techniques.We use a ATM cloud as the backbone for the network and high speed 1000 basex links are used to connect the headend servers to the headend gateway.These are the links that carry the high speed streaming MPEG-2 streams from the headend to the headend gateway. OC3 links are used to connect the headend and edge devices gateway. Three IP based servers were used. TheMPEG-2 streams can be encapsulated as SPTS so that each stream can be used to transport audio,video and sent to any desired IP destination at the lowest cost.GBE is used as a single GBE can provide a througput of 1000mbps.

CISCO 12016 Gigabit switch router is used as the gateway and some of the key features of this model are

- An IP forwarding rate of 60,000,000 packets/sec

- The router model implements a "store and forward" type of switching methodology.

Multiple GBE server outputs are aggregated into one GBE output and a payload of 900Mbps is recognized out of the GbE switch. The atm32_cloud node model represents a cloud through which ATM traffic can be modeled using 32 input/output physical links. Bandwidth management is highly critical in ATM networks as a delay of few milli seconds in a highly congested network can cause the cells to be dropped and lost. This results in retransmission of packets and it might compound congestion [6]. (Although voice and video are not retransmitted,this may cause a degradation in the performance of the enitre network.)

Some of the other parameters that are being used for the network are:

- Frame interarrival time information is assumed as 15 frames/sec.
- Frame size information is assumed as 128*240 pixels.

Type of service is assumed as high quality streaming multimedia. Simulataneous traffic generation is there in network until the end of the profile.Traffic is generated exponentially using a mean factor of 30.

Simulation Scenarios

Various simulation experiments were conducted and the performance of the VOD network under different test conditions were observed.There was a comparison of the different VOD architectures and the end-to end-delay, traffic of the network were analysed under two different scenarios namely load balancing and Server caching.There was a failure mode analysis of the network and the effects of the load balancer on the failure mode recovery are discussed below.

Scenario 1: Comparison of traffic in the network

Two different architectures namely the centralized and distributed architectures were simulated for heavy traffic conditions . Distributed architecture is preferred over centralized architecture in terms of all reduced congestion and delay, as the servers are located near the hubs, but management of media storage is a problem. This is the only disadvantage as the servers are located internally at different places unlike centralized architecture where servers are located at a single point.

Figure 2 shows the voice traffic sent for both the architectures. It is shown that voice traffic sent for distributed architecture increases linearly with time and voice traffic sent for centralized architecture remains steady after reaching certain time period. This shows that the distributed architecture has high throughput relatively compared to centralized architecture, which is evident from Figure 2.

Scenario 2: Load Balancing and Server Caching

a. No Load Balancing

In this scenario,no load balancing is applied on the servers and performance analysis of the network is done and the end-to end delays are measured. The end to-to-end delays are found to be relaitvely high when compared to other scenarios.The end-to end delay for this network using no loadbalancing is around 0.2 secs on an average and it is considered to be higher compared to load balanced and server cached scenario.

b.Load Balancing

In this scenario,a load balancer is used to control the load acting on the three servers.There are three policies that are being used .

The are classified as Round Robin,Server Load and the no of connections.The first server is loaded with a selection weigth of 10 and the other two servers are chosen with a selection weight of 5.The first server is loaded twice as that of other two servers and the end –to end-delay was found to be greatly reduced to 0.009 secs from 0.12 secs in a scenario that doesnot use load balancing.

c.<u>Server Caching</u>

In this scenario,an additional server is duplicated with contents of one of the head-end servers and is installed at the hub,and the end-to-end delays were analysed for this scenario.The end-to-end delay was found to increase,but the traffic in the network was decreased considerably as there was data duplication in the hubs.

d.<u>Failure Recovery</u>

Failure recovery analysis was conducted in VOD networks.One of the servers was set to fail while the servers are handling the high traffic in the network.The load balancer helps to isolate the server and it distributes the load among the other two servers in the networks.Towards the end of the simulation,the server recovers and it again couples with the other two servers to handle the load.The end-to end delays and the traffic of the network were analyzed and the graphs were obtained in Figure 6 .

When the server three fails,the end-to end delay increases and its 24.9 secs at the point of failure and it gradually decreases towards the end of the simulation as the server three recovers.Thetraffic in the network can be seen as decrease during the period of failure and gradually increase towards the end of simulation.The traffic dropped was analyzed in this scenario and it was found that there was no packet loss in the network.

CONCLUSION

This paper provides a good insight into the scalability isssues on VOD networks.We discuss the scalability issues in different perspectives as how disk scheduling and data placement can affect the performance of VOD networks. Three alogorithms namely EDF, round-based algorithm and QMPS are discussed under the scheduling disk issues. Data placement decides about the admission scheduling schemes and discusses the cluster placement issues. Our simulation study shows the overall packet-end-to-end delay in the VOD network under different scenarios. This shows that load balancing with server load policy can definitely be a good suggestion for high performance in the network. Delay is one of the major factors in high speed network and this shows that load balancing can certainly help in reducing the end-to-end delay of the network and it can also help with fault tolerance capabilities, if there is a problem of server failure during operation in the network. Server caching can be one of the possible solutions to reduce the traffic in the network, if streaming media is sent to longer distances,but it always has a drawback on the management and storage associated with the duplication of data.

<u>Future work</u>

Resilient Packet Ring (RPR) is a new transport standard and can be used for more efficient multiprotocol transport. This standard combines the best attributes of SONET,WDM,and GbE to provide with high quality of streaming content with redundant links. However this standard is at the initial stages of development and it could be expected to be more expensive than the other options that were discussed.

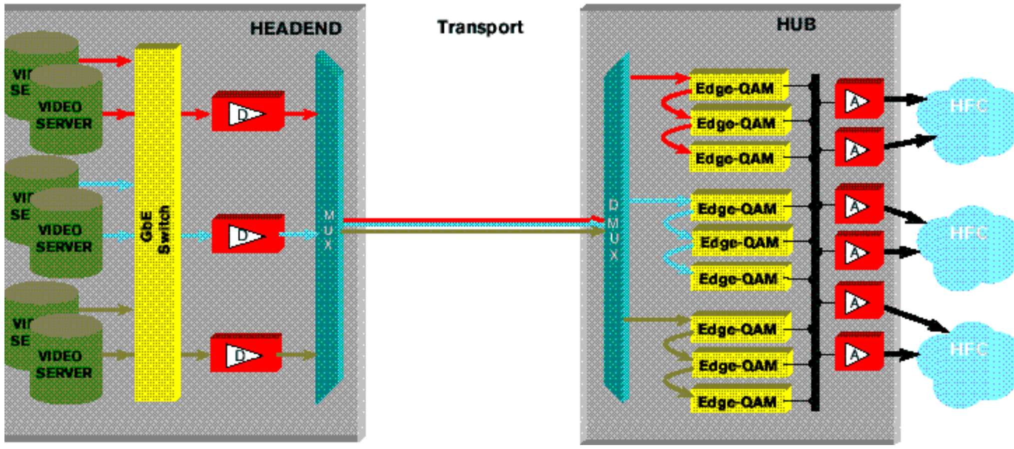


Figure 1.Components of VOD network [1]

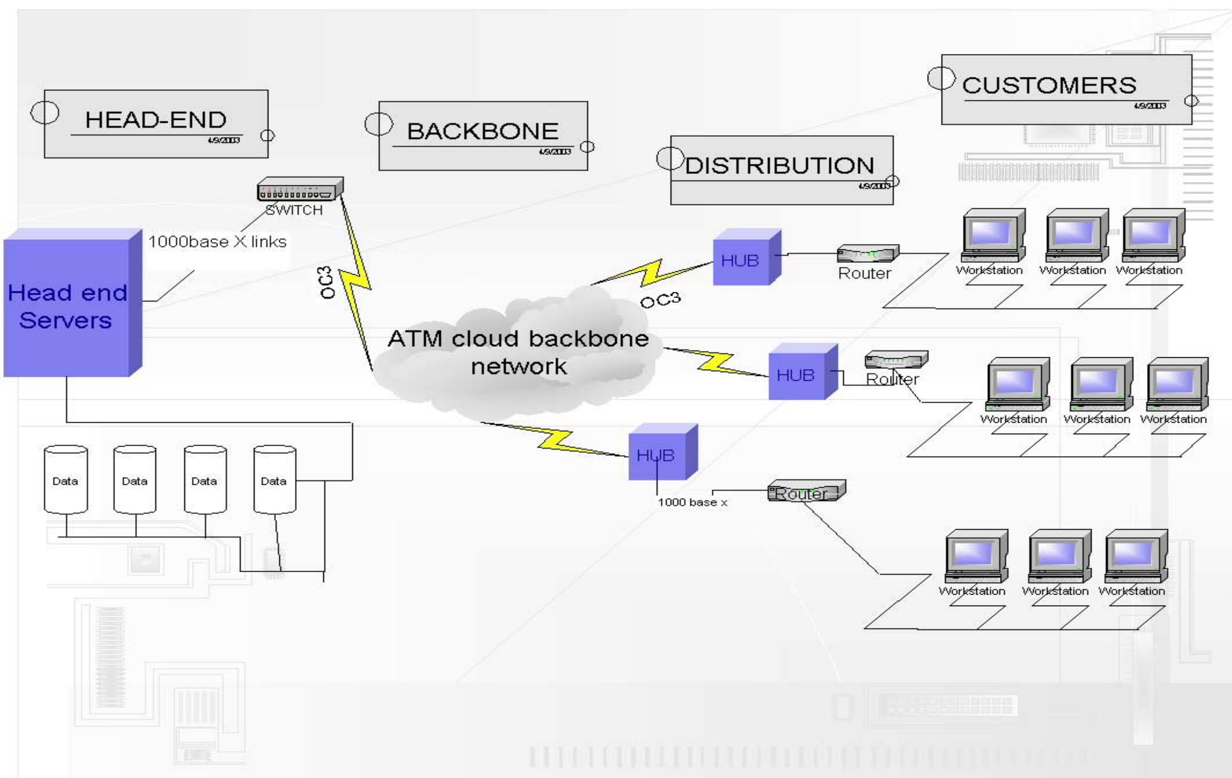


Figure 2 : Video on demand network - Topology

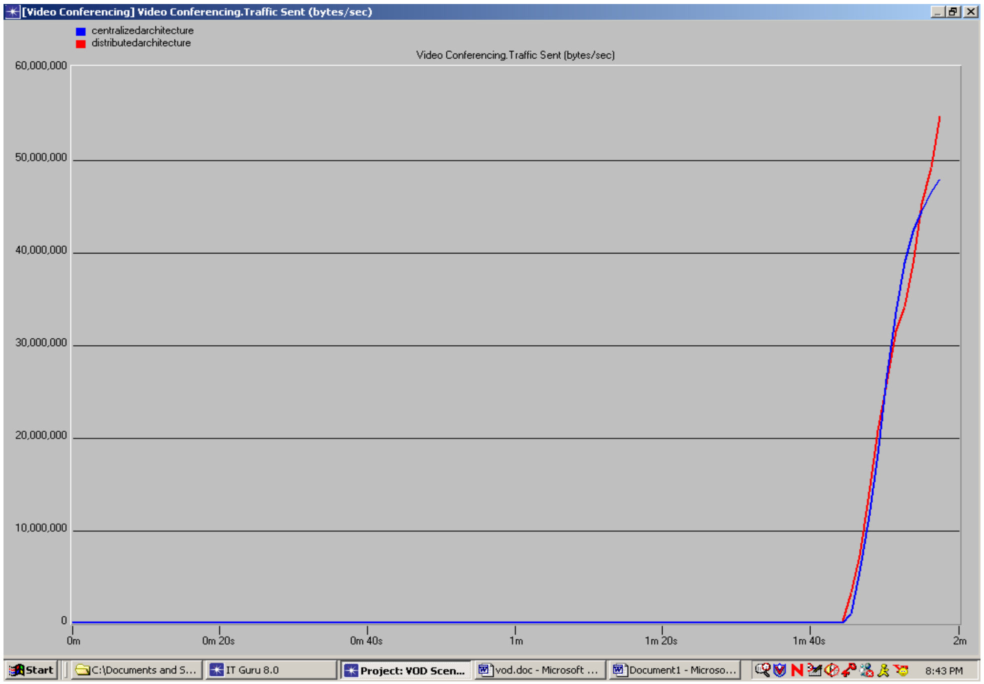


Figure 2 : Comparison of voice traffic received

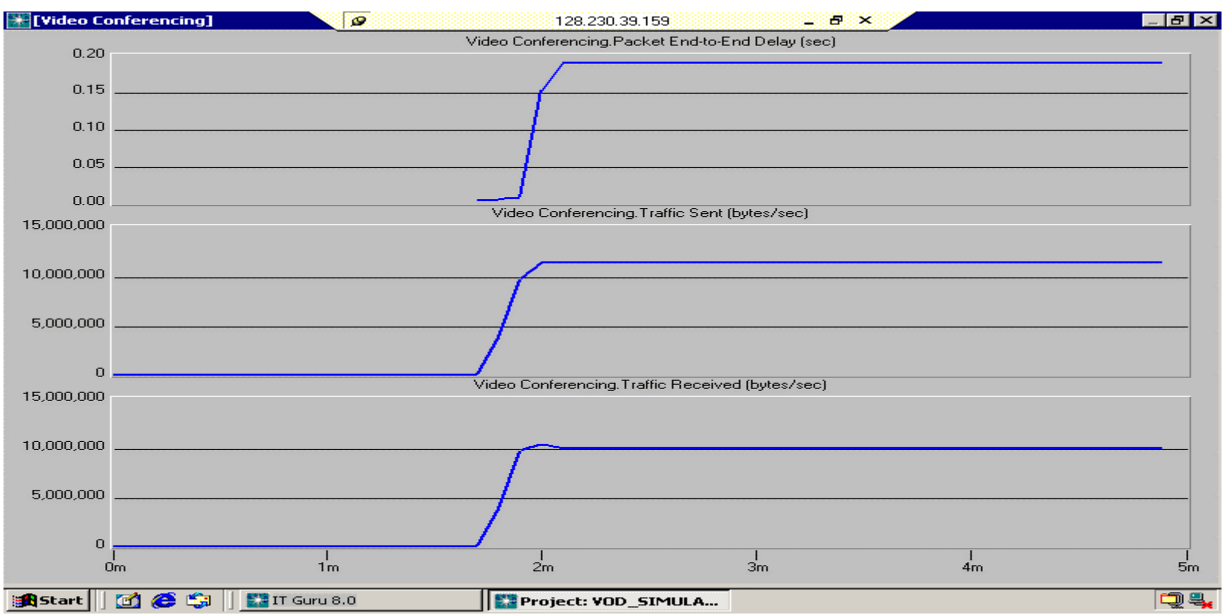


Figure 3: No Load Balancing Scenario

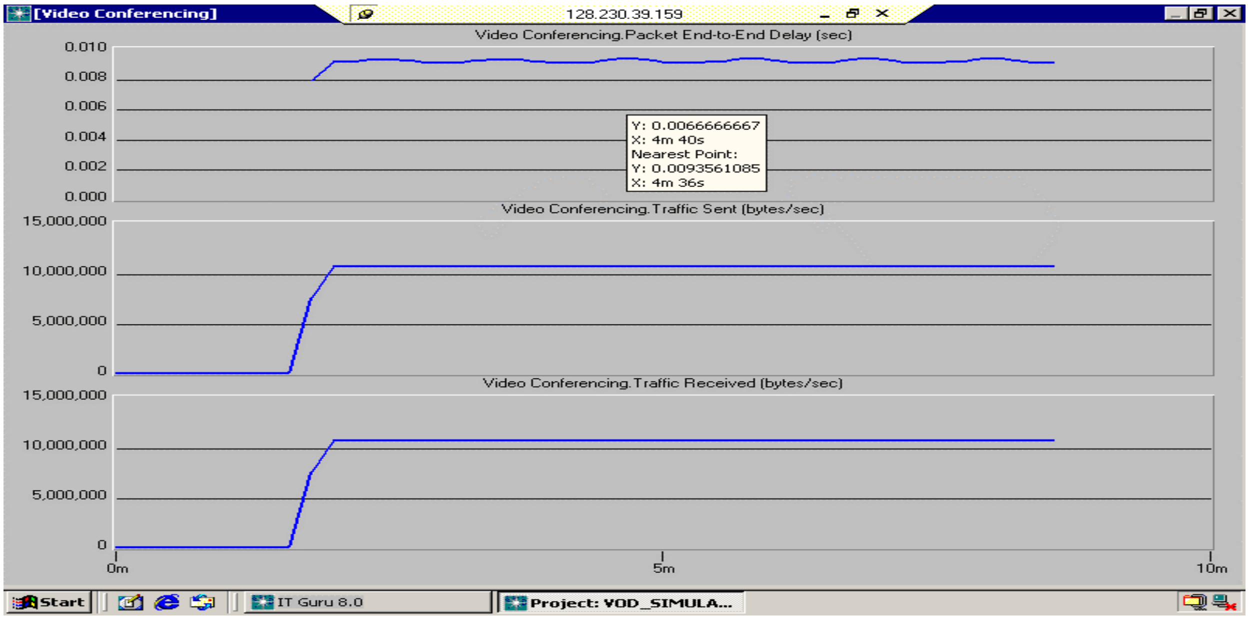


Figure 4: Load Balancing Scenario

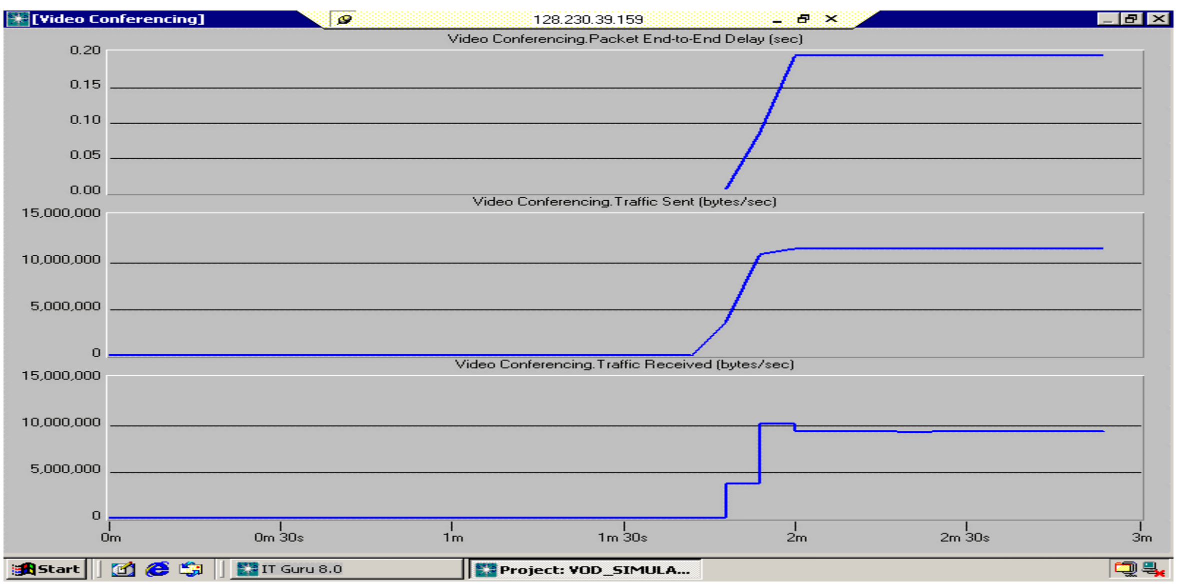


Figure 5: Server Caching

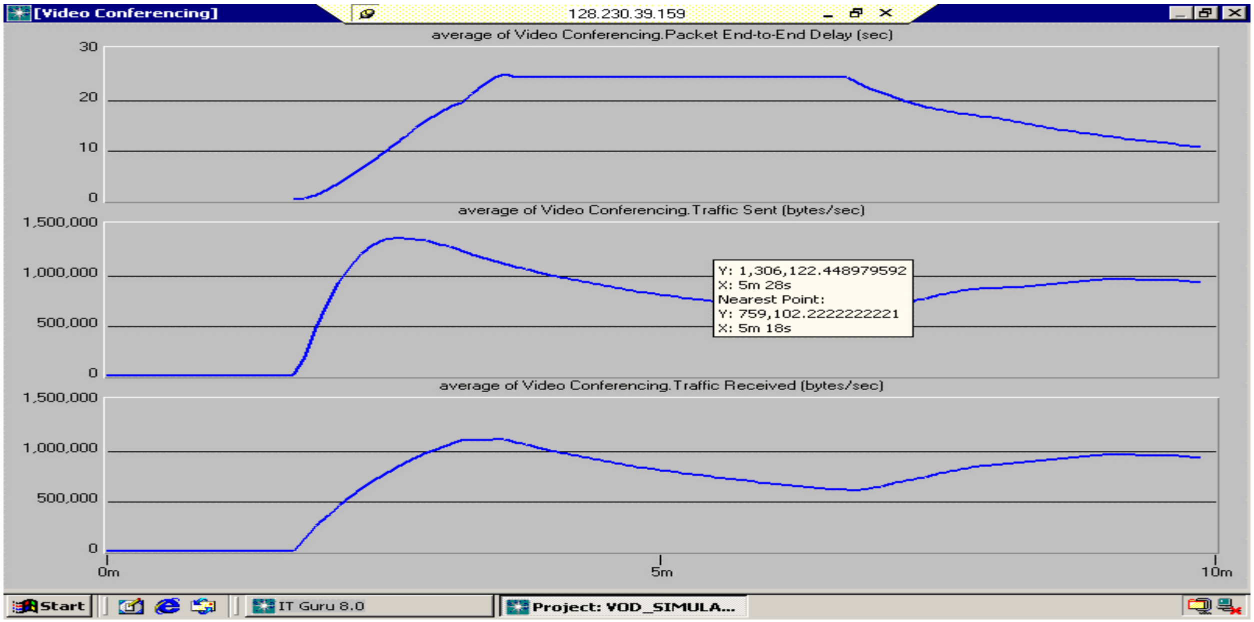


Figure 6 :Failure Recovery Scenario

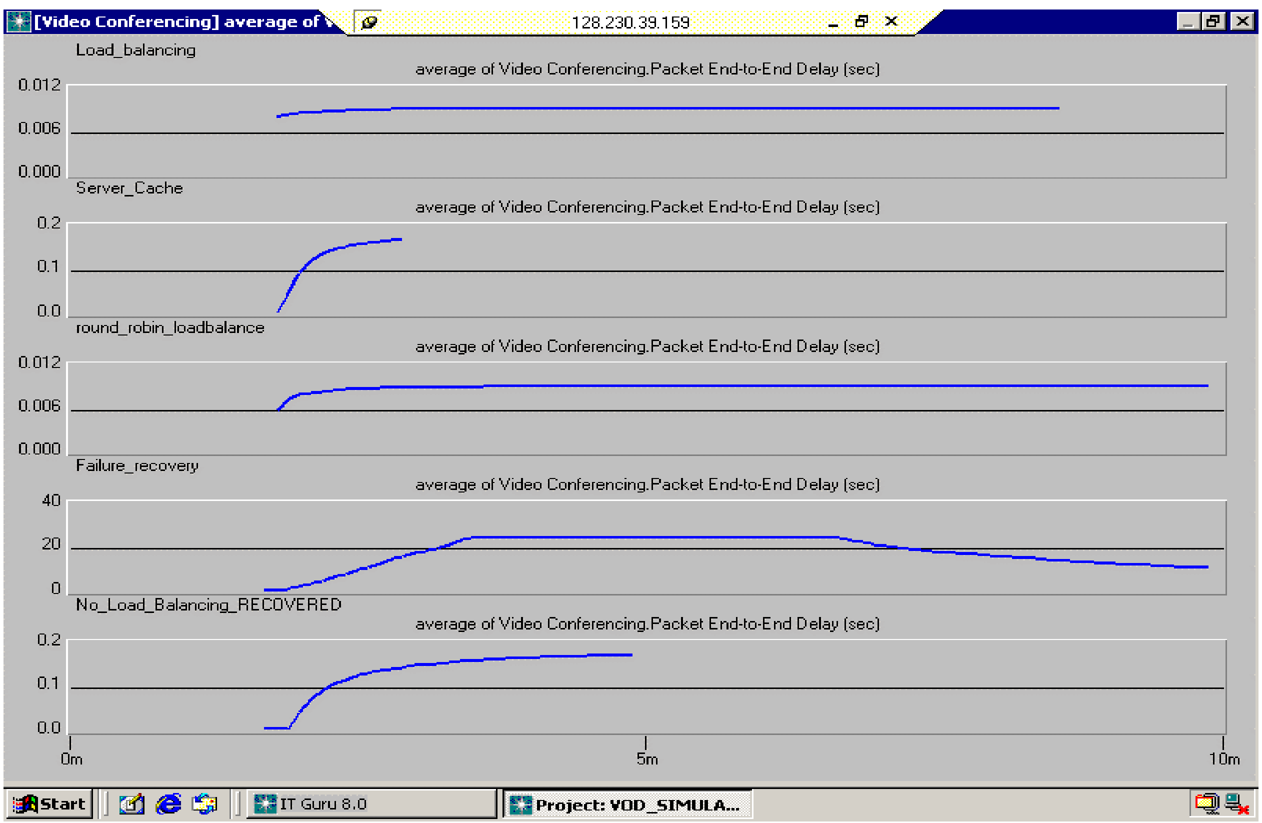


Figure 7 :Overall Scenario Summary

REFERENCE

[1] Fernando Amendola and Dr.Eric Schweitzer, Scalable VOD over DWDM Gigabit Etherent –Recent Deployment Experiences

[2] White Paper on Network and Access Architecture for On-Demand Cable Television . –Harmonic Inc March 2002

[3] White Paper on GigabitEthernet –over – DWDM Transport Architetures for On-Demand Services –Harmonic Inc September 2002

[4] Y.Doganat and A.Tantawi, Making a Cost Effective video server, IEEE Multimedia 1995

[5] Opnet Modeler and User Manuals

[6] Kyung Oh Lee,Heon y.Yeom,An Effective Admission Control Mechanism for Variable Bit-Rate Video Streams.Seoul National University,Korea

[7] Horng-Juing Lee and David Du Distributed Multimedia Center: Cluster Placement :A Data placement Scheme for a Mass Storage System on Storage System of Video-On-Demand Servers.

MULTI-ROOM DVR: A MULTI-FACETED SOLUTION FOR CABLE OPERATORS

Dave Clark
Scientific Atlanta, Inc.

Abstract

Now that your subscribers have come to depend on a Digital Video Recorder (DVR) for their TV watching experience, how do they enjoy that experience on other TVs in their home? Why limit DVR -- your subscribers' ability to watch what they want when they want -- to just one room in the house?

The Multi-Room DVR solution currently being developed by Scientific-Atlanta will give DVR users the ability to watch content recorded on their DVR set-top from other rooms in the house. As this paper illustrates, such a solution will give subscribers more flexibility in their TV viewing routine as it offers an attractive business case for cable operators to expand on their DVR success.

WHY MULTI-ROOM DVR?

The first question that must be answered when looking at Multi-Room DVR is, "Why?" What problem does Multi-Room address? Isn't DVR itself just ramping up? Is this just a solution looking for a problem?

Available research indicates that the need is real. In a recent survey of people who have Scientific-Atlanta's Explorer 8000 Home Entertainment Servers in their home, 67% indicated they were 'very interested' in being able to access a DVR from other TVs in their home. The survey also showed altered TV viewing habits with 65% saying they watch more TV programs now than before they had a DVR. In this same survey, 81% of the respondents said that "TV is more fun than it was before DVR." Clearly, DVR is a convenience that changes the way people watch video content.

Once consumers become accustomed to a DVR on their primary TV, they don't want to watch TV *without* DVR functionality. From that perspective, the ability to, say, start watching a program in the living room then watch the last half of it in the bedroom makes sense. That's one of the reasons why people have more than one TV in their homes in the first place. Multi-Room DVR moves the consumer experience of DVR to this highly desired, next level.

ENTERTAINMENT NETWORKING

While home networking is becoming a topic of intense discussion and analysis, it is important to distinguish between the two types of services that make up "home networking." "Data networking" includes multiple PC's and high speed data. "Entertainment networking" focuses on distributed video content and music. Multi-Room DVR is an important first step toward Entertainment networking and a successful Multi-Room DVR solution will need to be able to evolve as entertainment networking evolves.

Key Drivers for Entertainment Networking

Many factors must be considered and addressed when evaluating a solution for entertainment networking (including Multi-Room DVR).

The first factor to consider is whether or not average cable subscribers can understand and use an entertainment network. Will they embrace it or will it become just another novelty device whose appeal quickly comes and goes? If the consumer experience doesn't reflect a simple, easy to use solution, then the Multi-Room DVR concept itself becomes a moot point.

As important as consumers are to a successful Multi-Room DVR solution, other groups must also be considered. Content owners (e.g., programmers, movie studios, etc.) have a very real interest in how a solution is implemented. They want to be assured that the solution fosters a safe, secure environment that keeps content in the home, not illegally copied, for example, on the Internet. They need to know that the quality of their content will not be degraded. While keeping the content secure and safe, however, the system must allow the content to be able to be used freely by consumers, within the confines of their own homes.

Cost to both consumers and cable operators is another key factor to consider. Are there adequate revenue opportunities to offset costs associated with a Multi-Room DVR solution? If the cable operator establishes a Multi-Room DVR set-up, will revenue generating set-tops be displaced? Or will the operator be able to generate additional revenues throughout the home? And what about operational support costs? Does the DVR solution cause an increase in calls to your support center? Or will the solution be easy for consumers to embrace and operators to maintain? These are all important questions to consider when analyzing a potential solution.

ONE APPROACH TO MULTI-ROOM DVR

Scientific-Atlanta's Multi-Room DVR solution, currently under development, addresses all of these issues. From the beginning, the main development focus has been to provide a simple, low-cost solution that can be ready to release to the market in time for operators to take advantage of it. The solution also is designed to leverage the success and strength of the deployed Explorer 8000 Home Entertainment Server platform.

Ease of Use, Security, Quality and Value

The "ease of use" issue was the first issue to tackle in developing an effective solution. This was addressed by making sure that the user interface experience on the client set-tops mirrored that of the DVR or server set-top. So, no matter where subscribers access recorded content, the same, familiar look and feel greets them.

The next issue addressed was that of secure content. The goal was to leverage the cable operator's ability to deliver secure, digitally encrypted content from the headend to the subscriber's home -- in other words, to turn an Explorer 8000 Home Entertainment Server into a mini-headend. This enables the Explorer 8000 DVR set-top box to digitally transmit encrypted content over the home's coaxial wiring to other (client) digital set-tops in the home. By providing a safe, secure path from a DVR set-top to other set-tops within the home, content providers have the same technology that they rely upon today when digital content is delivered to digital cable set-top boxes.

The next area to address in developing a Multi-Room DVR solution was content quality. We had to make sure that content would not be degraded when transmitted from room to room. The Scientific-Atlanta solution assures content integrity because the home server set-top communicates with the client set-tops using MPEG-2 digitally encrypted signals over a fully integrated digital network in the home. So, not only can one room watch a recorded program from the home server's hard drive, but up to three client set-tops and the home server set-top itself can all simultaneously watch any recorded program. The system allows all four viewers to pause, fast forward or rewind different programs – or the same program -- independently, without affecting the other viewers.

One of the strengths of Scientific-Atlanta's Multi-Room solution is enabling operators to use existing and deployed entry level set-tops (e.g., the Explorer 2100) as client set-tops, minimizing capital costs for the operator. Older digital set-tops that might be churned back to the cable operator's warehouse can now be given new life as client set-tops in a Multi-Room DVR system. This solution also provides additional revenue opportunities through other TV's in the home because the client set-tops are fully functional digital set-tops, not just 'slave' units to a home entertainment server. So, they can also handle any applications that the cable operator has deployed (e.g., Video-on-Demand, Pay-Per-View, and any others).

Business Case

The current and growing success of DVR products and interest in home networking provide the subscriber momentum that will fuel Multi-Room DVR migration. Market research shows that subscribers want to access DVR in multiple rooms – and they are willing to pay for this service. The Scientific-Atlanta solution gives cable operators the ability to compete with upcoming Multi-Room DVR solutions from DBS providers. It also assures content providers that the quality of their content will remain intact. Combine these facts with a low-cost implementation that heavily leverages existing hardware and it is easy to see that the type of Multi-Room DVR service described here offers clear value to subscribers, content providers and cable operators.

SUMMARY

Scientific-Atlanta's digital cable Multi-Room DVR solution provides the ability to share content in multiple rooms without the expense of a hard drive at each TV. This solution leverages proven digital cable technology and security, while building on the success of digital cable DVR set-top deployments. The Multi-Room DVR system provides a low cost, safe, secure, high-quality content delivery method to both currently deployed and previously deployed digital cable set-tops from Scientific-Atlanta. Multi-Room DVR positions cable operators to get in on the ground floor of entertainment networking opportunities. Subscribers, content providers and cable operators all stand poised to benefit.

NETWORK DESIGN FOR A MULTIPLICITY OF SERVICES

Ran Oz
S. V. Vasudevan
BigBand Networks, Inc.

Abstract

Cable plants were initially built as a single-purpose infrastructure for the one-way broadcast of analog television. During the last ten years, in order to take advantage of emerging service opportunities and to head off competitive threats, cable multiple system operators (MSOs) have broadened their offerings to an expanding range of services, including digital broadcast and high-speed data. This progress continues with current trends towards additional growing services such as HDTV and various flavors of cable telephony and video on demand (VOD).

There is an increasingly holistic realization of the cable plant as an integrated multi-service and multimedia infrastructure. This realization holds promise to enhance the efficiency, scalability, functionality and ease of rollout of services, through activities such as the proliferation of open standards, sharing optimized resources, and balancing offsetting characteristics of different services and media.

The key to achieving these benefits is to evolve the cable network from a collection of isolated silos of vertically integrated components supporting particular services, towards more inherent fitness for multiple services and media. Greater inclusion of abstraction layers allows more best-of-breed providers of particular aspects of functionality, including the extension of shared functionality across multiple media and multiple services.

LEGACY CABLE SERVICE METHODS

The architecture of the cable plant traces directly to its broadcast legacy. The very concept of six megahertz channelization in North America exists because that's the amount of spectrum that conventionally carries a single analog NTSC program. Historically, relatively few programs were offered to all subscribers in a cable system. The only service consideration was selective subscription by some subscribers to premium or pay-per-view offerings, for which scrambling techniques were devised.

Not much architectural sophistication was required for a plant built for the singular purpose of pushing out relatively few programs to all subscribers in systems that generally were no larger than a very few hundred thousand homes passed. Because the cable operators initially constructing these plants were generally entrepreneurial and debt-financed, there was an imperative to optimize the economics of plant construction for the task at hand, which subsequently motivated technological innovations in selected areas such as downlinks, amplifiers and taps.

Besides its economic streamlining for the task at hand, a significant historical advantage of cable networks relative to other networks connected broadly to homes (most notably telecommunications networks), is the absolute transmission capacity of the coaxial cable which can carry so many rich media programs. However, telecom-munications networks historically boast other distinctive aspects of functionality such as line powering, two-way transmission, and dedicated switching of content to particular subscribers.

New services beyond analog broadcast television have emerged on the cable platform over the past ten years, and continue to do so today. Each of these services goes through its evolutionary stages of proof-of-concept, economic validation, widespread implementation and scaling over time. Because in part of the fact that mere functional viability must be established first, leading early proponents tend to provide complete vertical integration of the technologies required to provide such services. A byproduct of such integrations is the complete autonomy of that service over a discrete count of six megahertz channels, sometimes with granularity to produce content for the same channel distinctively by node in the hybrid fiber/coaxial (HFC) architecture. The fixed allocation of channels and associated resources from cable's analog broadcast legacy is thus extended to services that exist within their own dedicated silos.

MULTI-SERVICE OPPORTUNITY AND COMPETITIVE IMPERATIVE

From time to time, the telecommunications industry has touted its development of digital subscriber line (DSL) technology as a potential delivery method for video. Also during the last decade, the rapid subscriber growth achieved by direct broadcast satellite (DBS) operators has brought on another very real threat to the traditional cable business franchise. With multiple industries proposing to provide broadband networked connections to subscriber homes carrying video, data and voice content, cable has had to fortify its positioning in the face of emerging competition.

To its credit, and true to its entrepreneurial roots, the cable industry has always looked at expanding multimedia and multi-service offerings as a bona fide business opportunity and not merely competitive reaction. The earliest forays into interactivity and on-demand consumption go back several decades. But with progress made in the satellite and telecommunications industries, as well as the greater application viability of services like VOD, the time has emerged for cable to realize its multimedia, multi-service potential for defensive as well as offensive reasons.

Beyond analog broadcasting, the first services that the cable industry has widely deployed are digital broadcasting and cable Internet access. These services represent directly competitive spaces with DBS and DSL (which has grown as a high speed data service, but has yet to materially provide a video alternative), respectively. Because the deployment of these two digital media services over cable required various enhancements to the cable plant, the last ten years have seen several profound cable industry undertakings, including the development of the HFC architecture (with its node-level multicast granularity), return path capability in the plant, and several widely embraced digital standards such as MPEG-2 transport.

Competing effectively with impressive subscriber growth in both digital broadcasting and cable modem service, cable is currently embarking upon further service expansion, highlighting areas where it can leverage the investments made in the plant, and where it is uniquely or best positioned for distinction. Such services include high definition television (HDTV) and various packagings of VOD including movies on demand (MOD), subscription VOD (SVOD), free VOD (FOD), networked personal video recording (NPVR) and long form advertising (LFA).

Consumer adoption trends of HDTV and VOD offerings are encouraging. At the same time, the cable industry prepares itself for the widespread rollout of telephony services, which have achieved encouraging popularity in their isolated current deployments. Furthermore, the long anticipated interactive television (iTV) space also maintains its promise as technologies are further refined, bolstered by popular indicators such as the success of direct response advertising and audience participation in various forms of reality television.

OPTIMIZING TECHNOLOGIES IN MULTIMEDIA, MULTI-SERVICE ENVIRONMENT

As these trends play themselves out, the cable plant that was initially constructed for one-way broadcast of analog video becomes a much more complicated beast. We increasingly see a single, yet multi-faceted network that carries combinations of video, voice and data content; broadcast and personalized sessions; passive and interactive consumption – among other variations in subscribers' engagement with media and services.

While voice, video, and data may seem like extremely dissimilar services, the composition and delivery of these services share some technical similarities. The general technical elements that are integrated to deliver a service include the source of the service, the switching of the correct content towards the correct destination, the physical transport of content from source to destination, and the media processing of content.

For example, a VOD service requires servers with video storage and transactional and billing applications as service sources, switching capabilities to assure that the right programs are directed to the right service groups corresponding with subscriber sessions, physical transport to get the content to the node (or, in the case of server distribution at hub sites, to propagate the content to storage for later playback), and media processing such as de-jitter, QAM modulation and upconversion.

Some of the earliest VOD offerings integrated all of these elements in a solution offered by a single vendor. This severely limited choices by MSOs in how particular elements of the service could be configured, cost-effective service expansion could be achieved, and how flexible headend architectures over time could be accomplished. In the years since the early VOD deployments, there has been an increasing disaggregation of functionality, to the benefit of specialists in particular elements of functionality, enabling a decoupled system that allows different vendors to provide the server, transport, and modulation functions for VOD installations.

Lack of complete standardization or broad protocol flexibility continues to hem in choices. For example, supported transport protocols are based largely on the output format of a server. But with time, operators are increasingly able to migrate from legacy ASI transport to Gigabit Ethernet transport between facilities, with easier abstraction from which VOD server is selected. And modulation equipment at the network edge will be able to accommodate for the chosen method of transport, be it legacy or contemporary.

RESOURCE SHARING ADVANTAGES

The rich variety of services and media increasingly carried on the cable plant presents opportunities to leverage technologies that can be optimized for their offsetting as well as their common characteristics. Yet, the one-way broadcast legacy of the cable plant by and large maintains its influence, and services are introduced in sub optimal fashions as a result.

One clear example of this legacy is the continuing channelization of spectrum and associated resources on a service-by-service basis. In the same manner that the local affiliate of a major broadcast network is allocated a six megahertz channel, so is such a swath of spectrum is provided on a 24-by-7 basis to approximately ten digital broadcast programs, or VOD capacity for the anticipated peak consumption among approximately 100 digital cable subscribers in a node or service group, or for shared cable modem access by several hundred subscribers to that service.

This means that spectrum, and its associated resources such as QAM modulators and upconverters, are permanently allocated to particular services regardless of those services' use at particular points in time. Because resources and spectrum are generally capable of fungibility across services under common standards, this situation is suboptimal unless demand for services is completely static or correlated.

For example, if a major news story breaks, there may be a lot of demand for high speed data and live broadcasting and relatively little VOD activity. In such a scenario, the operator runs the risk of underallocating high speed data resources, which could deny transactional revenues or frustrate customers (resulting in churn vulnerability). At the same time, capital is being wasted on underutilized resources dedicated to VOD.

Because all of this digital media traffic is based on common standards, such as MPEG-2, it is relatively easy to reap the advantages of dynamic allocation of bandwidth and associated resources in response to real-time demand. Thus a six megahertz channel and its QAM and upconverter can be carrying a heavy load of data services at the moment described above, and more VOD at other times, such as weekends and prime time or particular spikes for such services as when popular titles first become available. Because of the more efficient utilization, fewer overall resources are required. Alternative methods of accommodating scenarios of potential demand peaks through overprovisioning require exorbitant capital expenditures and drive horrendous bandwidth and resource utilization.

Bandwidth and resource sharing across cable services leverages the advantages of statistical multiplexing in which various traffic loads, unlikely to all simultaneously peak, ride together. The result is greater efficiency of shared resources and less likelihood of denial when particular services spike in demand.

Consider a hypothetical situation of services A, B, and C which, respectively, are permanently allocated channels 1-3, 4, and 5-6 in a conventional cable plant as indicated in the image on the left, in Figure 1. The bandwidth available to subscribers of a service is limited by the channels allocated to that application. The left image shows all applications at exactly peak capacity, however, because of the periodic shifts in application consumption patterns (ex: night vs. day, weekday vs. weekend) and short term irregularities that occur (ex: a highly anticipated report is published via Internet, a popular movie becomes available on VOD), in reality a fixed-allocation system cannot be managed to achieve anything close to such high utilization of bandwidth, and fulfillment of demand.

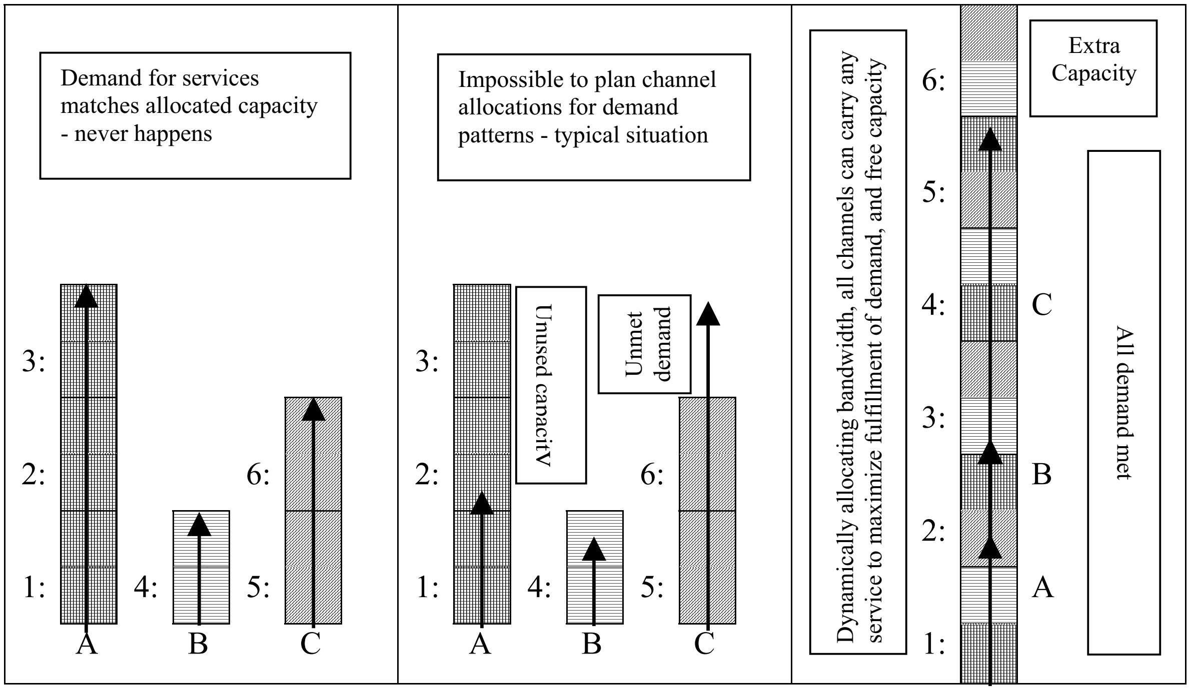


Figure 1: Dynamic allocation of bandwidth across services within channels.

The middle image shows a typical situation in which the demand for service A is less than half its three channels' capacity while demand for service C is well above the spectrum allotted. This results in lost revenues, angry customers, or both due to unfulfilled demand for service C, which occurs at a time when there is free bandwidth in channels 1-3.

By implementing dynamic bandwidth allocation, an MSO can achieve the situation in the right image in which assets are more flexibly managed as all channels, 1-6, can carry any of the services. The demand for A, B and C is combined and mapped onto this one big band. This fulfills all demand with even more capacity available, facilitating introduction of more revenue generating services.

Recent technological advantages have even introduced broadcast television to the pool of services dynamically sharing resources. As hundreds of simultaneous programs are made available to subscribers, it can be assured that not all of them would be watched at any particular point in time in any particular node. Instead of filling spectrum in all nodes with all broadcast programs all of the time, switched broadcast techniques can be utilized to dynamically respond to channel surfing and switch live programming only to the nodes where it is being actively watched, allowing broadcast programming to share spectrum and resources with other, interactive services.

When different services and media share the same bandwidth and resources, offsetting characteristics between them can be leveraged. For example, when video, a real time medium, shares bandwidth with data, which is time-sensitive but not necessarily real-time, peaks in video traffic can be accommodated by shifting data traffic into the troughs of video traffic. Another example is combining variable bit rate HDTV traffic with standard digital video, in which case rate shaping can be applied to the standard video when the HDTV requires more bandwidth, which assures the reliable and robust delivery of both services while maintaining HDTV quality. The combinations of various media within the same channels has the long term potential to enable new generations of services that richly combine media, such as multiplayer games that operate in conjunction with live video programming.

When plant resources in general, and bandwidth in particular, are dynamically and openly allocated across all services in response to real time demand, the processes of service launch and scaling become greatly alleviated. Traditionally, launching programs or services requires discrete determinations of which existing offerings must be stopped or reduced to accommodate spectrum. With open allocation, new services can be layered on, and if capacity contention arises, this can be addressed through implementation of intelligent implementation priorities by service, session or subscriber, or further bandwidth enhancements such as rate shaping to adapt video bit rates.

ARCHITECTURAL OPTIMIZATION

Many advantages are reaped with less vertical integration of services and more resource sharing among them. Specialization and innovation of functional components such as QAM, switching, or transport by particular vendors allows these vendors to hone best-of-breed deliverables in their areas of specialty.

Isolating and aggregating particular elements of functionality allows these elements to be designed for maximum economic scalability. And that scalability should be able to be achieved in an incremental fashion. This improvement, combined with decoupling of components, allows the operator to choose exactly how much of particular elements of functionality are required.

To exemplify the alternative, when systems are vertically integrated, storage may be combined with media processing for some services. Thus the operator's scaling of infrastructure is based on whichever is the bigger driver between total content provided (storage determination) or total session capacity (processing determination), and whichever component doesn't drive the installation is uneconomically overprovisioned as a result.

Decoupling functionality also enables the operator to optimize architectural configurations and not be constrained by vendor designs. In VOD, as an example, some operators express preference for centralized server consolidation, and some prefer to distribute servers to hub locations at the edge. In either case it is generally agreed that media processing such as QAM modulation should be located at the edge, to assure quality of content, with the most economic transport methodologies used to get content to the edge. By having independent QAM components that are separated from the server, this media processing can be maintained at the edge, with servers based in either location.

Servers could even be based both centrally and distributed in a hybrid configuration, which may be desirable either to leverage existing edge servers while moving towards greater centralization, or to allow a form of near-line storage so that more popular titles to be pushed towards the edge to streamline transport utilization, while larger libraries are centralized on economically scalable server resources. Content sourced from both locations share the same QAM resources, optimally placed at the edge, and utilized efficiently whether the temporal demand is relatively higher for the popular titles or the broader library content.

Breaking vertically integrated services towards greater specialization and abstractions between elements of functionality also protects the operator from technological obsolescence. With full vertical integration, modifying one element of functionality may not be viable as it would necessitate a complete forklift upgrade to the system. Alternatively, specialized components can be modified or upgraded with relatively minimal changes, perhaps only requiring modifications in the interface of existing components.

In fact, new and old components can coexist to drive innovation while maintaining existing plant investments. For example, in VOD there is a drive towards Gigabit Ethernet as the best transport method, and Gigabit Ethernet compatible QAMs are emerging to optimally accommodate this trend. Yet, many systems have already invested, in many cases only recently, in ASI QAMs. A solution that avoids any stranded capital is to use the Gigabit Ethernet transport for all VOD traffic, grow capacity with Gigabit Ethernet QAMs, and use Gigabit Ethernet to ASI protocol conversion to prolong the operational life of the already-purchased legacy QAMs.

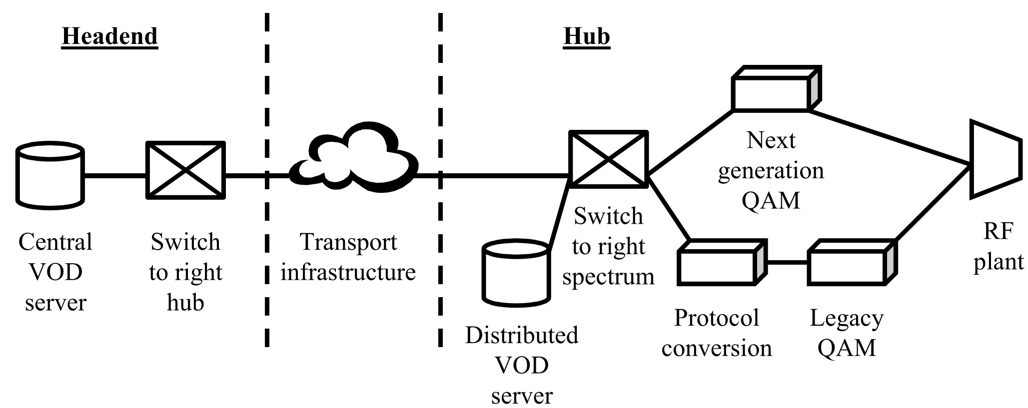


Figure 2: Flexibility for MSO to customize architectural deployment of best-of-breed components openly interfacing to each other (VOD example).

CONCLUSION

There is increasing volume and variety of multimedia services available to be provided over broadband networks. With inherently high bandwidth, increasing sophistication through trends such as migration to packetized digital content and node-level addressability, and established standards such as

MPEG-2, cable has an opportunity to emerge as the leading comprehensive provider of all services and all media.

Situations in cable systems are highly particular due to considerations such as what services are being emphasized, legacy investments already made in the networks, and overall unpredictability over how scenarios will play themselves out going forward. The one entity that can best determine how to architect infrastructure across media and services is the MSO.

The key to cable realizing its multi-service potential is for MSOs to have control over determination of deployments among open, interoperable components. Abstraction should be increasingly implemented among sources of service, switching, transport and media processing. This allows the cable operator to select best-of-breed components, optimize plant efficiency, economically scale resources, and determine precise architectures. So empowered, the cable operator is positioned to compete effectively and seize emerging opportunities.

Ran Oz
Chief Technology Officer
ranoz@bigbandnet.com

S. V. Vasudevan
Chief Architect
vasu@bigbandnet.com

BigBand Networks, Inc.
475 Broadway
Redwood City, CA 94063
650 995 5000
http://www.bigbandnet.com

NEW DEVELOPMENTS IN IEEE-1394 STANDARDS FOR THE CABLE SET-TOP BOX

Mark Eyer
Sony Electronics

Abstract

In December 2002, consumer electronics manufacturers joined with cable MSOs in announcing an historic agreement on digital cable compatibility. Among the provisions identified in the agreement was a commitment on the part of cable operators to provide, by July 2005, IEEE-1394 interfaces on high-definition digital cable set-top boxes acquired for distribution to consumers. Two standards related to the 1394 interface were cited in the agreement, ANSI/SCTE 26 [1] and CEA-931-A [2]. This paper describes the agreed-upon functionality of this high-speed serial bus network interface, and provides an in-depth discussion of the new CEA-931-A protocol.

INTRODUCTION

As envisioned by the engineers and policy-makers who forged the December 2002 Agreement, the primary function of the 1394 interface is to enable consumer recording or time-shifting of compressed MPEG-2 audio/video content, and delivery on a peer-to-peer home network of that digital programming. While the High-Definition (HD) digital cable set-top box outputs digital video for viewing via an uncompressed Digital Visual Interface (DVI) or optionally a High Definition Multimedia Interface (HDMI) port, an IEEE-1394 high speed serial bus interface must also be present.

The memorandum of understanding in the December 2002 NCTA/CEA Agreement [3] requires High Definition cable set-top boxes to:

> "… comply with ANS/SCTE 26 (as of 10/29/03) with transmission of bit-mapped graphics (EIA-799) optional, and shall support the CEA-931-A PASS THROUGH control commands: tune function, mute function, and restore volume function. In addition these boxes shall support the POWER control commands (power on, power off, and status inquiry) defined in A/VC Digital Interface Command Set General Specification Version 4.0 (as referenced in ANSI/SCTE 26 2001)."

We start with a system-level overview of the Digital Cable Set-top Box (DCSB) as it fits into a typical home network environment. We then move to a detailed review of CEA-931-A [2], published this year by the Consumer Electronics Association (CEA). We include along the way a discussion of a typical "IR Blaster" application to show how use of CEA-931-A can overcome the inherent limitations of that technique. Next, a brief summary of ANSI/SCTE 26 [1], first published in 1999 and updated in 2001, is presented. We then summarize the various normative references cited in ANSI/SCTE 26 and in the primary CEA standard for the display side, EIA/CEA-849-A [4], and conclude with a discussion of two implementation issues that may be of interest to designers.

SYSTEM OVERVIEW

Figure 1 diagrams a Digital Cable Set-top Box (DCSB) at the upper left. The video output flows from its DVI/HDMI port to a High-Definition Display. Audio connections are not shown in the simplified diagram.

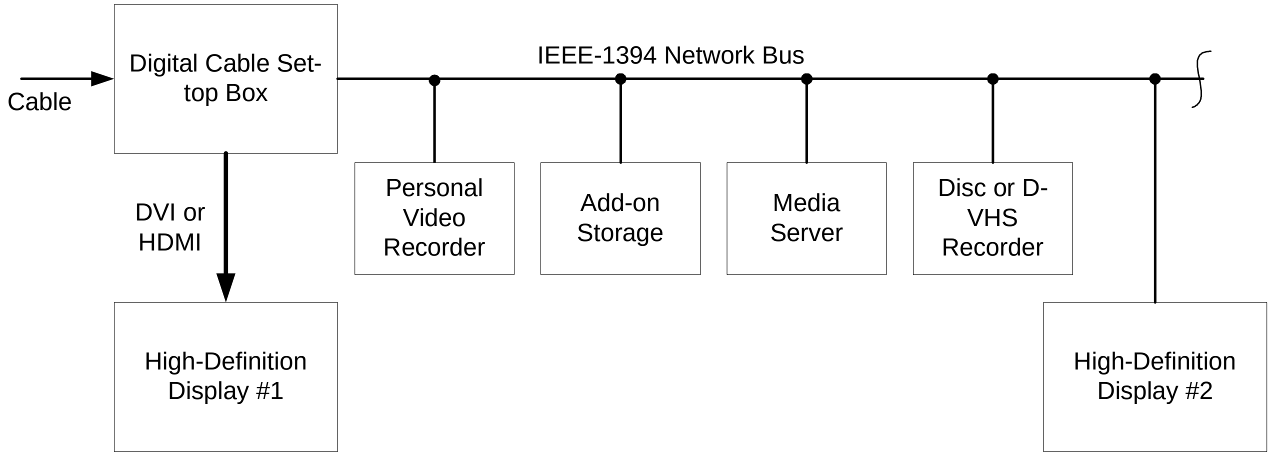


Figure 1. System Block Diagram

To the right of the DCSB in Figure 1, an IEEE 1394 Network Bus is shown interconnecting the set-top with a variety of audio/video devices. A hard-disk- based Personal Video Recorder (PVR) may have 100 GB or more of disk space available for temporary storage of audio-video programming. The PVR or DCSB may be able to make use of additional storage devices that may be present on the bus (represented here by the box labeled “Add-on Storage”).

A Media Server is also shown in Figure 1. While similar in many ways to the functionality offered by the PVR, the Media Server might include extra features such as archived musical recordings, access to Internet-based content, ability to manage prerecorded packaged media such as CDs and DVDs, and ability to accept and catalog personal content such as digital photos and home videos.

At the far right side of Figure 1, a second display device is connected via the 1394 bus. Compatibility with the DCSB is guaranteed for DTV displays supporting EIA-775-A [5] and the “MPEG profile” of EIA/CEA-849-A [4]. These two protocols are complimentary to ANSI/SCTE 26 [1]. Whereas the DVI/HDMI port on the DCSB provides a high-bandwidth pathway for graphics as well as for HD video, the 1394 path offers HD video but only standard-resolution graphics and a graphics frame rate that is limited by hardware capacity in either the source or display.

Digital Recording Functionality

How might a viewer use a setup like that of Figure 1 to make recordings of digital content? Several scenarios are possible. One possibility is that the DCSB may discover the Add-on Storage device and may be able to use it as a disk-based cache for audio/video programs. In this scenario, the set-top manages files and offers a suitable user interface to access them and to organize and manage disk resources. In another scenario, the PVR maintains its own file system, provides its own user interface to stored programs, and performs attended and unattended recording from selected source devices. It is with this latter scenario in mind (among other considerations) that CEA-931-A was developed.

As noted, an HD digital cable set-top box built to comply with the NCTA/CEA Agreement implements ANSI/SCTE 26 and certain provisions from CEA-931-A. Using these protocols, the PVR (or any other device on the 1394 bus) can identify the set-top box as a source of digital video, can turn it on and can tune it to any given digital channel. The format of data on the bus, including aspects of physical, link layer, transport, link encryption for copy protection, and audio/video compression formats are all specified. The precision and completeness of these provisions enables the development of this new category of consumer digital recording devices.

CEA-931-A STANDARD

CEA-931-A [2], entitled *Remote Control Command Pass-through Standard for Home Networking*, defines a standard method for communication on the network of simple "user intents" such as those typically represented by keys on a consumer Remote Control Unit (RCU). Additionally, the protocol recognizes and supports applications such as unattended recording that typically, in the analog world, must rely on devices capable of emulating the infrared pulses emitted by a given device's RCU. The new protocol offers a vast improvement in reliability and ease of use as compared with the analog techniques it replaces.

Remote Control Key Pass-through

Figure 2 illustrates in a simplified way the concept of remote-control key pass-through. The RCU in the figure is the unit sold with the DTV receiver. The format and carrier frequency of the infrared (IR) pulses it emits are recognized and accepted by that DTV receiver. Some of the commands, such as power on and off and picture controls (brightness, contrast, etc.) are directed at the DTV itself and are processed internally. Others, such as a "Record" key may not correspond to functions supplied by the display. Key presses such as these can be "passed through" the DTV and placed onto the 1394 bus. The function of the CEA-931-A [2] protocol is to define the standard method whereby these RCU keys are communicated across the network.

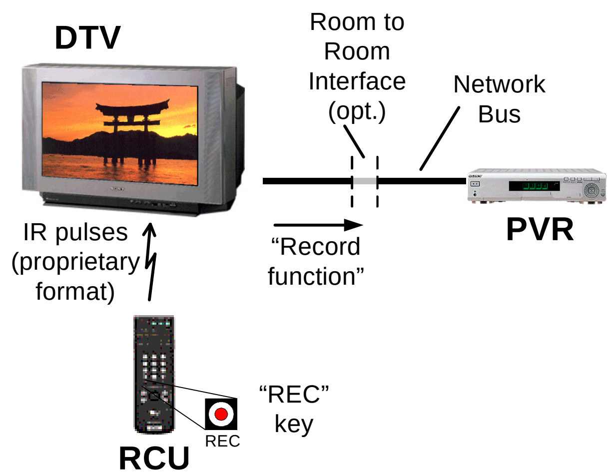


Figure 2. "Record" Key Example

Typically, the DTV will address the key commands not corresponding to internal functions to the specific network device currently selected as the current audio/video "input" or A/V source device. If several devices on the network could respond to a "Record" command, for example, the only one receiving the command will be the one currently selected as the A/V source.

A key feature of RCU key pass-through is that the IR pulses corresponding to the external function are proprietary to the manufacturer of the RCU itself, yet the functions themselves are mapped (in the device receiving the IR pulses) into standard key commands by the 931-A protocol.

What are the benefits of RCU key pass-through? The primary benefit is apparent if the device being controlled is not directly visible to the viewer who is holding the RCU. It may be in another room, for example, or simply inside a cabinet or behind a piece of furniture such that the RCU's IR pulses cannot reach its front panel. In such cases, the DTV receiver, whose IR receiver is in full view of the user, acts as a relay agent to deliver the commands to the hidden unit.

Another benefit of RCU key pass-through is that the networked A/V devices (if they all support CEA-931-A) may be controlled by a single RCU, thus eliminating the clutter and confusion of several remotes on the coffee table. This statement comes with a caveat: certain devices may include functions associated with dedicated keys on their native RCUs. These may not be mappable in a straightforward way to CEA-931-A key codes. The "universal" RCU remains elusive, yet CEA-931-A is clearly a step in the right direction.

Infrared "Blasters"

As mentioned, CEA-931-A [2] was designed with a scenario beyond simple key pass-through in mind. Typically, it is the viewer who operates his or her audio/video equipment via IR remote control by pushing keys on the RCU. Certain equipment, however, may take the place of a human operator in order to control a piece of A/V equipment when, for example, the user is not present.

Until the advent of digital home network busses and protocols such as CEA-931-A, analog methods were the only option. One approach employed "IR Blasters," devices capable of emulating the IR pulses recognized by the piece of equipment to be controlled. The controlling device typically includes an IR emitter attached to the end of a cable; the emitter is affixed in a position near the front panel of the device to be controlled. This approach, while workable, is fraught with difficulties.

To appreciate how CEA-931-A [2] may be used to overcome the shortcomings of IR Blaster approaches, we take a look at a typical application. This scenario involves a PVR using a cable or satellite set-top box as a video source and an analog VCR for archival backup of disk-based material to videotape. This setup is diagrammed in Figure 3.

As shown, the user operates the system primarily from the PVR's native remote control unit. When a channel change is desired, the PVR creates the necessary IR pulses to effect the channel change in the cable set-top box. For archival recording, it emulates the VCR's RCU to start and stop recording.

Beyond the large number of cables needed to wire these pieces of equipment together, one must allow for the limited length of the cables supporting the infrared emitters. If either of these emitters slips out of position, IR commands will not be received reliably.

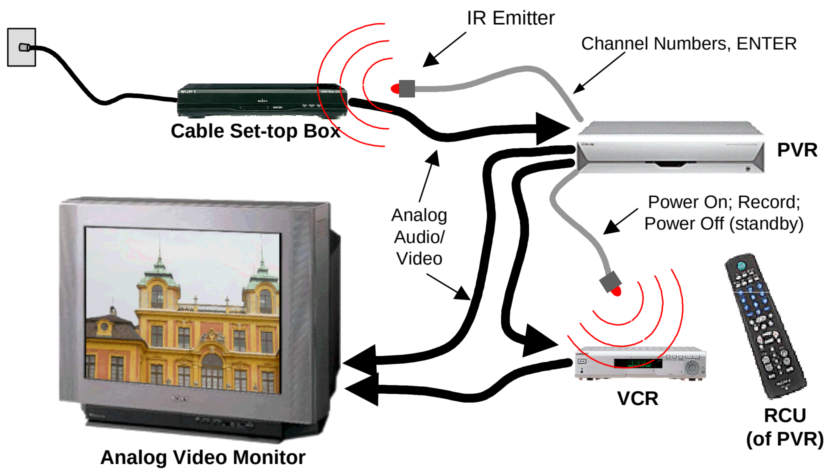


Figure 3. IR Blaster Example

Another reliability issue occurs because commands such as the one used to tune the DCSB to a given channel involve several key codes sent in sequence. If the codes are sent too close together, the set-top box may not recognize them. If they are sent too far apart, "channel surfing" is noticeably more sluggish.

Setting up the system in Figure 3 is a challenge by itself. The PVR must include a database covering as many models of audio/video equipment as possible. This database describes the modulation modes and infrared pulse formats corresponding to the needed commands for various models of various manufacturers' equipment, so that the PVR can control the user's specific set-top box and VCR. Typically, during the setup procedure the user will need to answer a series of questions such as "Does your set-top box require an ENTER key to terminate channel changes?" and "Does a channel number involve two digits or three?" With the setup procedure being time-consuming and prone to errors, it is no wonder many people are unwilling to attempt it.

Figure 4 shows the same application (PVR functionality with unattended recording and control of an A/V source device), but this time the devices handle digital audio/video on a 1394 bus and support standard protocols including CEA-931-A. 1394 cables interconnecting the pieces of equipment replace the bulky audio/video wiring.

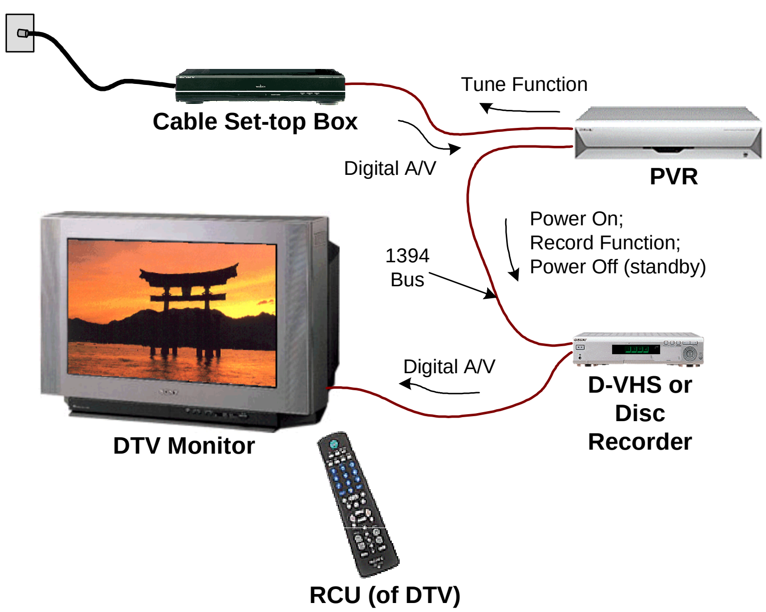


Figure 4. Using CEA-931-A for Unattended Recording

As soon as the equipment is powered on, each unit discovers the others on the network and can determine each device's type and function and what protocols are supported. No special setup is required at all!

The functions previously handled by the IR blasters are now handled by the CEA-931-A protocol. When the PVR wishes to tune the DCSB to a given channel, it may simply issue a "Tune Function" command providing the channel number to be tuned. Bus commands are acknowledged, so operation is reliable and predictable.

Whereas in the analog scenario the PVR's own RCU was used to control it, now the DTV's own RCU may be used to control the PVR, since the DTV is able to pass on to it key presses appropriate to its control.

When the PVR wishes to transfer a stored program to disc or tape, it can operate the recorder using other CEA-931-A functions. Power control is provided by commands defined in the *A/VC Command Set General Specification* [7], and the recording is initiated by a CEA-931-A "Record Function."

Now, when the viewer browses through the channel lineup to look for something entertaining, channel changes are fast: the time interval between channel number digits has been eliminated because the "Tune Function" involves a single command delivered on the home network bus.

CEA-931-A IN DETAIL

CEA-931-A [2] actually does not itself define or invent anything new. It is built on a portion of an existing industry standard developed by the 1394 Trade Association, called Panel Subunit 1.21 [6]. Whereas the Panel

Subunit document specifies both a "direct" mode and an "indirect" mode of operation, CEA-931-A uses only the simpler "indirect" mode. In this mode, whatever on-screen text or graphics constitutes the graphical user interface is either embedded into video or delivered via EIA-775-A bit-mapped graphics (defined in EIA-799 [8]).

The indirect mode of Panel Subunit was designed to allow a controller device to emulate the RCU of the device being controlled, hence its applicability to the needs identified by CEA for home networking.

The Panel Subunit command that conveys functions associated with RCU keys and "basic user intents" is called the PASS THROUGH control command. Each PASS THROUGH control command is carried within a standard AV/C Function Control Protocol (FCP) packet defined in [9], and identifies the particular RCU key or user intent by an "Operation ID" and (for some functions) one or more parameters.

Operation IDs

Table 1 lists the RCU keys and Deterministic Functions supported in Panel Subunit 1.21 and CEA-931-A. Those in the top portion of the table correspond to RCU keys, and have no accompanying parameters. Some of those in the lower portion labeled "Deterministic Functions" have associated data; these support applications like unattended recording and simple device control.

As can be seen by inspection of the top portion of the Table, all of the common RCU key functions are represented. Some are clearly aimed at specific types of devices. An example is the "Angle" key, used typically on DVD players to cycle among video tracks offering different viewing angles.

Each of the Operation ID types given in the top half of the table, when received by the target device, has exactly the same effect on the device as the corresponding key on its own RCU would have. That means that, for example, if repeated pressings of the PLAY key would cause the device to toggle between playing and pausing playback, reception of the

Tab1e 1. Defined Operation ID values.

Category	User operations
Navigation keys	Digits 0-9, Select, Up, Down, Left, Right, Right-up, Right-down, Left-up, Left-down
Menu selection	Root menu, Setup menu, Favorite menu, Exit
Media control	Play, Stop, Pause, Record, Rewind, Forward, Fast forward, Eject, Backward, Angle, Subpicture
Channel control	Channel up, Channel down, Previous channel
Miscellaneous	Power, Volume up, Volume down, Mute, Sound select, Input select

Deterministic Functions

Name	Function	Parameter
Play function	Start (or continue) playing content	Speed and direction of play
Record function	Start (or continue) recording	-
Pause-play, Pause-record	Pause playback or recording	-
Stop function	Stop playback or recording	-
Mute function	Mute audio	-
Restore volume function	Restore audio to previous volume level	-
Tune function	Tune to indicated channel (or virtual channel)	One- or two-part channel number
Select disk function	Select indicated physical media	Disk number (1-65,535)
Select A/V input function	Select indicated A/V input	A/V input (1-255)
Select audio input function	Select indicated audio input	Audio input (1-255)

"play" Operation ID would have the same toggling effect. Both "key down" and "key up" events are represented.

The Deterministic Functions listed in the lower half of Table 1, on the other hand, are defined such that the result of the command is entirely predictable. Toggling is not allowed. Accordingly, reception of the "Play Function" in the target device must result in playback either starting or continuing. Note that control of device power is not included here. CEA-931-A specifies that the POWER control command specified in [7] is to be used for deterministic control over device power state.

The benefit of Deterministic Functions is that the controller device no longer needs to try to keep track of the state of the device under control. With IR Blaster techniques, for example, the Power key on the RCU might toggle device power between "On" and "Standby." If the controlling device does not know the initial state, using the Power key might result in turning the unit off rather than on.

Deterministic Functions supported in Panel Subunit 1.21 and CEA-931-A include media control (Play, Record, Pause, and Stop), audio control (Mute, Restore volume), tuning control (Tune function), and functions to support selection of specific A/V inputs. The Play function is particularly powerful, in that it includes as a parameter the speed and direction of desired playback. All the trick modes are included, as well as fast-forward and rewind functions.

For reliability and error handling, the Panel Subunit specification describes methods any device on the network can use to determine whether a given target device supports the protocol. It also describes how a device issuing a PASS THROUGH command can determine whether or not the command is implemented in that device, and if implemented, whether it will be acted upon.

DTV 1394 INTERFACE

While the December 2002 Agreement does not cite specific requirements for a DTV receiver connected to the HD digital cable set-top via 1394, the applicable standards are well known in the industry. CEA has developed a logo program called "DTV Link" based on the "MPEG-2 profile" of EIA/CEA-849-A [4] (visit http://www.ce.org/dtvlink for details).

Any DTV display device that is compliant with the requirements of the EIA/CEA-849-A MPEG-2 profile is compatible with the HD digital cable set-top box. Simply put, "compatible" means the user will be able to interconnect the DCSB to the display and then view on-screen displays and program audio/video generated by the DCSB. In technical terms, this compatibility guarantees:

- The DTV display is able to discover and identify the cable DCSB on the 1394 bus as a compatible source (offering it up as a choice);
- The display is capable of decoding and displaying audio/video services including AC-3 audio and any of the allowed formats for compressed MPEG-2 video;
- The display is able to accept analog A/V output from the DCSB, and is able to switch between analog and digital DCSB outputs upon request by the DCSB (analog/digital source selection is discussed in more detail below);
- The display supports the MPEG-2 Transport Stream format delivered within an isochronous channel on the

1394 bus in accordance with IEC 61883-4 [10];

- The display determines which video program element to decode and display by examining the MPEG-2 Program Specific Information (PSI) tables: the Program Association Table (PAT) and the Program Map Table (PMT) it references. Whenever the DCSB includes a single program (service) in the PAT, the DTV display identifies the video component of that service and decodes it without need for viewer intervention. Likewise, it will find the appropriate audio track (perhaps based on its indicated language) and decode it without the need for user interaction.

For the same setup (DCSB connected to DTV), if CEA-931-A is supported by the DCSB, the DTV (or any other device on the network) can cause the DCSB to power on or go to standby power state, can cause audio outputs to be muted or un-muted, and can cause the DCSB to tune to any given analog or digital channel.

OTHER PROTOCOLS AND STANDARDS

CEA-931-A is the newest protocol applicable to home networking, but it is only part of the story. We provide a brief overview of protocols and standards relative to digital audio/video distributed by cable to an IEEE-1394-based home network. This discussion should not be confused with a 1394-based network primarily directed at A/V for the PC platform, as different video compression formats and protocols are involved there.

Figure 5 shows a digital cable set-top box on the left and a DTV display on the right. As mentioned, the primary standard defining requirements for the 1394 interface on the DCSB is ANSI/SCTE 26 [1]. The primary standard for the DTV receiver is EIA/CEA-849-A [4].

ANSI/SCTE 26

Entitled *Home Digital Network Interface with Copy Protection*, ANSI/SCTE 26 [1] specifies requirements for the 1394 interface on a digital cable set-top box. It states that a compliant cable set-top box must meet re-

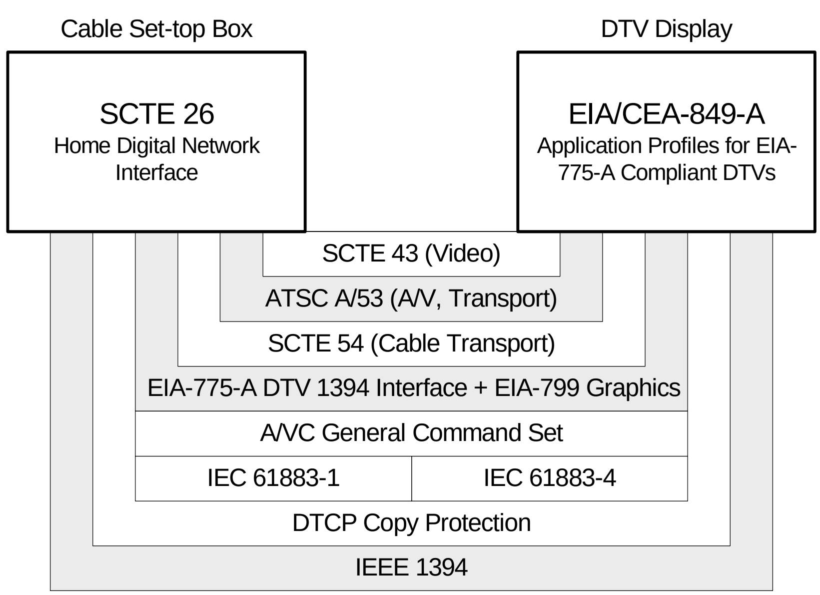


Figure 5. Protocol Dependencies

quirements for source devices given in EIA-775-A [5]. In addition, the set-top box must implement link encryption according to the "5C" method, also known as Digital Transmission Copy Protection (DTCP). The set-top box must indicate to the receiver whether to take its analog or digital output via the "analog digital source selection" method of EIA-775-A (discussed below).

EIA/CEA-849-A

Although EIA-775-A is an important element to 1394 compatibility between the set-top box and the DTV, it does not specify requirements for the higher protocol layers. For example, it does not state requirements for compatibility with various possible transport, or audio or video compression formats. EIA/CEA-849-A was written to address this need, defining a number of "application profiles" for EIA-775-A. The one relevant to our discussion here is called the "MPEG profile." If both a source device and a sink (display) device support a common EIA/CEA-849-A profile, 1394 interconnectivity (including the ability to decode audio and video) is assured.

Normative References

The ANSI/SCTE 26 [1] and EIA/CEA-849-A standards both cite a number of normative references, including:

- **ANSI/SCTE 43** [11], defining the allowable MPEG-2 video compression formats;
- **ATSC A/53** [12], the ATSC Digital Television Standard, defining audio, video, and transport aspects for sources originating from digital terrestrial broadcast;
- **SCTE 54** [13], defining transport-related characteristics for digital cable;
- **EIA-775-A** [5], defining how the IEEE-1394 protocol is used for discovery and connection management, delivery over isochronous channels of the MPEG-2 Transport Stream and over asynchronous channels of bit-mapped graphics defined in EIA-799 [8];
- **AV/C Command Set General Specification** [7], providing standard methods for device discovery and basic control, this document is the foundation for the family of AV/C protocols (learn more at http://www.1394ta.org/);
- **IEC 61883-1** [9], providing a robust foundation of lower-layer protocols, including common methods for encapsulating AV/C commands, and for formatting isochronous packets including timestamps;
- **IEC 61883-4** [10], specifying the standard method for carrying MPEG-2 Transport Streams on 1394;
- **Digital Transmission Copy Protection**, providing a standard method for link encryption to protect against unauthorized copying of high-value content (learn more at http://dtcp.com/); and
- **IEEE 1394** [14], the fundamental specification for the high-speed serial bus technology. The lower layers of the protocol stacks are defined here, including physical aspects of connectors and cabling.

IMPLEMENTATION ISSUES

We conclude with a couple of implementation issues for consideration by system designers.

Isochronous Channel Bit-Rates

Certain devices, such as for example disc or digital tape recorders, may not be able to handle the data rates as high as those delivered by a 256-QAM modulated carrier on cable. In 256-QAM mode, a given 6-MHz cable channel can deliver data at a rate exceeding 38 Mbps. A recorder capable of handling a maximum bit-rate of 20 Mbps, for example, would be able to record any single HD audio/video program, but would not be able to handle the rate of a full Transport Stream derived from a 256-QAM carrier on cable.

For this reason, and for the fact that from the user's perspective the desire is to record one program (not an arbitrary group of concurrently broadcast programs), the source device is expected to create a Single Program Transport Stream (SPTS). The process of creating a "partial" TS is straightforward, and is described in IEC 61883-4 [10].

Partial Transport Streams are structured like regular MPEG-2 Transport Streams except that not every 188-byte transport present in the original TS is present in the "partial" TS—only those corresponding to PID values of interest are included. For example, a partial TS may consist of TS packets carrying the Program Association Table (PID value 0), the Program Map Table section (PID value as identified in the PAT), and one audio and one video track (PID values identified in the PMT section). PIDs associated with Program and System Information Protocol (ATSC A/65) data may also be included.

The Common Isochronous Packet in [9] describes a time-stamp mechanism to enable the packet timing of the original partial Transport Stream to be accurately reconstructed at the receiving end of the isochronous channel. The important thing to note about partial Transport Streams is that the bus bandwidth needed to deliver them does not need to be any higher than the total data rate of the packets actually present. For example, a partial TS might include one standard-definition A/V programming service. While the full TS carrying that service might arrive in a 38.8 Mbps Transport Stream, the partial TS carrying the single program could be sent across a 1394 isochronous channel allocated with a much lower bandwidth (perhaps 6 or 8 Mbps).

Analog/Digital Source Selection

For the foreseeable future, HD digital cable set-top boxes will have analog outputs in addition to their digital ones. While it is possible (and becoming more practical every day) for the DCSB to digitally compress and encode services received via analog transmission channels, cost considerations may preclude this in the near-term. That means that when an analog channel is tuned, the digital video output from the set-top may cease. How is this situation handled?

ANSI/SCTE 26 requires the set-top to signal to the device connected to its digital output on the 1394 bus whether to take its digital or its analog output. This is described in Sec. 4.11 of [5], and it involves processing an AV/C CONNECT command identifying which output plug (analog or digital) should be taken as the current output from the box. Compliance with the MPEG-2 profile of EIA/CEA-849-A [4] (and hence DTV Link) also requires support for this analog/digital source selection mechanism.

A DTV display typically has several analog A/V inputs, perhaps labeled Video-1, Video-2, etc. In order for the AV/C CONNECT command to have the desired effect, the user must configure the DTV in a setup menu to identify which set of A/V inputs should be associated with a given set-top box.

CONCLUSION

This paper has explored the specifics of the historic December 2002 NCTA/CEA Agreement as it relates to the IEEE 1394 interface on the digital cable set-top box, and the ramifications of the provisions therein. The features and benefits of the CEA-931-A protocol published this year by the Consumer Electronics Association were outlined.

One hopes that the NCTA/CEA Agreement, along with the standardized protocols they reference, will spark a frenzy of creativity among manufacturers to bring the convenience and exceptional video quality of digital technology to a new family of home network-based Consumer Electronics products.

Acknowledgements

The author wishes to thank the members of the CEA R-7.5 Home Networking committee, chaired by James Williamson of Sony Electronics, for the development of the CEA-931-A protocol. The necessary extensions to the Panel Subunit specification were completed in the A/V Working Group of the 1394 Trade Association, chaired by Scott Smyers.

REFERENCES

[1] ANSI/SCTE 26 2001, *Home Digital Network Interface Specification with Copy Protection*, Society of Cable Telecommunications Engineers.

[2] CEA-931-A, *Remote Control Command Pass-through Standard for Home Networking*, Consumer Electronics Association, 2003.

[3] Consensus Cable MSO-Consumer Electronics Industry Agreement on “Plug & Play” Compatibility and Related Issues, Appendix B of Further Notice of Proposed Rulemaking FCC 03-3, January 10, 2003.

[4] EIA/CEA-849-A, *Application Profiles for EIA-775-A Compliant DTVs*, Electronics Industries Association and Consumer Electronics Association, 2001.

[5] EIA-775-A, *DTV 1394 Interface Specification*, Electronics Industries Association, 2000.

[6] *AV/C Panel Subunit Specification 1.21*, TA Document 2002009, 1394 Trade Association, January 2003

[7] AV/C Command Set General Specification, 1394 Trade Association.

[8] EIA-799, *On Screen Display Specification*, Consumer Electronics Association, 1999.

[9] IEC 61883-1, *Consumer audio/video equipment—Digital Interface—Part 1: General*, International Electrotechnical Commission, 1998.

[10] IEC 61883-4, *Consumer audio/video equipment—Digital Interface—Part 4: MPEG-2 TS Data Transmission*, International Electrotechnical Commission, 1998.

[11] ANSI/SCTE 43 2001, *Digital Video Systems Characteristics Standard for Cable Television*, Society of Cable Telecommunications Engineers.

[12] ATSC A/53B, *Digital Television Standard, Revision B, With Amendment 1*, Advanced Television Systems Committee, 2001.

[13] SCTE 54 2003, *Digital Video Service Multiplex and Transport System Standard for Cable Television*, Society of Cable Telecommunications Engineers.

[14] IEEE 1394-1995, *Standard for a High Performance Serial Bus*, Institute of Electrical and Electronics Engineers.

P2P, THE GORILLA IN THE CABLE

Alexandre Gerber, Joseph Houle, Han Nguyen, Matthew Roughan, Subhabrata Sen
AT&T Labs - Research

Abstract

There is considerable interest in peer-to-peer (P2P) traffic because of its remarkable increase over the last few years. By analyzing flow measurements at the border routers of a Tier-1 ISP backbone that carry broadband traffic, we are able to study its properties. P2P has become a large part of broadband traffic and its characteristics are different from older applications, such as the Web. It is a stable balanced traffic: the peak to valley ratio during a day is around two and the IN/OUT traffic balance is close to one. Although P2P protocols are based on a distributed architecture, they don't show strong signs of geographical locality. A broadband subscriber is not much more likely to download a file from a close region than from a far region.

It is clear that most of the traffic is generated by heavy hitters who "abuse" P2P (and other) applications, whereas most of the subscribers only use their broadband connections to browse the web, exchange emails or chat. However it is not easy to directly block or limit P2P traffic, because these applications adapt themselves to their environment: the users develop ways of eluding the traffic blocks. The traffic that could historically be identified with five port numbers is now spread over thousands of TCP ports, pushing port based identification to its limits. More complex methods to identify P2P traffic are not a long-term solution, the cable industry should opt for a "pay for what you use" model like the other utilities.

INTRODUCTION

P2P (peer-to-peer) file sharing applications have grown dramatically over the past few years and contribute a significant share of the total traffic in many networks. In this paper, we analyze flow-based measurements of broadband traffic spanning several months, gathered in the backbone of a large ISP network. We first develop an understanding of P2P traffic behavior from the viewpoint of broadband provider networks (earlier studies were based on a Tier-1 ISP backbone viewpoint [1] and on a University edge-network viewpoint [2]). The study then describes some key issues and challenges in handling/controlling this traffic, and presents a potential solution approach. We begin with a description of these P2P systems

File Sharing Applications

Many popular P2P applications such as KaZaA and Gnutella are organized as application-level overlay systems in which large numbers of computers (called peers) across the Internet link together in a decentralized manner via application-level connections. The predominant use of these systems is for sharing large data files (particularly music and video) among the connected users. The data files and associated metadata information (useful for searching content) are distributed across the different peers. A key difference with traditional client-server systems is that each host in a P2P system acts as both a client and a server of content. In contrast to the stable configurations of traditional distributed systems, the individual peers can frequently join and leave the P2P system.

The process of obtaining a file can be broadly divided into two phases – query search followed by object retrieval. First, a user specifies a query (e.g., a combination of name, genre, artist name etc.), and the P2P protocol searches for the existence of file(s) that match the query. The requesting peer receives one or more responses, and if the search is successful, identifies one or more target peers from which to download each file. The search queries as well as the responses are transmitted via the overlay connections. The details of how the search is propagated through the overlay is protocol-dependent. In earlier P2P protocols exemplified by Gnutella version 4.0, a peer initiates a query by flooding it to all its neighbors in the overlay. The neighboring peers in turn, flood to their neighbors, using a scoping mechanism to control the flood. In contrast, for newer protocols like KaZaA, as well as for newer versions of Gnutella, queries are forwarded to and handled by only a subset of special peers (called SuperNodes in KaZaA, and UltraPeers in Gnutella). A peer transmits an index of its content to the ``special peer" to which it is connected. The special peer then uses the corresponding P2P protocol to forward the query to other such peers in the system.

Once the search results are in, the requesting peer directly contacts the target peer, typically using some variant of HTTP (the target peer has a HTTP server listening by default on a known protocol-specific port), to get the requested resource. Some new systems use *swarming download*-- a file is downloaded in chunks from multiple peers.

Although the earlier P2P systems mostly used default network ports for communication, there is strong evidence to suggest that substantial P2P traffic nowadays is transmitted over a large number of non-standard ports. This seems to be primarily motivated by the desire to circumvent firewall restrictions as well as rate–limiting actions by ISPs targeted at such applications - we shall discuss this more later in the paper.

Another recent occurrence has been the development of tools that allow an end-user to explicitly select the SuperNode it connects to [3]. This appears to be an attempt to improve the quality of the best-effort search process in the P2P system, for files that are not widely distributed, but are geographically localized. For instance, connecting to a SuperNode in Brazil may increase the chances of locating Samba-related content.

Data Collection

We have access to “flow-level” data about broadband traffic at the border routers of a large ISP. Flow-level data is considerably more detailed than data sets such as SNMP, and at least this level of detail is needed to perform application classification. When looking at these flows we can make a very educated guess about whether the flow is associated with a Broadband consumer and from which region it originates. A region typically ranges from an extended metropolitan area to a state. For the remainder of this paper we focus on traffic that appears to be associated with broadband subscribers.

By flow, we mean a sequence of packets exchanged by two applications. More precisely we define a flow to be a series of uni-directional packets with the same IP protocol, source and destination address, and source and destination ports (in the case of TCP and UDP traffic). The flow measurements used here are called Cisco Netflow [4]; they are implemented in many of Cisco’s routers. The data collected about a flow (apart from the information above) are the duration, the number of packets, and bytes transmitted, and which header flags (SYN, ACK, …) were used in the flow. Measured flows are also constrained in time (Cisco Netflow collection sends flows from the router at 30 minute intervals), so there is a need to reconstruct the actual traffic from a single "connection". After reconstruction there will be one flow per connection – a potentially enormous volume of information.

In order to minimize any performance impact on the routers collecting the flow measurements the measurements are based on sampled packets collected on the routers, which then export the flows to aggregators. To reduce the huge data volume the aggregator further samples the flows using the smart sampling algorithm [5] that is better suited for heavy tailed distribution, such as typically found in Internet flows. In addition, there is also an uncontrolled sampling due to measurement packet losses. These three types of sampling can be estimated and corrected and don't affect our results that are based on the weekly or monthly average traffic generated by hundreds of thousands of broadband subscribers between May 2002 and February 2003.

Identifying Applications

There are a number of ways one could go about identifying individual applications within IP traffic. However, as noted, Netflow only keeps data on some aspects of flows. The most useful of these for application breakdowns are the source and destination port numbers, and the IP protocol number. The protocol numbers used are well documented [6], with TCP being protocol 6, and UDP being 17. TCP, and UDP traffic also define (16 bit) source and destination port numbers intended (in part) for use by different applications. The port numbers are divided into three ranges: the Well Known Ports (0-1023), the Registered Ports (1024-49,151), and the Dynamic and/or Private ports (49,152-65,535).

A typical TCP connection starts with a SYN/ACK handshake from a client to a server. The client addresses its initial SYN packet to the server port for a particular application, and uses a dynamic port as the source port for the SYN. The server listens on its port for connection. UDP uses ports similarly though without connections. All future packets in the TCP/UDP flow use the same pair of ports at the client and server ends. Therefore, in principle the server port number can be used to identify the higher layer application using TCP or UDP, by simply identifying which port is the server port (the one from the well-known, or registered port range) and mapping this to an application using the IANA list of registered port [7].

There are many barriers to determining applications from port numbers. For instance, well know and registered ports are not defined for all applications and this is typical of P2P applications. Further more, in some cases server ports are dynamically allocated as needed (for instance, one might have a control connection on which a data port is negotiated). Finally, the use of firewalls to block unauthorized and unknown applications from using a network has spawned work arounds that have made the mapping from port number to application ambiguous.

Despite this, a great deal can be said about the mapping of port to application, though obviously there will still be some ambiguity, and chance for errors. Note that both ports must be considered as possible candidates for the server port, unless other data is available to rule out one port.

The algorithm that we have adopted here chooses the server port by (1) looking for a well known port, (2) a registered port, or (3) an unregistered port which is known (from reverse engineering of protocols) to be used by a particular (unregistered) application. If both source and destination port could be the server, then we choose the most likely one through ranking applications by how prevalent they are in detailed (packet level) traffic studies – for instance, WWW is considered a high ranking application, as are email, and P2P applications.

The result is a mapping from flows to applications, that while not perfect, has been shown to be reasonably effective. The biggest problem is that there are still a substantial number of flows which cannot be mapped to an application. We further classify these unknown flows by the size of the flows: the category of most interest here is "TCP-big", which consists of unknown flows that transmit more than 100kB in less than 30 minutes.

We shall argue in this paper that the TCP-big traffic is primarily P2P traffic that is using unregistered ports unknown to us. P2P applications already use unregistered ports, and the structure of P2P protocols (with separate control and data traffic) allows data traffic to be assigned to arbitrary ports. In the past the major applications have typically used default ports (for instance 1214 for KaZaA) but in the recent past many efforts have been made to constrain P2P traffic through rate limiting single ports or by blocking some ports at firewalls, with the result that P2P users commonly use work-arounds. Where-ever we refer to P2P traffic we are using the traffic on the ports known to be directly associated with P2P applications: we shall keep this separate from TCP-big except where explicitly noted. Also note that some P2P traffic may be misclassified into other application classes and so our estimates of the total volumes of P2P traffic are conservative.

We should note that we are not collecting any information about URL's, or individual subscribers usage: IP addresses measured are not related to individual subscribers, and we only view the bulk properties of the traffic, such as its distributions.

APPLICATION COMPOSITION

Overview

Table 1 shows the application traffic composition for 2 broadband regions in May 2002 and January 2003. For each of these regions, we examine both the traffic coming from outside the region to some IP address within the region (referred to as IN) and the traffic sourced within the region and destined for outside the region (OUT). For each time period and region, we display the per-application traffic volume in each direction as a percentage of the total traffic in that direction. For a given application we also show the traffic normalized by dividing by its IN traffic volume for May 2002, in order to show the In/Out ratio, and the growth between the two periods.

We note that in either direction the P2P traffic forms a much smaller percentage of the overall traffic in January 2003 than in May 2002. TCP-big registered dramatic increases in traffic contribution in both directions (10.5 times for Outgoing and 6 times for Incoming) over the same period. The normalized figures show that the P2P incoming and outgoing traffic are very similar for either of the 2 months considered. Note also that the TCP-big traffic in the 2 directions becomes much more balanced recently than earlier. For example for broadband region X, the ratio between incoming and outgoing TCP-big traffic volumes changes from 1.94:1 in May 2002 to 1.12:1 in January 2003.

	Broadband Region X								Broadband Region Y							
	Applicationx Mix (percentage)				Normalized Consumption				Applicationx Mix (percentage)				Normalized Consumption			
	May 2002		January 2003		May 2002		January 2003		May 2002		January 2003		May 2002		January 2003	
	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN	OUT	IN
All	100.0%	100.0%	100.0%	100.0%	1	1.65	1.97	3.2	100.0%	100.0%	100.0%	100.0%	1	2.19	1.83	4.08
ESP/GRE	0.4%	0.5%	0.6%	0.5%	1	1.98	3.12	4.3	0.4%	0.5%	0.3%	0.4%	1	2.71	1.7	4.67
OTHER	4.4%	3.7%	5.7%	4.5%	1	1.37	2.54	3.23	4.6%	3.2%	5.4%	3.4%	1	1.53	2.16	2.97
TCP-BIG	8.9%	10.5%	47.5%	32.5%	1	1.94	10.5	11.68	9.5%	11.8%	45.3%	32.1%	1	2.71	8.71	13.72
AUDIO/VIDEO	0.2%	1.6%	0.2%	1.6%	1	16.61	2.77	32.64	0.1%	1.5%	0.2%	1.5%	1	23.71	3.1	44.29
CHAT	0.7%	1.3%	1.0%	1.7%	1	3.08	2.93	7.93	0.7%	1.2%	0.7%	1.4%	1	3.81	2.02	8.67
FTP	1.0%	1.3%	1.0%	0.7%	1	2.22	1.91	2.4	1.4%	1.4%	0.4%	0.9%	1	2.24	0.56	2.64
GAMES	1.6%	1.2%	3.6%	2.5%	1	1.29	4.54	5.15	1.3%	1.2%	3.4%	2.4%	1	1.92	4.73	7.43
MAIL	1.7%	0.6%	1.1%	0.7%	1	0.6	1.26	1.28	1.0%	0.5%	0.9%	0.5%	1	1.13	1.71	1.88
NEWS	0.3%	7.3%	0.2%	5.3%	1	38.52	1.51	54.55	0.7%	17.5%	0.7%	14.6%	1	54.99	1.76	85.33
P2P	75.2%	45.6%	32.9%	20.6%	1	1	0.86	0.87	75.1%	38.5%	36.7%	19.5%	1	1.12	0.9	1.06
WEB	5.6%	26.4%	6.2%	29.4%	1	7.8	2.2	16.88	5.2%	22.8%	5.9%	23.5%	1	9.53	2.06	18.27

Table 1: Application Composition of two broadband regions in May 2002 and January 2003.

Time of Day Pattern

We next examine the diurnal behavior of P2P traffic. Fig. 1 plots the time series of the incoming and outgoing traffic volumes (P2P, web and TCP-big) for a given broadband region across a week in February 2003. For each application, all the data values are normalized by the mean per-hour incoming data volume for that application, averaged across that week.

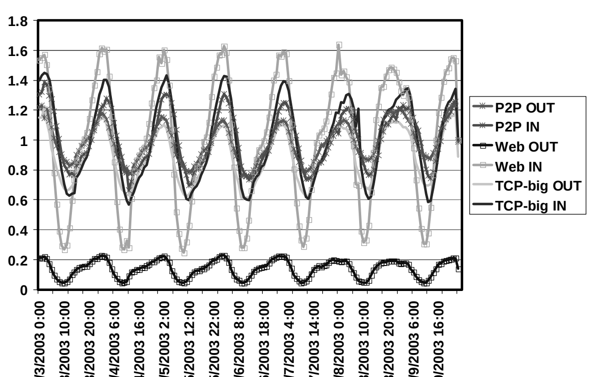


Fig. 1: Time of day pattern of P2P and Web traffic.

All three applications exhibit similar diurnal behaviors with peak loads (in either direction) around 2.00 AM GMT (10.00 PM EST, 7.00 PM PST). The P2P traffic exhibits less variability across a day than Web traffic. The peak load is about 2 times the minimum as opposed to 5 times for Web traffic. The smaller variance in P2P traffic across a day may be a function of the programmed download feature in P2P applications that allow users to specify multiple files in advance, that can be downloaded asynchronously by the P2P application.

For Web, the outgoing traffic is significantly smaller than (at most 20% of) the incoming traffic, suggesting that the broadband subscribers are mostly consumers of web data. In contrast, for P2P, the traffic in the 2 directions track each other much more closely, across a day and across the week. Another notable here is that the TCP-big traffic distribution across time is very similar to the P2P traffic. Also, just like P2P, the TCP-big traffic in the 2 directions are similar. These behavioral similarities are another indicator that the TCP-big traffic includes some P2P applications. Finally for all 3 applications, we do not see significant variations across days and between weekdays and weekends.

P2P LOCALITY

One of the potential advantages of P2P applications is that by distributing content, they provide the ability to download this content from locations closer to a user. It is therefore interesting to consider whether this really happens, and moreover to consider the question of locality in P2P traffic in general.

We approach this question by considering the simplest possible counter examples to localized traffic: the simple gravity model [8]. In this model, a packet entering the network at S, makes its decision about its destination D independent of the arrival point. That is, the packet is drawn (as if by gravity) to destinations in proportion to the volume of traffic departing at those locations.

The gravity model can be used to make predictions of the traffic volumes between two regions based purely on the volumes entering and exiting at those two regions, by the formula

$$T^{S,D} = \frac{T_{in}^{S} T_{out}^{D}}{T}$$

where T is the total volume of traffic across the network, T_{in}^{S} is the traffic entering the network at region S, and T_{out}^{D} is the traffic exiting the network at region D. Fig. 2 below shows a comparison of the gravity model predictions for inter-regional traffic of a broadband ISP. The plot is based on Netflow traffic collected during one week in September 2002; it shows traffic traversing the backbone between regions. The figure shows a scatter plot of the real inter-regional traffic versus the gravity model prediction, for both P2P traffic, and the total traffic to the broadband regions. One can see that in both cases the gravity model predicts the true traffic within about ±20%.

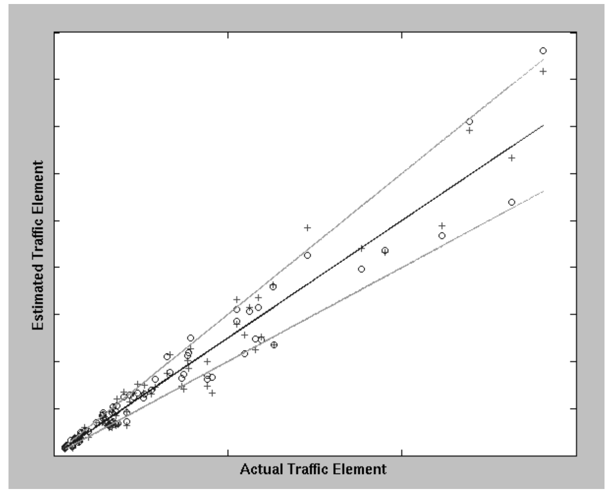


Fig. 2: Comparison of the real matrix elements to the estimated traffic matrix elements for a broadband ISP. The circles represent purely P2P traffic and pluses represents the total traffic. The blue solid diagonal line shows equality and the green dashed lines show ± 20%.

What does that tell us? Well the main point is that the gravity model above explicitly excludes any notion of geographic, or topological distance. Therefore, as the measured traffic fits this model to some extent, we may believe that neither P2P traffic nor the traffic overall exhibit strong locality at the regional level. A further, somewhat subjective conclusion one might drawn from the graph is that P2P traffic actually seems to fit the gravity model slightly worse, and so we may hypothesize that P2P traffic shows more locality than other traffic sources.

To examine these hypothesis in more details we present Table 2, which shows the normalized traffic volumes between regions for the P2P traffic. The table shows the normalized probability that traffic originating from a particular region in one broadband network, will depart from each region in the same broadband ISP (given it stays on the same broadband network). Table 2 can be seen to have a number of almost identical rows (for instance the group of regions R1, R2, and R5 are very similar, as is the group R6, R7 and R8) indicating a complete lack of locality of traffic with reference to these regions. Other regions (specifically R3 and R4) are not dramatically far away, but rather fall somewhere in between the other two groups.

However the table also shows some disparity between the groups of rows. This disparity is at its height when comparing the regions in the Eastern Standard Timezone (EST), with those in the Pacific Timezone (PST). This is an indication of some degree of weak locality in P2P traffic, at the "super-regional" level.

From/To	R1 (PST)	R2 (PST)	R3 (MST)	R4 (MST)	R5 (CST)	R6 (CST)	R7 (EST)	R8 (EST)
R1 (PST)	-	0.18	0.14	0.126	0.174	0.128	0.124	0.127
R2 (PST)	0.172	-	0.141	0.126	0.19	0.132	0.118	0.12
R3 (MST)	0.132	0.12	-	0.189	0.135	0.145	0.139	0.14
R4 (MST)	0.107	0.111	0.182	-	0.124	0.163	0.155	0.158
R5 (CST)	0.161	0.18	0.136	0.132	-	0.135	0.127	0.129
R6 (CST)	0.107	0.108	0.145	0.155	0.125	-	0.187	0.173
R7 (EST)	0.107	0.106	0.137	0.157	0.127	0.182	-	0.184
R8 (EST)	0.109	0.111	0.127	0.161	0.128	0.178	0.185	-

Table 2: Normalized inter-regional traffic matrix of broadband ISP X weighted by P2P+TCP-big traffic (Longitude defined by the Timezone).

This super-regional locality could arise for a couple of reasons (other than P2P applications explicitly taking advantage of content locality to improve performance). Firstly, because of usage patterns (specifically the times at which a user is connected to the P2P network), there is a slight increase in the likelihood that a search will find content in a local time zone. Secondly, there may be a group of people within a super-region with content that is slightly more relevant to the local super-region. However, the data so far suggests that both of these effects are not dominant, and certainly there is no strong locality influence such as might be seen if the main P2P applications exploited locality information.

In both of the above examples the monitoring location limits our data to seeing only inter-regional traffic. Thus, one might argue, we are missing the key component in any study of traffic locality: the intra-regional traffic. While the data limitations prevent us from seeing the intra-regional traffic on a single broadband ISP, we can gain a good view of this data by considering the traffic between broadband ISPs. If locality were being exploited in P2P applications, then one would expect traffic from ISP Y, region R to prefer going to ISP X, region R, rather than the alternative regions.

Table 3 shows an example, giving the normalized probabilities that traffic from ISP Y to X will go from regions M to R. Although the regions for the two broadband ISPs are slightly different, regions M3 and R7 are very closely matched as are M4 and R8. However, we see only very minor bias towards traffic from M3 to R7 (compared to other EST regions), and similarly from M4 to R8.

From / To	R1 (PST)	R2 (PST)	R3 (MST)	R4 (MST)	R5 (CST)	R6 (CST)	R7 (EST)	R8 (EST)
M1 (MST)	0.133	0.121	0.157	0.125	0.118	0.111	0.089	0.146
M2 (CST)	0.121	0.095	0.114	0.158	0.117	0.145	0.094	0.156
M3 (EST)	0.12	0.114	0.12	0.138	0.119	0.128	0.14	0.122
M4 (EST)	0.11	0.115	0.109	0.137	0.135	0.119	0.133	0.142
M5 (EST)	0.117	0.115	0.133	0.135	0.129	0.12	0.121	0.129

Table 3: Normalized traffic matrix from broadband ISP Y to broadband ISP X weighted by P2P+TCP-big traffic.

Our conclusion is that, although there is some evidence for weak locality at a large spatial scale, P2P applications do not yet exploit such information on a large scale, and consequently, P2P traffic does not show strong signs of geographic locality. Recent developments such as the KazuperNode tool [3]) provide methods for selecting the super-node to which one connects. On the one hand this could potentially increase locality if users tend to connect to nearby supernodes. On the other hand, there could be less locality if users connect to supernodes in different locations in their attempts to locate content.

HEAVY HITTERS AND P2P

It is well known in the broadband industry that some heavy hitters consume most of the bandwidth. We shall divide subscribers into classes by their total usage, and analyze their consumption characteristics such as the application composition and the traffic balance per class. We define three groups of users: the heavy users who consume more than 1 Gbytes/day in average over a week, the medium users who consume between 50 Mbytes/Day and 1 Gbytes/Day and the light users who consume less than 50 Mbytes/Day.

User Distribution

We first compare the distribution of traffic per subscriber. In order to see if there are consistent patterns we compare three regions, all at two different points in time: during the week ending June 26th 2002 and during the week ending February 9th 2003. By subscriber, we mean an active IP address. Even though the IP address is not statically assigned (the user obtains an IP automatically via DHCP), in the networks we examined it is "sticky". That is, over a week a subscriber maintains the same IP address in practice, because the DHCP lease expires only after 4 days and it is reassigned to him if it is still available. However, the IP address distribution doesn't reflect exactly the subscriber distribution since it misses the inactive subscribers and the subscribers with a very low usage that may not be sampled.

The six distributions in Figure 3 and 4 are quite consistent. In each case, the top 1% of the IP addresses account for 18.6 — 24.4% of the total traffic and the top 20% of the active IP addresses account for slightly more than 80% of the traffic.

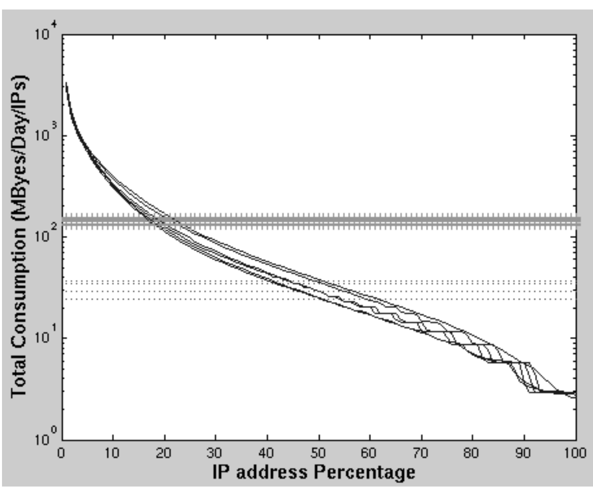


Fig. 3: Consumption per percentile of IP addresses of three regions during a week in June 2002 and a week in February 2003. The mean consumptions are around 140 Mbytes/Day/IP and the medians are roughly 30 Mbytes/Day/IP.

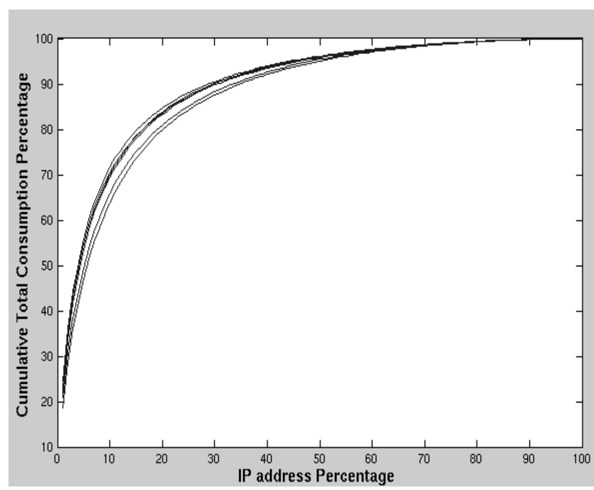


Fig. 4: Cumulative Consumption of three broadband regions during a week in June 2002 and a week in February 2003.

Consumption Characteristics

Since the median consumption is 4 to 5 times smaller than the average consumption, it is clear that the average consumption doesn't reflect the behavior of most of the subscribers. This still holds if we compare the application composition of each group of users, as defined earlier, with the average application composition that was studied earlier in this paper. Indeed, in a close look at one of these regions Table 4 shows that the light user group (67% of the IP addresses) is still mainly browsing the web, exchanging email and chatting online. Its traffic balance – the

IN/OUT ratio – is 4.8, which is far from the traffic balance of the heavy and medium user groups at 1.4-1.7 and 1.8, respectively. Table 5 makes it clear that this class of subscriber is not familiar with P2P or News since only 12.6 % of that group is lightly using one of these applications and it generates 1.1 % of the outgoing News traffic and 1.8 % of the outgoing P2P traffic.

	Week ending June 26th 2002								
User Type	Heavy		Medium		Light		Heavy	Medium	Light
Direction	OUT	IN	OUT	IN	OUT	IN	IN/OUT	IN/OUT	IN/OUT
Normalized Traffic per Sub	266.8	445.5	27.0	48.9	1.0	4.8	1.7	1.8	4.8
AUDIO/VIDEO	0.1%	0.3%	0.1%	1.9%	0.4%	2.7%	3.2	26.4	29.8
CHAT	0.2%	0.4%	0.6%	0.8%	2.9%	2.0%	3.2	2.4	3.4
NEWS	1.1%	34.9%	0.5%	13.5%	0.2%	2.1%	53.6	54.1	55.1
MAIL	0.4%	0.1%	1.5%	0.4%	8.3%	2.3%	0.5	0.5	1.4
FTP	0.7%	0.9%	0.6%	1.1%	0.8%	0.3%	2.2	3.5	1.7
GAMES	0.4%	0.5%	1.5%	1.5%	2.8%	1.0%	2.0	1.7	1.7
ESP/GRE	0.0%	0.2%	0.7%	1.1%	5.3%	2.8%	6.9	3.0	2.6
P2P	87.4%	44.0%	82.3%	43.2%	18.5%	6.8%	0.8	1.0	1.8
TCP-BIG	6.9%	8.4%	3.3%	6.3%	2.4%	2.5%	2.0	3.4	5.1
WEB	0.9%	5.3%	5.1%	26.6%	46.2%	71.6%	10.1	9.5	7.5
OTHER	2.0%	5.1%	4.0%	3.7%	12.2%	5.7%	4.3	1.7	2.3

	Week ending February 9th 2003								
User Type	Heavy		Medium		Light		Heavy	Medium	Light
Direction	OUT	IN	OUT	IN	OUT	IN	IN/OUT	IN/OUT	IN/OUT
Normalized Traffic per Sub	288.3	415.1	26.1	47.8	1.1	5.2	1.4	1.8	4.8
AUDIO/VIDEO	0.1%	0.5%	0.2%	2.2%	0.4%	2.6%	4.9	17.3	28.4
CHAT	0.3%	0.6%	0.7%	1.2%	2.6%	2.3%	3.0	3.0	4.1
NEWS	1.0%	32.8%	0.4%	10.5%	0.1%	1.4%	49.6	46.6	46.2
MAIL	0.1%	0.3%	1.3%	0.7%	8.1%	2.7%	2.7	0.9	1.6
FTP	0.8%	0.7%	0.5%	0.8%	0.6%	0.2%	1.4	2.8	1.9
GAMES	3.3%	1.9%	4.1%	2.7%	2.9%	1.0%	0.8	1.2	1.7
ESP/GRE	0.1%	0.3%	1.0%	1.4%	6.0%	3.1%	5.6	2.5	2.5
P2P	37.7%	22.9%	29.5%	14.0%	7.0%	2.3%	0.9	0.9	1.6
TCP-BIG	51.2%	30.5%	47.6%	29.3%	13.1%	6.8%	0.9	1.1	2.5
WEB	1.6%	6.5%	6.4%	31.5%	46.7%	72.3%	5.7	9.0	7.5
OTHER	3.9%	3.1%	8.2%	5.8%	12.5%	5.3%	1.1	1.3	2.1

Table 4: Comparison of the application composition of the heavy, medium and light user groups of a typical region.

On the other hand the heavy user group is mainly generating file sharing traffic. That group is actually providing content to the rest of the P2P community since its P2P traffic balance is below 1. Even though that subscriber group accounts for only 2.9% of the subscriber population, it generates almost half of the P2P traffic (table 5). What is more surprising is that these P2P applications are not the only way for the heavy hitter class to download files. Only 83.6 % of that group of users installed one of these major P2P applications. This percentage goes up to 96.7% if we take also Netnews into account. Finally the remaining 3.3 % chose other solutions that include FTP and downloads from the Web. It is interesting to notice that Netnews and the Web are only means to download content but not to share it and so the traffic balance for these applications is very large: up to 50 bytes received for one byte sent.

	Week ending June 26th 2002					
Direction	OUT			IN		
User Class	Heavy	Medium	Light	Heavy	Medium	Light
IP address Percentage	2.9%	30.1%	67.0%	2.9%	30.1%	67.0%
Traffic Percentage	46.6%	49.4%	4.1%	41.6%	47.9%	10.5%
NEWS	68.6%	30.4%	1.0%	68.4%	30.5%	1.1%
P2P	49.6%	49.5%	0.9%	46.2%	52.1%	1.8%
TCP-BIG	64.9%	33.1%	2.0%	51.5%	44.5%	4.0%
WEB	8.5%	52.2%	39.3%	9.8%	56.6%	33.6%
P2P Users in that Class	83.6%	63.4%	10.1%	83.6%	63.4%	10.1%
News Users in that Class	25.8%	12.4%	2.6%	25.8%	12.4%	2.6%
News or P2P Users	96.7%	71.6%	12.6%	96.7%	71.6%	12.6%

Table 5: P2P and News Users in a region having more than 100 000 subscribers.

Looking at the evolution of the traffic balance of Web traffic of the heavy users also leads to the conclusion that a more complex phenomenon is happening. Indeed in June 2002, the web traffic balance of the heavy users – 10.1 - was clearly higher than the web traffic balance of the light users whereas, in February 2003, that heavy hitter web traffic balance went down to 5.7, i.e. even lower than the one of the light users. This suggests that web traffic starts to be contaminated by a more balanced traffic, namely P2P applications. Furthermore, the traffic balance per application is another evidence that most of the traffic classified as TCP-big this year was actually what was classified as P2P last year. While the TCP-big traffic of the heavy hitters increased enormously, its traffic balance shifted from 2.0 to 0.9 and is now equal to the traffic balance of the P2P traffic that is still classified as P2P. It is now high time to understand why we are reaching the limits of port based identification of P2P traffic.

LIMITING P2P TRAFFIC

The ability to accurately identify P2P traffic is a crucial requirement for appropriately handling this traffic in the network - through either traffic engineering, provisioning, rate-limiting or pricing. However, P2P applications have evolved rapidly in a direction which makes accurate accounting of the traffic more difficult. In particular, previously the applications used default TCP ports, and it was possible to account for the bulk of the P2P traffic by monitoring a relatively small number of ports. However, the current widespread use port-hopping makes such mapping exceedingly impractical. We next present specific evidence of this trend and then discuss the implications for managing this traffic.

<u>KaZaA Rate Limiting Experiment</u>

We first show an interesting case study which graphically illustrates how difficult it can be to limit P2P traffic. In Fall 2002, a particular broadband region began rate limiting traffic on port 1214 (the default port for KaZaA). Fig. 5 shows the IN traffic for web, p2p and TCP-big for that region before and after the rate limiting was initiated. Note that the P2P traffic decreases significantly after the rate-limitation was initiated. However, the TCP-big starts increasing and in 2 months has tripled compared to its value just before rate–limiting began. The web traffic (port 80, 8000, 8080) also increases over the same period. A reasonable explanation for the jump in the TCP-big traffic coincident with the rate limiting action on the KaZaA port is that the traffic spurt was caused by KaZaA traffic migrating to other ports that were mapped to TCP-big.

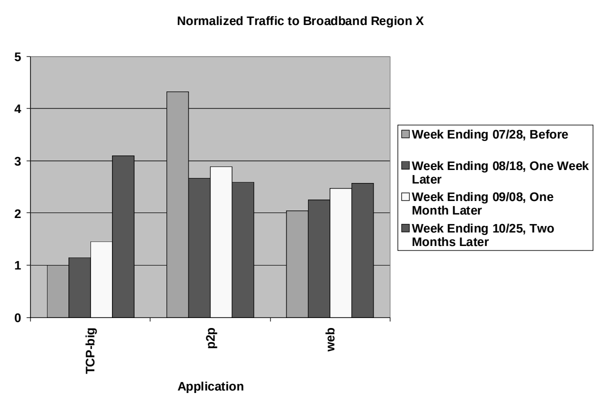


Figure 5: Mutation of P2P traffic into TCP-big traffic.

This conclusion is supported by the previous findings of this paper, but we shall investigate in even more detail. Fig. 6 plots the per-port traffic distribution for June 2002 and February 2003, for the P2P or TCP-big ports for the 2 time periods. Note that in 2002, 60 % of that P2P and TCP-big traffic was contributed by only three ports. However, in February 2003, the traffic was much more uniformly distributed among a larger number of ports – the top 3 ports now account for only 20 % of the traffic. To get 60 % of the traffic we would need to monitor a larger number (1000) of ports.

Much more difficult is the task of mapping the traffic on these heavy-hitter ports to specific applications. Given the use of port-hopping by bandwidth-intensive applications like P2P, an important unanswered question is how much of the traffic on these ports can be attributed to the IANA-registered applications, and how much is P2P.

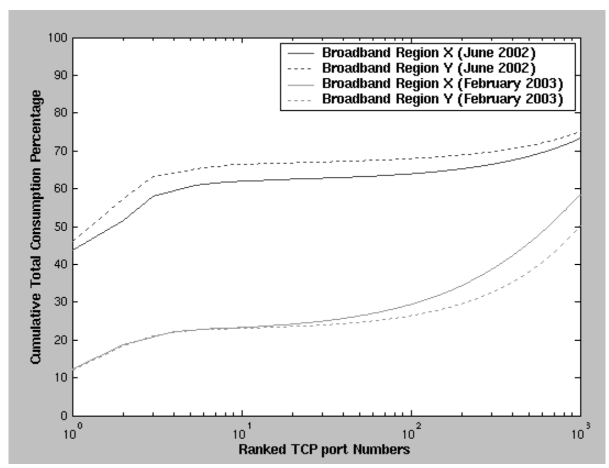


Fig. 6: Distribution of traffic by TCP port numbers classified as P2P or TCP-big

Given the limitation of port-based accounting, one might try to develop alternative techniques to accurately identify P2P applications. For example, additional information such as packet-level data, identification of SuperNodes etc. could help in developing signatures of P2P traffic. However, P2P applications have exhibited remarkable ability to rapidly evolve to evade detection and control. For example, many P2P applications now encrypt their communications, making it more difficult to reverse-engineer and/or monitor such systems at the application-level.

The above trends have important implications for port-based traffic control of P2P applications. If the rate control is targeted to a few well-known P2P ports, a significant fraction of the P2P traffic will evade the limit by hopping to other ports. The alternative is to track a larger number of ports that contribute significant traffic volumes and that are suspected to carry P2P traffic. The problem with this approach is that (i) it may not be feasible to track such a large and potentially dynamic set of ports, and (ii) such a widespread rate control may adversely affect the performance of many non-P2P users running valid applications on these other ports – this would be undesirable for the broadband providers.

SERVICE EVOLUTION TO TAME THE P2P GUERILLA

There are an assortment of approaches to address the "problem" of P2P traffic. Let's review a few that may be applicable to the cable industry.

Over the past few years many Multiple System Operators (MSOs) have incorporated "caps" into their service definition. These service caps tend to be implemented by controlling the rate at which data can flow into or out-of the network. The effect of these caps is to limit the instantaneous peaks of on-demand transactions. This has started us down the path of keeping bandwidth hogs in check. Some MSOs are now adding "tiered caps". This allows the bandwidth hogs to identify themselves as such and pay a price for the enhanced service they are receiving.

Caps have been good to the industry and take us part of the way to where we want to go. However, P2P traffic is a relatively "passive" phenomena. The requester can queue-up a set of requests for files then walk away. The file provider does not even need to be at the serving PC. In this situation rate capping will make the requests take longer, but will likely not change the behavior of the P2P participants. Fig. 1 enforces this point with the lower correlation between P2P traffic with the times users tend to be at their PCs.

Attempts to manage P2P traffic explicitly have met with little success. As illustrated in Figure **5**, attempts to block standard ports of one P2P application only cause the user population to shift their behavior so that the traffic reappears on other ports. Devices inside the network to block or significantly throttle specific port numbers have questionable economic return given the "slipperiness" of ports that P2P applications use and the risk that valid applications also are using those ports.

Not that we should treat High Speed Data Services as a classic utility, but let's look at how other "utilities" handle the problem of consumption hogs. Water, power, landline phone utilities all have a "pay for what you use" model. There is no attempt in these industries to limit the usage besides the economic consequence of paying for what is used. Cell phone providers put an additional twist on this model and provide usage bands. These bands allows a subscriber to sign-up for a usage band that best represents their need, but then gets charges for usage beyond what is included. With these revenue models consumption hogs are not "bad", they are just big consumers.

User response to these revenue models may not be as bad as we may fear. Users will be concerned that this will raise their rates. Surveys suggest that many users, on the average, feel they themselves are heavy hitters. But Figure 4 suggests only 5% of the users are creating 50% of the traffic. With strategic selection of banding, the users will be pleasantly surprised to find that they can buy one of the lower bands. There will be a small percentage of users (maybe the 1% that is causing the 20% of the traffic) that will not be happy with their new rates and will balk to other broadband services, but those are the ones that the cable industry can afford to lose.

CONCLUSION

In this paper, we examined a large set of flow-based measurements of network traffic associated with broadband consumers, spanning several months. Our analysis reveals several interesting features. Firstly, they illustrate that broadband consumer traffic is dominated by P2P applications. We further look into the properties of the various application classes, in particular the traffic patterns, and IN/OUT ratios, noting that P2P traffic has a much more balanced traffic pattern and IN/OUT ratio than applications such as the web. In addition we show that geographic locality is not yet a dominant feature of P2P traffic.

The paper then considers the traffic patterns of user groups, showing that the well known 80-20 rule (80% of the traffic is generated by 20% of the users) applies here, but moreover that the group of heavy users actually tend to use different applications : they tend to generate more P2P and Netnews traffic, while the group of light users tend to use more web, email and chat applications.

Finally the paper considers how one might control the large volumes of P2P traffic that currently flood the broadband networks. The more obvious controls, such as rate limiting traffic on particular ports are shown to be ineffective, because they simply push the traffic onto alternate ports. A more practical approach is to adopt a usage-based pricing approach, where the customers are billed for the resources they use.

REFERENCES

[1] "Analyzing Peer-to-Peer Traffic Across Large Networks", S. Sen and J. Wang, Proceedings of ACM SIGCOMM Internet Measurement Workshop, Marseilles, France, November 2002.

[2] "An Analysis of Internet Content Delivery Systems", S. Saroiu, K. P. Gummadi, R. J. Dunn, S. D. Gribble, and H. M. Levy, Proceedings of the 5th Symposium on Operating Systems Design and Implementation (OSDI), Boston, MA, December 2002.

[3] KaZaA supernodes: http://www.fasttrackhelp.com/kazupernodes/en

[4] Cisco Corp. NetFlow Services and Applications: http://www.cisco.com/warp/public/cc/pd/iosw/ioft/neflct/tech/napps_wp.htm

[5] “Charging from sampled network usage”, N.G. Duffield, C. Lund, M. Thorup. ACM SIGCOMM Internet Measurement Workshop 2001, San Francisco, CA, November 1-2, 2001.

[6] Internet Assigned Numbers Authority,protocol-numbers: http://www.iana.org/assignments/protocol-numbers

[7] Internet Assigned Numbers Authority, port-numbers: http://www.iana.org/assignments/port-numbers

[8] “Fast Accurate Computation of Large-Scale IP Traffic Matrices from Link Loads", Yin Zhang, Matthew Roughan, Nick Duffield and Albert Greenberg, to appear in ACM SIGMETRICS 2003.

PACKET NETWORK TOPOLOGIES FOR NEXT GENERATION VIDEO ON DEMAND AND SWITCHED BROADCAST SERVICE DELIVERY

John Amaral and Paul Pilotte
Artel Video Systems

Abstract

This paper focuses on packet network architectures that are optimized for the delivery of next generation Video-on-Demand and Switched Broadcast. The paper explores the behavior of switched video delivery networks that satisfy the growing user demands for unique orthogonal sessions. A detailed analysis of video delivery infrastructure composition is undertaken. The paper discusses packet switching systems, optical transport, Layer 2 forwarding, QAM modulation and storage. A hypothetical 300,000-subscriber VoD network is employed as the basis for describing network behavior under several scenarios. The analysis culminates in a cost-effective, extremely high capacity network that dramatically increases bandwidth resource utilization and provides dynamic and agile program delivery. The disclosed topologies possess effective redundancy and resiliency. Several practical examples are considered with regard to the disclosed topologies; the examples include "Everything on Demand" (EoD) and Switched Broadcast Services. The analysis is predicated by the feasibility and practicality of the described topologies. Considerations such as interoperability, cost, ability to deploy, and ease of use is taken into account as important factors when describing the topologies.

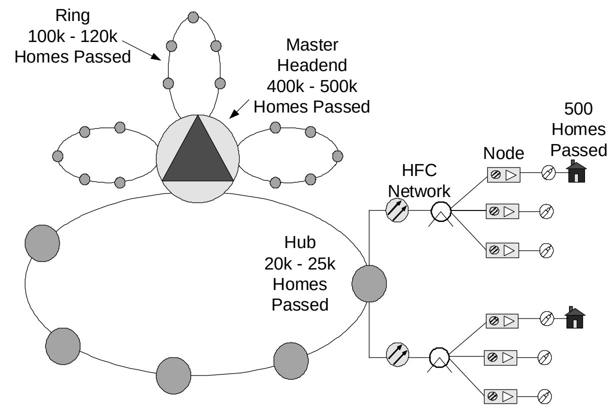


Figure 1 - Typical Cable Network

THE MOVE TO VOD

The deployment of next generation video on demand, EoD, and diverse content offerings by the MSO's in one form or another is regarded as a foregone conclusion. MSO's must deploy VoD services to counteract the competitive threats of Digital Broadcast Satellite (DBS). Most regard the near-term rollout of these services critical to both reducing digital churn and to increasing subscriber revenues. To date nearly all of the North American cable operators have deployed video on demand. The fantastic success of trials and early deployments has motivated cable operators to accelerate rollouts.

These service types rely on the ability of the cable networks to deliver unique, orthogonal video streams established by dynamic user control. This facility behaves in much the same way as the public telephony network; a user connects to the network (lifts the receiver), signals a unique switched communication path (dials the number) and conducts a unique session (talks to the desired party). One-to-one connectivity is necessary in telephony, and in data networks, because nearly every transaction context is unique; i.e. different content, at a different time, for a different reason. Based on conservative estimates, 10 million unique and simultaneous "one-to-one" streams will be deployed by North American MSO's in the next five years.

The move to session based video delivery will require a cost effective, high capacity switched packet network infrastructure, two way digital HFC plant and sophisticated content processing, storage and management facilities. Fortunately, the 50 billion dollars spent for HFC plant upgrades and bidirectional digital television capability underpin this endeavor. MSO's are left to focus additional spending to enable the head-end to hub network infrastructure for VoD.

Today, the cost of deploying end-to-end networks capable of delivering such a large number of independent streams continues to be prohibitive. The cost to provision a single stream (server, switch, transport, QAM, RF) is around five hundred dollars (assuming a Gigabit Ethernet based delivery infrastructure). High equipment demand, and cost pressure from MSO's coupled with technical advances and a growing competitive landscape is rapidly decreasing the per provisioned stream price.

Accelerated deployment of advanced services relies on further advancement and cost reduction of network components. This paper assumes equipment will inevitably reach acceptable price points. However, low component costs are not sufficient to enable large scale and effective VoD networks. A comprehensive architecture must be adopted to effectively provision and manage networks of sufficient scale and density necessary to support thousands of interactive and unique content sessions, a volume of sessions requiring hundreds of Gigabits per second of bandwidth.

The embodiment of such architecture must yield a system that is subjectively easy to use. It should also be scaleable in both size and capability. An optimal solution must have low capital and operational expense, high asset utilization, and manageable complexity.

EVOLUTION OF VOD NETWORKS

Early Video on demand networks deployed video servers at the edge of the network in a distributed model. The network edge is the location in the hub where the QAMs interface to the HFC plant. At the time, low utilization, scarce content and limited investment in equipment made it acceptable if not preferable to utilize resources in this fashion. In fact, some MSO's still operate large distributed VoD networks. Early deployments helped prove the business cases for VoD. In addition, important lessons about market behavior and consumer preference were learned. Experience gained during deployment of early-distributed VoD systems drove technical initiatives to further optimize VoD network design.

Distributed VoD systems are technically simple in that the delivery network is inherently localized. VoD servers sit in hubs, connected to QAM modulators and feed channel groups dedicated to serve content. Generally, servers deployed for this purpose had ASI output connections; in some cases direct QAM outputs. Asset distribution to servers varied from an intensely manual process, a technician driving from hub to hub with a TK-50 tape in his hand, to more automated systems having servers connected by an out of band channel (ATM or IP network). This channel allowed content to be inserted at a central location and copied to the remainder of the VoD servers via FTP.

MSO's rapidly realized the shortcomings of distributed VoD systems as they began to deploy en masse. In a distributed architecture there is invariably a mismatch between the playout capacity of the VoD server and the number of provisioned streams. This mismatch is due to the lack of granularity inherent in most VoD servers. The unused playout capacity drove the cost per stream unacceptably high. Storage utilization was also poor because each server in the network needed to be loaded with exactly the same content. In addition operational expenses were unacceptably high because service personnel needed to travel to each hub to maintain and upgrade VoD servers. An apparent solution was to centralize video server assets in a common head end and transport the streamed sessions to the hubs. The change to this "centralized" approach better matched server playout capacity to stream demand, thereby reducing the amount of unused server capacity. Load sharing allowed storage to be arranged such that the amount allocated to a title is proportional to the number of simultaneous sessions demanded of that title. This resulted in improved cost per stream because fewer servers were needed to service the same number provisioned streams.

ASI optical transport is routinely deployed for VoD and SVoD by MSO's today and has been an effective choice for building low to medium scale VoD systems. Rings and point-to-point ASI over DWDM networks connect VoD servers to QAM modulators in remote hubs.

There are two inherent disadvantages of ASI transport systems. The first disadvantage is that payload capacity of ASI is only about 216 MbpS. That comprises about 40 VoD streams per ASI. When deploying small-scale VoD networks with low peak load, provisioning services at this granularity is acceptable. As peak load demand increases, it is necessary to provision unacceptably large numbers of ASI links. The low ASI transport capacity undersubscribes server resources and physical transport assets (fibers, lasers etc). Replacing ASI transport with Gigabit Ethernet transport can increase the payload capacity by 500% to 1000% for the same cost. The second disadvantage is that the payload containers carried by ASI are MPEG-2 transport streams. MPEG-2 transport streams are not intended, nor capable, of forming the basis for a switched network interconnect. The virtues of a packet switched interconnect for VoD will be discussed in the next section. The effect of these two properties is unacceptable cost per bit transported. In order to reach the cost points necessary to proliferate VoD and move toward EOD, most believe that a switched packet interconnect based on Gigabit Ethernet is necessary.

GIGABIT ETHERNET AS A BASIS FOR VOD NETWORKS

The benefits of Gigabit Ethernet technology in transport systems for current and next generation VoD networks begin with its ubiquity. Quite simply, lots and lots of Ethernet equipment is bought and sold each year. This assures the contributing electronic components will remain commodity. Ethernet’s plug-and-play interoperability makes networks deployed with Ethernet easy and cost-effective to install, manage and use. Gigabit rate optical and electrical interfaces prevalent in the data communications world provide a high-speed, robust interconnect between servers, edge devices and processing elements connecting to the transport infrastructure. Contemporary servers and storage systems are optimized to operate in increments of 1 GbpS and make excellent use of off the shelf network adapters. A 1GbpS Ethernet link can transport up to 240 VoD streams, a suitable increment of streams to deploy in large scale VoD networks. Another advantage of Ethernet is its media access control (MAC) layer. The Ethernet MAC layer is a data link protocol with sufficient properties to form a sophisticated, and extremely scaleable packet switched network. Switching allows effective bandwidth management thereby reducing per stream costs.

INTRODUCTION TO NEXT GENERATION VOD TOPOLOGIES

What does the ideal VoD network look like and why? VoD networks provide service to a geographical region approximately the size of a large city. They are high in transmission capacity, moderate in complexity and provide robust, reliable interconnect. They are typically deployed as overlay networks, and are not intended to replace existing broadcast infrastructure or data/voice networks. The motivation is to develop low cost per stream networks tailored for the unique properties of VoD traffic. The systems envisioned have basis in Ethernet as the data link protocol, no different from a LAN. The departure between a standard Ethernet LAN and VoD network is the allowed distance between end terminals, intermediate path capacity, and optical route usage (DWDM). In order to develop a system optimized for VoD delivery the following capabilities must be employed.

Layer 2 Aggregation for Bandwidth Recovery

Layer 2 (Ethernet MAC layer) switching can be used to combine multiple partially filled Ethernet links to build fully utilized optical transmission paths between headend assets and hubs. Fully utilized optical links ensure each transport laser is operated at maximum capacity. System cost is reduced by virtue of needing fewer lasers. This is important because laser cost is the single largest contributing factor to overall VoD transport costs. Layer 2 capability enables aggregated links to be demultiplexed and delivered to the correct destination. Packet switching provides full mesh connection between content sources and HFC destinations allowing servers to load balance.

Figure 2 illustrates the benefits of Layer 2 aggregation over Layer 1 transport. In the Layer 1 example, content from each VoD servers can only be directed to the corresponding downstream QAM.

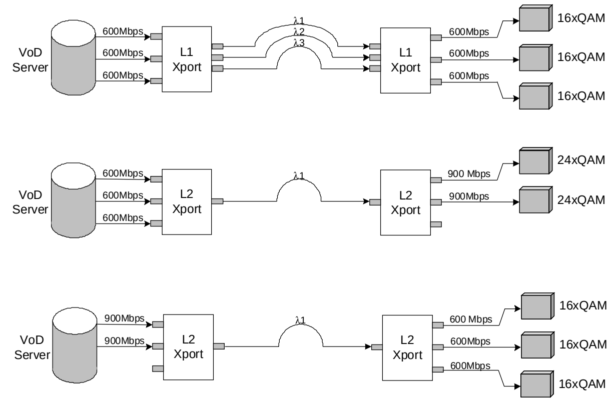


Figure 2 – Layer1 vs. Layer2 Aggregation

In the Layer 2 examples, content from all three VoD server ports can be aggregated to form a fully utilized wavelength, thus reducing fiber and laser costs. In addition video streams can be directed to any of the QAM devices from any server port on a packet-by-packet basis.

Layer 2 Forwarding and Shared Wavelength Topologies

Layer 2 forwarding allows the contruction of shared wavelength topologies. Video streams entering multiple inputs to a head end transport device can be aggregated and tagged with information which allows them to be discerned by hub end devices connected a shared optical path. Each device on the shared path can selectively receive any of the video streams available on that path and forward them to the QAM. This optimizes bandwidth utilization by allowing streams to be delivered to a number of QAMs using only the bandwidth they instantaneously require. This results in the lowest cost per video stream.

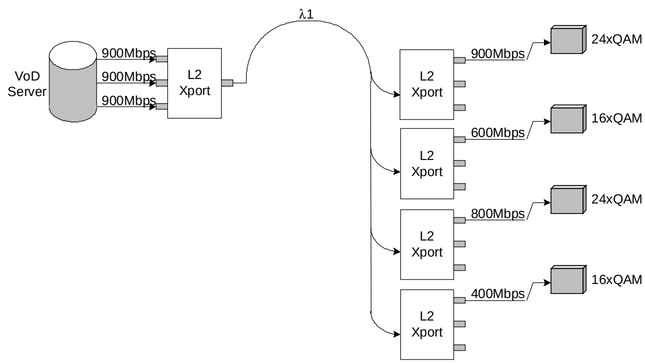


Figure 3 - Layer 2 Shared Ring Topology

An example of these benefits is evident in Figure 3. In this case Layer 2 switching aggregates three server outputs and places the video streams on a single wavelength of the shared ring. Multiple hub end devices are connected to this single wavelength. The hub end devices are configured to receive only the video streams that are destined to their associated QAMs. Ethernet allows this to be done automatically with no user intervention. Bandwidth to the QAM devices is allocated in arbitrary proportions as required. This is extremely useful in multicast environments where single copies of a video stream are presented to all hub devices, but selectively switched to QAMs based on user demand.

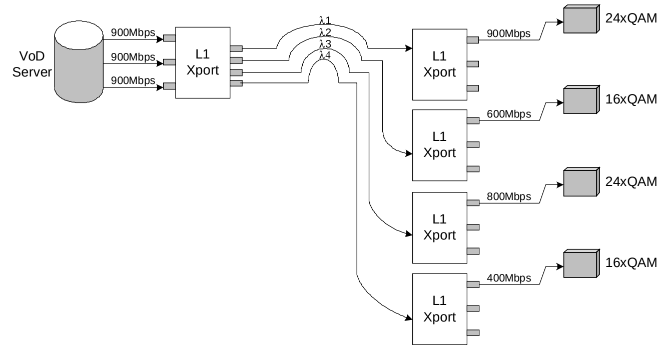


Figure 4 - Layer 1 Point-to-Point Topology

Figure 4 shows a logically equivalent topology using Layer 1. Note that the Layer 1 topology is point-to-point and requires four underutilized lasers while the Layer 2 topology with the shared ring uses only one fully utilized laser. This results in a much higher cost-per-stream compared to the Layer 2 solution.

Asymmetric Reverse Path

Up to now the topology diagrams have shown the forward video path only. There may be a need for a reverse path to carry control and management traffic from the hubs to the head end. This reverse path typically requires an order of magnitude less bandwidth than the forward path. In a Layer 1 system, bidirectional links must be used to support a reverse path. In this case the reverse path would have the same cost as the forward path even though it does not carry video content and is underutilized.

Figure 5 shows an example of asymmetric reverse path for hub to head-end interconnect. Since VoD traffic is predominantly from head-end to hub-end with comparatively low-bandwidth control traffic required for the reverse direction, Figure 5 shows a topology where a single wavelength transports the reverse path management information for all of the downstream QAMs to the head end over a single wavelength.

This preserves all but one wavelength for forward path traffic.

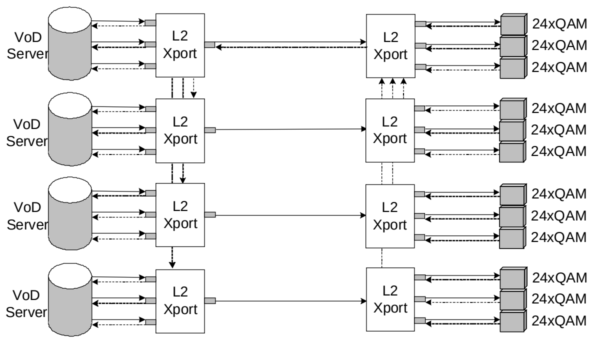


Figure 5 - Asymmetric Reverse Path

Layer 2 aggregation concentrates the reverse path control traffic from the hub onto a single gateway element, which then places this reverse path traffic onto an available optical wavelength in the ring. All other wavelengths in the hub are receive-only and do not require transmitter optics. Likewise, at the head-end the wavelengths not used for reverse traffic are transmit only and do not need receivers. This tailoring of transport optics to better match the forward and reverse path bandwidths of VoD traffic can provide significant cost-per stream savings.

The optimal VoD network makes extremely efficient use of high capacity DWDM optical components and in particular optimizes the use of cost-effective solutions available today. Efficiency is gained through fully utilizing each optical wavelength by a combination of link overhead minimization and Layer 2 traffic aggregation. Essentially each wavelength is operated near theoretical maximum capacity.

Scalability

MSO network planners require networking equipment to be cost effective and scaleable for both large and small market regions. They also require that the equipment can satisfactorily support the deployment of diverse implementations within a single large network. Furthermore, deployments may begin with a few Gigabit Ethernet links and grow to many more as the number of provisioned subscribers increases. Equipment used for the early deployments must be cost-effective for the initial small number of links yet scale and remain cost effective in the growing network.

Switched broadcast

Switched broadcast is an application intended to solve the problem of delivering limitless content choices within the finite limitations of the provisioned infrastructure by minimizing bandwidth utilization in the transport and HFC networks. By this means only program content being watched is delivered over the network. Furthermore, content watched by multiple users is transported as a single copy over the optical network. Distributed Layer 2 switching performed by transport devices in the head-end and hubs effectively utilizes bandwidth by directing content needed by that hub over a shared optical ring. Without a switched broadcast solution, video content demand will eventually exceed the available bandwidth of the cable infrastructure. Deploying systems that can support switched broadcast services which future-proof the VoD network for this eventuality.

TRANSPORT OPTICS

The optical components and topologies of the VoD transport network are designed to provide the lowest cost while fulfilling the requirements for bandwidth, reach/distance, resiliency, and ease-of-use. Figure 7 shows a typical transport network for hypothetical 5-hub model.

Video streams from the VoD servers are passed through an Layer 2 Ethernet switch, video gateway elements which perform Layer

2 aggregation of video streams onto a DWDM optical network, optical multiplexors; single-mode optical fiber to connect to an adjacent hub, optical splitters, optical protection switches, optical demultiplexors, video gateway elements performing Layer 2 deaggregation, and Gigabit Ethernet QAM devices.

Bandwidth

The VoD transport network is designed to carry VoD streams of personalized video content for each viewer from VoD servers to destination QAM devices. Each VoD stream requires 3.7Mbps of payload bandwidth quantized as 188-byte MPEG-2 video packets. Digital packet switching using Ethernet has become the defacto standard transport mechanism for VoD because of its ubiquity and low cost. The typical packet encapsulation scheme adds 3.4% overhead and uses 7 MPEG-2 packets (1316 bytes) encapsulated over Ethernet (18 Bytes) over IP (20 bytes) over UDP (8 bytes). Table 1 shows the number of VoD streams, which can be carried over a single fiber using various transport mechanisms.

Table 1 illustrates the number of Gigabit Ethernet ports provisioned based on service offering and provisioned peak load. Note that for all services above 100 Titles + SVOD and a peak load greater than 1.5%, Gigabit Ethernet links serve as ideal containers. Furthermore, 3G optics are sufficient for up to 5000 Titles while 10G optics remain highly underutilized and therefore result in higher cost per stream.

	SVOD	100 Titles	100 Titles + SVOD	1000 Titles	5000 Titles	NPVR
Peak Load% of HP	0.5%	1%	1.5%	2%	4%	10%
Streams	100	200	300	400	800	2000
#GigE's@16QAM	0.6	1.3	1.9	2.5	5	12.5
#GigE's@16QAM	0.4	0.8	1.2	1.6	3.3	8.3

Table 1 - Optical Technology Comparison for "5-Hub" Model Assuming 10% Subscription Rate

Reach / Distance

The fiber distances between the head-end node and hub nodes in the VoD transport network introduce optical loss and dispersion which need to be accounted for to maintain error-free transmission. The optical losses are introduced at connector boundaries, splices, mux/demux elements, and predominantly in the fiber itself. Network engineers compute an optical link budget by subtracting the optical receiver sensitivity and optical losses from the optical transmitter minimum output power. A link budget with adequate margin ensures adequate optical power at the receiver; while a link budget with no or negative margin indicates the need for an optical amplification device (an EDFA) in the optical path. EDFAs (Erbium Doped Fiber Amplifiers) use active Erbium-doped fiber and a laser pump source to boost the optical signal in a fiber. EDFAs become essential elements in a VoD transport network with medium to long spans between hubs.

Dispersion in optical fiber will eventually result in unacceptably low bit error rates. Because dispersion cannot be compensated by EDFAs, the optical distance (also called reach) limit is determined by the dispersion characteristics of the optical transmitter. When the VoD distance requirements exceed the laser optical dispersion limit, network architects must either add O/E/O regenerators or add another head-end, both costly alternatives.

Resiliency

Resiliency is the networks' ability to recover and sustain traffic in the presence of fiber cuts or equipment failures. VoD optical transport networks should be designed with resiliency in mind to eliminate single points of failure and provide for redundancy of critical components and a means for automatic detection and failover. The asymmetric nature of VoD and cost per stream pressures favor optical protection architectures which allow differing levels of resiliency which can be chosen to optimally balance the cost per video stream against optical protection coverage and switchover times.

The telecom industry has provided very mature network topologies (SONET/SDH and 2-fiber and 4-fiber BLSR) and elements for achieving full redundancy of optical equipment and fiber and very fast protection detection and switching times (less than 50 milliseconds) that exceed the times needed for VoD. While these can be used to provide effective resiliency for VoD transport they do so at a high cost per video stream, add additional management complexity in managing an additional transport layer, and dictate a costly symmetrical ring network topology that does not match the inherent asymmetric nature of VoD transport traffic. Unlike telecom networks, cable networks do not have strictly defined redundancy requirements and the ability to choose and optimize redundancy cost/performance is desirable. VoD transport networks that are designed for the same telecom resiliency goals (for example hybrid networks designed to carry both voice and VoD traffic), these topologies can be very effective. However, while the burgeoning VoD network infrastructure demands favor the lowest cost approach, other resiliency schemes become more favorable.

An effective strategy for optical resiliency is to design protection for those elements whose failure would affect the greatest number of VoD users. Optical protection costs for VoD can be lowered by moving the protection for fiber cuts from the Layer 1 electrical layer (as is done with SONET) to the optical layer. This eliminates the added cost of 1:1 or 1:N electrical interface protection. Ideally, individual optical transmitter and receiver redundancy protection could be optionally provided, allowing network designers to make the cost/resiliency tradeoff per wavelength while still maintaining 100% protection for fiber cuts.

Ease of Use

The operational expense of the VoD transport network is minimized and system uptime maximized by providing an integrated network which is easy to install, replicate, and maintain. The interoperability and ubiquity of Gigabit Ethernet has made the interface to the VoD server, video gateway elements, and QAM devices inexpensive and easy to maintain. The optical mux/demux, splitters, and ADPs are passive optics requiring no on-line management and have very high reliability. The EDFAs and optical switches can either be deployed as unmanaged devices which power-up in a default state or can be on-line managed; for example via SNMP or a Network Management System.

Optical Elements

The optical elements used in the transport network are well understood and have been used in existing cable networks for years. The laser transmitter performs the electrical to optical conversion of data streams. To achieve the bandwidth required for VoD, several optical channels are used per fiber using Dense Wavelength Division

Multiplexing (DWDM). The key parameters for laser choice are cost, optical reach, and overall bandwidth. 1G lasers achieve the lowest cost per laser, but the limited bandwidth results in a relatively high cost per stream. 10G lasers have high bandwidth, but are costly and have distance limitations due to optical dispersion. 10G optics also have coarse cost granularity, requiring high incremental cost as additional bandwidth is added to the network. An optimal compromise of cost per stream and optical reach can be achieved using 3G lasers with extended optical reach.

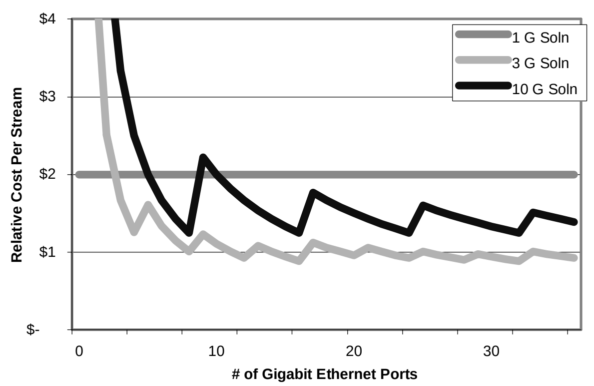


Figure 6 - Optical Laser Cost Comparison

Optical receivers are used for optical to electrical conversion of data streams. EDFAs (Erbium Doped Fiber Amplifiers) are used to provide amplification of all wavelengths on an optical fiber to offset fiber and connector losses and ensure that sufficient optical power is available for downstream receivers. Low cost passive optical mux/demux and splitter elements are used in the optical fiber distribution plant from head end to hub. Optical switches are used as protection devices to automatically switch to alternate fibers in a ring topology.

PACKET FORWARDING AND SWITCHING SCHEMES

Previous sections discussed and highlighted the subjective usefulness of Layer 1 and Layer 2 constructs and their relative benefits in developing a suitable VoD infrastructure. The following sections provide a detailed description of how traditional packet based constructs are used to fulfill the key capabilities described earlier.

Layer 1 Transport

For VoD network equipment not capable of Layer 2 and higher packet switching, Layer 1 provides a “direct wire” interconnect between the head-end and hub elements.

Layer 1 transport provides physical path connectivity between two or more endpoints on an optical link. Link information entering the input ports of a Gigabit Ethernet Layer 1 transport device are interleaved and encoded on an optical carrier. Optical wavelengths may accommodate one or more Layer 1 signals. Time Division Multiplexing generally accomplishes this. Layer 1 switching, commonly referred to as cross bar switching, offers some path flexibility by which source and destination Layer 1 ports can be cross connected arbitrarily. Layer 1 devices, however simple to provision, do not support fractional link aggregation and therefore typically underutilize optical transport capacity. For example, a Layer 1 device connected to a 600Mbps VoD server will have 60% optics and fiber utilization, while a Layer 2 device can achieve near 100% utilization by aggregating multiple fractional links. In addition, Layer 1 function does not provide shared wavelength-forwarding capabilities necessary for such schemes as switched broadcast. Layer 1 devices can only provide full duplex or simplex links. Operational complexity grows unreasonably with scale because transport assets cannot be dynamically reallocated or shared. Also, because there is no fractional rate support, asymmetric traffic patterns cannot be utilized for reverse trunking.

While Layer 1 interconnect simplifies the network topology for small-scale VoD networks, it fails to provide the benefits of bandwidth recovery, shared rings, and switched broadcast.

Layer 2 Capabilities

The Layer 2 behavior of the proposed high capacity transport network differs slightly from a conventional Ethernet LAN but is implemented in such a way that equipment costs are low and Ethernet interoperability is preserved. The following sections are a technical tutorial on the salient Ethernet functions used to achieve the desired VoD network features such as switched broadcast and shared rings.

Layer 2 classification and forwarding

Layer 2 classification and forwarding are the principal operations by which packets entering the transport domain are identified, organized and delivered to one or more destinations. The destination address contained in Ethernet packets are identified by the classification process. The result of this classification process is a list of forwarding destinations. The Layer 2 learning process ascertains the forwarding destinations. The Ethernet frame under consideration is encapsulated on the optical link and identified as a frame addressed to the downstream device listening on the selected optical wavelength. The act of identifying Ethernet frames based on their destination address, aggregating them with equivalent flows, and presenting them to downstream devices is called forwarding. By virtue of selective forwarding to interfaces based on destination information resident in the Ethernet packets, switching is accomplished. Switching forwarding and classification are used in the presented topologies to develop several classes of flows. The flow types are the following.

Path routes

Flows are grouped based on port affiliation and act like a virtual wire. Packets entering a physical port on a Gigabit Ethernet transport device are grouped together and transferred to a corresponding destination port over the fiber optic plant.

Layer 2 groups

Layer 2 groups are formed by the aggregation and dissemination of traffic to and from multiple sources. Filtering is performed and decisions are made about which packets go where on a packet-by-packet basis. The use of multicast addressing provides capability to forward the single packet to one or more destinations.

Tunneling

Tunneling is a mechanism used to trunk equivalent flows to a common end destination. Tunneling allows Layer 2 devices to forward aggregated flows on a shared virtual medium. Packets are grouped together and transported through the network as "equivalent" flows. The terminal transport node disseminates the flows through classification and delivers each packet to the described destination.

Layer 2 filtering

Layer 2 filtering works in conjunction with the classification process and provides a mechanism to scope and restrict traffic flows. Filtering can be configured to drop frames based on destination address or matched filtering criteria. A common use is to manage unknowns in the network. Unknowns are packets for which the destination is not present, or is unreachable in the network. Unknown filtering is useful in defeating broadcast storms and forwarding loops that can result in service disruption due to excessive bandwidth consumption.

Layer 2 add, drop, pass

Layer 2 add, drop, pass capability allows the formation of packet rings and provides a basis for multicast and broadcast over a single wavelength. Layer 2 flows are injected and terminated by members of the optical ring. Multiple members of the ring can receive a singular flow. Shared optical wavelength paths, used in conjunction with "star over ring" paths, provide a way to organize transport bandwidth based on steady state and transient load. Star topologies deliver the basis bandwidth which is the steady state load. Ring topologies provide a mechanism for bandwidth leveling and a shared medium for delivery of content assets, and switched broadcast services. This hybrid approach offers the ability to manage and mitigate transient bandwidth demand with little or no over provisioning. Both star and ring connections may exist within the same fiber or within tunnels on the same wavelength

Layer 2 Path and Flow Aggregation

Layer 2 Path and Flow Aggregation are used to concatenate traffic and fully utilize optical paths. They also provide the ability to utilize aggregation to virtual domains for transport of more granular traffic flows such as reverse path traffic. Asymmetric reverse path forwarding is an example of Layer 2 flow aggregation that exploits the traffic patterns in VoD transport networks.

Layer 2 asset provisioning and load balancing

A significant advantage gained through Layer 2 switching is the ability to ideally match content delivery assets to transport infrastructure capability. Switching allows equipment providing session fulfillment, such as video servers, to reach any endpoint in the network. This allows servers to share the workload in delivering VoD sessions.

Layer 2 address learning and aging

This function utilizes Ethernet source address awareness and classification to determine destinations for packets impinging the network (i.e. you don't have to know which downstream port to plug the QAM modulator into learning will find which port its plugged into). Learning also provides information to forwarding/filtering functions to automatically determine which end points are reachable by the network.

SUMMARY

The VoD architectures presented in this paper can be summarized by considering the high capacity 5-hub 300,000 homes passed VoD network model presented earlier. The key architectural and cost advantages of this model are high VoD stream capacity per fiber, long optical reach, automatic optical protection switching, and a low-cost asymmetric reverse path.

The optical transport system is constructed as a hybrid ring/star architecture. A northbound ring and a redundant southbound ring connect the head-end and five hubs. Each ring has 100GHz spaced

DWDM 3Gbps wavelengths, which in aggregate carry, over 33,000 VoD streams. Transmitter/receiver optical link budgets exceeding 32dB are achievable with low-cost laser transmitters and six spans each exceeding 25km without the need for regeneration and using two EDFAs per ring. A lower cost per stream point could be achieved with shorter reach optics, but would add more network design complexity and limit the flexibility of deploying the same architecture in nearly all VoD deployment areas.

The duplicate rings provide 1:1 protection for fiber cuts anywhere along the transport ring as well as failure of any of the optical splitters and EDFAs used on the ring. By monitoring the recovered optical power at the destination, the optical protection switches can failover without any user intervention. The topology can be constructed with an asymmetric reverse path where the head-end uses mostly transmitter-only optics and the hub-end uses mostly receiver-only optics. This reduces the optical transceiver components (which can be 20% to 40% of the overall transport network cost) nearly in half.

In the hybrid ring/star topology, the Ethernet switch at the head-end forms a star Ethernet connection between the VoD servers and the QAMs. The gateway elements then aggregate the VoD streams onto an optical wavelength, each of which is connected via the optical transport ring. A corresponding QAM at a hub-end demultiplexes the VoD streams from the optical wavelength and switches individual VoD streams to the appropriate QAM device. This allows any VoD stream from any VoD server to be directed to any head-end QAM output.

Gateway elements vary in complexity from simple Layer 1 devices to fully featured Layer 1-4 devices capable of advanced features such as switched broadcast. A Layer 1 device functions as a “wire” connecting entire Gigabit Ethernet ports (with all VoD streams remaining intact) from the head-end Gigabit Ethernet switch to a hub-end QAM. Adding Layer 2 capabilities into the gateway elements allows individual Ethernet frames (and the VoD streams they are carrying) to be individually switched and aggregated which gives much finer granularity in provisioning traffic from the VoD servers to the hub-end QAMs. This allows advanced features such as traffic aggregation and load balancing which can save VoD costs by reducing the required number of Gigabit Ethernet switch and QAM ports.

In conclusion for a VoD transport network to be cost effective and scaleable, it must be more than just a large pipe. Optimal use of fiber bandwidth and reducing the cost of transport optics are essential to reducing per stream costs. Layer 2 VoD stream aggregation reclaims fallow bandwidth. Asymmetry in the reverse optical path reclaims fallow fiber bandwidth and cuts the cost of DWDM transmitters and receivers nearly in half.

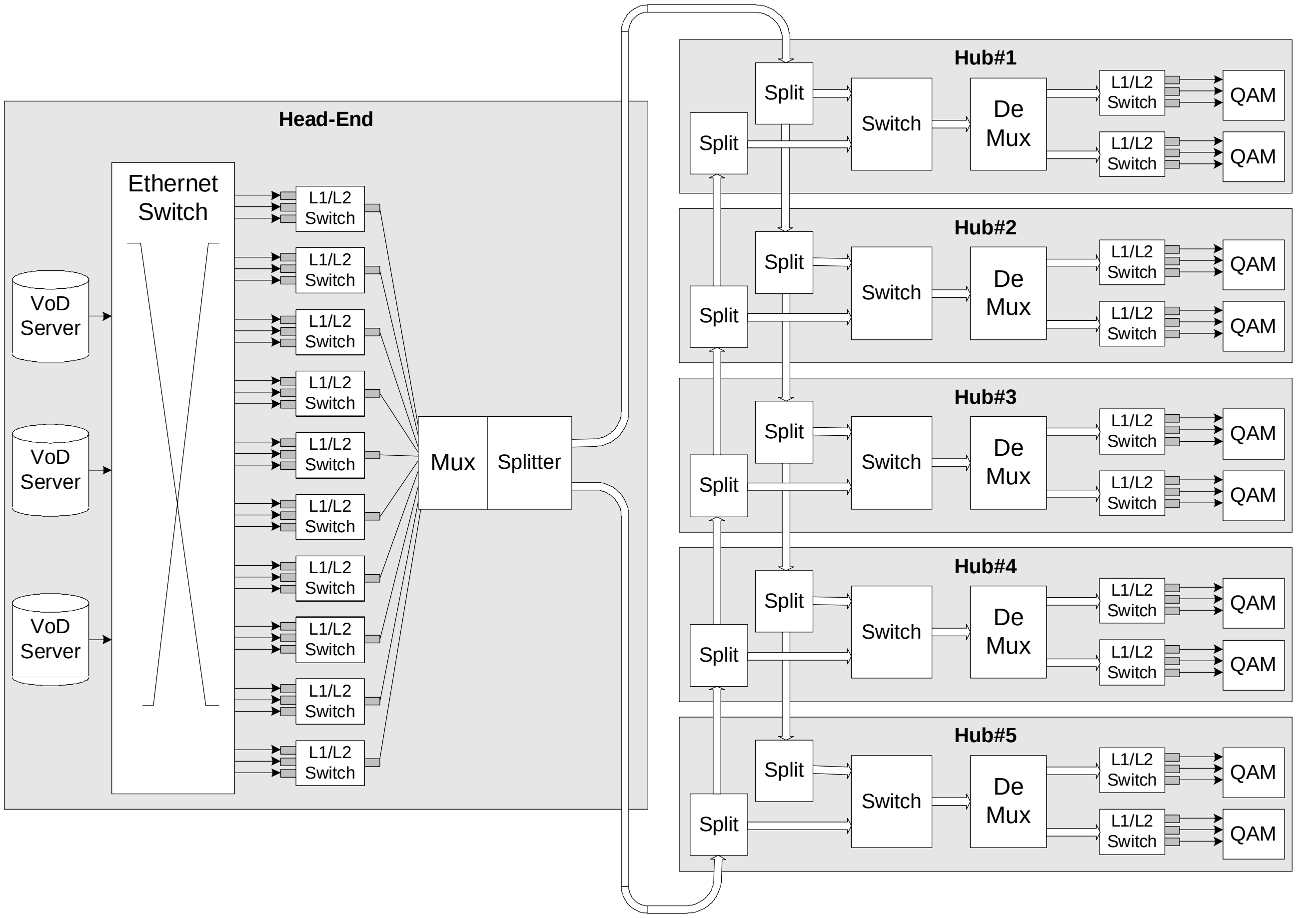


Figure 7 - 5-Hub VoD Network Model

CONTACT INFORMATION:

John Amaral
Artel Video Systems
237 Cedar Hill Street
Marlborough, MA 01752
jamaral@artel.com
508-303-8200 x285

Paul Pilotte
Artel Video Systems
237 Cedar Hill Street
Marlborough, MA 01752
ppilotte@artel.com
508-303-8200 x304

QUALITY OF SERVICE OVER HOME NETWORK USING CABLEHOME™ 1.1

Amol Bhagwat
Cable Television Laboratories, Inc

Abstract

CableHome™ 1.1, the latest version of the CableHome specification by CableLabs®, has defined a Quality of Service (QoS) solution for home networks. The key challenges in designing the CableHome 1.1 QoS system were: varying degrees of QoS support from home-networking technologies, support for legacy home LAN devices, and backward compatibility with CableHome 1.0. The CableHome team specified a priorities-based QoS system in the CableHome 1.1 specification that addresses these key challenges. The main functionalities of this QoS system are prioritized queuing, prioritized media access, and provisioning of application specific priorities. The first functionality resides only in a residential gateway whereas the later two are part of both a residential gateway as well as home LAN devices. Provisioning of application specific priorities is a very simple process. Hence the CableHome 1.1 QoS solution is easy to deploy and implement by cable operators. In addition, since this design allows legacy home LAN devices to co-exist with complaint QoS-enabled devices it is a convenient solution for consumers too.

INTRODUCTION & BACKGROUND

CableLabs, a research and standards development consortium for the cable industry, has initiated the CableHome project at the direction of its member cable television companies. The project is aimed at developing a managed infrastructure that enables cable operators to offer high-quality, value-added broadband services to their subscribers over any available home-networking technology in a seamless and convenient manner. The CableHome 1.0 Specifications were released in April 2002 and certification testing of the products began in October 2002. CableHome 1.0 specifications also gained international acceptance via International Telecommunications Union (ITU) in 2002 when ITU document J.191, that adopted CaleHome 1.0 specifications almost entirely, was consented as a fully approved ITU recommendation.

The CableHome 1.0 specification standardizes functionality for a residential gateway device that simplifies manageability of subscriber's home network [1]. The following are some of the key features offered by CableHome 1.0:

1. Remote configuration and management of residential gateway in a secure manner.
2. Hands-off authentication and provisioning of residential gateway.
3. Application and cable friendly standardized NAT/NAPT
4. Secure download of software images
5. Firewall management and rule set download.
6. Remote home LAN devices visibility and connectivity tests
7. Local Name Service
8. Protection of cable network from home network traffic

A follow on version of the specification is CableHome 1.1. In addition to enhancing capabilities of the residential gateway, CableHome 1.1 extends its reach beyond the residential gateway to devices in the home LAN. CableHome 1.1 specifies a new set of functionalities for home LAN devices that enable several new key features including: 1. Quality of Service (QoS) over home networks, and 2. Device and services discovery. In general new capabilities that CableHome 1.1 enables are as follows:

1. Standardized firewall configuration
2. Configuration file authentication
3. Simple Parental Control
4. Static Port Mapping
5. VPN Support
6. QoS over the home network
7. LAN Management Messaging
8. Device and Services Discovery

This paper focuses on the home network QoS functionality designed in CableHome 1.1. The paper first discusses key challenges involved in designing a generic QoS system that could be overlaid on any OSI layer-2 home-networking technology in the second section. The third section describes in detail the CableHome 1.1 QoS solution and how cable operators can implement it. The implications of this QoS solution are discussed in the fourth section and the last section presents the conclusions.

CHALLENGES IN DESIGNING A GENERIC QoS SYSTEM OVER HOME LAN

A quality of service system over home networks can be provided via three main functionalities:

1. **Management of shared media access:** When multiple devices are sharing the same transmission media some mechanism is required to manage the access to this media. This involves manageability of various traffic QoS characteristics such as traffic priorities, bandwidth, jitter, and latency. In order to implement such management a certain set of functionality needs to reside in a residential gateway as well as in home LAN devices to be able to manage and obey these characteristics.
2. **Packet Forwarding and Queuing**: This is a functionality of a residential gateway or a bridge in which packets arriving at multiple interfaces are to be retransmitted through another outgoing interface. This functionality needs to be enhanced so that packet forwarding is performed to meet the necessary QoS requirements.
3. **Management of QoS Characteristics**: This functionality deals with assignment of QoS characteristics to various devices and applications in the home and remote manageability of these characteristics. This functionality is a part of both residential gateway and home LAN devices.

There are two main QoS paradigms that can be utilized to provide the aforementioned functionalities: **parameterized** (planned, guaranteed) QoS and **prioritized** (differentiated) QoS.

- **Prioritized QoS**: The prioritized QoS paradigm entails providing differentiated shared media access to the traffic based on priorities and prioritized queuing and forwarding in a residential gateway and in a bridge. This mechanism does not provide performance guarantees for QoS parameters such as bandwidth, jitter and delay.

- **Parameterized QoS**: In this paradigm, performance guarantees for QoS parameters can be provided to the traffic over the network. This is a planned approach for allocating resources on a network. Such planning is done based on the prior knowledge of resource requirements of various devices and applications in the network.

There are pros and cons for each of these paradigms and it was necessary that the methodology chosen for CableHome 1.1 QoS solution satisfy the requirements set forth by cable operators for CableHome 1.1. The key cable operator requirements for CableHome 1.1 QoS solution were:

- It should be able to support legacy home LAN devices and best effort traffic such that they can coexist with new QoS-enabled devices.
- It should be OSI layer-2 home-networking technology independent
- It should be software upgradeable from CableHome 1.0

There were several challenges in fulfilling these requirements. The rest of this section is dedicated to discuss these challenges.

Varying Degrees of Qos Support From Different Standards Based OSI Llayer-2 Home-Networking Technologies

The requirements for the CableHome 1.1 QoS solution mandated that cable operators should be able to overlay the QoS system on any standards based OSI layer-2 technology. This requires that the QoS system should be designed strictly at OSI layer-3 and above. Due to this fact such a system is dependent on the underlying home-networking technology for its QoS support at the MAC layer. However, the support for QoS in different standards based home-networking technologies varies from technology to technology. It is essential to assess this support in order to design a QoS system that could be OSI layer-2 home-networking technology independent and is still realistic. *(See Appendix 1 for information on QoS support in leading standards based home-networking technologies.)*

Shared vs. Point-to-point Media

With respect to QoS considerations different OSI layer-2 home-networking technologies can be categorized into two main categories: point-to-point technologies and shared media technologies. For a point-to-point technology there is a direct connection between two devices that are communicating with each other, e.g. Switched Ethernet. However, in case of shared-media technologies all of the devices share the same media for all of their communications. Most of the home-networking technologies such as 802.11 a/b/g, HomePNA, HomePlug, are shared-media technologies. For such shared media technologies some mechanism is required to control how and when devices transmit data on the media. This can be achieved by employing either parameterized or prioritized QoS paradigm.

Support for Prioritized QoS

Most of the standards based shared media technologies- 802.11 a/b, HomePNA and Powerline (HomePlug) have support for priorities based QoS scheme. 802.11 a/b and HomePNA supports 802.1p/q [2] priorities while HomePlug has native priorities support. In general, for these technologies, prioritized media access is accomplished by providing preferential media access for higher priority traffic. The highest priority traffic gets first opportunity to transmit its data on the shared media and then, depending upon the bandwidth availability, lower priority traffic gets subsequent opportunities to send their data.

Support for Parameterized QoS

The amount of bandwidth consumed by higher priority traffic cannot be controlled by using a prioritized scheme. A parameterized scheme is necessary for such a control.

Parameterized QoS requires that the underlying PHY/MAC technology be able to deliver constant bandwidth and jitter. It is very difficult to achieve this for home networking technologies based on wireless, phoneline, and powerline as underlying throughput and jitter can be strongly influenced by rapidly changing interference. Perhaps due to these reasons, at the time when CableHome 1.1 QoS system was being designed, none of the standards based home-networking technologies supported a truly parameterized QoS scheme.

Special Case of Ethernet:

Most existing Ethernet hubs in home LANs today do not support either a 802.1p/q priority scheme or a parameterized scheme and it is more than likely that Ethernet will not support these capabilities in future. However, when CableHome 1.0 is deployed in a consumer's home, existing hubs are likely to be replaced with switches that are integrated in the CableHome 1.0 residential gateway devices. For switched Ethernet, differentiated media access is not of much value, in many cases; since traffic is essentially point-to-point and it is likely that such a link is less contentious. Finally, 100Mbps bandwidth seems to be sufficient to address most of the needs of home networking applications, especially when it is for each point-to-point link. Hence QoS functionality adds very little value in the case of CableHome residential gateways that have Switched Ethernet interfaces. Thus while designing CableHome 1.1 QoS solution Ethernet was considered as an outlier among other available shared media home-networking technologies.

Supporting Legacy Home LAN Devices and Best Effort Traffic

A key cable operator concern was that newly designed CableHome 1.1QoS system should not incur substantial inconvenience to either customers or to cable operators when it has to co-exist with legacy devices. Additional cost for hardware or software upgrades of legacy home LAN devices so that they can coexist with QoS-capable CableHome complaint devices was considered highly undesirable, e.g. requiring a "QoS adapter" for best effort devices adds additional cost and is inconvenient for the customer. Thus the challenge was to devise a QoS solution that will not require an upgrade to the legacy devices and will make sure that best effort traffic from these legacy devices will not interfere with the traffic from QoS-enabled devices in the home.

Prioritized Approach

A prioritized media access system can be overlaid on existing shared-media home networks. Even though most of the current standards based home-networking technologies support prioritization of traffic, in general, these prioritization schemes are not consistent and there is no central entity managing the priorities in the home. A residential gateway in the home LAN can perform the function of priority assignment, on behalf of a customer, for various applications and devices, at the direction of cable operators. Thus, if priorities based QoS functionality is added to a residential gateway and to new compliant home LAN devices, then traffic originating from the these devices can utilize priorities to take advantage of the prioritized media access capabilities of underlying OSI layer-2 home-networking technology. Traffic from legacy non-compliant home LAN devices will continue to use best effort priority and therefore typically will not interfere with the media access opportunities for prioritized traffic from compliant devices. Thus with a prioritized QoS system compliant CableHome devices as well as legacy non-compliant devices can co-exist in the home network without compromising the integrity of QoS for the applications that are taking advantage of the QoS system.

Legacy (non-compliant) devices do not have a means of requesting and using media access priorities for the packets. Thus these devices cannot perform prioritized media access while transmitting their data. However, with manual (operator or consumer) set-up of priorities for legacy devices in a residential gateway, it can be instructed to perform prioritized media access for traffic that is destined to legacy device. Thus prioritized QoS can be provided for a sink-only legacy home LAN device, sinking traffic from a residential gateway. Also with such manual settings a residential gateway can perform prioritized queuing for traffic to and from a legacy device going through the residential gateway.

Parameterized Approach

The parameterized QoS paradigm entails planned opportunities for media access and queuing. This requires that all of the devices and applications on the home network convey their requirements for various parameters such as bandwidth, jitter, and delay to a centralized network controller such as residential gateway. When a device or an application needs to transmit data over the network it sends a request to this centralized controller (termed as Admission Controller) [3]. Based upon the set policies and available network resources, an admission controller either accepts or rejects the request in such a manner that guaranteed QoS could be maintained over the network. However, such QoS guarantees can be provided only if all the devices in the network obey decisions of the admission controller.

Through specifications, new complaint home LAN devices can be instructed follow the process of sending their QoS parameter requirements to the centralized admission controller (implemented in the residential gateway) and to follow its decisions before sending traffic over the network. However, using available standards based OSI layer-2 home-networking technologies residential gateway cannot have any control on legacy home LAN devices as to when they should send the traffic and how much; unless the legacy devices are upgraded with hardware or software addition that instructs them to obey the admission controller in the network. Thus, without this hardware or software adapter, legacy home LAN devices will interfere with the planned transmitting intervals of complaint devices and as a result will compromise parameterized QoS system over the home network. Without appropriate support for legacy devices from underlying OSI layer-2 home-networking technologies this limitation could not be overcome while implementing parameterized scheme. This particular limitation of parameterized QoS paradigm is very undesirable from cable operator and consumer convenience point of view.

<u>Software Upgradeability of Existing Cablehome 1.0 Devices</u>

One of the key overriding requirements for CableHome 1.1 was that it be software upgradeable from CableHome 1.0. This would enable cable operators to upgrade CableHome 1.0 residential gateway devices in the field to CableHome 1.1 via remote download of a new software image, thus avoiding the need for a truck roll. This gives substantial cost advantages to cable operators, allows new complaint functionality to evelove, and enables them to offer new products and services with CableHome 1.1. Thus a CableHome 1.1 QoS system should be such that any newly specified residential gateway features could be added to the CableHome 1.0 device using just software implementation.

Upgradeability of Prioritized Approach

If a prioritized QoS paradigm were to be employed for CableHome 1.1 additional features that need to be added in the CableHome 1.0 residential gateway would be: prioritized media access, prioritized queuing and management of QoS priorities over the home LAN. To perform prioritized media access the residential gateway could be software upgraded to set priorities for packets transmitted on the shared interfaces. Similarly, prioritized packet queuing and forwarding could be accomplished through a software upgrade. For queuing and forwarding functionality, additional network protocol stacks are not required but additional processing steps would be required for existing CableHome 1.0 packet forwarding process. Management of QoS priorities would require additional MIBs in the residential gateway to store priorities and additional software to manage and communicate those priorities to various complaint home LAN devices connected to the residential gateway. The software footprint of this functionality can be minimized by using the same communication protocols as those used for the in-home device management and discovery features of CableHome 1.1.Thus, the various required features of a prioritized QoS system can be implemented applying software upgrade to a CableHome 1.0 residential gateway. Preliminary estimates indicate that the footprint of such implantation doesn't seem to be substantial enough to warrant a hardware upgrade.

Upgradeability of Parameterized Approach

If a parameterized QoS system were to be specified for a CableHome 1.1 residential gateway, implementation of an admission controller feature would be required as well as a communication protocol to communicate parameter requirements and network admission decisions. Additional MIBs would also need to be implemented to store and manage the QoS parameters required for applications on the home network. The admission controller feature could be implemented using Subnet Bandwidth Manager (SBM) [3] and QoS parameter communication and reservations could be performed using the RSVP [4] network protocol stack. Accurate estimates of the software footprint of these additional protocol stacks for residential gateway weren't available. However, it was clear that comparatively it is heavier than the functionality required for prioritized QoS system. Thus it was uncertain if parameterized QoS system could be implemented on existing CableHome 1.0 residential gateways by just upgrading software.

THE CABLEHOME 1.1 QoS SOLUTION

After analyzing different challenges in providing in-home QoS as well as key requirements for CableHome 1.1 QoS system, prioritized QoS paradigm was chosen as the most appropriate solution. Due to the lack of adequate support for parameterized QoS from underlying OSI layer-2 home-networking technologies parameterized QoS based solution for CableHome 1.1 seemed unrealistic. In addition, a QoS system based on a parameterized scheme would potentially require additional hardware or software upgrade to legacy best effort devices to maintain the integrity of in-home QoS. Taking into consideration these facts a priorities-based QoS solution was specified in the CableHome 1.1 specification. The later part of this section describes in detail various elements of the CableHome 1.1 QoS solution.

CableHome 1.1 QoS Architecture

The CableHome 1.1 QoS architecture consists of various logical elements and sub-elements as shown in Figure 1:

1. Portal Services Element (PS): This is a logical element in the CableHome architecture [1] that represents CableHome specified functionality within a residential gateway device.

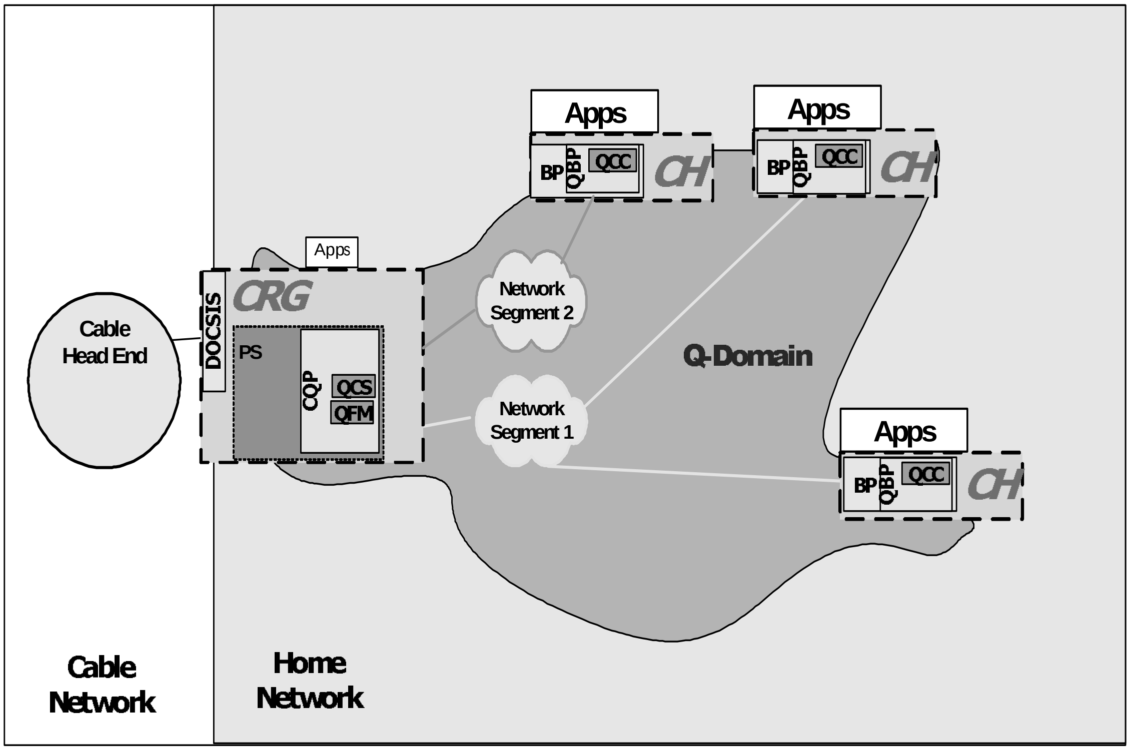


FIGURE 1: CableHome 1.1 QoS Architecture

2. Boundary Point Element (BP): This is a logical element in the CableHome architecture that represents CableHome specified functionality within a Home LAN device.

3. CableHome QoS Portal Sub Element (CQP): CQP is a sub element of the PS logical element. The CQP acts as a CableHome QoS portal for CableHome compliant applications. Its primary function is to enable priorities based QoS for the devices within the home network. It performs priorities based queuing/forwarding and media access for the traffic originating from the PS as well as for the traffic transiting through the PS. It is also responsible for communication of QoS characteristics to various devices within the home (described later in this section).

4. QoS Boundary Point Sub Element (QBP): QBP is a sub element of the BP logical element. It performs priorities based media access for the traffic originating from the BP. It is also responsible for the reception of QoS characteristics from the PS.

In addition, these logical elements described above contain QoS related functionalities (QFM, QCS, and QCS) that are described later in this section.

CableHome Priorities

CableHome 1.1 defines the following three different types of CableHome QoS priorities:

1. CableHome Generic Priorities
2. CableHome Queuing Priorities
3. CableHome Media Access Priorities.

CableHome Generic Priorities:

CableHome 1.1 introduces the concept of Generic Priorities. This is primarily due to the fact that OSI layer 2 priority approaches are not consistent as the number of priority levels supported varies from technology to technology. A generic priorities scheme gives cable operators a consistent approach, which is abstracted from the particular OSI layer 2 home-networking technology. In addition, this single generic priority can serve to indicate both media access priorities, as well as queuing priorities (described below).

CableHome 1.1 defines eight CableHome Generic Priority levels, 0 through 7, 7 being the highest and 0 being the lowest. Cable operators assign one of these eight priorities to an application. Application is identified using an application ID, which could be an IANA assigned port number for the application [5]. Of the three types of priorities defined by CableHome, a cable operator sets only the CableHome Generic Priority value for an application based on its ID. The other two priorities - CableHome Queuing Priorities and CableHome Media Access Priorities - are derived from this CableHome Generic Priority depending on the capabilities of the hardware and software in the device.

CableHome Queuing Priorities:

Packets can be transmitted from multiple incoming interfaces to single outgoing interface in the residential gateway. Hence each interface implements a queuing function. In order to provide prioritized QoS for in-home traffic passing through the PS, CableHome specifies prioritized queuing functionality per physical interface in the PS. A physical interface will have one or more queues associated with it and each individual queue is designated with a certain queuing priority. This is defined as the CableHome Queuing Priority. The CableHome Queuing Priority needs to be identified for each packet to be transmitted on each PS interface so that the packet can be placed in an appropriate queue. This CableHome Queuing priority is derived from the CableHome Generic Priority using the number of queues supported per interface on the PS. Implementation of number of queues per interface is vendor specific.

CableHome Media Access Priorities:

This is the media access priority of a packet and is derived from its CableHome Generic Priority based on the number of media access priorities supported by interface's layer-2 shared media technology. Since the number of priorities supported by different OSI layer-2 home-networking technologies varies, such mapping is necessary. CableHome Media Access Priority values are logical levels that represent a level of preference that a packet should receive for media access.

CQP QoS Functionality:

The CableHome QoS Portal (CQP), which resides in the PS element, consists of two main functionalities as shown in figure 1: QoS Forwarding and Media Access (QFM) and QoS Characteristics Server (QCS).

QoS Forwarding and Media Access (QFM):

The QFM element provides the PS with a mechanism to order and transmit packets out of a PS interface to a LAN host according to assigned priorities. The PS exercises QFM functionality on any packet that is transmitted out of the PS on any LAN interface. The QFM performs following three actions on the packet once it is received in the PS:

1. **Packet Classification:** The PS examines the destination IP address and destination port number of the packet. Using these values the PS looks up a corresponding CableHome Generic Priority for the packet from the classifier table stored in the PS database. If no matches are found for that destination IP and port, then the PS assigns priority 0 to the packet.

2. **Prioritized Queuing:** The PS then maps CableHome Generic Priority of the packet to CableHome Queuing Priority based on the number of queues implemented for the interface on which the packet is to be transmitted. Multiple queues implemented for the interface are designated with different CableHome Queuing Priorities. The PS puts the packet in an appropriate queue based on its queuing priority. The QFM polls all of the queues on each interface according to their priorities to extract packets. The packets are extracted from the queuing system by employing a methodology of First in First Out with Priorities, Highest Priority Queue First.

3. **Prioritized Media Access:** After the packet is extracted from the set of queues associated with an interface, the packet needs to be transmitted on the shared LAN media with the appropriate media access priority. The QFM performs the mapping of the CableHome Generic Priority value of the packet to the CableHome Media Access Priority. The packet is then transmitted on the shared media with the appropriate level of preference as indicated by the CableHome Media Access priority.

QoS Characteristics Server (QCS):

The QCS element provides a mechanism for the cable head-end to communicate desired QoS Characteristics (for particular applications) to the PS and then further to BPs in the home. In CableHome 1.1 QoS characteristics refer to priority information for different applications over the home network. The overall functioning of the QCS is explained below.

1. **Application Priority information to the PS**: The cable head-end provides mapping of application IDs to CableHome Generic Priorities to the PS either using a PS configuration file or via SNMP MIB interface. This mapping in the PS serves as a master table in determining priorities for various applications or services on the home LAN.

2. **BP Application Information to the PS:** The QCS receives information about the applications associated with a BP in the form of an XML message, called the BP_Init message, which is sent using SOAP over HTTP [6]. This message contains the list of application IDs that a BP supports. It may also contain a list of destination IP address and port number pairs for which a particular application on the BP likes to request destination specific priority. Such a request for destination IP and port specific priority is sent by a BP to the PS after an application session has been established.

3. **Application Priority Information to the BP**: Upon receipt of the application

information from the BP, the QCS consults the priority master table provided by the cable operators and determines appropriate priorities for different applications on the BP. If there is no entry for a particular application in the priority master table, then the QCS assigns a default priority of 0 (best effort) for that application. QCS also determines the destination specific priority information as requested. Both these priorities are determined using applications IDs. The QCS sends this updated priority information to the BP in the XML format using BP_Init_Response message (SOAP over HTTP). The QCS also stores all of this updated priority information in the PS database (which is accessible to cable operators via the MIB interface).

Thus through these three main processes the QCS manages and communicates priority information to various applications on the home network

QBP QoS Functionality:

The QBP is a logical sub element of a BP that resides in a CableHome compliant home LAN device, termed as CableHome Host. The QBP consists of only one QoS functionality: QoS Characteristics Client (QCC).

QoS Characteristics Client (QCC):

The QCC has two main responsibilities: obtaining application priority information from the PS and using this priority information for prioritized media access. These two functions of the QCC are explained in detail in the subsequent paragraphs.

1. **Requesting priority information to the PS:** As explained earlier in the QCS section, the BP sends its application information to the PS in the BP_Init Message. The QCC entity in the BP is responsible for that message exchange. Also if an application needs a specific destination IP address and/or port specific priority, then the QCC sends a request for such destination IP and port priority in the BP_Init Message, after the application on a BP establishes a connection with another application. In addition, the QCC is responsible for communicating to the PS any updates (addition/deletion) to the application information in the BP. After the PS sends updated application priority information to the BP, the QCC makes sure that priority information for applications on the BP gets updated appropriately.

2. **Prioritized Media Access**: Once the application on the BP starts communicating, the QCC uses the priority assigned to it for prioritized media access. If a destination IP and port specific priority is requested then QCC uses destination specific priority otherwise it uses the default priority assigned to the application. The QCC maps the CableHome Generic Priority to CableHome Media Access priority based on the number of media access priorities supported by underlying layer-2 home–networking technology and then delivers the packet on the shared media.

IMPLICATIONS OF THE CABLEHOME 1.1 QoS SOLUTION

As explained in the earlier sections, the CableHome 1.1 QoS system is a simple and elegant solution that cable operators can provide on CableHome devices easily. The only additional provisioning step that cable operators need to perform is provisioning of the application priorities master table in the residential gateway. Applications on compliant CableHome devices will receive appropriate priority information that they can use for subsequent management of prioritized traffic flow. With CableHome 1.1 QoS, legacy home LAN devices can co-exist with complaint QoS-enabled devices without collapsing QoS of the entire network. Thus it is a convenient solution for both consumers and cable operators. Minimal additional functionality is required to implement CableHome 1.1 QoS on residential gateway and home LAN devices.

The benefits of the CableHome 1.1 QoS solution are significant and compelling, however there are a number of implications associated with the chosen approach. Since this QoS solution is based on a prioritized paradigm it does not provide absolute guarantees for QoS parameters such as bandwidth, jitter, and delay. It provides preferential media access to certain traffic, classified as higher priority. Thus two traffic streams that have the same priority would contend with each other for media access and thus they might get best effort treatment between themselves. Also, if a top priority application consumes the entire bandwidth of the home network then it is possible that access to the shared media could be denied for all applications. Considering the bandwidth provided by typical home-networking technologies today and services that CableHome 1.1 plans on enabling it seems that this scenario is unlikely. However, in future if cable operators decide to offer bandwidth intensive services such as video distribution over the home network, this scenario may occur and in that case the applicability of a prioritized QoS scheme may need to be revisited. [See Appendix 2 for typically expected QoS requirements for various applications and services]

CONCLUSIONS

CableLabs has defined priorities based QoS solution for its home networking project-CableHome 1.1. The CableHome 1.1 QoS solution can be deployed on any layer-2 home-networking technology that supports priorities. Cable operators and consumers can take advantage of this solution in a convenient and seamless manner, as it does not require any additional hardware or software upgrade for legacy devices in order to maintain QoS over the home network. Additional functionality required to implement this solution is minimal. Hence it is attractive and cost effective for vendors to build products for this specification. Cable operators can provision QoS for their applications in the home network by a simple configuration of application specific priorities in the CableHome 1.1 based residential gateway. Thus, CableHome 1.1 QoS is a simple, cost-effective and easy-to-use solution that enables cable operators and consumers to take advantage of QoS over the home networks.

AUTHOR CONTACT INFORMATION

Amol Bhagwat
OSS Engineer, Broadband Access
Cable Television Laboratories, Inc
400 Centennial Parkway
Louisville, CO 80027
P: 303-661-3333
F: 303-661-9199
E: a.bhagwat@cablelabs.com

ACKNOWLEDGEMENTS

The author wishes to acknowledge the following individuals for their valuable contribution, inputs and assistance.
Ralph Brown, Steve Saunders, Kevin Luehrs, Stuart Hoggan, Jennifer Doran, Liz Weeks

APPENDICES

Appendix 1: Features summary of various leading standards based OSI layer-2 home-networking technologies

Home Networking Technology	**Specifications and Standards Group**	**PHY Layer Modulation**	**Data Rates**	**QoS Capabilities**
Ethernet	IEEE 802.3	Baseband	10Mbps,100Mbps, & 1Gbps	None
Wireless LANs	IEEE 802.11a	OFDM	54Mbps	802.11 e Working group (Prioritized & Parameterized Proposals)
	IEEE 802.11b	DSSS	11Mbps	
	IEEE 802.11g	OFDM	54Mbps	
Powerline	HomePlug 1.0, Home Plug Powerline Alliance	OFDM	10Mbps	Prioritized
Phoneline	HomePNA 1.0 & HomePNA 2.0, Home PhoneLine Networking Alliance	FDQMA	1Mbps & 20Mbps respectively	Prioritized

Appendix 2: Expected QoS requirements for various applications and services

Service	Input Parameters of Performance Testing				Output Parameters of Performance Testing		
	Number of Streams	Payload Rate (per stream)	Header Type	Packet Size (bytes)	Max PER	Max Latency (ms)	Max Jitter (ms)
HQ Voice Calls	2 per call	64 kb/s	IP/UDP/RTP	120	$1.5*10^{-3}$	10	+/-5
MQ Voice Calls	2 per call	8 kb/s	IP/UDP/RTP	80	$1.5*10^{-3}$	30	+/-20
HQ Video Conference Call	2 per call	1.5Mb/s	IP/UDP/RTP	228	$3.6*10^{-5}$	10	+/-5
HDTV	1	19.68 Mb/s	IP/UDP/RTP	228	$3.6*10^{-5}$	90	+/-10
SDTV	1	3 Mb/s	IP/UDP/RTP	228	$3.6*10^{-5}$	90	+/-10
CD Quality Audio	1	256 kb/s	IP/UDP/RTP	360	$5.8*10^{-5}$	100	+/-10
High Speed Data	1	10 Mb/s	TCP/IP	1540	0	>100	>100
Med. Speed Data	1	2 Mb/s	TCP/IP	1540	0	>100	>100
Low Speed Data	1	500 kb/s	TCP/IP	1540	0	>100	>100

NOTES:

1. Voice Packet = (IP/UDP/RTP Header) + (voice payload)

IP/UDP/RTP Header: 40 bytes (20 bytes IP Header + 8 bytes UDP Header + 12 bytes RTP Header) without RTP header compression. If RTP header compression is applied header reduces to 2-4 bytes. In this table we assumed no RTP header compression.

Voice payload: variable size depending on codec, considering the end-to-end latency budget, typically 10-40 ms voice samples can be used. Given the Max Latency/Max Jitter in HN portion and to keep packet overhead to its minimum, we assume 10 ms voice samples for HQ voice and 40 ms voice samples for MQ voice.

Video Packet = (IP/UDP/RTP Header) + (video payload)

IP/UDP/RTP Header = 40 bytes (20 bytes IP Header + 8 bytes UDP Header + 12 bytes RTP Header).

Video Payload size = MPEG packet size which is 188 bytes.

Data Packet = (IP/TCP Header) + (Ethernet payload)

Packet Error Rate (PER): is measured at MAC-SAP for packets delivered from MAC layer to higher layer. For data, since packets in error should be discarded and only error free packets are passed to the MAC-SAP, then for data PER=0.

REFERENCES

[1] CableHome 1.0 Specifications, Cable Television Laboratories, April 2003, http://www.cablelabs.com/projects/cablehome/specifications/

[2] IEEE Std-802.1Q, IEEE LAN MAN Standards Committee, December 1998

[3] SBM, A protocol for RSVP-based Admission control over IEEE 802-style networks, IETF RFC 2814, May 2000, http://www.ietf.org/rfc/rfc2814.txt

[4] Resource reSerVation Protocol, IETF RFC 2205, September 1997, http://www.ietf.org/rfc/rfc2205.txt

[5] IANA Port Numbers, http://www.iana.org/assignments/port-numbers

[6] SOAP Version 1.2, W3C Working Draft, World Wide Web Consortium (W3C), June 2002, http://www.w3.org/2000/xp/Group/#drafts.

ROUTERLESS AGGREGATION: CONVERGING DATA AND TDM NETWORKS

Benoit Legault
Dir. Technology Advisory Board
ADC

Abstract

MSOs interested in remaining competitive and deploying new services are faced with two network architecture options. Traditional routing-intensive metro networks offer clear benefits for data services, but fall short on the ability to converge data and TDM networks. Innovative MSOs interested in offering voice and other delay-sensitive services should deploy a network architecture with fewer router: "routerless aggregation".

INTRODUCTION

The current economy dictates that MSOs increase earnings and free cash flow. This can be done by increasing revenue through the introduction of new services and increasing the level of profitability of current services through reducing CapEx and OpEx.

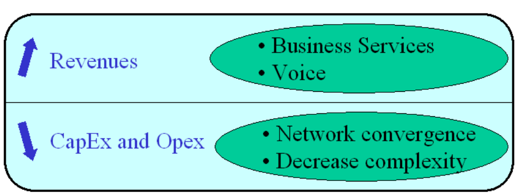


Fig 1 - MSO Opportunity

Increasing revenues

MSOs can improve revenues through increasing the market share of products and services with a potential for growth (i.e. high-speed Internet access), protecting market share of mature products and services (i.e. video), and expanding the product portfolio. Business services, historically an ILEC monopoly, are one possible area of growth or portfolio expansion, which have not been aggressively exploited by MSOs.

Reducing CapEx and Opex

An MSO's network represents the largest portion of its investment. It is composed of three different segments: the access network, which starts at the hub or headend and terminates in the subscriber's home, the metro network, which interconnects the hubs of a metropolitan or regional area, and the backbone network which interconnects the metro networks.

Today, MSOs own and operate three different metro networks: video, data and TDM. Therefore, the opportunity exists for operators to significantly reduce their network CapEx and OpEx through converging the data and TDM networks.

The Requirements

Current metro data network architectures can handle the requirements of Internet Access, both residential and commercial. To allow convergence in the future, metro networks must be able to support the requirements necessary to deliver the following services:

- Residential voice (individual POTS lines)
- Business private networking
- Business voice (T1s and T3s for PBX applications)

Voice Services

Voice services, both for residential and commercial customers, require high-availability networks with low latency, delay and jitter. The main attributes of high-availability networks are:

- No single point of failure, both in the signaling path and in the call path
- Fast network convergence to avoid dropping calls upon failures
- Transparent software upgrades to avoid downtime associated with network maintenance

Commercial private networking services require that the MSO be capable of delivering Layer 2 "pipes" across the metro network. This requirement is driven by two factors. First, many business customers still carry legacy protocols, such as IPX, SNA, LAT, DECnet, Appletalk and others, on their networks. In addition, businesses do not expect or desire that their address plan be impacted by the carrier's network.

Current Metro Network Architectures

There are two main problems with the traditional routed metro network architecture. First, it is composed of enterprise-class elements, which do not support high-availability. Second, it is exclusively composed of routers at every hop, which complicates the offering of private networking services.

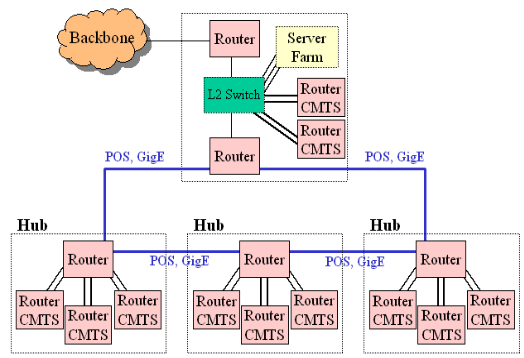


Fig 2 - Current Metro Network Architecture

Voice services, whether individual POTS lines or T1s and T3s for PBX applications, require high-availability networks that guarantee service availability and sub 50ms automatic switchover. Since the enterprise-class routers currently used in traditional metro network architectures do not offer carrier class availability, the classical method of increasing the availability of these networks consists of installing redundant edge routers at every hop. This architecture has two major drawbacks when offering telephony services:

- Since high-availability is provided at the IP layer through routing protocols, TDM services can only be offered through circuit emulation over IP.
- Since OSPF's convergence time is far above the traditional 50ms recovery time of voice networks, circuit emulation requires that operators implement MPLS-TE for its fast recovery features.

In addition, the traditional metro network routed architecture lacks native support for business private networking services, which can only be supported through the introduction of new networking protocols. The only solution is to emulate Layer 2 over Layer 3, which requires the configuration and management of a number of protocols (L2-MPLS/VPNs (Martini), MPLS-TE, OSPF-TE extensions, CR-LDP, etc.) across the metro. The result is an extremely complex and costly network to own and operate.

Routerless Aggregation

As stated previously, routed metro networks pose two main problems: the lack of native support for TDM services because of the absence of sub-IP layer path protection and the complexities of offering private networking services.

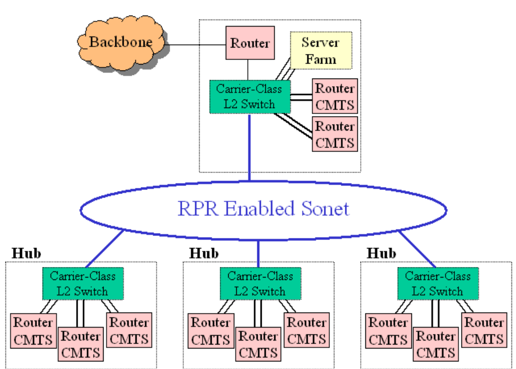


Fig 3 - Routerless Aggregation Architecture

The *routerless aggregation metro network architecture* solves both these problems. This architecture uses carrier-class Layer 2 switches as the aggregation element in each hub and pushes the routing function to the edges of the metro network. It essentially applies the networking principle of *routing at the edge and switching at the core* through the use of Layer 2 switching as opposed to the more complex MPLS switching. In a typical regional network, this architecture is composed of Layer 3 CMTSs at the subscriber edge (located in the hub), and edge routers at the provider or MSO backbone edge (located in the regional head-end or regional data center). All other elements in between are Layer 2 switches.

This architecture provides a number of benefits and, as we will discuss, a number of shortcomings that must be addressed. Lets start with the benefits.

The Benefits

Faster – Cheaper

Routerless aggregation is far more cost effective than enterprise-class Layer 3 aggregation for mainly two reasons. First, carrier-class switches only require half the number of elements and interfaces, compared to the traditional network architecture, by eliminating the need to duplicate the elements. A number of manufacturers now offer Layer 2 switches with built-in redundancy at all levels, and support for transparent software upgrades. Note that, when combined, these features allow a single Layer 2 switch to provide equal or better overall system availability than a pair of enterprise-class routers. Second, on a side-by-side comparison, any particular router port is generally more expensive than its equivalent on a Layer 2 device.

Simple support for private networks

VLANs have been used for years to support private networking services. Most PTTs and ILECs around the world have been using this technology for well over a decade, and continue to do so with much success. Through routerless aggregation, MSOs operate metro networks that can support Layer 2 point-to-point or multipoint-to-multipoint "pipes" from anywhere to anywhere within the region. The private networks are simple to configure and manage, and allow the MSO to support any Layer 3 protocol without getting involved with the subscribers' Layer 3 address plan or even being aware of the transported Layer 3 protocols.

PHY-layer protection

The carrier-class nature of these switches, when combined with interfaces that provide native support for TDM and packet-based services, allow the routerless aggregation architecture to provide native TDM services from anywhere to anywhere in the metro. Sonet and ATM are good examples of such interfaces. This architecture truly allows the convergence of the data and TDM networks in the metro, further reducing CapEx and OpEx.

<u>The "Gotchas" And The "Fix-Its"</u>

Avoiding spanning tree

Layer 2 networks usually rely on spanning tree to manage redundant paths in the network. Spanning tree provides slower convergence and is far less intelligent than routing to control and manage redundant paths in a network. Spanning tree is known to cause outages through broadcast storms, constant flapping to administrative mode, and other problems that derive from its basic operation. It is, in most cases, the main reason why many network architects have previously dismissed Layer 2 networks as viable network architectures. Most of the reasons why spanning tree was considered inadequate still exist, and therefore, the author shares the view that if spanning-tree cannot be avoided in a routerless aggregation architecture, the architecture should be considered incomplete and problematic.

On the other hand RPR, which is a new Layer 2 protocol that creates fault tolerant rings as an overlay of point-to-point GigE or Sonet links, allows the use of Layer 2 devices without resorting to spanning tree to manage redundant paths. Elements on an RPR ring are provided with a single Layer 2 path to all other elements on the ring, such that spanning-tree is never required to manage the ring's redundant paths. The RPR MAC layer handles interface and link failures transparently, such that changes to the network's links' status are never apparent to any element's Layer 2 (or Layer 3, for that matter) forwarding table.

VLAN scalability limits

The maximum number of supported VLANs on any given interface, per the standard Layer 2 header, is 4096. In some cases, this limit poses a scalability problem for MSOs, especially in medium to large size regions.

The routerless aggregation network architecture proposes to solve this problem by creating multiple Layer 2 RPR aggregation rings in the metro and to joint these rings through the use of edge routers implementing Layer 2 MPLS VPNs (Martini). This approach addresses the scalability issues of VLANs without introducing the complexities associated with implementing MPLS throughout the metro network. The result is a very scalable Layer 2 VPN solution that is manageable and has a level of complexity that grows with the services' level of success.

Impact of Layer 2 aggregation on OSPF

The routerless aggregation network architecture essentially flattens the metro network from a routing perspective. Flattening the metro has impacts on OSPF, or any other routing protocol. The most significant impact is that it increases the number of OSPF adjacencies maintained by each router in the network. If not factored into the design, an oversized growth in OSPF adjacencies will cause problems in the operation of the network. Routers will suffer from performance problems, convergence will be slow, and network stability will be negatively affected. Note that the maximum number of adjacencies supported by any given router is vendor-specific.

The scalability solution for VLANs also solves the OSPF scalability issues associated with routerless aggregation. In large metro networks, where routerless aggregation would cause too many OSPF adjacencies if a single Layer 2 network was created, the network should be partitioned into multiple sub-networks through the use of edge routers. This goal can be reached either by a physical implementation, or by creating the partitions through a logical overlay. An example of a physical implementation is to create two rings in the metro and place an edge router at their intersection point, physically separating the two MAC domains. A logical overlay uses VLANs to create two separate MAC domains over a single physical network through the use of an edge router that can be attached anywhere in the ring.

<u>RPR-Enabled Sonet</u>

This final section describes how routerless aggregation is a network architecture that can truly enable the convergence of TDM and data networks in the metro. As described earlier, RPR plays an important role in routerless aggregation metro network architectures. Also, Sonet and ATM are essential to the native coexistence of TDM and packet-based services over a single infrastructure. Given the fact that RPR can either use GigE or Sonet as its underlying physical network layer, RPR-enabled Sonet networks are, if implemented properly, the true enabler of converged networks.

A proper implementation of RPR-enabled Sonet allows the MSO to carve STS-1s out of the Sonet bandwidth allocated to the RPR interface to natively support T1 and T3 services across the metro. Products that support this feature can provide pure circuit switched T1 and T3 interfaces for telephony applications, along with GigE and 10/100BT interfaces for data applications on the subscriber side, and use a single uplink on the network side to carry all services across the metro.

<u>CONCLUSION</u>

The routerless aggregation metro network architecture provides a number of key benefits over current metro networks architectures, which allow

MSOs to *a la fois* increase revenues and reduce CapEx and OpEx.

The architecture benefits MSOs by enabling:

- Data and TDM network convergence, with native support for TDM services (as opposed to circuit emulation)
- Private networking services in a simpler and just as scalable manner as MPLS

The architecture avoids the age-old issues associated with link redundancy in Layer 2 networks through a new protocol: RPR.

SEAMLESS, SCALABLE HDTV ROLL-OUTS OVER TODAY'S HEADENDS

Sylvain Riviere
BigBand Networks, Inc.

Abstract

Offering HDTV programming is a necessity for cable operators competing with satellite and terrestrial broadcast television alternatives. However, digital cable faces several challenges in broadly replicating the HDTV services currently available via alternatives:

- *Maintenance of equal video quality, satisfactory to broad-casters, content creators and subscribers, while maintaining cable plant bandwidth efficiency*
- *Compliance with carriage of data associated with programming according to the PSIP standard and openly accessible by consumer electronics devices*
- *Responsiveness to programming changes by on-air broadcasters whose content is being mapped onto cable plant, such as when sudden shifts are made in use of bandwidth from an HDTV feed to several SDTV feeds*

This paper describes a comprehensive way to roll out full HDTV experiences in existing headends. It suggests channel line-up scenarios that achieve bandwidth efficiency with high quality content through advanced video bit rate adaptation techniques, effective multiplexing of PSIP data, and real-time intelligent responsiveness to broadcaster changes.

MAINTAINING CABLE PLANT EFFICIENCY WHILE ROLLING OUT HDTV

This section justifies bit rate adaptation techniques as a solution to overcome bandwidth challenge while scaling up HDTV roll-outs. It suggests channel line-up scenarios and implementations that maintain video quality over cable.

Existing bit rate constraints prevent scaling up of HDTV service

The Advanced Television Systems Committee (ATSC) gave birth to HDTV and 8-VSB (8-level vestigial sideband) modulations standards. At that time ATSC adopted the VSB transmission because of its "large" bandwidth, which is needed to transmit HDTV programming off air. In December 1996, the FCC approved those standards to replace the analog standards of the NTSC. The 8-VSB mode supports up to 19.4 Mbps of content and drives the maximum bit rate allowed for HDTV streams.

MPEG-2 technologies have brought encoders a long way during the last decade to offer similar video quality at lower bit rates benefiting from statistical multiplexing techniques. Such techniques take advantage of the inherently variable bit rate of video feeds, such that when multiple feeds are combined, it is highly unlikely that most or all will experience intensive action simultaneously, and in fact bandwidth peaks for some will usually correspond with troughs for others. Unfortunately, such techniques are predominantly employed for SDTV and little is expected from statistical multiplexing of HDTV feeds because only one or two feeds are carried per multiplex.

The following table compares the maximum number of HDTV versus SDTV feeds carried today using various modulation approaches. It also shows how broadcasters efficiently carry one HDTV feed versus other multiplexing alternatives.

	Modulation	Theoretical Rate	HDTV Carried at Near Constant	SDTV Carried with Statistical Multiplexing
Broadcaster	8-VSB	19.4	1	
Cable	64 QAM	27.0	1	7-8
	256 QAM	38.8	2	10-12
Satellite	QPSK	27	1	8-12
	8-QPSK	40	2	12-20

Cable operators face a severe challenge in scaling up HDTV while accommodating for over 19Mbps per HDTV stream, as defined per ATSC standards. This is on top of the challenges of bringing together content sourced from broadcast and satellite feeds, with their distinctive formats, onto a single plant.

Unless massive efforts such as cable plant upgrade or aggressive analog reclaim are undertaken, the HDTV constant bit rate approach won't scale up in a world of fast growing double-digit available HDTV programs. MPEG-2 bit rate adaptation techniques, also called rate shaping, can address those problems.

<u>HDTV rate shaping optimizes statistical multiplexing techniques – it's not all about crushing bits!</u>

Rate shaping describes bit rate adaptation techniques applied to MPEG-2 encoded streams, to further enhance bandwidth efficiency. This technique can substitute for decoding-encoding operations that are expensive, space consuming and ultimately harmful to content quality.

Accommodating various transport alternatives, rate shaping also adjusts the necessary bit rate to "bridge" HDTV bit rate from satellite and off air delivery onto cable plant. By doing so, HD rate shaping removes fixed bandwidth allocation constraints imposed by ATSC standards. The technique considers and accommodates cable plant transmission capabilities for greater bandwidth efficiency, without the harmful effects of decoding and re-encoding. HD rate shaping does not blindly steal bits to squeeze more into a channel, so video quality does not suffer for bandwidth efficiency.

In the case of three HDTV into one 256QAM channel, the rate shaping operation does not reduce all streams 33% to squeeze one more program. Instead bit rates are dynamically driven by incoming content complexity. Depending on content, economic bit rate reduction per program can be as low as 10% and up to and beyond 50%, while maintaining identical perceived video quality to original sources.

By taking multiple outputs of this process from multiple sources and packing them together, varying bit rate results in cumulative bit rate efficiencies at desired video quality. This process, also called statistical re-multiplexing, outlines how cable operators can accommodate their own bandwidth efficiency by moving away from the near constant HDTV bit rate expectations traditionally imposed on them.

Statistical re-multiplexing is utilized by cable operators for SDTV feeds already and can be applied to HDTV feeds similarly. However because of the nature of available HDTV content and its video quality emphasis, caution is required when implementing HD rate shaping in the headend.

<u>Applying content intelligence in rate shaping.</u>

When deploying rate shaping technology in cable headends to gain bandwidth efficiency, it is important to stay competitive with alternate sources. As mentioned earlier, content intensity drives bit rates in statistical re-multiplexing operation. Because cable operators have the freedom to choose specific channel line-ups, they can optimize for certain scenarios that will maintain video quality compared to alternatives.

As an example, sports, movie and news content introduce different complexities that drive the bit rate reduction allowances differently in HDTV. The higher the complexity is, the less likely is bit rate reduction to occur, and vice versa. As a general rule, avoiding excessive content within a re-multiplexing pool will allow optimum video quality and greater bandwidth efficiency.

Certain statistical re-multiplexers have the ability to combine both HDTV and SDTV feeds. Benefits include offering granularity at which the bandwidth efficiency is reached independently from content complexity. Combined with priority mechanisms, there are cases where the HDTV bit rate can remain untouched while bandwidth efficiency is gained by removing stuffed packet (nulls) for SDTV channels. This technique is called pass through mode as it exactly replicates the on screen HD content and quality level as at the peak bit rate, while finding unused bandwidth in stuffed packets that can be allocated to other traffic.

Fig. 1 below outlines HD and/or SD channel line-ups and their respective bandwidth efficiencies in the example of 256QAM cable plant. The efficiency gain is the amount of content carried on a QAM channel versus a constant bit rate alternative of two 19.4 Mbps HDTV feeds. For purposes of this analysis, the general guideline applied is that six SDTV feeds conventionally consume the same bandwidth as one HDTV feed.

The figure also suggests deployable scenarios under the condition justified by identical subjective video quality comparison to alternatives sources available in the field.

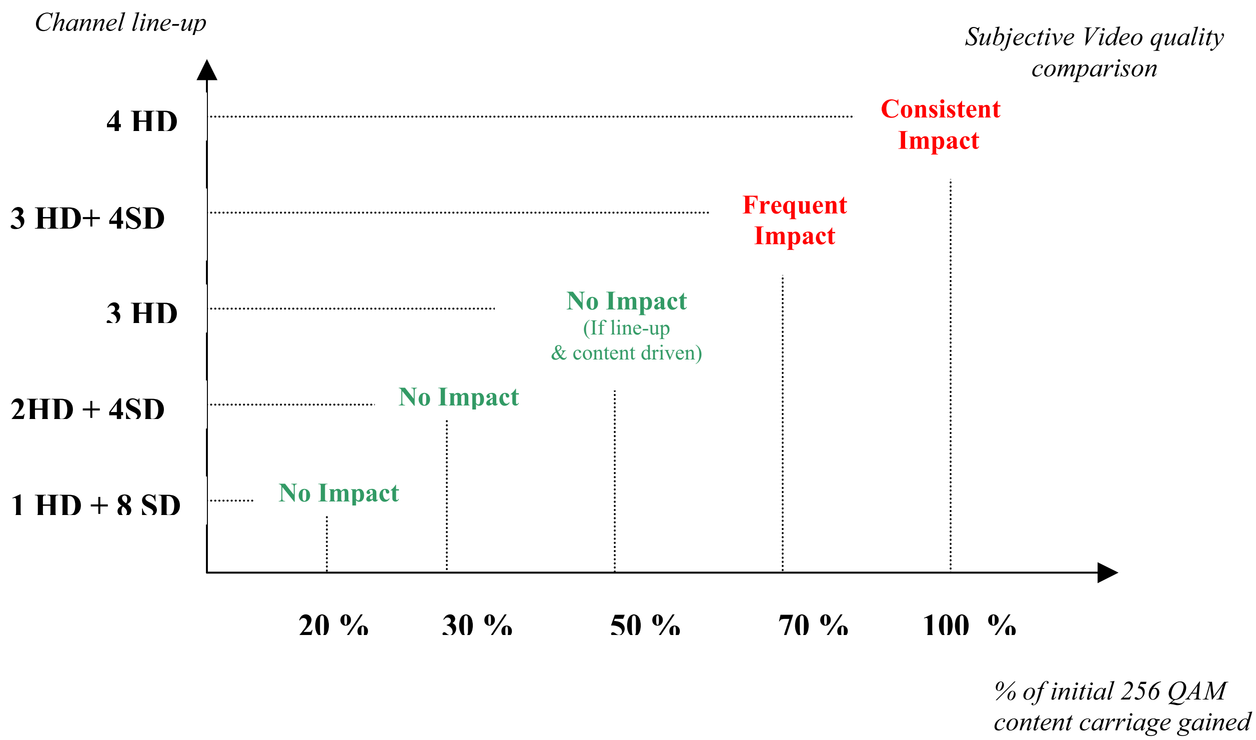


Fig. 1: Bandwidth and quality impact of rate shaping and statistical multiplexing alternatives.

By appropriately implementing rate shaping with proper channel line-up, cable operators can achieve up to 50 % bandwidth efficiency gains, while preserving similar video quality to alternatives. Beyond 50%, video quality can be impacted, and it is recommended that operators make considered decisions about deployment.

PSIP SUPPORT ON THE CABLE PLANT

This section addresses cable-ready television support through the implementation of the PSIP format, including compliance with industry standard agreements.

What is PSIP and why is it needed?

PSIP (Program and System Information Protocol) has been standardized by the ATSC to allow tables of information to be transmitted along with the associated video programming. PSIP data are originated by broadcasters and are required by some ATSC receivers to tune to the correct digital channel.

Set-top boxes used by cable operators do not require PSIP to tune to a channel, but the growing availability of retail market cable-ready digital televisions requires PSIP presence on the cable plant to guarantee proper tuning. Besides tuning, PSIP tables contain other important information about the programming, such as branding and electronic program guide data.

Consider a local station that offers three different shows during the day. Digital televisions allow viewing the analog service as channel 65, for example, and the three others as channels 65-1, 65-2 and 65-3, even tough the digital broadcast is technically on channel 66. The main number indicates that those channels belong to the same broadcaster whether they are analog or digital. In this way the analog branding is preserved in the digital world.

The electronic program guide (Fig. 2) is the navigation interface provided to the user to tune to a channel, specifying timing and event description. PSIP carries tables that can be used by cable-ready televisions to build the electronic program guide. This function is often seen as a benefit of converting to digital by the user, and will be provided automatically by cable-ready televisions as long as PSIP tables are available. The ATSC PSIP standard requires that a minimum of the next 12 hours of program information be available in advance, although PSIP can offer up to 16 days of programming.

Fig. 2: Electronic program guide representation.

<u>Complying with NCTA-CEA agreement.</u>

The emergence of cable-ready televisions, with their PSIP support, call to question digital cable interoperability with ATSC PSIP standards. There is a large number of standards, agreements, specifications and FCC rules that relate to the interoperability of consumer electronic products with cable-ready televisions (like the one shown in Fig. 3), but the cable and consumer electronics industries have taken important steps towards achieving interoperability by establishing the NCTA/CEA technical and PSIP agreement.

NCTA and CEA negotiations resulted in an agreement that provides a consistent set of standards that enable cable-ready televisions to be connected directly to a cable plant without the need for a set-top box. The agreement section relating to PSIP addresses television receivers that do not have a security module (Type 1 Television).

During the negotiation, the cable industry made clear that carriage assumes the availability of PSIP data from the content provider, and that it would be prepared to support carriage of PSIP information when made available from the content provider in accordance with the agreement. CEA agreed with the document. The agreement also specified mandatory/optional PSIP tables to carry on the cable plant while recommending standards to overcome implementation issues. For more information on NCTA\CEA agreement specific to PSIP, refer to ATSC A/65 and SCTE DVS-097 standards.

Fig. 3: Mitsubishi cable ready TV – WS55909.

PSIP agreement implementation in the headend.

Cable operators will need to obtain necessary hardware and software to implement the NCTA/CEA agreements. Whether two broadcast signals or more are combined in a single transport stream for delivery using 256QAM or 64QAM, the re-multiplexing operation (with or without rate shaping) will require rebuilding the PSIP information.

Sent in-band with the video, PSIP table implementation requires a platform that can combine both video and data. In some cases, rate shaping can be required to make room for data when bandwidth of incoming video consumes what's available.

ACCOMMODATING BROADCASTER MULTICASTING HABITS

This section addresses how off-air multicasting can impact cable service, and proposes an alternative approach to multicasting through automatic response.

What is multicasting?

Multicasting occurs when broadcasters suddenly shift use of bandwidth between various combinations of HDTV feeds to SDTV feeds. The broadcaster, in an effort to accommodate formats and leverage content availability, often disrupts the multiplex several time a day switching from one channel line-up to another. Stream characteristics also change on the fly depending on the operation done at broadcaster sites. While those sudden changes are transparent to an ATSC receiver, cable architecture does not accommodate multicasting transparently, possibly impacting service.

Possible impact on services with multicasting.

Depending on broadcaster and site, impacts on cable services and their consequences varies, and can include the following:

Loss of service even though content is present.

Stream characteristic changes disrupt overall equipment performance, starting with the decoder that may need to be retuned to the channel to update stream characteristics. This can be the case when SDTV feeds replace HDTV feeds. In some cases headend equipment may lose the incoming identification information (same program number but different packet ID), preventing automatic restoration even though the content is restored.

Customer satisfaction issues.

If a channel disappears, its session already mapped on the cable plant is still maintained whether the program is there or not. The user tunes to a black screen although it had content earlier, increasing call center activity dramatically. Even worse can be the subscriber who becomes frustrated by the service to the point of cancellation, without ever placing a call. Generally the cable operator is blamed for service loss although it is a consequence of broadcaster multicasting.

Channel line-up confusion.

The switch between SDTV and HDTV format streams brings confusion as far as channel line-up structure and how HDTV streams could be grouped together in logical channel line-up numbering.

SDTV format and HDTV format with upconverted SD content (black bar).

A particular manifestation of channel line-up confusion is when the user sees standard definition 4:3 ratio on a 16:9 ratio television. In one case the user can stretch the image of its television to avoid the black bar caused by SDTV content ratio (see Fig. 4), in the other case the television will not allow it because the feed is already in HDTV format although SDTV content was upconverted. Too much time with black bars on screen also risks image burn-in on the television.

Fig. 4: Example of standard definition stream stretched by television to 16:9 screen.

Satellite and broadcaster alternatives provide a clear indication of what type of feed is broadcasted to avoid this confusion. An alternative for cable is to organize channel line-up per stream format to avoid the confusion, independently from the content when multicasting involves both types of streams on same channels.

Solution to accommodate multicasting in existing cable architecture.

An intelligent re-multiplexing platform capable of real time responsiveness to broadcasters' sudden changes is needed prior to the usual cable session management systems. Strategic positioning of this re-multiplexing can preserve static channel line-ups independently from multicasting events and act as a shield during eventual stream characteristic change.

In response to program loss, a graphic message can be substituted for the program disappearing to offer an indication to the user that a channel will return or is not available, and suggest alternative programming locations. Also, smart session management can redirect streams dynamically to the proper channel number in a line-up whether the format is HD or SD, indicating whether or not television stretching function is enabled.

SUMMARY

Because of the realization of FCC timetables for terrestrial digital broadcasting launch and satellite progress and aggressive marketing, cable is particularly challenged to compete effectively in providing HDTV services. Elements such as constant bit rate encoding, PSIP data tables and real time broadcaster switching between formats are inherently unfriendly to the cable plant.

This paper has shown that an intelligent multiplexing platform can overcome obstacles through HDTV rate shaping; capabilities to combine HDTV and SDTV within the same channels; regenerating PSIP tables for cable in compliance with the NCTA/CEA agreements; and recognizing and accommodating broadcaster format shifts. These techniques enable the cable industry to provide subscribers with a complete and fulfilling HDTV experience. The cable industry can then leverage its own inherent advantages, such as its balanced access to both local and national feeds. Cable operators can be liberated from defensive positioning in HDTV and industry competitiveness can be enhanced.

Sylvain Riviere
Director of Product Marketing
BigBand Networks, Inc.
475 Broadway
Redwood City, CA 94063
650 995 5024
sylvainr@bigbandnet.com
http://www.bigbandnet.com

SECOND GENERATION POINT OF DEPLOYMENT (POD) INTERFACE FOR MULTI-TUNER CABLE RECEIVING DEVICES

Joseph W. Weber
David Broberg
CableLabs

Abstract

Under the current agreement between CE companies and cable operators, uni-directional "digital cable ready" televisions may soon be offered to consumers. Cable customers could be able to receive premium digital content without the need for a set-top box through the use of a conditional access point-of-deployment (POD) device.

As these single-program POD devices begin to be deployed, CableLabs along with its members and the vendor community is developing a next-generation POD capable of providing multiple streams of premium digital content. This will enable new devices and expanded services such as picture-in-picture, watch-and-record PVR, and home networking of multiple displays within the home.

This article looks at the features of the second generation POD and some of the technical details of how it will operate.

THE POINT-OF-DEPLOYMENT (POD) MODULE

The Point-of-Deployment (POD) module, as currently defined by SCTE 28 [1], SCTE 41 [2], and OpenCable™ specifications [3,4], provides a common format for decrypting premium MPEG-2 content delivered via a cable network. As a result of the use of open standards the consumer premises equipment can be independent of the conditional access system used on that particular cable plant.

The operation of the POD module is shown in **Figure 1**. Digital content is received via a QAM tuner and sent to the POD module. Premium content that the customer is entitled to view is decrypted using the network's conditional access system. A dedicated out-of-band communication channel (either one-way from the cable network to the device, or two-way) is required in order for the POD to connect with the conditional access system.

Premium content is then re-encrypted within the POD module with the open standard POD copy protection method defined is SCTE 41 [2]. Authenticated devices are able to decrypt the POD copy protected content for display or recording via a 1394 interface with 5C (DTCP) copy protection.

CableLabs has defined a number of different consumer premise devices that interface with POD modules [5]. These include one-way digital televisions and sophisticated two-way set-tops with the OpenCable middleware (OCAP). Appendix A is a list of the currently defined OpenCable defined devices. In the future corresponding versions compatible with a multistream POD will be defined as well.

Single Streams

The current POD specifications were built upon the National Renewable Security Standard (NRSS-B) [6]. As such they were designed to work on a single multi-program transport stream received via a single QAM tuner. While it might be possible to multiplex multiple transport streams into a single transport stream, it would require sophisticated hardware on the part of the receiving device. In addition, the current out-of-band signaling methods used on the POD make it difficult to share a common out-of-band transmitter. Therefore it is necessary to define a new POD device and interface capable of supporting multiple transport streams from multiple QAM tuners.

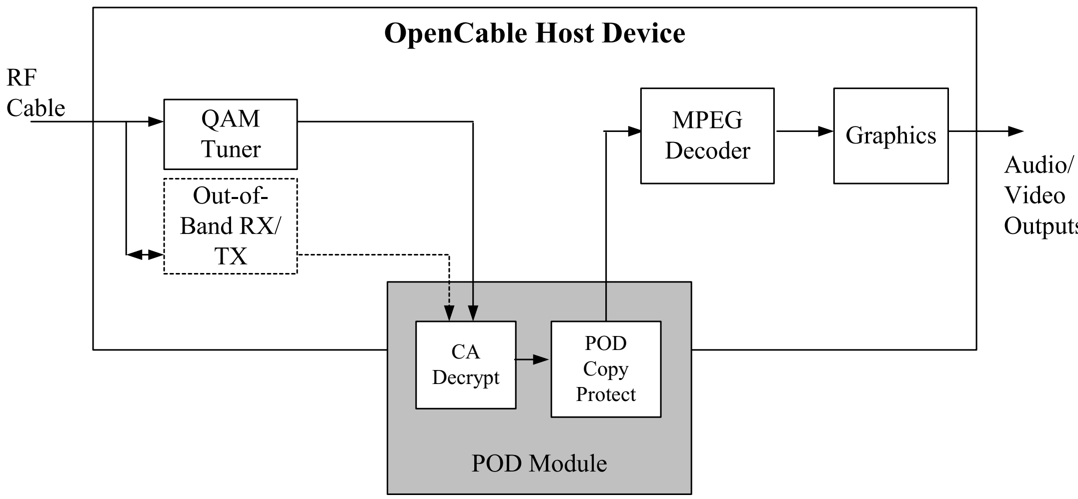


Figure 1 - Current POD and Host Operation Diagram

MULTIPLE STREAMS

Expanded Features

A POD capable of supporting multiple transport streams would enable a number of new digital cable devices. A multi-stream POD would be capable of decrypting multiple premium programs located on different multiplexes received by multiple QAM tuners. Perhaps the simplest feature enabled is picture-in-picture (PIP) where one program is displayed in a window overlaid on another program. Multiple tuners and a POD capable of decrypting multiple programs are required for a device offering this feature.

If the host device contains a hard disk drive (HDD) for temporary storage of MPEG streams, then multiple tuners and a multi-stream POD enables the ability to watch-and-record, or to record two shows simultaneously.

Even more sophisticated host devices might serve multiple displays within the home via a home network. A central home server device may have multiple tuners (at least one per display device) as well as a HDD for storing content for later viewing. Content is then delivered via the home network to the various display devices.

Requirements

In collaboration with the cable operators, CableLabs has defined a tentative list of requirements for the second generation POD device that supports multiple transport streams. These requirements are subject to change as we develop the technical specifications.

1) The Multi-Stream POD specification should encourage the development of retail digital cable set-top and terminal host devices through the use of OpenCable and other publicly available specifications. Just as the

current open standards for the POD module are enabling multiple companies to offer digital-cable ready devices, the multi-stream POD should enable even more innovative products utilizing multiple tuners.

2) The interface shall provide sufficient bandwidth for a maximum data rate which supports the payload from up to six simultaneous 64 QAM transport streams, or up to five 256 QAM transport streams, or any combination that is below the maximum data rate.

3) The Multi-Stream POD shall be able to decrypt multiple Programs from a single Transport Stream as well as multiple Programs from multiple Transport Streams, up to the resource limitations of the POD (see Requirement number 6 below as well).

4) The Multi-Stream POD shall be backward compatible with the Single-Stream POD. It shall appear as a Single-Stream POD in a Single-Stream host device. This will enable the second generation PODs to be used in current single-stream POD devices. Cable operators can transition to the new PODs and still support the deployed single-stream devices.

5) Multi-stream PODs shall support both traditional QPSK out-of-band methods and out-of-band data delivered via cable modem. The POD will indicate to the Host which OOB method to use depending on what the cable plant supports. As a result, second generation PODs will be able to be used on any digital system in North America now and in the future.

6) The Multi-Stream POD interface shall provide a discovery mechanism for a Multi-Stream capable host to discover how many simultaneous Transport Streams and PID decrypts the POD scan handle. This allows the Host to manage POD resources and prioritize which programs to send to the POD.

7) The multistream POD specification shall allow for the use of multiple PODs in a single device. This would allow for any future products that might need more stream support than a single multistream POD can handle. This extensibility ensures any future devices that contain even more tuners than anticipated today could be supported.

8) All two-way multistream Hosts shall contain a cable modem that supports the DSG specification [9]. All multistream PODs shall support DSG out-of-band.

9) Every transport packet that enters the POD from the Host will be returned to the Host in the same order, and with a fixed delay.

Given these requirements, a POD-Host interface specification is being developed that meets the requirements while also creating a viable commercial product. CableLabs is working with cable operators and consumer equipment manufacturers to create these specifications.

Second Generation POD Operation

The multi-stream POD builds on the current POD by providing a command interface, out-of-band communication interface, and transport stream ports. The command interface will use the same layering of objects as the current POD.

Like the current POD, the multistream POD will provide a single transport stream input port and a single output port. As a result the transport streams from multiple tuners will be multiplex into a single stream before entering the POD. The POD will decrypt any selected and authorized premium content programs. Transport packets are returned to the Host device in the same order they are received. The Host then de-multiplexes the various streams and provides them to the various services requesting them. A block diagram is shown in **Figure 2**. In the multiplexer section, the sync byte within the MPEG transport packet header will be modified by the Host in order to uniquely identify the different transport streams. The host de-multiplexer would then use the sync byte to identify packets returning from the multi-stream POD.

The multi-stream POD will use the same HOST-POD communication method as the present POD. The Host and POD communicate via a series of Application Protocol Data Units (APDUs). Most of the APDUs for the multi-stream POD will be identical to the ones for the current POD as defined in [1,3]. This will allow developers to re-use most of the software from single stream PODs and devices. Some APDUs will be modified to include identifiers to indicate to which transport stream they apply. For example, the CA_PMT() APDU is used to indicate which encrypted programs are to be de-crypted by the POD. In the multi-stream POD the CA_PMT() APDU will be modified to indicate which program as well as which transport stream contains that program.

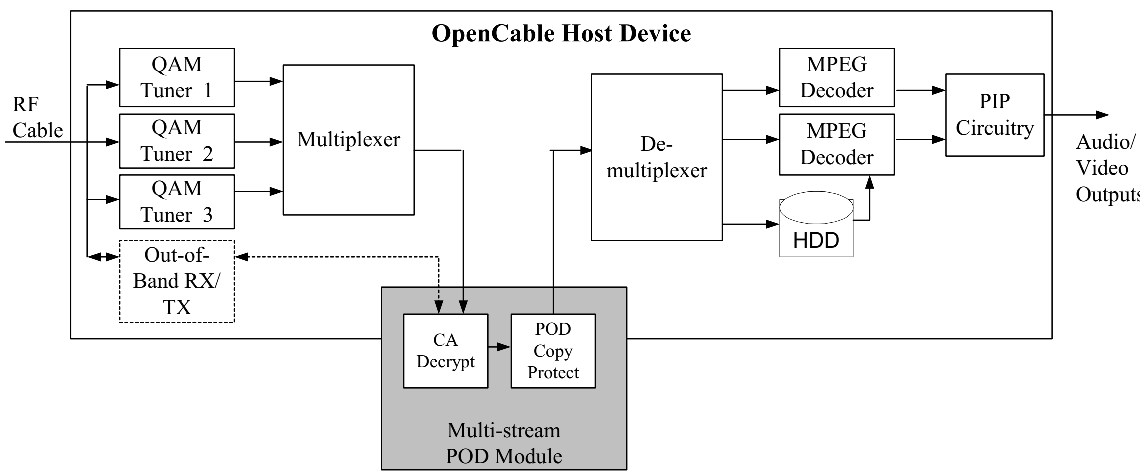


Figure 2 - Sample Multistream POD and Host Diagram

Further Study

CableLabs will continue to work with its member and interested vendor companies to complete and publish the multi-stream POD specification. Interested companies can contact CableLabs for information on how to become part of the drafting team for the multi-stream POD specification.

Appendix A – OpenCable Host Devices

CableLabs defines the following OpenCable host devices to operate with the current POD:

OCAP: OpenCable Application Platform [7,8]. A common middleware specification based on Java for interactive television applications.

OpenCable Set-top Box: Cable box with ability to decrypt digital tiers. Includes one-way and two-way capable boxes, requires POD for decryption of cable provider services. Outputs include RF, DVI, Component outputs and 1394.

OpenCable TV: Cable ready TV with ability to decrypt digital tiers. Includes one-way and two-way capable TVs, requires POD for decryption of cable provider services.

Advanced OpenCable Set-top Box: OpenCable Set-top Box with a cable modem for two-way services.

Advanced OpenCable TV: OpenCable TV with a cable modem for two-way services.

OpenCable HD Set-top Box: OpenCable Set-top Box that supports decoding of High Definition TV. Can be either one-way or two-way, includes new outputs such as DVI and HDMI.

OCAP 1.0 Set-top Box: Supports all OCAP compliant applications, and is two-way.

OCAP 1.0 TV: Supports all OCAP compliant applications, and is two-way.

OCAP 2.0 Set-top Box: Supports all OCAP compliant applications, two-way, with a cable modem (Advanced OpenCable Set-top Box).

OCAP 2.0 TV: Supports all OCAP compliant applications, two-way, with a cable modem (Advanced OpenCable TV).

References

1. ANSI/SCTE 28 2001, "Host POD Interface Standard (formerly DVS 295)"
2. ANSI/SCTE 41 2001, "POD Copy Protection System"
3. "OpenCable Host-POD Interface Specification", OC-SP-HOSTPOD-IF-I12-030210
4. "OpenCable Point-of-Deployment (POD) Copy Protection System", OC-SP-PODCP-IF-I09-030210
5. "OpenCable Host Device Core Functional Requirements", OC-SP-HOST-CFR-I12-030210
6. EIA-679-B, "National Renewable Security Standard"
7. "OpenCable Application Platform Specification (OCAP) 1.0", OC-SP-OCAP1.0-I05-030210
8. "OpenCable Application Platform Specification (OCAP) 2.0", OC-SP-OCAP2.0-I01-020419
9. "DOCSIS Set-top Gateway (DSG) Interface Specification", SP-DSG-I01-020228

Reference Acquisition

SCTE: Society of Cable Telecommunications Engineers, 140 Philips Road, Exton, PA 19341
Telephone 800-542-5040;
E-mail: standards@scte.org
URL: http://www.scte.org

CableLabs: Cable Television Laboratories, Inc., 400 Centennial Parkway, Louisville, CO 80027
Telephone: 303-661-9100
URL: http://www.cablelabs.com

SWITCHED BROADCAST CABLE ARCHITECTURE USING SWITCHED NARROWCAST NETWORK TO CARRY BROADCAST SERVICES

Gil Katz
Harmonic Inc.

Abstract

Bandwidth is a precious resource in any cable network. Today, Cable MSOs broadcast hundreds of digital channels over HFC networks to all cable subscribers. These channels occupy a sizable part of the plant RF spectrum, yet at any given moment, most channels remain unviewed. Significant RF spectrum can be reclaimed by switching "less popular" broadcast channels according to user demand.

A narrowcast switched video network for VOD services is already in place in many large cable systems today. This switched network provides digital video content to subscribers on demand, occupying bandwidth only when a title is requested and sent over the HFC.

This paper will discuss the existing switched narrowcast network architecture as a scalable, cost-effective, flexible and "switched broadcast ready" network.

INTRODUCTION

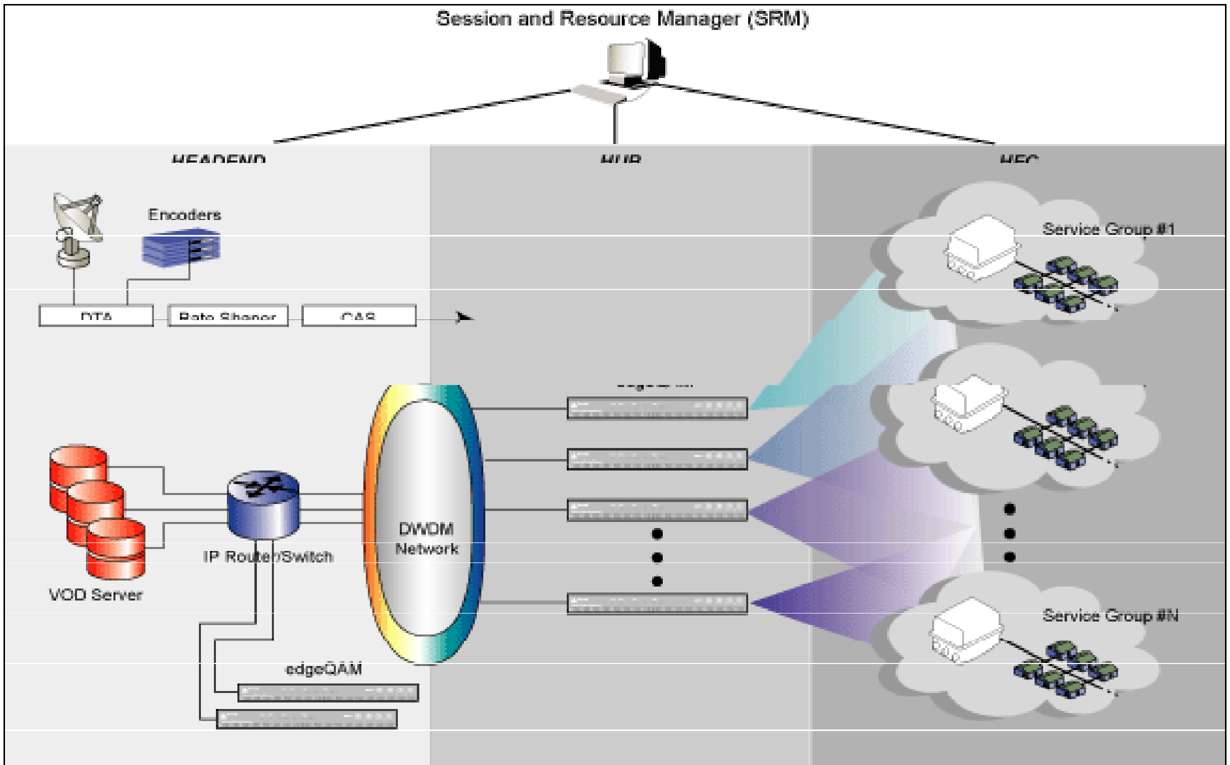
Bandwidth is a valuable resource for cable operators. They are constantly trying to leverage their existing infrastructure to expand their broadcast channel line-up and offer as many revenue-generating services as possible. Possible new services include video-on-demand (VOD), high-speed data or voice-over-IP (VoIP).

A switched broadcast architecture allows operators to offer a virtually unlimited number of broadcast programs while freeing costly bandwidth for revenue-generating services. The switched broadcast model dynamically switches a broadcast channel "on", via the narrowcast channel to which the subscriber is connected, when a subscriber attempts to tune in the channel. Operators can also add an array of specialty or targeted channels without increasing their systems' broadcast channel spectrum or capacity.

The theory behind switched broadcast relies on typical Pay-TV viewing behaviors. Within a particular service segment or node, only a handful of channels are being accessed at any given time. In other words, numerous channels are not being watched and there is a lot of bandwidth that could be saved or used for other services.

Many cable systems already have a narrowcast video network for VOD services in place. The most common network architecture for VOD uses standards-based commodity GigaBit Ethernet (GbE) switches to build a cost-effective switched network. Edge QAM devices with standard IP interfaces serve as gateways between the standard Ethernet/IP network and the HFC, and enable QAM sharing for multiple services. Switched broadcast is just another type of service that can use that narrowcast video network infrastructure and **share** the IP network and QAM resources.

The diagram below describes a typical VOD system running in parallel to a broadcast network.



SRM parallel Architecture

This paper presents a solution using an ***existing*** VOD infrastructure to enable a switched broadcast services overlay at ***nominal cost*** to the operator. The discussion will outline any required additions to an existing broadcast infrastructure in order to support a switched broadcast application.

THE SOLUTION

The solution presented here has been adapted from a standard Multicast IP solution. For several years, Telco operators have been providing video services over their standard IP networks. The video content may be provided on demand and transmitted to a specific Unicast address, or broadcast in Multicast IP groups into the IP network. Each client, in the standard IP solution, simply needs to join a multicast group in order to get the broadcast service. The IP router will send the service to the client once it is part of the group.

Using the same approach, the IP edge QAM device will join the multicast group following a subscriber's request to receive a switched broadcast service, and will forward the service to the QAM feeding the set-top box's Service Group.

A single edge QAM device can serve multiple service groups. Its MPEG-2 multiplexing core enables service **multicasting** at the edge of the network by duplicating the content and streaming it simultaneously to multiple destinations or QAMs.

THE SWITCHED BROADCAST SERVICES

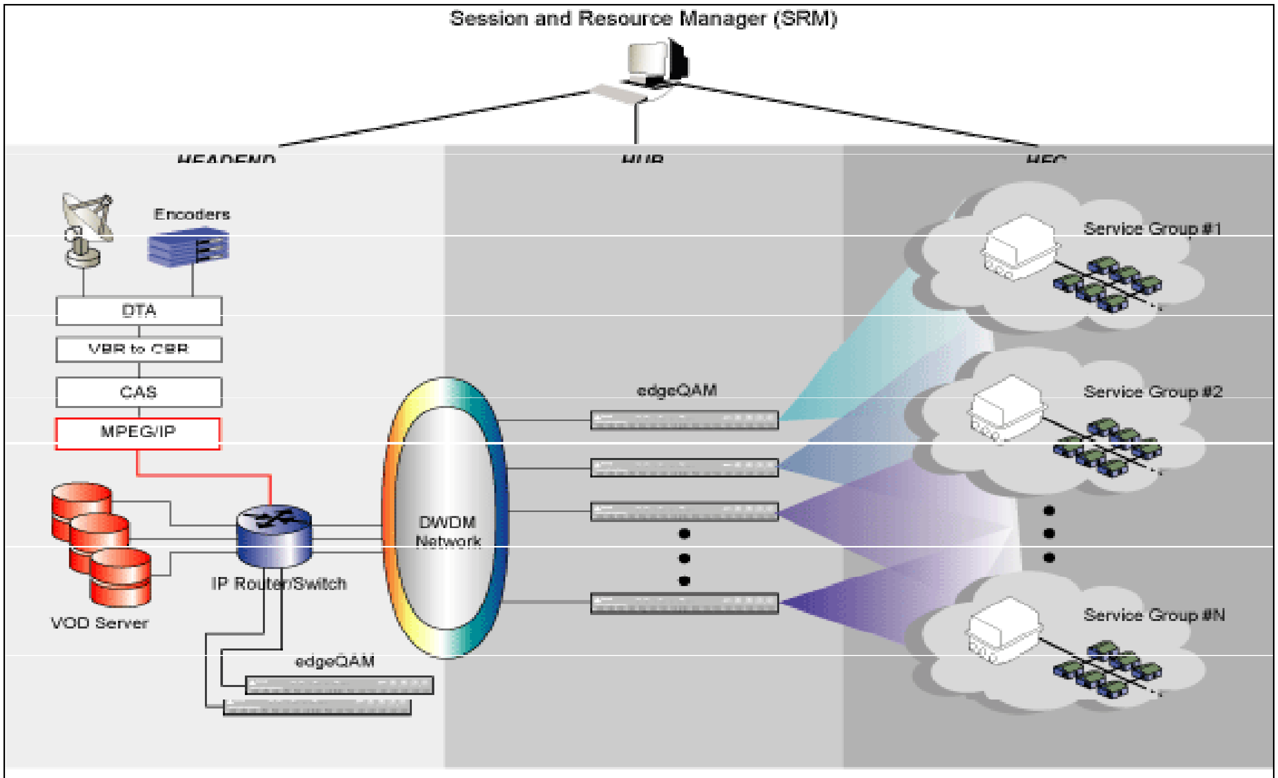
Content intended for switched broadcast can be any standard definition and/or high-definition video programming, or data content such as games or electronic program guides. The programs may be locally encoded or received off the air. The channels should be carefully selected as "less-viewed" programs relative to the entire broadcast domain in order to guarantee efficient use of network resources. For example, these channels could be niche programming, ethnic programming or local interest channels.

ENABLING SWITCHED BROADCAST

The concept proposed in this paper suggests the use of existing components owned by the MSO to keep the capex investment nominal. Looking at the diagram below, which represents the modified architecture, it is clear that not much has been added. Using the same broadcast feeds, some will be forwarded to broadcast channels as they are currently, and some will be sent to the switched narrowcast network.

The same devices that were used before for rate shaping will now be used for VBR to CBR conversion. A new device should be added, converting the switched broadcast streams into SPTS (Single Program Transport Stream) over Ethernet/IP/UDP frames. The IP/UDP frames are identical to those being generated by the Video Server. A single 1-RU converter box costing less than $10,000 supports hundreds of SPTSs!

Most of the modifications needed to enable switched broadcast services happen in software, such as the Switched Broadcast application on the STB and the SRM (Session Resource Manager). The SRM is responsible for sharing the network and QAM resources between the VOD services, the switched broadcast services and other future services. The diagram below shows the suggested architecture.



SRM Suggested Architecture

SO, HOW DOES IT WORK?

Each channel selected for switched broadcast will be transmitted into the streaming IP network as an SPTS over a dedicated IP/multicast group (multicast address). The QAM edge device emulates the RF portion of the network to an IP network and treats the switched broadcast channels as **standard** IP/multicast services throughout the network - from the video source to the subscriber's set-top box (STB).

Once a "switched broadcast" program is selected from the program guide, the STB forwards the request to the SRM. The SRM identifies the service group (SG) where the request originated, and checks the bandwidth availability of the QAMs feeding this SG (or zone).

The SRM provisions the appropriate edge device for the relevant multicast address, and sends information regarding which QAM should receive the stream. Upon provisioning, the edge device will "join" a multicast session and re-multiplex the stream into the appropriate MPEG transport stream/QAM. The SRM sends an acknowledgement to the STB, providing the QAM channel and the program ID.

In general, the process is nearly identical to the way a subscriber selects a VOD service. The primary difference between a switched-broadcast stream and a VOD stream is that streaming does not originate directly from a server or other storage device. Rather, the content is simply streamed off the broadcast services.

LEVERAGING THE BROADCST SERVICES

Other applications can be integrated into a switched broadcast infrastructure. An operator could easily provide all available locally-encoded content, generated in different regions, to all systems within a cable network. Doing so would require no extra bandwidth. Instead, only standard IP connectivity between the systems is required.

Virtual VOD (V-VOD)

Virtual VOD can be seen as an improved version of NVOD services using the narrowcast network. It still enables a level of VCR functionality, while dramatically reducing the bandwidth consumed by the NVOD service over the broadcast network. Another significant benefit is that V-VOD requires less streaming capacity than regular VOD services.

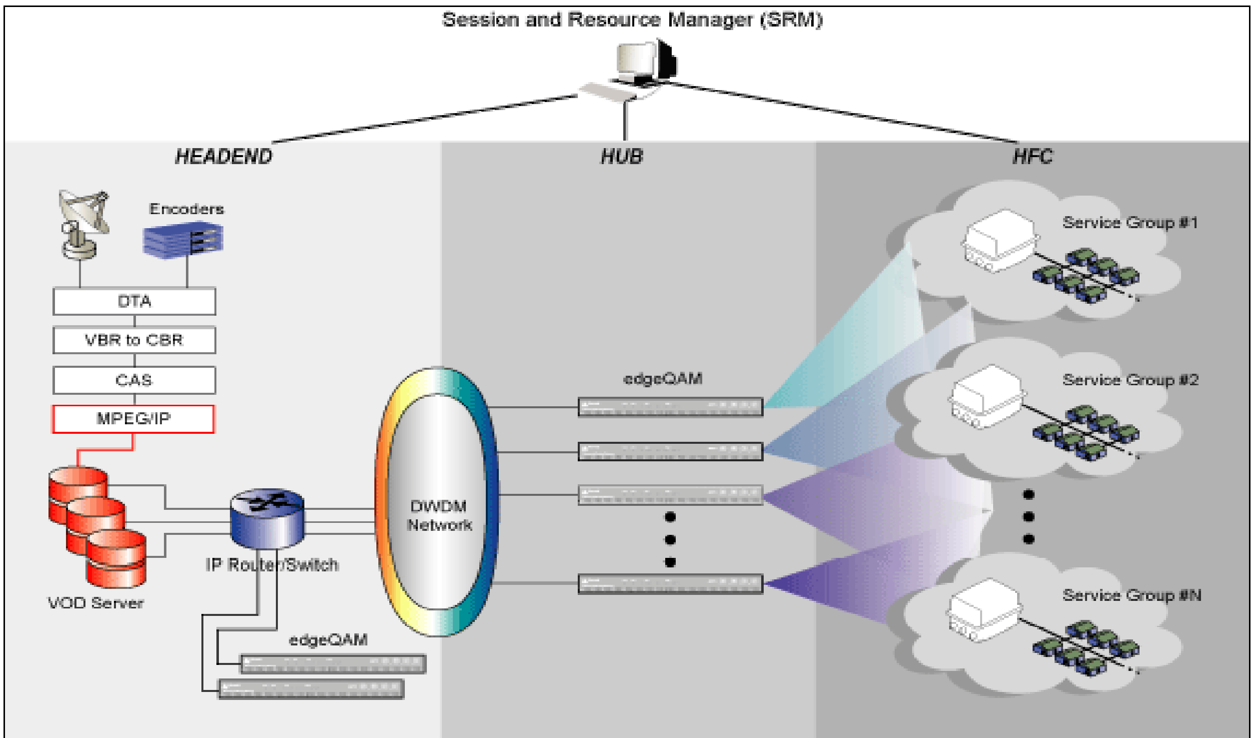
For example, ten two-hour movies that start every five minutes (or twelve times each hour) require 10 x 24 = 240 streams. At 3.75 Mpbs per stream, 240 streams require a total of 900 Mbps bandwidth. A single standard VOD server will suffice. The VOD server will stream all 240 streams regardless of actual user demand. However, as in the case of switched broadcast programs, the streams will be dropped at the IP switch connected to the VOD server and will not consume HFC resources unless requested by a subscriber.

In this architecture, switched V-VOD leverages the narrowcast QAM access to carry NVOD streams based on demand. While the concept does not require any in-band channels for NVOD, it does enable V-VOD so that each subscriber will not have to wait more than five minutes for a movie to start.

nPVR APPROACH

Network PVR (nPVR) is already providing broadcast channel programming per user demand. In the nPVR model broadcast channels are being recorded on Video Servers, the **same** Video Servers providing VOD services. The content is being provided to the user using the **same** narrowcast network infrastructure used for VOD services. The advantages of the nPVR solution are clear: it enables VCR control as well as very targeted Ad insertion. So, why wouldn't we use the nPVR model, which provides prime broadcast content on demand, for the more niche programming? The answer relates to storage cost and capacity, and to the narrowcast network cost. Nevertheless, the migration to Everything on Demand (EOD) will drive lower prices on Video Servers as well as the narrowcast network, enabling niche programming over nPVR infrastructure.

Again, The argument here is about minimal investment for the niche programming, and reuse of existing infrastructure. We are proposing a migration path from broadcast to switched broadcast (provided as another source into the switched narrowcast network) to nPVR.



SRM nPVR Architecture

SUMMARY

A switched broadcast architecture enables an unlimited expansion of the broadcast channel line-up while freeing up precious bandwidth for other revenue-generating services. New services could include Virtual VOD and local content distribution.

The solution introduced in this paper relies on standard "off the shelf" IP/GbE devices such as switches and IP edge QAM devices. Existing VOD systems based on these devices are scalable, cost-effective, flexible and "**switched broadcast ready**."

Minimal capital investment is required to enable switched broadcast services on a switched narrowcast network built for VOD services. **The same QAM and IP network resources can be shared between the different services.**

TAMING THE PEER TO PEER MONSTER USING SERVICE CONTROL

Michael Ben-Nun
P-Cube Inc.

Abstract

This document explains the increasing bandwidth and network capacity planning challenges peer-to-peer file exchange applications cause Internet Service Providers. It discusses how Service Control – the concept of statefully tracking network usage and enforcing advanced subscriber, application and destination differentiated policies – is key to resolving the peer-to-peer traffic issues within existing network infrastructure.

PEER-TO-PEER AFFECT ON NETWORK CONGESTION

The Evolution of Peer-to-Peer (P2P)

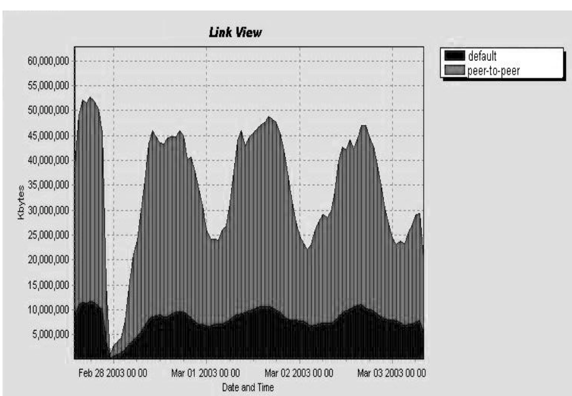
Understanding the relatively short history of P2P applications and its underlying technologies is critical to the comprehension as the impact it has on broadband IP networks. Internet based P2P is a relatively new technology, which allows for the creation of decentralized, dynamic, and anonymous logical networks for information exchange using the public Internet. In "traditional" client/server model a well-known source provides content and information to requesting clients, whereas in P2P, applications utilize various techniques to allow users to search and share content between themselves. There are several different P2P technologies and architectures that evolved from the most basic type – one that has a central "coordinating" server utilized for content searches between clients (e.g., Napster). Completely decentralized P2P has no central server (e.g., Gnutella) to provide search capabilities due to the fact that the clients search amongst themselves. Other variations of P2P provide application specific networks (e.g., KazaA) and some utilize an open standard (e.g., Gnutella and OpenNAP) to allow clients share all sorts of content. All of these applications allow individual users (conveniently shielded by the anonymity of the network) to share files over the Internet. These files often contain copyrighted materials (e.g., songs, movies, software, etc.) that no commercial content provider could legally afford to publish.

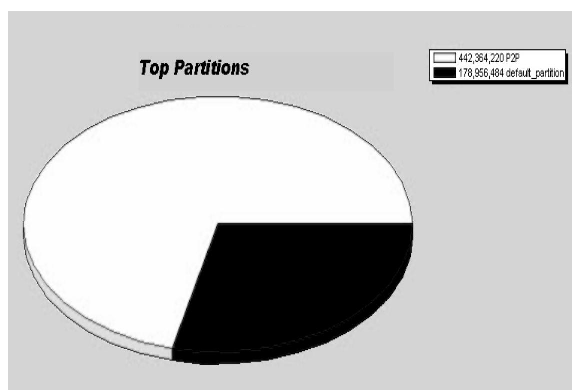
Due to this simple file sharing method, Napster, which is considered to be the first P2P application with mainstream appeal, was an immediate success among Internet users, especially those with high-speed Internet connections. A court ordered shutdown of the Napster service did little to decrease the amount of P2P file swapping activities, rather it can be argued that the added publicity probably achieved the opposite effect and the popularity of P2P applications has increased ever since. With new P2P clients and applications released to provide more functionality and ease of use, P2P traffic comprises a large part of Internet bandwidth usage. The popularity and use of different P2P clients is varied and can be determined by a variety of factors. Some clients are more popular in certain geographies (such as Winny which has wide spread acceptance in Japan), while others have a strong following among the "distributors" of specific types of material.

Peer-to-Peer Incurred Congestion

P2P clients, due to their numbers and intensive need for network bandwidth are causing significant network congestion. With less bandwidth left for other network traffic, this results in a reduction of the overall broadband experience for other subscribers on the network, and raises network capacity, planning, and management issues. Every IP network is built with assumptions about usage, which in turn is used to analyze and compute the necessary amount of network capacity and resources needed to support a given subscriber base. P2P applications are different from traditional client/server applications in the way that users run them and how the applications use the network. The table below provides a glimpse of some of the parameters used by service providers, their importance for planning the network, and the influence P2P technologies have on these parameters. P2P applications are increasing in popularity and constitute a growing percentage of network traffic. These applications are so popular that a new term has been coined to describe the more avid users of these technologies. Often referred to as “bandwidth hogs” or “abusive subscribers,” these users are using their broadband network connections to generate a disproportional amount of network traffic and significantly contributing to network congestion.

The following charts, produced from analyzing the usage of a particular network, serving HSD cable subscribers uncovers the alarming truth: Approximately 70% of network bandwidth is being used by P2P applications.





CONTROLLING PEER-TO-PEER TRAFFIC: TECHINCAL REQUIREMENTS

With the growing amount of P2P traffic, there is a clear need to address the link congestion and bandwidth issues it creates. To solve the problem, service-providers must use a solution that is able to:

(a) Identify, account and report on P2P usage.
(b) Control the bandwidth these applications consume.

The following section provides detailed technical requirements that a solution must provide.

Technical Requirements

When attempting to identify and control P2P traffic, it is important to remember the underlying technical requirements from a proposed solution. Once the requirements are fully understood they could be used to evaluate possible solutions. The unique technical requirements that need to be addressed are:

IDENTIFY:

- Ability to classify traffic based on layer3-7 parameters: Peer-to-peer applications do not utilize well-known port numbers, and thus cannot be classified by simply looking at IP packet headers (IP addresses, TCP port-numbers, etc.). Rather, deep inspection of packets, including the identification of layer-7 patterns and sequences *must* be supported.
- Ability to maintain bi-directional flow state: In order to identify a particular flow of packets as peer-to-peer, carriers cannot inspect each packet within that flow to make the identification. The solution that performs proper identification of P2P traffic *must* ensure that once a particular flow (e.g. a TCP connection between two hosts) is identified as P2P, all packets on that flow are tracked, and treated as such. Of critical importance is the ability to tie between both directions (i.e. upstream & downstream) of a flow, since in many cases the initial identifying pattern resides in a packet sent from one host, yet the majority of traffic can flow in the other direction.

- Ability to provide quick turn-around for new P2P applications: As peer-to-peer applications constantly change, and new ones emerge, the underlying protocols used to carry the peer-to-peer traffic change frequently. The solution *must* be quick to adapt to new protocols, and provide new identification mechanisms.

Note that the importance of the above-mentioned identification requirements increase in complexity and number with the growing speed of the development of new peer-to-peer applications/protocols. Even today, P2P applications use well-known ports, assigned to other network uses (such as port-80 for web-browsing), and they are constantly migrating to these port numbers in an attempt to masquerade as 'traditional' network activities and thereby avoid detection. Hence, simple analysis based on port-numbers leaves most of the P2P traffic unaccounted for, and will not truly address the problem.

CONTROL:

- Ability to control bandwidth at various isolation levels & granularities: To control the bandwidth impact of P2P applications it is necessary to provide a network control mechanism for different levels of isolation and control. The solution *must* provide the means to control bandwidth at "subscriber granularity", whereby it limits the total amount of bandwidth each subscriber can consume. It *must* be able to control the bandwidth of particular flows, so as only the P2P identified traffic of a particular subscriber is limited, while the rest of that subscriber's traffic is left unaffected.

- Ability to enforce time, destination and subscriber differentiated policies: To control the bandwidth congestion cause by P2P, and enforce various control policies, while maintaining the necessary flexibility to actually implement these on real-life subscribers, the solution *must* provide the means to create differentiated enforcement schemes (or policies) based on time of day, destination and subscriber. Specifically, the ability to create different enforcement packages for different subscribers *must* be supported.

- Ability to maintain subscriber level quotas: In order to control P2P traffic in a persistent manner for each subscriber, the solution *must* provide the infrastructure to maintain a usage state for subscribers, and account for the total amount of P2P traffic over time. As an example, the ability to maintain the total amount of P2P traffic each subscriber has consumed on a daily/weekly/monthly basis, and apply different bandwidth quota based consumption restrictions based is key to moderating the use of the network.
- Note that while the issue of controlling and enforcing P2P bandwidth consumption is crucial for maintaining a congestion-free and predictable broadband network, it can cause customer expectation issues, as the current subscriber-base is unaccustomed to imposed limitations on its high-speed data access. Therefore the above flexibility is mandatory as service-providers create the policies best suited for their subscriber-base.

- Support high-speed network rates, and subscriber-capacities: As today's broadband networks are built to sustain significant traffic loads, the solution *must* support today's network interfaces and traffic rates. Typical broadband networks use Gigabit Ethernet and OC interfaces with high throughput. In addition, the solution *must* have the capacity to support the total number of subscribers served by the network links, for both existing subscriber numbers today, and for forecasted growth.

APPROACHES TO CONTROLLING PEER-TO-PEER TRAFFIC

With the technical requirements in mind, the following section explores possible solutions to identifying and controlling peer-to-peer traffic.

Using Router/Switch QoS Mechanisms

Existing routers, switches or similar network devices contain various types of traffic classification and QoS mechanism, which could potentially be used to control P2P bandwidth.

However, as these devices were not designed to address these issues, they do not provide the following capabilities:

- They do not provide Layer 3-7 traffic classification. Nor do they maintain state across packets flows.

- They are not "subscriber-aware" and cannot provide subscriber differentiated enforcement

As a result, switches and routers do not provide the means by which the peer-to-peer traffic can be identified, and network usage policies be applied to it. Additionally, as the QoS mechanisms in switches and routers attempt to deal with link congestion and bandwidth distribution, they do not provide the necessary subscriber-differentiated policies, required to control the peer-to-peer traffic once identified.

Using DOCSIS 1.1

The DOCSIS 1.1 specifications, contains many features and capabilities to control bandwidth utilization, and offer differentiated services to subscribers. However, by itself the DOCSIS 1.1 specifications cannot fully address the issue of controlling P2P applications. This is due to the fact that DOCSIS 1.1 does not:

- Provide the mechanisms to classify traffic based on layer-7 capabilities, or maintain state for bi-directional network flows.
- Provide the required bandwidth control isolation and granularities. DOCSIS 1.1 provides the means to control traffic at a defined flow specification (typically a combination of layer3-4 parameters). However, as mentioned above, to fully control P2P bandwidth consumption, there is a need to implement various layer of bandwidth control, which the DOCSIS 1.1 specifications does not attempt to address.

As a result, while DOCSIS 1.1 is a potential key component in service differentiated high speed data networks, it does not provide the mechanisms to control the peer-to-peer abuse problem.

Using Service Control Platforms

A Service Control Platform is defined as a platform that is able maintain state for each network flow, classify it according to layer3-7 parameters, and implement various bandwidth shaping and control rules, based on the classification of the traffic and the subscriber it is mapped to.

The following diagram depicts the internal operations of a service control platform.

On step (1), the platform classifies each packet received into a stateful, bi-directional flow.

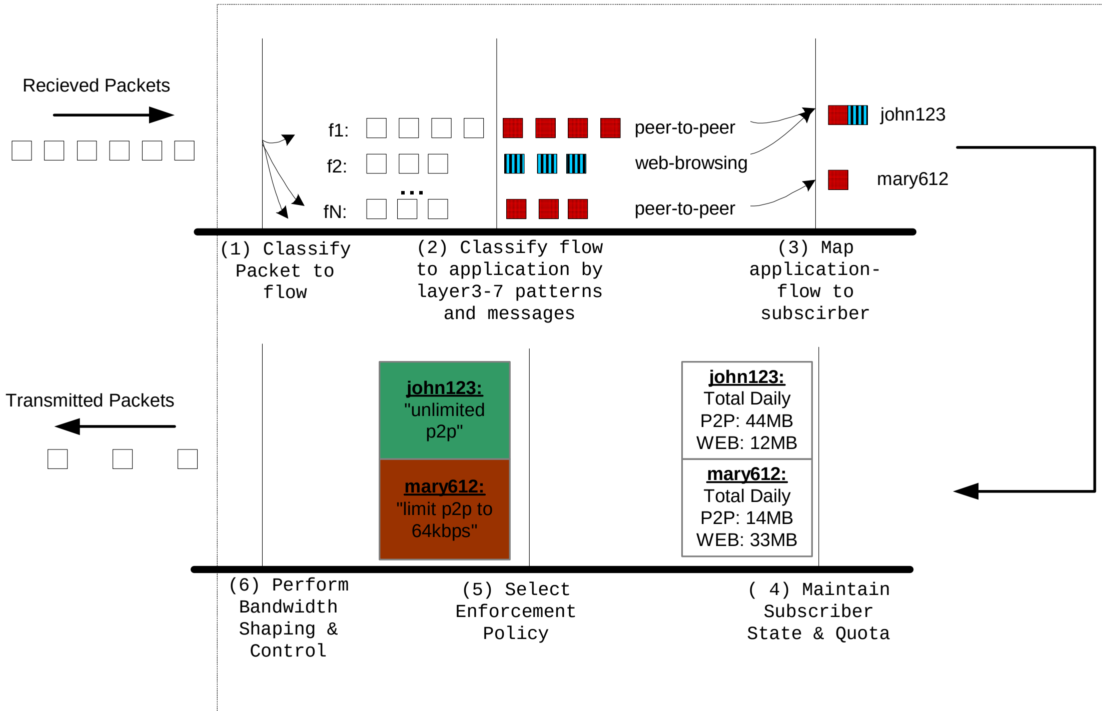
On step (2), the platform performs dynamic stateful reconstruction of the application (layer-7) message exchange in the flow, and identifies the application used by each (peer-to-peer, web, mail, etc.)

On step (3), the platform maps each such flow into a particular subscriber. Typically there is a many-to-many relationship, in which many application-flows are mapped to many subscribers.

On step (4), once the traffic has been classified, identified and mapped, it is accounted for on a subscriber basis. Subscribers' state is updated according to the traffic they transmit or receive, and this impacts (along with the their assigned policies) the final bandwidth enforcement policy (5) applied.

On step (6), the selected policy is translated into packet level decisions, indicating how the actual implementation of the bandwidth restriction is performed.

Ultimately the total bandwidth consumed is reduced through control implemented in the service control platform and the overall network congestion is reduced to a level acceptable to the network provider.



CONCLUSION

The combination of Peer-to-Peer applications' aggressive use of network resources and the growing popularity of P2P is straining broadband networks, and causing congestion, operational costs, and user satisfaction issues. Using Service Control, P2P traffic can be precisely identified and controlled, so as to contain its affects on the network without influencing other applications and network users. Furthermore, the P2P consumption has a different network behavior and usage pattern than typical "common" network applications that require differentiated bandwidth (such as enterprise based SLA and QoS).

Users utilizing the network for P2P traffic are typically residential subscribers, unaccustomed to enforced bandwidth restrictions. As such the control mechanisms required to contain the affects of P2P, while avoiding subscriber alienation due to rigid policies, are not provided by standard QoS mechanisms. This means that commonly deployed switched and routers cannot act in the same capacity as Service Control Platforms for the purpose of P2P monitoring and control. A complete and effective solution requires the combination of P2P identification flexibility, traffic control, quotas, and subscriber awareness – the main building-blocks of a Service Control Platform.

Parameter	Importance for network planning	Influence of traditional applications	Change caused by P2P applications
Upstream / Downstream Traffic Ratio	Networks are asymmetrical in nature: the amount of traffic that a network can sustain upstream (i.e., from the subscribers to the network), is different from the amount it can sustain in the opposite direction. The ratio required between these two directions is in direct correlation to the requirements of the applications using the network. Networks are built with a specific ratio which, if incorrect, may cause high rates of congestion and unutilized capacity.	A typical residential user uses the network for downstream applications. These applications (e-mail, web browsing, etc.) generate a larger amount of downstream traffic for each corresponding upstream request, and service providers have come to rely on this ratio to model network capacity	P2P applications encourage users to share files, and a typical peer serves gigabytes of files. This causes a drastic change in the upstream/ downstream ratio, and as a result congestion on the upstream link (due to individual users' increased uploading of files).

Parameter	Importance for network planning	Influence of traditional applications	Change caused by P2P applications
Time of Day and Percentage of Activity	Service providers typically assume an average duration of network use per subscriber per day, and (based on subscriber profiling) peak use periods. A service provider would typically be able to predict and account for network "rush hours" and "lulls" periods of network use. This subscriber profiling is based on assumptions that residential home users primarily use the network during weekends and evenings, and that telecommuters and small offices use it primarily during business hours. Sudden or sporadic changes in these patterns may cause congestion during certain hours that were not evident before.	The time of day and percentage of activity expected for residential broadband subscribers is rooted in the premise that a typical residential customer uses the network only when the subscriber is physically present and actively using the connection. Such is the case when web browsing, reading e-mails, etc.	As P2P applications are usually used to upload or download large, multi-megabyte files, they are typically left unattended for days at a time while the application constantly attempts to download a list of files. At the same time it can serve as a search node for the P2P network and serve multiple file requests of other peers. This creates a never ending, high volume stream of network activity throughout the day. For example, a student's computer with a broadband connection can compete with telecommuters for vital network resources during business hours while the student is at school.

Parameter	Importance for network planning	Influence of traditional applications	Change caused by P2P applications
Traffic Destination and Peering points	The costs associated with serving each network packet and connection can depend on the location of the peer of the subscriber. Carefully crafted peering agreements with other network providers can help reduce the amount of traffic, and hence the cost of expensive transit connections. Furthermore local traffic (often referred to as OnNET) that does not leave the service provider's own backbone network, is significantly lower in cost than traffic that does (OffNET).	Traditional uses of the data network are mainly OnNET (email, nntp, web-proxies), with a small percentage being OffNET. This small percentage of traffic is for content that is located at sites external to the network providers domain.	P2P traffic has increased the amount of traffic between users in a significant way. When two or more P2P clients start using the network they form a direct connection to exchange the file. Whether the clients use the same or different providers is not a determining factor in how the P2P connections are made. P2P file exchange has significantly increased the potential for OffNET traffic.
Estimated Traffic Volume	No matter the topology and architecture of the network, there is a finite amount of bandwidth available for all its users, and certain over-subscription assumptions are used when planning the capacity of the network	Traditional applications have a large "time-to-consume" factor: A small web-page can take several minutes to read, a single e-mail message might take a number of hours to process. This determines how the traffic volume for each type of content served.	P2P applications are mainly used to share large binary files that have a much lower "attention-per-byte" ratio. A three-minute song is usually 3-5 megabytes. A 10-minute movie can be hundreds of megabytes long. Each piece of content that is served is traffic/bandwidth intensive.

THE THREE DIMENSIONS OF HOME NETWORKING

Doug Jones
YAS Broadband Ventures

Abstract

Home networking is a lot more than "wired vs. wireless." Home networking also includes both the command/control language to search for and play content within the home and defining how applications are run and managed. Keeping an eye on all three dimensions is needed to promote plug and play interoperability of home equipment.

INTRODUCTION

Business Case

It started with Personal Video Recorders (PVRs). Now those PVRs can both be networked within the home, and over a high-speed connection to sources outside of the home. The home is becoming a source of stored content and services.

With content in the home, users will need not only better and faster connectivity within the home, but also the means to search the home for content and play it back to the audio/video device of their choosing over a reliable home network.

Home networks will interconnect both entertainment devices and general computing devices. This will allow services like a home calendar and shopping list to be integrated onto the same platform as entertainment services.

Just like cable is providing many services over one network (video, voice, data), it makes sense that users will want a single network in their homes to serve their needs. This network will connect many and various devices for services such as entertainment, communication, energy management and home control.

To accomplish all this, the home network will need to be easy to install, easy to connect devices to, easy to upgrade, adaptable, low cost, and secure. Quite a wish list indeed, but we are just at the beginning of home networking.

This paper discusses in technical terms specific functions needed on the home network in order to have different devices and services work together. All of the topics discussed in this paper are available today from more than one supplier, though not always in an interoperable fashion. In order to have widely available plug and play interoperability, the industry will have to converge around a select few of the initiatives currently available.

Three Dimensions

This paper discusses three distinct dimensions of home networking, all of which are critical to understand from both a technical and a business perspective. Understanding what happens in these three dimensions is important to understanding how home networking really works.

The three dimensions are:
Datalink Technology, e.g., 10/100Base-T, IEEE 802.11a/b/g, HomePNA, IEEE 1394, HomePlug, etc. These are all technologies

that move electronic bits. Some are wired and some are wireless; some are synchronous and some are asynchronous; some work over existing wires and some require new wires. Each has inherent advantages and disadvantages that will be discussed.

<u>Interoperability Software</u> to do functions such as device discovery and content search and playback. All of these functions are sometimes known as "plug and play." This very important dimension defines how the home devices learn about each other without prior provisioning and how they work together to offer useful services. In order to use these "plug and play" protocols, the devices are connected using some kind of datalink (wired or wireless) to carry the bits that make up the message used by the interoperability software. Suites of protocols such as UPnP™ (Microsoft), Rendezvous (Apple), JXTA™ (SUN), DENi, etc., solve these issues in their own ways. Other companies such as Ucentric, Motorola and Scientific-Atlanta are creating solutions in this area. While many claim to be "standards" or standards-based, the reality is there is a very fragmented marketplace right now with several solutions vying for the top position. This is the key area to ensure service interoperability among devices on the home network.

<u>Applications Frameworks</u> to allow the development and execution of machine-portable applications within the home.

A well-known framework is PersonalJAVA™ although there are many others. These environments allow an application to be run on any of several computing platforms, for instance a Personal Digital Assistant from one supplier and a multimedia personal computer from a different supplier.

<u>Overall System View</u>

The discussion begins with a simple network consisting of two devices, one that has content (PVR) and one that wishes to search through and play a piece of content (Set Top Box [STB]). These devices are shown in Figure 1.

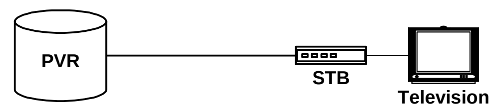


Figure 1 – Simple System for Home Networking

Note that a television is also shown in Figure 1. The STB could be embedded in the television but the two are shown as separate devices here because STB/TVs are not generally available (yet). As well, the PVR could have been embedded in the STB (or visa versa). Note also that the STB is not necessarily connected to a HFC Cable network, rather, the STB is just connected to the PVR and converts content to a form that can be viewed on the television.

This example alone is useful as new paradigms for devices on the home network will arise as multiple suppliers enter the market and consumer choice begins to dictate what devices do. Traditional suppliers and traditional devices will have to adapt.

For the time being, assume the PVR and the STB are connected via some kind of datalink. The datalink technology is important from the standpoint that maybe the STB can only accept analog video, or perhaps it can only accept digital video. The datalink technology has to support carrying data in the format that is needed to be useful. If digital video is used, the system has to know if the datalink supports a 3.5 Mbps

Standard Definition (SD) bit rate or a 20 Mbps High Definition (HD) bit rate.

If the datalink is a dedicated connection between the PVR and the STB (that is, not shared with any other device) then maybe it does not need to support Quality of Service (QoS). If the datalink is shared by other devices, perhaps QoS is needed to ensure the bit stream sent to the STB arrives there uninterrupted and without loss. Raw bit rate, support for QoS, and the ability to be shared are some of the characteristics to consider when studying datalinks.

When the PVR and the STB are first networked together, they will use some method to discover each other. That is, devices on the home network should be able to learn of each other without the need to be explicitly programmed by the user. In this way devices from various suppliers can be placed on the same network and learn about each other without user intervention. This complex task is handled by interoperability software running on all the devices on the network.

Once discovered, the STB has the capability to search the content on the PVR and play some of it back, including pause, fast-forward, and rewind capabilities. Highlights of this process are shown in Figure 2.

To complete the exchange shown in Figure 2, the PVR and STB have to talk the same interoperability protocols. Regardless of supplier, if the devices on the network use the same interoperability software then they can coexist on the same home network and function in a plug and play manner.

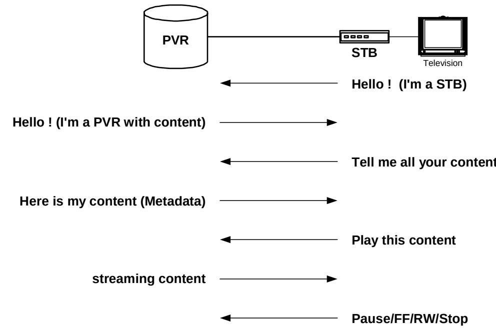


Figure 2

The first two lines in Figure 2 (the "Hello's") are what allow devices from different suppliers to recognize each other on a home network. While these two lines are highly simplified, they show that when devices are connected to a datalink they advertise what kind of device they are and what kind of services they offer. This information is generally broadcast around the home network to allow all the other devices on that network to learn about each other without user intervention.

The latter exchanges shown in Figure 2 look somewhat similar to a Video on Demand (VOD) session; searching content and controlling playback. The home will have its' own sources of content and devices on the home network need the capabilities to search for content and play it back. An issue to be solved includes allowing a user to seamlessly search for content on both in-home devices as well as devices in the service providers network.

In order to introduce the third dimension of home networking, applications, a new piece of equipment is added to the system diagram as shown in Figure 3.

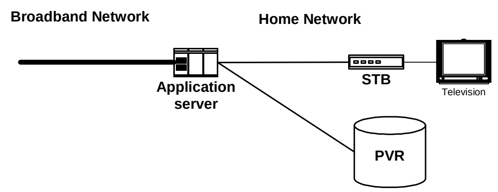


Figure 3

This new piece of equipment is labeled "Application Server" for lack of a better term. It could be a STB, or a home media center, but it has the characteristic that applications from a service provider can be placed there and function as expected regardless of what the physical device is. The subscriber can use these applications to add value to their home networking experience. Also out of convenience, the high-speed connection to the home is shown connected to the application server.

If one wanted too, the Application Server, STB, and PVR could all be collapsed into one device. The cable industry has an example of such a device and it is called the OCAP hardware platform. This is a monolithic, fully integrated, higher cost solution. What if the hard disk in the PVR needs to be changed out? What if more memory is needed to run applications? What if the user wants to add HDTV outputs on the STB? In the monolithic case, the entire device has to be switched out. With a modular home networking approach, the user can mix and match equipment from several suppliers and upgrade the equipment as driven by either personal choices or changes in available technology. The three dimensions of home networking provide the necessary connectivity tools and environments to allow devices from multiple suppliers to interoperate on a home network.

DATALINK TECHNOLOGIES

Introduction

Datalinks are technologies that move bits on a wire, or in the case of wireless, across the air. The bits could represent MPEG frames, Ethernet frames, IP packets, ATM cells, or some other type of protocol data unit. The datalink does not care what it carry's, rather it just makes sure the bits get from one end of the connection to the other. There are dozens of datalink technologies available; every day examples include wired 10Base-T and wireless IEEE 802.11b. Additional examples are given in later sections.

Home networks can have more than one type of datalink, for instance using a wired connection to get to various rooms within a home but then using wireless within the room.

The datalink has to carry two basic types of information, interoperability software messages and content. The datalink does not care about this type of information, it is just moving bits around and the bits can be just about anything. But there are some datalinks that will do a better job than others. For instance, if the datalink can only support 1 Mbps, it will not be able to carry digital video at a rate of 3.5 Mbps.

When all said and done, the two key items of concern with a datalink are having enough throughput (bits per second) to carry the services and easily being able to make connections in the rooms where subscribers want them.

Two of the most debated aspects of datalinks are "wired vs. wireless" and bit rate. These will be discussed in the following sections.

Wired Datalinks

One of the key questions is, which wire? Datalinks exist for home power wiring (HomePlug™), for home telephone wiring (HomePNA), for home data wiring (Cat-5 10/100Base-T, USB, IEEE 1394) and for optical fiber (SPDIF). There are even several companies working on turning the in-home coax cable into a datalink.

The pro's of wired datalinks include higher speeds, more consistent speeds, and the knowledge that there is a physical connection. Con's include the possibility of having to install new wires.

Wireless Datalinks

These technologies include Bluetooth™, HomeRF, the IEEE 802.11 series, Magis Networks, etc. These datalinks have the benefit of "no new wires," however, they may not have the bit rate needed to support several streams of HDTV along with high-speed Internet applications like peer-to-peer and gaming.

The pro's of wireless datalinks include no new wires and relatively higher speed throughputs that are coming to market e.g., IEEE 802.11g. Con's include speeds that can fluctuate based on distance and concerns over security.

Datalink Bit Rate

There is a debate of how much bandwidth is needed on the home network to support services that consumers want. A key driver to get at the answer will depend on the compression used for video services. Entertainment quality video takes a fair amount of bandwidth, in the megabits per second range, that must be delivered consistently and reliably.

If there is not enough raw throughput at the datalink, QoS technologies may be needed. QoS is needed when there is the need to give one service better treatment than another when there is congestion on the datalink. With enough raw throughput available, an arguably simpler datalink can be offered, one that does not need QoS. On the other hand, there are different types of QoS too, specifically prioritized versus parameterized. Parameterized QoS is more complex, having many parameters to guarantee exactly the QoS needed for that particular service. Prioritized QoS is relatively simpler, giving certain services higher priority than other services on the datalink. Prioritized QoS does not give guaranteed bandwidth, but sometimes a simple higher priority is sufficient to support the needed service. Of course if QoS is needed, that means some services will get better treatment than others, and the ones that get the "less better" treatment may not be happy.

INTEROPERABILITY SOFTWARE

Introduction

Interoperability software enables plug and play architectures where devices and services can be introduced into a network without configuration hassles. In addition, interoperability software is an important step toward eliminating manually installed drivers, relying instead on standard interfaces to put devices in touch with other devices and the services they offer to the network.

Interoperability software is a very important step, one that will free the consumer from having to understand the technical details of each piece of equipment and manually provision them to interoperate. The protocols and procedures included with the interoperability software allow devices on a home network to learn each other's capabilities and to interact in a way to provide services.

Interoperability software can work completely in the background; the user does not have to configure the equipment to make it work. These protocols are device-to-device, and can occur autonomously once a device is connected to the home network. Other devices on the home network that understand and talk the same interoperability software will respond autonomously as well.

There are two main architectures for interoperability. One is peer-to-peer, where all devices communicate with all other devices on the network to learn about services. Individual devices create their own internal database of the devices and services available on the network. A second architecture is centralized, where a device is "elected" to be the centralized repository of information. This centralized device then periodically broadcasts its presence for new devices coming on the network and all the while aggregates information about all other devices and services on the home network. Other devices then query the centralized device when they need specific information about other devices and services on the network.

Home networking interoperability software provides several key functions, including device discovery, service discovery, and playback control. These issues will be described in the following sections.

<u>Device Discovery</u>

Device discovery is the process by which a device learns about other devices on the home network. At the device level, the information exchange is on the order of Ethernet Media Access Control (MAC) addresses and Internet Protocol (IP) addresses. By learning the addresses of each device on the network, the devices are able to contact each other and request content and services.

<u>Service Discovery</u>

Service discovery provides a means for devices to advertise the services they can offer to other devices on the home network. For instance, a device could advertise that it stores content, or that it is a printer, or that it is an audio and/or video playback device.

Consider the case where a user wants to playback audio to a set of speakers in the kitchen. Service discovery information is used to create a list of playback devices that is categorized by audio speakers. The user would choose the speakers in the kitchen. In the case of looking for a networked printer, the user would call up a list of all printers. From there, the user could further differentiate the printers based on service characteristics such as black and white printers versus color printers.

<u>Playback Control</u>

Once all the devices are connected, and all are discovered (including their services), the fun can begin for the consumer. Interoperability software also includes the mechanisms to search and play content. Searching can be implemented seamlessly across devices. When a user requests a list of all the digital photographs available, they are presented a list of all the digital photos on the network, regardless of the specific device on which they are stored. This is the scenario that is closest to the VOD service of today, except that the service is entirely in the home. With content on various storage devices in the home, other devices can search that content and play it back using normal controls such as fast-forward and rewind. To continue the VOD analogy, its like having movies cached both locally and centralized. When the user searches the VOD titles, they are not aware of where a particular title is stored, just that it is available for viewing.

Industry Initiatives

As stated earlier, various industry initiatives, such as UPnP and Rendezvous, are promoting their solutions for interoperability software. There are no less than eight initiatives out there and several suppliers are developing additional solutions.

The key about all these initiatives is that they are not interoperable. While they generally all use the same underlying standard protocols and procedures (e.g., DHCP, SLP, SOAP, XML, etc.), they do so in ways that are not interoperable. As explained a bit later in this paper, CableHome is positioned to take advantage of the innovations in this space.

APPLICATION FRAMEWORK

Description

The application framework is the third dimension of home networking and allows applications to be device independent. That is, an application can run on any appropriate device regardless of supplier, operating system, underlying hardware, etc. The relationship of the Application Framework to the rest of the device is shown in Figure 4.

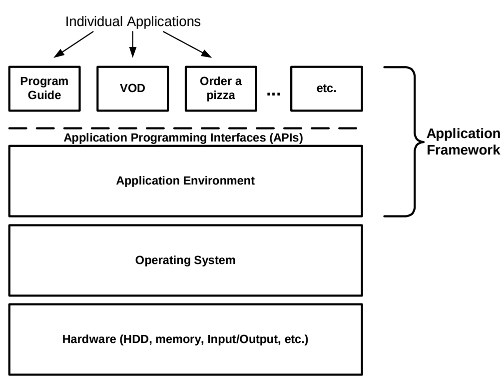


Figure 4

The application framework consists of the Application Environment, the APIs, and the individual applications. This software is portable to different types of devices running different operating systems.

Individual applications are written to run within a specific application framework. The framework itself is ported to various operating systems. In this way, the application, which the service provider cares about, works on various supplier boxes that have implemented the application framework. For instance, a particular application framework (including all the applications) should be able to run on a device with an Intel processor running a Microsoft operating system or a device with a SUN processor running the Solaris operating system. It is this device portability that makes an application framework so powerful. The operator no longer has to worry about the underlying hardware device, rather, the operator just manages the applications that run on that device.

Application Environment

The application environment is a place where software applications are allowed to run. The environment supports both one or more programming languages that developers use to write applications and various tool to manage those applications. A widely recognized application environment is PersonalJAVA™, which is an environment for applications written in the JAVA™ programming language.

The application environment includes tools that are specific to the programming language. This support can include complex technical tasks such as transaction management, state management, resource pooling and security checks for all the applications running within the environment. This is very technical stuff the consumer should never know about, but it ensures the integrity of the applications running within the environment, especially since several applications can be running simultaneously.

Application Programming Interfaces (APIs)

APIs can be thought of as standard subroutine calls that developers use to create applications. These calls include everything from mundane tasks such as writing to memory to more exotic tasks such as drawing a graphics overlay on a display device. The APIs are the “Rosetta Stone” of portable applications. The underlying operating system and hardware speak a low-level machine programming assembly language. The applications speak a high-level programming language. The APIs translate between these two languages allowing humans to write in a high-level programming language and allowing the machine to operate on low-level assembly language.

Security and Rights Management

No discussion of home networking would be complete without including security and rights management.

Security is fairly straightforward, the data carried on the network should be private such that it cannot be snooped, especially in the case of wireless datalinks.

Rights management is an important topic that is currently being debated not only within the cable industry, but also by other industries as well including broadcasters, consumer electronics, content producers and satellite. There are several examples or rights management technologies available, including DFAST (Dynamic Feedback Arrangement Scrambling Technique) and HDCP (High Definition Copy Protection).

For the purpose of this paper, suffice it to say that rights management is needed and there are many groups working on a consensus solution that hopefully can be presented in the not too far distant future.

CableHome™ Positioning

The CableHome initiative, which seeks to extend cable-delivered broadband services throughout a customer’s home, issued key specifications for the 1.0 version in April 2002. Also in 2002, CableHome achieved international standardization at the International Telecommunications Union (ITU).

Whereas most Interoperability Software industry initiatives are concentrating specifically on connectivity within the home, CableHome started with a focus of first connecting the home to a high-speed data connection over cable using a residential gateway. This is illustrated in Figure 5.

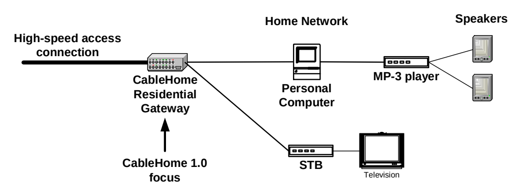


Figure 5

It is expected that the other initiatives will eventually focus on the high-speed connection as well, and CableHome in turn is beginning to focus on solutions for device and service discovery and the accompanying security issues of rights management and content protection.

Among the residential gateway features provided by CableHome 1.0 are mechanisms for secure remote management and configuration of residential gateway capabilities that include DHCP, DNS, NAT/NAPT, LAN test tools and a firewall.

CableHome key features:

- CableHome helps ensure customer privacy because it does not provide cable operators with the ability to probe or configure consumer devices, such as PCs, within the home.

- CableHome also helps protect customer privacy by providing cable operators with the tools that assist in mitigating unauthorized snooping of Wi-Fi based home networks.

- Consumers only use this additional CableHome functionality (e.g., remote diagnostics) if they have voluntarily chosen to subscribe to a cable operator's home networking service. CableHome equipment will be deployed in those households where consumers have elected to pay for this additional functionality because they see value in having the cable operator manage the technical complexity associated with the deployment and operations of a residential gateway.

CableHome is a tool to extend cable-delivered services to the home.

Summary

Home networking is a complex topic. There are three clear and distinct dimensions that comprise home networking, and each dimension provides a new set of technologies and a new set of suppliers.

Author Contact

Doug Jones
YAS Broadband Ventures
300 Brickstone Square
Andover, MA 01810
voice: +1 978.749.9999 x226
doug@yas.com

THEFT OF SERVICE IN HIGH SPEED DATA SERVICES: A WAY TO DEAL WITH THIS DIFFICULT PROBLEM

Jonathan Schmidt
PerfTech Bulletin Services

Abstract

The residential HSD/Internet services environment is experiencing rapid technology changes. The providers' management tools have not kept pace. That has exacerbated the vulnerability and extent of service theft. Consequently, the tools must change. Some have, as in the newer set-top boxes; DOCSIS 1.1 CMTS/modems use strong certificate-based authentication to prevent service theft through modem cloning and spoofing. At the same time, home networking equipment has become very inexpensive, is common even in MSO offerings, and has filled computer store shelves. With this equipment, the home network devices become anonymous and can exist in large numbers behind a single modem. Similar to the downstream portion of the analog video cable plant, high quality Internet access can be replicated in a tree-and-branch architecture behind that modem and serve many users. WiFi products enable this architecture to be implemented with invisible, wireless links that are also open to opportunistic taps. These links are not susceptible to on-site inspection to check for redistribution.

This paper presents an in-band communication channel that can target all workstations on suspected accounts with unblockable screen alerts, which will utilize these invisible links to become an essential tool in combating such theft.

BACKGROUND

A lot of effort has been devoted to the control of HSD network components to eliminate service theft from modem cloning and spoofing. That has left one remaining service leak — the connection behind the cable modem.

Stopping Modem Spoofing and Cloning

Control of the subscriber modem has been achieved through certificate-based authentication in both configuration and data transfer in DOCSIS 1.1.

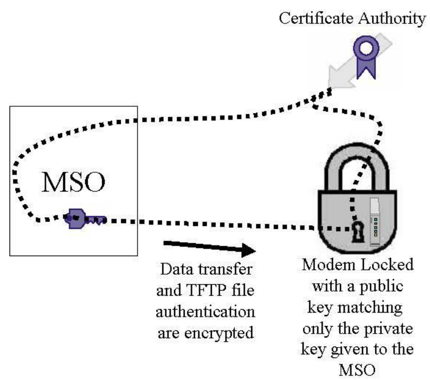


Figure 1: Certificate-based security in modem data transfer and BIN file authentication

Stopping Address Theft and Disruptive CPE Address Configurations

Control of the subscriber CPE network address configuration has also been implemented. With Cable Source-Verify, the DOCSIS MAC domain is protected against rogue CPE configurations for unassigned IP addresses or duplicate IP/MAC addresses.

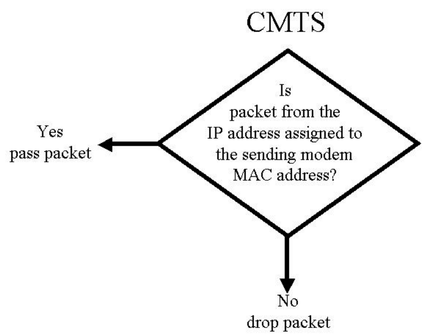


Figure 2: Cable Source-Verify control of Address assignment on DOCSIS network domain

Theft of Service Behind the Modem

A single home networking NAT gateway can be inserted between a network of PCs and the HSD modem and can appear to the network management as a properly configured, single user. However, it can serve high quality Internet access anonymously to many users both inside and outside the account residence through a tree of wired and wireless links. Those users who are outside the primary residence are engaging in theft-of-service according to the Terms of Use clauses in most MSO service agreements.

THE RAPID GROWTH OF HOME NETWORKS BEHIND THE MODEM

The home gateways provide security, anonymity, and the support of multiple workstations through many-to-one address translation, NAT.

Early MSO Rejection Became Acceptance

Early MSO service agreements included wording that would preclude the use of multiple computers through a gateway.

In the first years of HSD Internet provisioning, some attempts were made to detect the number of computers behind a gateway. An example can be found in the paper "A Technique for Counting NATted Hosts," by Steven M. Bellovin of AT&T Labs Research. The paper concludes that the technique is imprecise in the single case, and may only be of use in estimating workstation populations behind home gateways.[1] Indeed, the mechanism doesn't work at all with some gateways, such as Nortel's Instant Internet. It is also easily confused by internal LAN activity such as Windows intra-network activity. The difficulty in definition and detection of the NAT gateways caused very casual, if any, enforcement of anti-gateway provisions.

Today, for example, Time Warner Cable offers sales, support, and installation of multiple user home gateways and home networking equipment. Home networks are a source of new subscribers and of additional product revenue.

In Comcast's recently acquired AT&T Broadband networks, they promote the sale of multiple IP addresses (up to 4) for extra cost as a solution to the problem of using multiple computers on a single modem. The Terms of Use agreement doesn't appear to forbid the use of NAT gateways. Computers that aren't "directly connected to the modem" are not supported. However, one of their current FAQs provides a configuration for a LinkSys NAT gateway.

Gateways Pass the Knee of the Adoption Curve

The rapid rise in the rate of installation of home gateways has followed the drop in prices ($500 to $50 in 5 years) and the improved simplicity afforded by a synergistic and automatic configuration of the gateway-HSD and the PC-gateway interfaces.

Microsoft is Focusing on Home Networks

Microsoft now adds protocols to their operating systems and to their hardware gateways to expand the services that can pass through the address translations.

Microsoft intends to incorporate everything from the refrigerator to multi-media components into the home network behind such gateways. This would indicate a continued, increasing population of home networking installations.

WiFi Increases Home Networking Use

Additional demand for gateways has been fueled by the similar reduction in price and greater simplicity of wireless LAN components, mostly the standard called WiFi.

Additional home computers that had been beyond easy reach of wired Ethernet are easily and inexpensively integrated with the home network with these wireless LAN components.

WiFi Is Encountering a Very High Growth Rate

- Shipments of WLAN products increased more than 100 percent in 2002, and will continue growing as the stellar wireless market performer for the next several years.
- Average equipment prices overall dropped by almost 28 percent, while revenues increased by more than 50 percent in 2002 to $2.6 billion[2]
- The WLAN equipment market will continue growing at a 43 percent compound annual growth rate (CAGR) to $10.3 billion in 2006.[3]

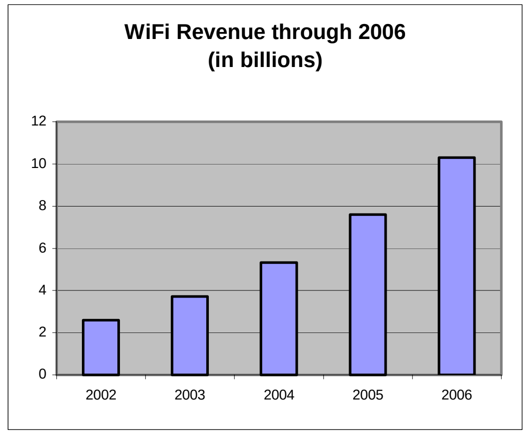


Figure 3: Estimated growth in revenue in WiFi market through 2006

AN ENVIRONMENT INVITING REDISTRIBUTION

Redistribution is Simple to Wire up and with WiFi, Easier than with Analog TV

Home gateways, WiFi components, hubs, switches, cables, and PCs are very simple to configure in a home network to give all PCs access to the Internet. In fact, for the most part, the components configure themselves. An infrastructure required for redistribution within a multi-tenant building, for example, can be constructed entirely with parts and technical assistance from a Radio Shack® store.

Redistribution Can Support Many Users

The high inherent bandwidth of the HSD Internet connection and the bursty data demands of the typical browsing or e-mail user allow these environments to redistribute high quality Internet access to dozens of simultaneous users.

Wireless LANs Offer More Opportunity

WiFi access points can be connected to home gateways entirely within the broadband service agreement. These access points, when connected to a home gateway, can provide anonymous and uncontrolled access to the account's Internet connection over a wide area as large as hundreds of feet around it.

The gateway user's assumed network anonymity combined with a widespread "the Internet is free" attitude adds encouragement to those who are inclined toward service redistribution or an opportunistic wireless tap. Users of WiFi have easy access to programs such as "NetStumbler" and "MiniStumbler" that provide the user with a graphic display of all the locally accessible, unencrypted WiFi networks and the necessary "SSID" to sign on (SSID: ***S****ervice* ***S****et* ***Id****entifier*, a 32-character unique identifier attached to the header of packets sent over a WiFi network). NetStumbler is used with PCs and MiniStumbler is used with pocket computers. These programs are used by ordinary hobbyist computer users and are not restricted to hackers. Open WiFi networks in populated areas don't remain a secret for long.

The Billing Method Doesn't Discourage Redistribution

The billing method for most HSD Internet access accounts is a fixed rate for a fixed bandwidth high-speed connection. There is no limit on consumption and, therefore, no extra cost to the account holder for extra use. This method is unlike that for water and electricity, for example, but is similar to the downstream portion of an analog TV cable plant and can be expected to similarly encourage theft. Redistribution subjects the account holder to no additional billing if not caught.

MSO are expected to offer tiered services based on consumption and that may reduce the illegal practice both because of the potential for extra charges to the tiered services account holder as well as the conversion of some participants to being subscribers of the more attractive, less expensive, low consumption service.

However, collecting additional revenues from high-bandwidth users accustomed to "free" access could be problematic unless addressed before tiered services are introduced.

DETECTING REDISTRIBUTION

Redistribution is, indeed, a problem since each instance represents one or more potential subscriptions that are lost.

It does not appear that there is a method to attack the problem by prevention either in configuration or warnings in the terms of service. Therefore detection and response to detection represent the remaining option.

Detection with Network Equipment

Detection of likely incidents of service theft should be much easier than with analog cable TV theft since the multiple users behind the gateway create a protocol stream that has a variety of signatures that indicate the number of individual users.

Although, as shown in the Bellovin AT&T Labs paper, identification of individual accounts with multiple users is imprecise with standard network devices that observe at L3, specialized devices could be used to identify such accounts when inspecting other layers. For example, data passing through a device similar to a firewall could identify accounts exhibiting many simultaneous connections or the use of many versions of browsers, or the checking of many different e-mail accounts. Although it is possible within the character of the protocol, equipment to perform such checks has not been brought into such service.

Detection of likely redistribution situations is not necessarily an indication of a definite redistribution event. It is only a violation if users outside the account residence are accessing the Internet through it.

Drive-by Detection of Open WiFi Points

The default installation of most WiFi components establishes the network without encryption. These networks can often be detected with simple PC equipment and freely available software such as NetStumbler. When detected in this way, the drive-by inspection PC can also detect if the network has an Internet connection through the HSD network of the MSO.

Interestingly, unsecured WiFi equipment is not in violation of the Terms of Use agreement in force with the subscriber and, therefore, is not an indication of a definite event of redistribution. It is only a violation if users access the Internet through the network from points outside the account residence.

Other External Indicators

There are also other external flags to indicate potential redistribution activity such as:

- Blatant advertising of wireless "WiFi hotspots" at the location of a residential broadband subscriber
- Mass service cancellations at a multi-tenant location.
- Alert service representatives who spots trends of returned modems in a concentrated area.
- A serendipitous tipster

REAL SERVICE THEFT, VAGUE INDICATIONS, INVISIBLE LINKS

The Vulnerable Revenue is Real

"Comcast has found that the Internet business has become even more profitable than providing basic cable service. The capital costs are lower: a cable modem costs $50 compared with a television set-top box at $225. And churn - the rate at which customers cancel their service - is far lower for Internet service than for video."[4] The HSD service and subscriber base is a major asset. Vulnerabilities can chisel away at the revenue and do it well before the loss is measurable with current tools.

The Service Theft Problem is Real

The NCTA "2000 Report on Cable Industry Lost Revenue" indicates theft of video services is 10% throughout all service levels from basic to premium.[5] These figures occur despite the fact that it is illegal to possess cable descramblers in 32 states.

HSD networks face corner-store availability of the complete set of equipment, legal to sell and own, required to implement an extensive wired or wireless re-distribution system.

The Indications are Frustratingly Vague

The network indicators of service re-distribution, although potentially more varied than that for cable TV, can be difficult to extract and would often flag a benign situation.

Ambiguous evidence for potential service redistribution requires a more controlled response than the traditional service disconnection and personal confrontation.

Invisible Links are Difficult to Spot

On-site inspection, the major tool of cable TV service theft control, is clearly ineffective in the case of HSD propagation through wireless links.

REAL TIME MESSAGING TO ALL SCREENS BEHIND A TARGET MODEM

This paper describes the utility of an in-band communication channel that can target all workstations on suspected accounts with unblockable screen alerts. It can utilize these invisible links to reach all hidden workstations. It is an effective new tool for this new problem.

Unlicensed Users Will Know They Have Been Discovered

Informational, personalized warnings are effective in dissuading service theft without the unpleasant, confrontational contact that normally accompanies service theft situations. By using the providers' own channel, it becomes a way to communicate with both the unlicensed users as well as inadvertent opportunists who grab service through the anonymous Ether of WiFi.

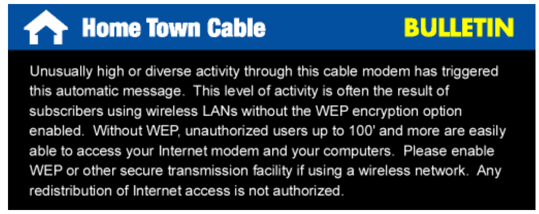


Figure 4: Sample bulletin

Immediate Messaging to Every PC User Is Not Available with Traditional Means

Up to now, there has been no reliable mechanism to reach anyone on a PC with or without a gateway. Most of a provider's subscribers don't have a known e-mail address and of those who use the provider's e-mail address, few of them check it regularly. A recent Forrester report determined that over one-third of all e-mail users change their address over the course of a single year. The variety of incompatible instant messaging systems is not useful, and not just because of the incompatibility.

Many providers are opposed to maintaining a software component on subscriber PCs. Software components rapidly degenerate to nonfunctionality. And, a sizable portion of subscribers will blame any malfunction of their computer on the provider-installed software.

Other Communications Channels are Unreliable and Untimely

Telephone calls go unanswered, door knocks are expensive and have many other problems, and bill stuffers are mostly thrown away, unread.

Creating a Mechanism to Allow the MSO to Use their Own Channel to Communicate with Subscribers

The normal Internet Protocol does not provide for any communications to a browsing subscriber except for the pages that are requested from the destination site or from pages, in turn, linked from the destination site.

The constraints of the carrier-grade requirement of the MSO limits the opportunities for a solution and defines the requirements:

- Absolutely no software or configuration change required at the subscriber PC
- Operational on any workstation, PC, Macintosh, Linux, or other device with a browser even through wireless links
- Transparent to DOCSIS versions
- Operational in the presence of any gateway or firewall or pop-up blocking software
- Installation in the provider network while the network is operating
- Failsafe operation in that any failure in the delivery system does not affect the normal operation of the network
- Non-disruptive to the network operation and throughput
- Security controls to guarantee access functions from only authorized personnel and locations
- Delivery to individual targeted users or described groups with a characteristic determined by a database or subnet
- Proof-of-delivery receipt
- Delivery of fully interactive screens in order to provide abuse warnings and, in the case of virus contamination alerts, a screen that also includes the remedy

The facility to accomplish this type of communication that meets the above requirements was developed in 2002 and deployed testing has begun by Perftech Bulletin Services and the MSO WideOpenWest.

The installation consists of router-attached devices called Bulletin Directors that are distributed to aggregation routers

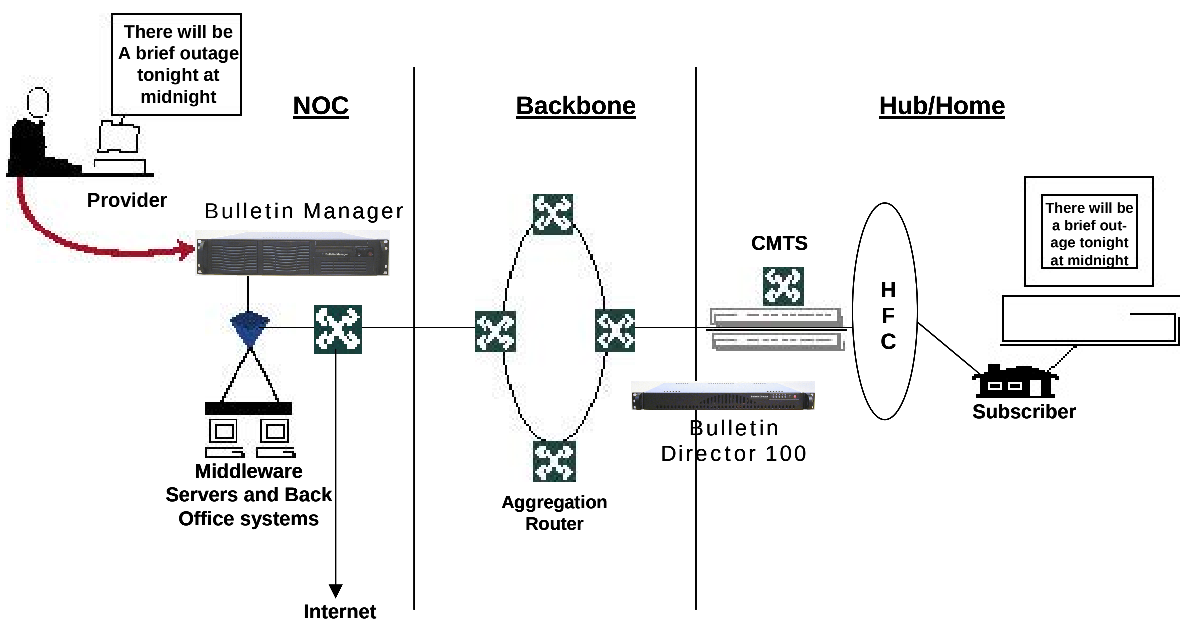


Figure 5: Sample network configuration

and a management device that administers the policies for bulletin delivery.

Installation takes place while the network is up and running using a fail-safe layer 4 redirection only of upstream port-80 traffic. None of the downstream traffic or other upstream traffic passes through the director devices assuring both performance and privacy. There is no measurable impact on network throughput.

As expected in the initial deployment, installation was made into an operating network and the network performance was unaffected. Test bulletins are successfully delivered to the targeted accounts.

<u>How It Will Be Used</u>

The communication can be used to warn and stop likely abusers by dispelling the illusion that “they don’t know I’m here.” Every unlicensed screen will display a selected alert whether inadvertently or intentionally involved in the theft of service. Warnings can become sufficiently insistent that surfing becomes difficult.

The result can be expected to turn many unlicensed users into paying subscribers. An analysis by Time Warner Cable as presented in "Cable And Broadband Security Case Study: The Broadband & Internet Security Task Force"[6] demonstrated that nearly 1/3 of all discovered unlicensed cable TV connections were turned into paying subscribers. Unauthorized re-distributors can also be turned into commercial accounts when they need to continue the re-distribution.

<u>An Economical Solution</u>

The cost of this solution is a small. Amortized over several years, the cost per subscriber is in cents per year. That is very small when compared to the cost of on-site visits and personal confrontation with the account holder.

<u>A Broadly-Applicable Support Tool</u>

Unlike cable TV or telephone service where the subscriber's premises device is a fixed, passive entity, HSD Internet broadband access is a complex technical partnership between the provider’s network and the subscriber's computer and configuration — all of which is required to sustain Internet accessing activities. Immediately available provider-subscriber communication is clearly a requirement for proper support of such a system.

For those subscribers who do not know that their personal WiFi network is being abused and hacked and, therefore, insecure, the revelation is welcome.

Advance-warnings of system outages would be more than welcome by users who have incorporated the Internet access into critical work-at-home activities, stock trading, or auction participation. It would also be welcome at the provider’s support call center where the call floods that accompany such outages would be significantly diminished.

Alerts should be issued about temporary outages in subsystems such as e-mail servers. Informed users are not likely to be participants in call floods and are likely to be less agitated and more understandable when kept informed.

Instances of disruptive network operation due to virus contamination in a subscriber’s PC can demonstrate the value of immediate, interactive communication to handle this difficult support problem. In this case, the subscriber can be delivered an alert immediately upon being seated at the PC and browsing. The alert would contain an explanation along with a button that would be a remedial link to decontamination facilities.

CONCLUSION

Theft of service is going to get worse. Support problems are going to get more severe and complex. The subscriber population is going to keep growing at a very high level. The provider's own communications channel is the key tool in the control of all these issues.

CONTACT INFORMATION

Jonathan Schmidt
Vice-President Business Development
210-349-7152
jon@perftech.com

REFERENCES

[1] Steven M. Bellovin, "A Technique for Counting NATted Hosts," from AT&T Labs Research, Nov. 2002.
[2] Forward Concepts, Feb. 2003
[3] InStat/MDR, Feb. 2003
[4] Saul Hansell, "In Broadband, Comcast Lets Users Find Their Own Flourishes, " New York Times, March 17, 2003
[5] "2000 Report on Cable Industry Lost Revenue," NCTA, March 20, 2003
[6] "Cable And Broadband Security Case Study: The Broadband & Internet Security Task Force, Tap Audit," Time Warner Cable, Syracuse, NY Division

TIERED DATA SERVICES TO ENHANCE CUSTOMER VALUE AND REVENUE: WHY ARE DOCSIS 1.1 AND ADVANCED QUEUING AND SCHEDULING REQUIRED?

Gerry White
Senior Director of Advanced Technology
Network Infrastructure Solutions
Motorola Broadband Communications Sector

Abstract

High-speed data services are making significant revenue contributions to Broadband Operators as service penetration rates are now averaging over 10% and exceeding 30% in some markets. Still, there are large potential customer segments that can be captured by offering data service tiers that provide a better match between customers needs and income levels and the features of the high-speed data service required. For example, business customers are willing to pay much higher rates for a service with greater bandwidth and service level guarantees. At the other end of the spectrum, customers who are paying $25-30 per month for dial-up services would jump at the chance to have an "always on" connection at speeds that are double standard dial-up rates.

This paper will discuss how measurable data service tiers can be delivered to customers through Quality of Service (QoS) control enabled by the use of the DOCSIS 1.1 specifications and advanced queuing, scheduling and congestion control techniques.

Additionally the paper will present a mechanism to support overbooking of network resources in a tiered environment, enaling cost effective deployment while ensuring that guaranteed service levels are met for each of the service tiers.

INTRODUCTION

Initial DOCSIS deployments have generally followed a "one size fits all" model. To date the success of this approach in terms of service penetration and revenue generation has been considerable such that cable modems have become the preferred means of Internet access for a significant section of the population.

In order for Multiple System Operators (MSOs) to attract a wider customer base and reap additional rewards from their investments in Hybrid Fiber Coax (HFC) and IP infrastructure it will be necessary to go beyond this initial audience and provide more tailored services to specific target populations. Thus high-end service offerings can be created for business customers offering guaranteed services at bandwidth (and price points) comparable to T1 and business DSL offerings. At the other end of the spectrum there are many dial up customers for whom current cable modem service is too expensive. Providing a rate limited entry-level service for these users can add to the customer base with an immediate revenue impact and also create a pool of broadband addicts, which can be targeted for later upgrade to premium services.

Ultimately all networks are a shared resource and rely on statistical multiplexing between users to provide cost effective service. In the case of the HFC network the shared resource extends to the customer premise, while for a DSL network the customer link may be dedicated copper pair. In both cases the infrastructure network is shared and statistical multiplexing is used to reduce costs. Thus to ensure fairness between users (even for a best effort single service tier network) QoS must be provided to some degree, although it may be relatively simplistic. When tiered service offerings are enabled QoS becomes more complex as it must not only provide fairness between users within a service tier but also differentiate between tiers.

The MSO community is in the fortunate position of having the tools to provide tiered services readily available. In fact in many cases they are already deployed, waiting to be enabled.

DOCSIS 1.1 provides the mechanisms to deliver tiered services over the HFC network and when this is combined with suitable backbone technologies the MSO can deliver end to end service tiers which compete effectively with any in the market.

TRAFFIC FLOWS

At the customer level the focus of interest is on individual applications, the user wants to make a phone call, connect to a corporate network or simply surf the web. In general each application will involve multiple traffic flows. For example a VoIP client will exchange control packets with a call management system and voice traffic with the far end called system. These traffic flows may take different paths through the network and each will require QoS to be provided. The concept of traffic flows and of providing QoS to each flow is central to both DOCSIS and backbone QoS mechanisms.

AGGREGATE VS PER FLOW QoS

Two distinct mechanisms are typically used to provide QoS on an end-to-end basis over large IP networks. These mechanisms have been the subject of research and standardization efforts resulting in the differentiated services (DiffServ) model for aggregated QoS and the integrated services (IntServ) model for per flow QoS [References 1 and 2].

At the edge of the network bandwidth is typically scarce and expensive to provide. Devices in this domain, such as a CMTS, deal with a bounded number of traffic flows (e.g. to a maximum of 8000 flows in each direction in DOCSIS). In this region QoS can be provided most effectively at a per flow level. For each flow a traffic specification and a flow specification are defined. The traffic specification defines a classifier to identify the packets that belong to a specific flow. Typically this is a masked set of fields based on the content of the packet header such as IP addresses, port numbers, DSCP markers, etc.. The flow specification defines the QoS parameters to be applied to the flow (bandwidth, latency…). This mechanism was defined by the IETF as the IntServ architecture and was adapted to provide the basis for DOCSIS 1.1 QoS [Reference 3]. DOCIS 1.1 provides for a signaling mechanism to set up new flows, for admission control functions and for isolation between flows. The CMTS has the primary responsibility to provide QoS in the DOCSIS realm by implementing admission control and providing isolation between flows based on the upstream and downstream scheduling mechanisms.

In the core of the network bandwidth is relatively abundant (and cheaper) but the infrastructure is shared between tens of thousands of clients. Thus core switches and routers must support hundreds of thousands of flows. If these devices were to operate at an individual flow level the amount of data to be maintained would be massive and systems would simply not scale so that an aggregated mechanism is needed. To achieve the scaling required packets entering the DiffServ domain are classified into one of a limited number of behavior aggregates (64 max.). DiffServ defines a field in the IP header of the packet known as the DiffServ code point (DSCP). Systems at the edge of the DiffServ domain mark packets with the code point desired before transmission into the domain. All packets with the same DSCP to be transmitted on a given link are considered part of the behavior aggregate and are to be treated in the same manner. The DSCP defines the required behavior for each packet so that no per flow state must be maintained. Routers within the DiffServ domain use the DSCP to determine the QoS desired by the packet and apply the appropriate queuing and scheduling algorithms to achieve the required QoS.

Multi Protocol Label Switching (MPLS) [Reference 3] has recently emerged as another mechanism to provide aggregated QoS in metropolitan and core networks. It defines a mechanism to set up label switched paths (LSPs) between endpoints at the edges of the MPLS network. Packets entering the network are assigned to a particular path and a label added to the packet identifying the LSP to be used. This label is used by the MPLS routers in the core of the MPLS network to forward the packet to its destination endpoint. The core network does not need to examine the packet headers beyond the label and can therefore focus on switching traffic to its destination as quickly and efficiently as possible. Each LSP can be associated with a defined forwarding equivalency class (FEC) which defines specific QoS parameters so that the MPLS network can provide a mesh of QoS enabled paths.

In a typical MSO network environment the transition point between per flow and aggregated QoS domains occurs at the intersection of the HFC/DOCSIS and metro/IP networks. The HFC access network is based on the DOCSIS 1.1 protocol, which provides QoS to applications on a per flow basis. In the upstream direction an application flow is mapped to a service identifier (SID) and in the downstream direction to a service flow identifier (SFID). The cable modem (CM) and CMTS cooperate to assign the required QoS to each flow and to ensure that it is met. The MSO metro networks that connect the CMTS systems are typically based on gigabit Ethernet, Sonet or RPR physical infrastructure. Historically an IP routing infrastructure ran on top of this possibly providing QoS based on the DiffServ model. Newer networks replace the pure IP routing model with one based on MPLS infrastructure. Both types of network provide the MSO with the capability to provide aggregated QoS based on traffic engineering and provisioning. In either case the CMTS/ER provides the transition point between the QoS domains. The major issues, which must be resolved at this demarcation point, will be considered later in this paper.

QoS MECHANISMS

The ability to deliver QoS involves four key functions:

- Classification of packets to determine which flow a packet is part of and the appropriate service level for each traffic flow
- Policing of traffic to prevent flows from getting higher than agreed upon service levels
- Buffering to ensure that queues are created to contain packets during periods of congestion
- Scheduling to enforce packet transmission in accordance with QoS policy.

In order to provide QoS successfully the CMTS/ER must provide this functionality for both HFC and metropolitan network environments.

QoS IN THE HFC NETWORK

The DOCSIS 1.1 specifications provide QoS for the cable access network. They define enhancements to the Media Access Control (MAC) protocol of DOCSIS 1.0 to enable more sophisticated access methods over HFC access networks by adding the following:

- Packets are classified into service flows based on their content. Thus each application can be mapped to a unique service flow.
- Network access (upstream and downstream) is scheduled per service flow using one of a number of defined scheduling mechanisms including constant bit rate, real-time polling, non real-time polling and best effort.
- Service flows may be configured through management applications or created and deleted dynamically in response to the starting and stopping of applications.
- Fragmentation of large packets is required to allow low latency services to operate on lower-bandwidth upstream channels.

These features provide the basic tools for QoS management. They allow applications to request QoS changes dynamically and allow providers to isolate multiple data streams from each cable modem, set-top box or MTA. DOCSIS 1.1-based systems can therefore potentially deliver the ability to allow application-specific QoS treatment within the HFC access network for each traffic flow.

Packets transmitted into the network from a host system are treated as follows. The interaction between the CM and the CMTS upstream scheduler is shown in Figure 1:

1. The CM filters each packet and classifies it into a service flow identified by a unique SID.
2. The CMTS schedules upstream transmission for the SID based on its QoS parameter set and the traffic history of the SID. This is the most complex and the critical step for providing upstream QoS.
3. The CM transmits the packet to the CMTS
4. The CMTS receives the packet and reclassifies it based on CMTS configuration. In general the CMTS is the first trusted device in the network and should not rely entirely on user or CM packet classifications.
5. The CMTS maps the packet into the QoS scheme for the MAN (mark traffic for differentiated services forwarding, map traffic to MPLS LSP tunnels, map traffic to physical interface)
6. The CMTS queues the packet to the egress link based on the required QoS
7. The CMTS implements its network interface scheduling algorithms and transmits the packet to the link when it reaches the head of queue.

Packets received by the CMTS from the network are treated as follows and shown in Figure 2

The CMTS receives the packet and classifies it into a downstream service flow identified by a unique SFID.

1. The CMTS enforces policing on maximum rate if required.
2. The CMTS queues the packet to the egress link based on the required QoS
3. The CMTS implements its downstream scheduling algorithms and transmits the packet to the link when it reaches the head of queue.
4. The CM forwards the packet to the host.

QoS IN THE MAN/WAN

There are two primary methods for providing QoS control in the regional network, packet based and connection-based.

In the packet based case, individual flows are policed and marked at the edge of the network using a DiffServ DSCP marker in the IP header, so that aggregated flows are delivered to the network core with each flow tagged for the appropriate QoS treatment. The DiffServ standard defines the code points to use and the per-hop forwarding behavior to be applied to each marked packet.

Connection-based QoS can be implemented using MPLS. In an MPLS network a number of paths are established between the end points of the network. Each path can be traffic engineered to provide a defined level of QoS. All packets on the path share the same forwarding equivalency class (FEC) consisting of the MPLS end point and the QoS parameter set. As with the DiffServ case the packets from the individual flows are policed and marked at the edge of the network. In this case the marker is an MPLS label that is prepended to the packet and identifies the path through which the packet will traverse the MPLS network. Each path is referred to as a label switched path or LSP. A single LSP, with a defined FEC can support multiple flows and thus provide the necessary aggregation mechanism.

QoS AT THE BOUNDARY

In order for applications to see real benefits, QoS must be provided on an end-to-end basis. Thus the QoS-enabled traffic flows from the HFC access network must be mapped to the QoS mechanism(s) used in the regional or backbone networks. The CMTS/ER at the boundary must be able to perform this mapping and implement the QoS mechanisms for the two domains. To deliver QoS at a per service flow level this must take place for multiple flows at wire speed. The mechanisms employed within the CMTS/ER must maintain the QoS during this transition. The problem is complex due to QoS requirements constantly changing as service flows are created and deleted dynamically.

The metro to HFC boundary is also a capacity transition point. In the downstream direction packets received from gigabit speed optical links must be transmitted onto megabit capacity DOCSIS networks. Thus the CMTS/ER must implement congestion management based on the conformance of the subscribers and applications to their Service Level Agreements (SLA's).

The queuing and scheduling mechanisms used within the CMTS/ER to implement the transition between the per-flow HFC and aggregated metro domains will determine how successfully QoS can be delivered. The key concepts required to provide this transition successfully are:

- Queuing and Scheduling
- Congestion Control

QUEUEING and SCHEDULING

Three queuing and scheduling mechanisms will be considered; FIFO, class based and per flow.

FIFO Queuing

First-In-First-Out (FIFO) queuing is both a queuing and a scheduling mechanism. In FIFO queuing, all packets are stored in a single queue and are transmitted in the order that they are received. FIFO queuing is easy to implement and it requires little configuration. Unfortunately it does not provide any support for the differing QoS levels required by diverse applications. In a FIFO scheme packets for a low-latency service can be queued behind those from high bandwidth services and must wait for these to be transmitted.

Class-Based Queuing

Class-based queuing (CBQ) attempts to avoid this problem by sorting the traffic into different classes by examining the packet and trying to determine the type of traffic to which it belongs. Once the packet has been classified, it is placed in a FIFO queue that contains only other packets of the same type. Each per class FIFO queue can be serviced according to configured policy to provide the behavior required. Thus a FIFO that is serviced frequently so that it is usually empty of packets could provide a low latency, low loss service such as VoIP. Similarly a FIFO that is kept full can provide a service such as bulk file transfer, which requires high bandwidth but can tolerate moderate packet loss.

In theory, this allows the FIFO of each class to provide the desired type of service, but in practice there are a number of problems with this approach. It requires a constant, heavy configuration burden because the operator has to configure the allocation of service to the different classes e.g. 1/10 for e-mail, 1/10 for voice, 1/3 for Web traffic, etc. In a dynamic network environment, the class-based queuing method of allocating service is impractical because the allocation is independent of the number of users of a given class. CBQ does not provide application isolation, as although each queue contains traffic from a similar application type (e.g.VoIP), flows from multiple users are mixed within the queue.

Per-Flow Queuing

Per-Flow Queuing (PFQ) solves the problem of providing isolation between application flows by assigning each packet stream to its own queue. Those queues for flows with QoS reservations are served at their guaranteed rate while flows without reservations are served in a round robin or fair-share manner. Thus the queue for each flow is served at the rate defined in the service level agreement.

To assign flows without reservations to a queue a method known as stochastic queuing is used. In stochastic queuing the parts of the packet header that are the same for all packets of a flow, such as the source and destination IP addresses and source and destination port numbers are fed to a hash function that is used to map the packet to a queue. This ensures that all packets for the same flow are mapped to the same queue (to avoid miss ordering). It also eliminates the need to configure bandwidth shares per-class, and consequently avoids the miss-allocation caused by varying usage patterns. If the system can support more queues than there are flows then most flows either have their own queue or share it with a small number of other flows. To support stochastic per-flow queuing the CMTS/ER must support thousands of flows (with DOCSIS 1.1 a CMTS can support 8000 flows in each direction per DOCSIS domain).

CONGESTION CONTROL

Given that the data rate in the MAN will be significantly greater than the capacity of an HFC link it is also important to consider congestion control mechanisms, especially as applied to the downstream traffic. Three mechanisms to handle congestion will be discussed; tail drop, Random Early Detection (RED) and Longest Queue Pushout (LQP).

Tail Drop and RED

(Refer to Figure 3)

With all packets sharing the single queue for simple FIFO or all packets of the same type sharing the same queue for CBQ there are limited options for congestion control. The simplest scheme is simply to drop packets from the tail of the queue when no buffers are available for them. This takes no account of the content of the packet and hence it’s potential value or of whether the flow to which it belongs is in compliance with its service level agreement. Mechanisms such as RED attempt to solve this problem by randomly dropping some of the arriving packets when the queue starts to fill. The intent is that the end systems using a windowed flow control, such as TCP will notice the packet loss and slow down. Thus it is dependent on the end systems of the flow to slow down transmissions to solve the congestion problem. While this might work in a well-controlled environment such as an enterprise network it is unlikely that the end systems will be as cooperative in a public network. FIFO queuing does not provide a mechanism to ensure isolation for well-behaved flows from miss-behaved applications generating heavy loads.

As the number of flows sharing the FIFO queue increases this problem becomes worse. The limitations of class-based queuing and RED often result in the inappropriate discarding of packets. Since traffic flows are lined up in shared queues, it is impossible to isolate and discard those flows that are exceeding service level agreement (SLA) guarantees before discarding traffic flows that are staying within their SLAs. Packets are therefore discarded randomly from shared queues as they become full.

Longest Queue Pushout

(Refer to Figure 4)

In a per-flow queuing environment longest queue push out (LQP) is the mechanism that best meets the congestion control requirements. LQP allocates buffers to the individual flows as required until 100% of the buffer pool is used. When no buffers remain and a new packet is received, LQP discards traffic from the flows, which are the longest queues. The scheduling system is transmitting from these per-flow queues at a rate that matches the QoS assigned for each flow. Thus by definition the longest queue is that which is exceeding its allocation by the greatest amount so that traffic is automatically discarded from those applications which are non-compliant with their SLAs. This occurs without the need to configure congestion control.

Further details on PFQ performance can be found in [Reference 5].

OVERBOOKING TIERED SERVICES

In order to create a commercially viable network the operator must rely on statistical multiplexing. Not all users are active at the same time so that the network can be over subscribed to reduce the cost per user. In a single tier best effort network overbooking is relatively simplistic. Subscribers are added to the network until a local heuristic, such as the number of cable modems per upstream or a defined traffic load, is reached. In a tiered network overbooking becomes a more complex process and should be applied at each service tier as well as on each network interface. The amount of over booking is dependent on the types of services to be offered (e.g. guaranteed vs. best effort) and on the traffic patterns of the users within each service tier.

Service Classes

Responsibility for providing QoS to a service flow resides with the CMTS, which must be configured with the parameters necessary to control this operation. In order to simplify this configuration and help operations staff retain their sanity the concept of service classes has been introduced. DOCSIS 1.1 has defined a set of QoS parameters, including maximum sustained and minimum reserved traffic rates, and a way for associating specific QoS parameter values to service flows. It has further incorporated the concept of a service class name so that service flows, when being created, may be assigned their QoS parameters by referencing a service class name.

Thus an operator may define a number of service classes. Individual service flows will be assigned to a service class and all flows belonging to that class provided with a defined quality of service. Service classes can be supported for both downstream and upstream directions.

In order to facilitate overbooking the concept of the service class can be extended by introducing two additional service class parameters, Maximum Assigned Bandwidth (MAB) and Configured Active Percent (CAP). With these additional parameters the operator can explicitly control overbooking for each service tier and for the network interface in total. With this scheme all service flows must be assigned to a service class. If a service class name is not present in a Registration Request message generated by a CM, then the CM service flows are assigned to a default service class.

Maximum Assigned Bandwidth

Maximum Assigned Bandwidth (MAB) specifies the amount of bandwidth a service class is permitted to consume on an interface. It is expressed as a percent of the total interface bandwidth capacity. The MAB of a service class is applied during admission control to determine whether to admit a new service flow and again by the scheduling algorithms to provide a class-based weighting to the scheduler. Any unused portion of a class' bandwidth may be used 'on demand' by other classes which have a traffic load in excess of their own MAB.

Configured Active Percent

Since not all service flows are active simultaneously the service level classes feature permits customers to overbook service classes. Overbooking means admitting service flows to a service class such that the sum of their guaranteed minimum reserved rates are in excess of the configured MAB for the service class. To control the amount of overbooking, a configurable overbooking factor the Configured Active Percent (CAP) is provided. The CAP is an estimate of how many service flows, expressed as a percentage, are likely to be active simultaneously. For example, if the CAP for a service class is set to 20 percent then it is estimated that only 20 percent of the service flows belonging to that class will be active simultaneously. Therefore, 5x (1 / 0.2) overbooking would be allowed. A CAP of 100 percent means that no overbooking will be allowed. A CAP of zero percent means that unlimited overbooking is allowed.

ADMISSION CONTROL

Admission control is a process wherein the bandwidth requirements of a service flow are checked to verify that admission of the service flow to a service class does not exceed the class' MAB after accounting for the allowed level of overbooking. Service flows are created during modem registration or through dynamic service messaging. A CM registering with primary service flows should be permitted to register regardless of whether the admission of its service flows would exceed its service class' MAB. In this case however the service flow would be admitted in a 'Restricted' state meaning that the service flow will not be provided any guaranteed minimum reserved rate. Service flows created via dynamic service messaging will be rejected if admission of the service flow would cause its service class to exceed its MAB.

Examples of MAB and CAP operation can be seen in Figure 5 and Figure 6. In both examples the bandwidth on the interface is shared between two service tiers a best effort class and an enhanced service class such as a business service tier. In the first example three flows are active in the enhanced class with no flows active in the best effort class.

Initially the total bandwidth is below the MAB for the class so that all flows receive their guaranteed bandwidth. At this point the MAB is reached so that requests for additional flows to be set up would be rejected by admission control (or could receive best effort bandwidth as defined by operator policy). As flow 3 bursts to a higher data rate, above the guarantee, it will share available bandwidth from the best effort service class if this is available.

The second example shows the same 3 flows active with all MAB consumed so that no further flows could be admitted. Flow 3 then terminates at which time bandwidth consumption falls below the MAB for the class and additional flows could be added.

SUMMARY

Providing tiered service offerings has the potential to extend the target audience and revenue stream for high-speed data services.

Extending existing best effort DOCSIS to support service tiers with QoS guarantees requires

- The use of the DOCSIS 1.1 protocol extensions
- Sophisticated scheduling and congestion control mechanisms in the CMTS.
- A means of overbooking, which is service tier aware.

All of these features are available in CM and CMTS systems, including much of the currently installed equipment base.

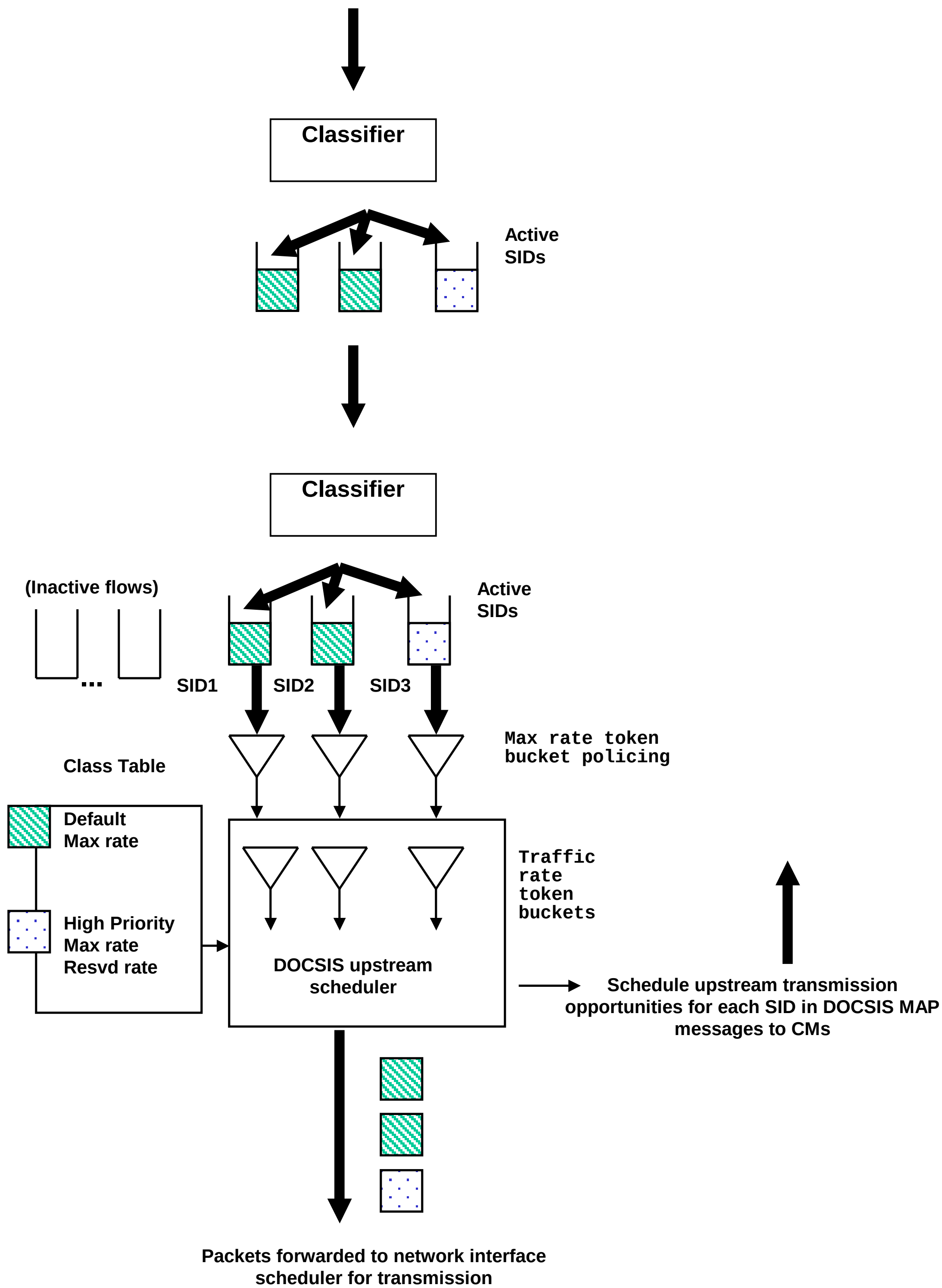


Figure 1 Interaction between CM and CMTS upstream scheduler

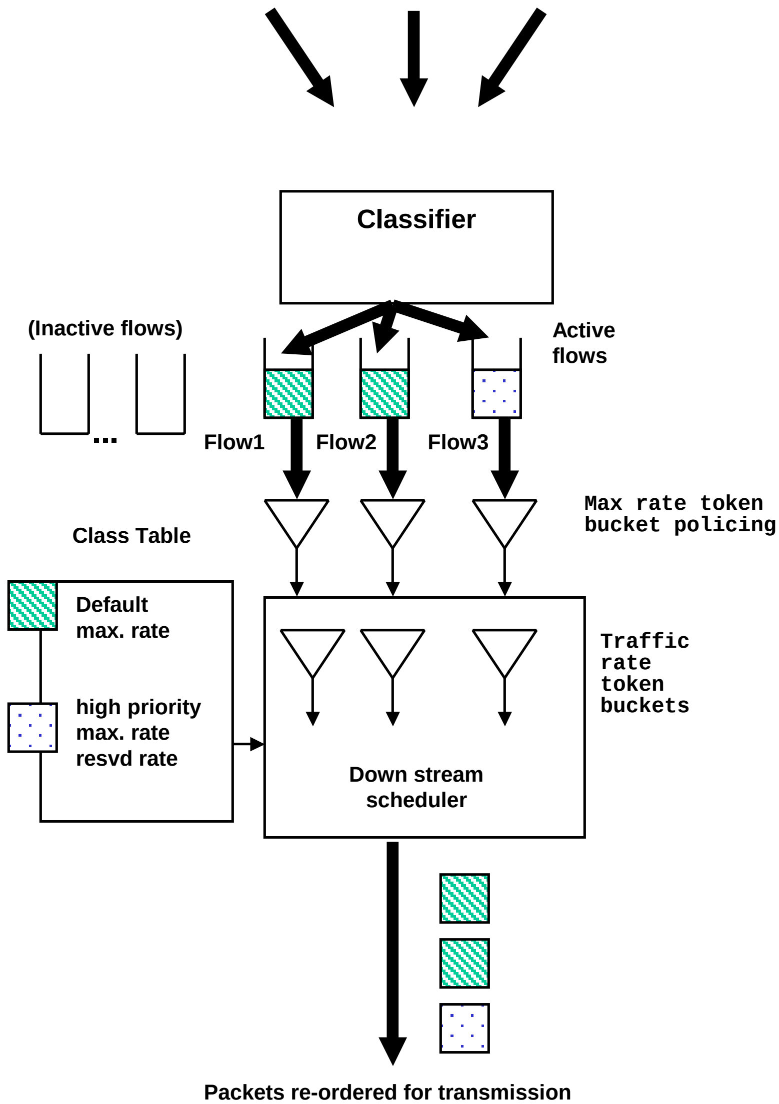


Figure 2 Downstream Packet Scheduling

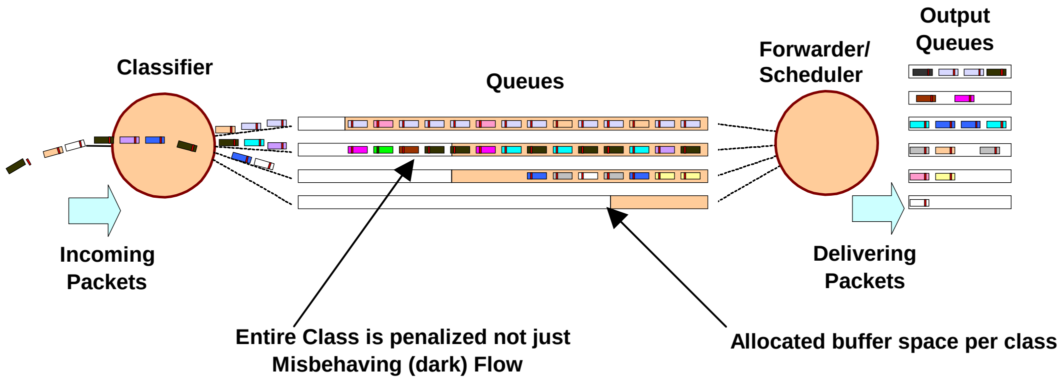


Figure 3 Class Based Queuing with RED

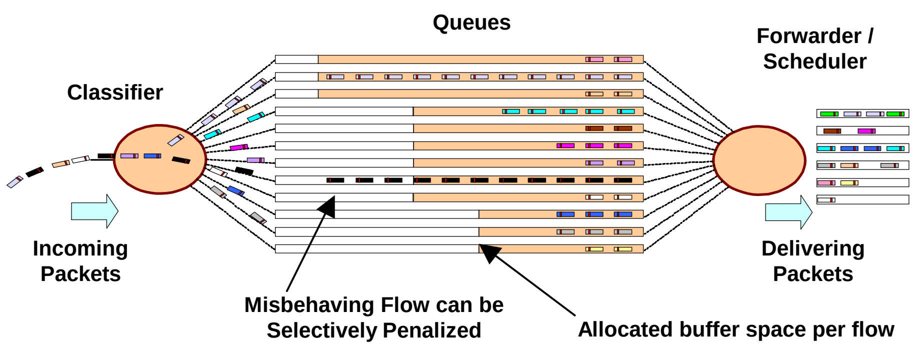


Figure 4 Per Flow Queuing and Scheduling with Longest Queue Push out

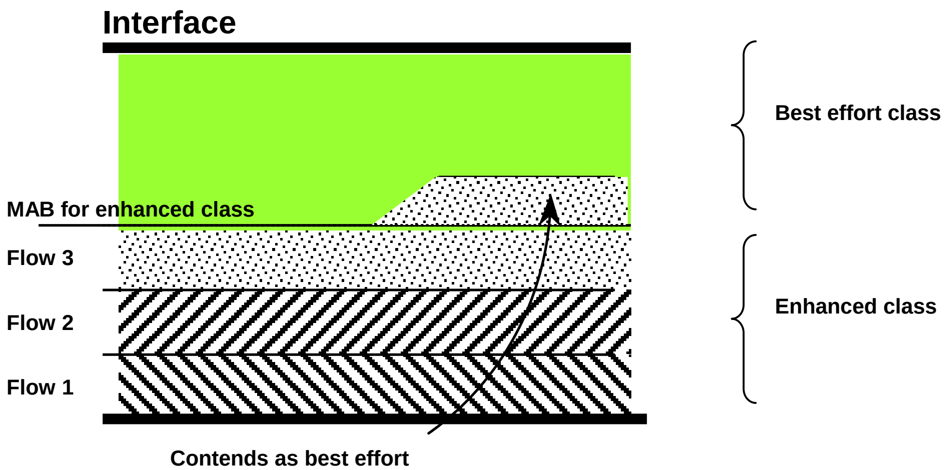


Figure 5 MAB Operation -flow 3 exceeds reserved rate

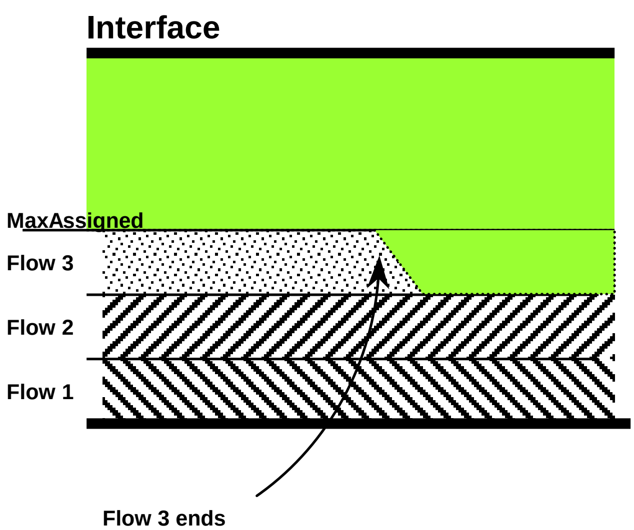


Figure 6 MAB Operation -reserved rate flow 3 ends

REFERENCES

1. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss. An Architecture for Differentiated Services (RFC 2475).
2. R. Braden, D. Clark, S. Shenker Integrated Services in the Internet Architecture: An Overview (RFC1633).
3. DOCSIS 1.1 Radio Frequency Interface Specification.
4. E. Rosen, A. Viswanathan, R. Callon. Multiprotocol Label Switching Architecture (RFC 3031)
5. J.C.R. Bennett and H. Zhang Hierarchical Packet Fair Queuing Algorithms

###

About the Author

Gerry White is Senior Director of Advanced Technology within the Network Infrastructure Solutions Group of the Motorola Broadband Communications Sector, and he was previously CTO of RiverDelta Networks, Arris Interactive, and LanCity. He has presented papers and spoken at dozens of cable industry events worldwide and has published many trade press articles on cable industry issues. White is the co-author of several patents and articles on data communications technology, and he holds a BSc. (Honors) from University College in London.

Author Contact Information

Gerry White
Senior Director of Advanced Technology
Network Infrastructure Solutions Group
Motorola Broadband Communications Sector
3 Highwood Drive
Tewksbury, MA 01876
(978) 858-2300
gwhite@motorola.com

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