# THE COMPLETE
# TECHNICAL PAPER PROCEEDINGS

FROM:

# ADVANCED MONITORING OF SWITCHED BROADCAST SYSTEMS

Ludovic Milin and Ran Oz
BigBand Networks

## *Abstract*

*Switched broadcast has become a viable technology for reclaiming bandwidth and optimizing spectrum, allowing richer program lineups and increasing personalization of content. By leveraging switched broadcast cable operators are able to offer subscribers hundreds of high definition programs and long tail content, such as coverage of local sporting events, without requiring plant upgrades to expand capacity.*

*This paper presents data gathered from multiple switched broadcast deployments to illustrate the imperative of monitoring system performance and the resulting benefits to operators.*

*The authors assert that performance monitoring should be applied at three key stages of deployment of a switched broadcast system:*

*1) Prior to Deployment – during this stage an operator must assess which programs in its broadcast lineup are viable candidates for putting on a switched tier. A non-intrusive analytical tool is, arguably, the best way to collect viewership statistics for all available programming, allowing identification of "long tail" content.*

*2) During Deployment – while a switched broadcast system is being implemented an operator will benefit from leveraging the same monitoring platform used in the prior characterization effort, potentially still gathering viewership data on broadcast programming.*

*3) After Deployment – once the installation has been completed the monitoring focus can be broadened to allow identification of trends, such as an unanticipated growth, that may lead to insufficient bandwidth being available to support all channel requests. Early identification of such trends can allow pro-active remedies, minimizing negative subscriber impact.*

*The authors also highlight the imperative of protecting consumers' privacy and describe how the techniques used to collect viewership data can be made to comply with the 1984 US Cable Privacy Act.*

## INTRODUCTION

Switched broadcast came of age recently as two major North American cable operators deployed the technology in several markets. Switched broadcast currently supports over six million homes passed and switches one million set-top boxes. With other cable operators evaluating the technology and undertaking field trials, the footprint of deployed switched broadcast systems is set to grow rapidly.

By reclaiming the bandwidth that would otherwise be consumed by delivering unwatched content, switched broadcast provides cable companies the opportunity to expand the amount of programming they offer subscribers. However, the oppprtunity to provide thousands, even tens of thousands, of programs requires operators to evolve the methodologies used to track viewership of that content. The ability to collect statistics, in real-time, about which programs subcribers are watching is necessary for successful capacity planning and deployment of switched broadcast systems. The insights gained from collection of this data facilitates rapid and smoother deployment of switched systems.

Valuable statistics include the viewership of linear and switched broadcast programs, the number of active STBs (set-top boxes) per service group, and the bandwidth utilization within service groups.

Additionally, the ongoing collection of information enables round-the-clock performance monitoring, leading to opportunites for pro-active network maintenance. Data about blocking events, failures for STBs to tune to the requested channel and network response times can all be analyzed, leading to enhanced performance and, ultimately, increased subscriber satisfaction.

In addition to the operations benefits obtained from deploying a statistics monitoring tool, the ability to build accurate and complete records of viewership behavior introduces new revenue possibilities. Personalized news is one example of ways that precise viewership records collected by a statistics monitoring tool could be leveraged for new revenue generation. Another is addressable advertising, which matches ads more closely to subscribers' interests, by leveraging insights into which content viewers are habitually tuning to.

Finally, while the use of third-party market research firms has generally been sufficient to obtain insights into subscriber viewing habits, the plethora of new programs that will be available to subscribers in switched environments makes this task more daunting. The value of this information is high because it provides insights into the viewing patterns of all subscribers on the switched tier, not just the subset of viewers that have been enlisted and whose viewing habits may not necessarily represent those of the majority.

This paper describes a method for collecting statistics that supports the applications described above, namely:

- Capacity planning;
- Diagnostics;
- Content personalization;
- Addressable advertising;
- Market research.

This paper discusses each of these applications in depth and cites the specific benefits of a statistics monitoring platform in each instance. It begins, however, with a description of how the platform functions.

## ARCHITECTING NETWORKS FOR STATISTICS RETRIEVAL

The statistics monitoring / diagnostics platform proposed by the authors consists of the following components:

- Collection Engine – gathers STB activity logs from the servers, known as SBSSs (switched broadcast session servers), used to manage a switched broadcast network;
- Analytical Engine – processes the STB activity logs gathered from the network;
- Data Warehouse – stores the data and metadata obtained from the both the network and analytical components, and supports long-term storage for trending analysis;
- Webserver – provides a user-friendly interface to customize and view reports.

Figure 1 shows how these elements are combined in a generic switched network.

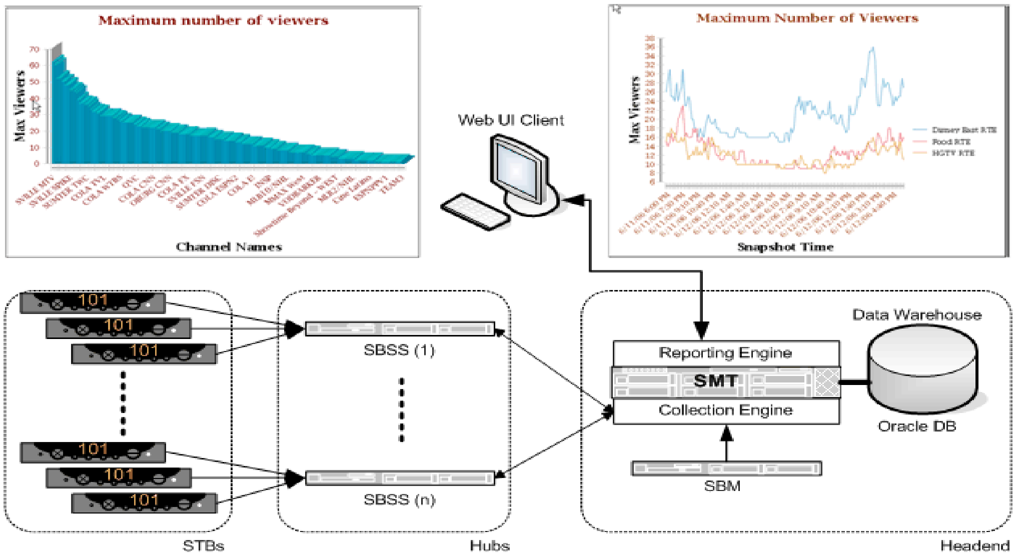


Figure 1: Components of a statistics management platform

Rather than the traditional store and forward approach in which each STB dedicates a portion of its memory footprint to recording activity and, on a regular basis, sends large UDP (user datagram protocol) packets over the OOB (out-of-band) upstream path to some collection server, the authors recommend an approach that has no impact on the STB memory footprint or the OOB network. This feature stems from each message requiring only about 64 bytes. This approach, which leverages standard switched broadcast protocols, is not sensitive to packet loss between an STB and the data collection engine. It is not tied to a specific switched broadcast client or headend environment – any LOB (load on boot) or native switched broadcast client can be utilized to send STB activity messages to the SBSS.

An example of these STB messages is presented in Figure 2.

The same protocols can also be leveraged to monitor linear broadcast channel activity via simple reconfiguration of the switched broadcast client on an STB. This enables activity messages for all channels to be sent upstream instead of only those on the switched tier.

The real-time metrics derived from the protocol messages can be grouped into general audience metrics and specific switched broadcast performance metrics.

General audience metrics (for both switched and linear programming) include:

- Number of viewed channels per service group;
- Overall bandwidth utilization per service group;
- Number of active STBs per channel and per service group;
- Number of tune-ins and tune-outs per channel and per service group.

Switched broadcast performance metrics include:

- Switched broadcast QAMs occupation ratio per service group;
- Number of tune-ins to previously unmapped channels per service group;
- Time to map previously unwatched channels per service group;
- Upstream bit rate per servce group;

- Number of blocking events (bandwidth not available) per channel and per service group;
- Number of other errors per service group (unresponsive STBs, tuning errors, and so on).

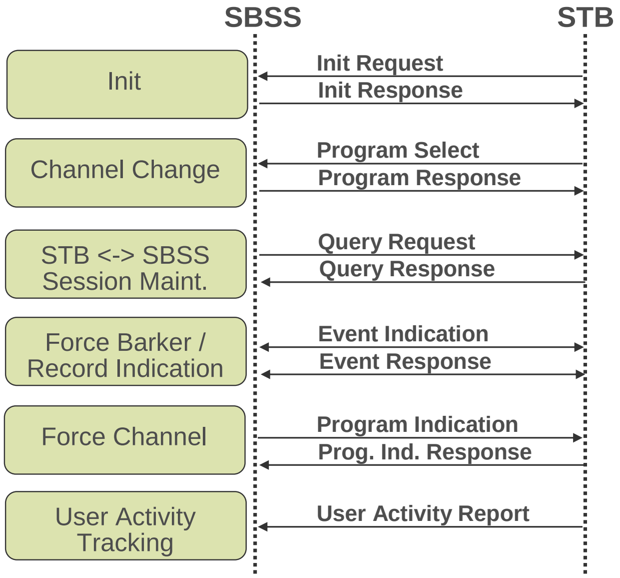


Figure 2: Example of protocol messages communicated between the STB and the SBSS

The following assumptions were made during development of the statistics collection and analysis tool described in this paper:

- A service group is a group of nodes;
- Each node serves multiple households
- A hub serves multiple service groups;
- An SBSS supports multiple service groups;
- All nodes in a service grup are served by common switched RF spectrum;
- The SBSS shares the RF spectrum;
- The SBSS logs all STB activity in real-time (channel changes, keep alives, last user activity, and so on.)

A common set of assumptions simplified and accelerated creation of a diagnostics tool and enables the analysis it performs to be more effective / broadly applicable.

The figures presented throughout this paper are screen captures from the web-based user interface of the diagnostics tool. The reports can also be configured to output pdf files for easy sharing amongst a working group, and CSV (comma–separated values) files so the raw data in the reports can be analyzed further. The data was gathered from deployed switched broadcast systems in multiple cable networks.

## APPLYING PERFORMANCE MONITORING: THREE SCENARIOS

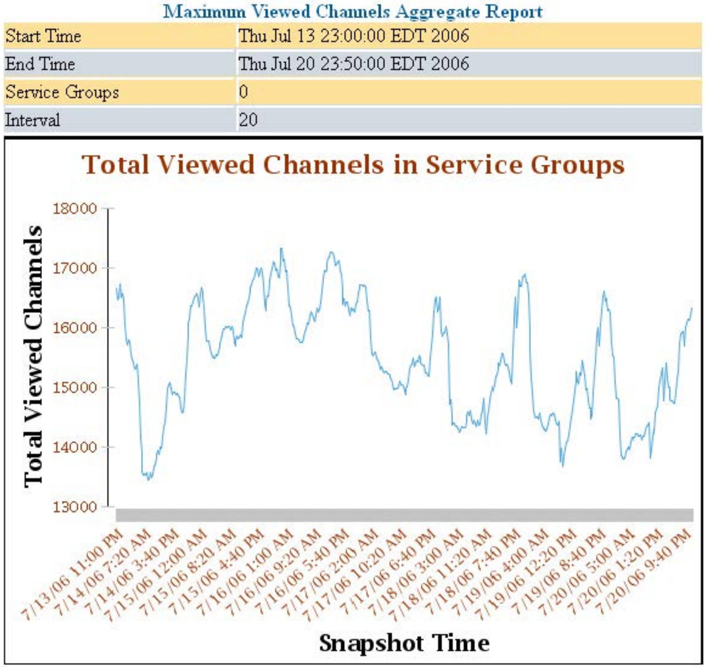
There are at least three scenarios where the ability to collect viewership statistics and performance diagnostics can be especially valuable:

- Before deployment;
- During deployment;
- After depoloyment.

Each of these scenarios is discussed in the following sections.

## BEFORE DEPLOYMENT

To ensure best-of-breed network performance a cable perator should assess which programs to put on the switched tier. A non-intrusive analytical platform is, arguably, the best way to collect viewership statistics for all available programming, and enable identification of "long tail" content. This type of platform can also provide the parameters necessary for modeling switched infrastructures, allowing for optimization of Edge QAM numbers, service group size and over-subscription ratios.

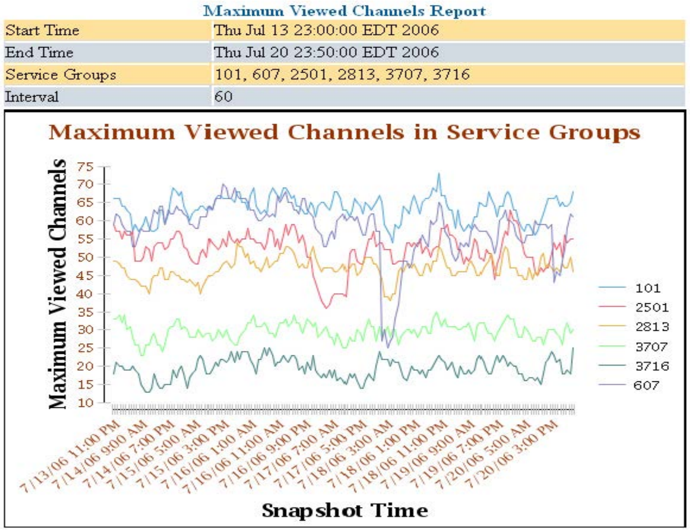


Maximum Viewed Channels Aggregate Report

| | |
|---|---|
| Start Time | Thu Jul 13 23:00:00 EDT 2006 |
| End Time | Thu Jul 20 23:50:00 EDT 2006 |
| Service Groups | 0 |
| Interval | 20 |

Figure 3: Typical weekly busy hour report (total number of viewed channels for all service groups)

Figure 3 provides the total number of viewed channels (active streams) for all service groups for a week. Notice the daily spikes between 10PM and midnight and more consistent viewership throughout weekend. It is possible to generate similar reports for each service group to check whether the conclusions made for the overall population apply, to compare service groups that are geographically distributed or whose population is different, and also to assess the impact of service group size. Examples are shown in Figure 4 of the maximum number of active streams, varying from single to triple at constant service group size, indicating that the number of switched broadcast QAMs per service group can be optimized, instead of being a constant.



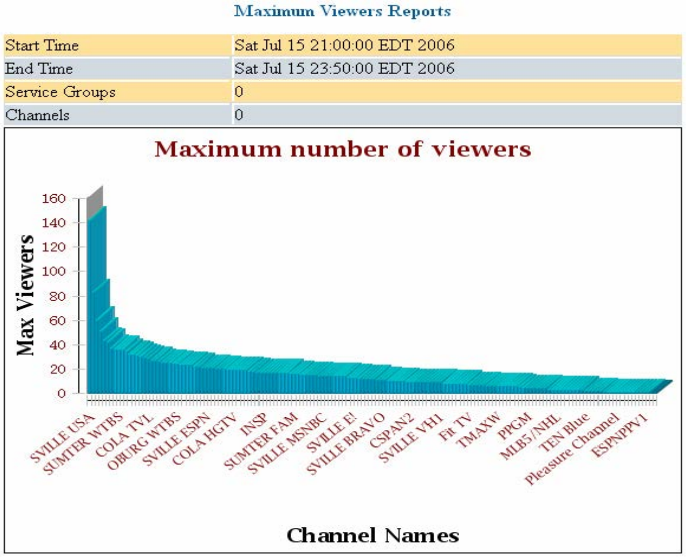
Maximum Viewed Channels Report

| | |
|---|---|
| Start Time | Thu Jul 13 23:00:00 EDT 2006 |
| End Time | Thu Jul 20 23:50:00 EDT 2006 |
| Service Groups | 101, 607, 2501, 2813, 3707, 3716 |
| Interval | 60 |

Figure 4:  Example of weekly busy hour report for an individual service group

Once the busy hour has been identified one should analyze the viewership of individual channels audience to allow “long tail” content candidates suitable for being placed on the switched tier to be identified.  It is important to narrow down this analysis to the busy hour as this is where the most significant bandwidth saving benefits can be accomplished.

Figure 5 shows a typical report for all service groups aggregated.  Because this example provides insights into channels that are already on the switched broadcast lineup “long tail” programs can be identified easily

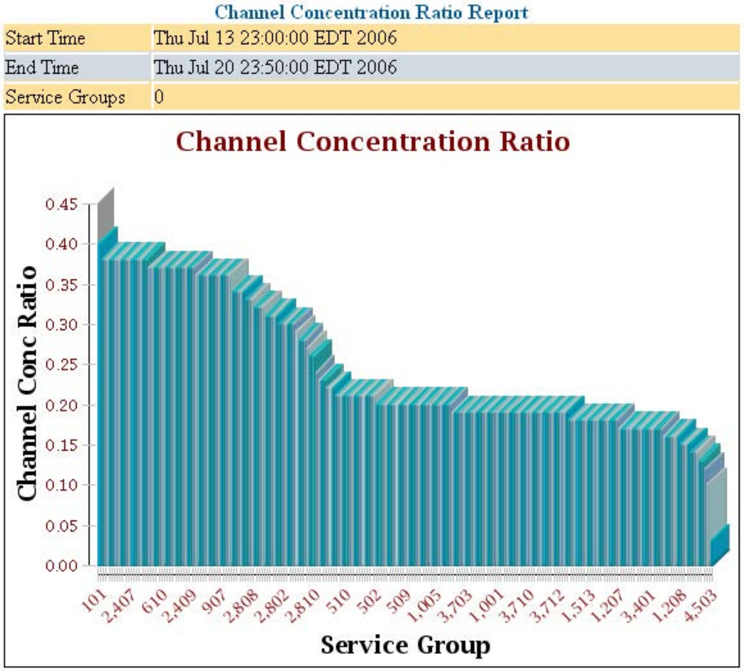


Maximum Viewers Reports

| | |
|---|---|
| Start Time | Sat Jul 15 21:00:00 EDT 2006 |
| End Time | Sat Jul 15 23:50:00 EDT 2006 |
| Service Groups | 0 |
| Channels | 0 |

Figure 5:  Distribution of “long tail” content

The channel audience in service group can also be used to predict the switched broadcast over-subscription ratio. This is achieved by examining the channel concentration ratio, (i.e. the ratio of unique viewed channels to the total number of available channels in the SWB tier.) Figure 6 shows that this ratio is at most 40%, leading to a minimum 2.5:1 oversubscription ratio in the switched tier. Using that data, one can again optimize the number of switched QAMs assigned per service group.



Channel Concentration Ratio Report

| | |
|---|---|
| Start Time | Thu Jul 13 23:00:00 EDT 2006 |
| End Time | Thu Jul 20 23:50:00 EDT 2006 |
| Service Groups | 0 |

Figure 6: Channel concentration report

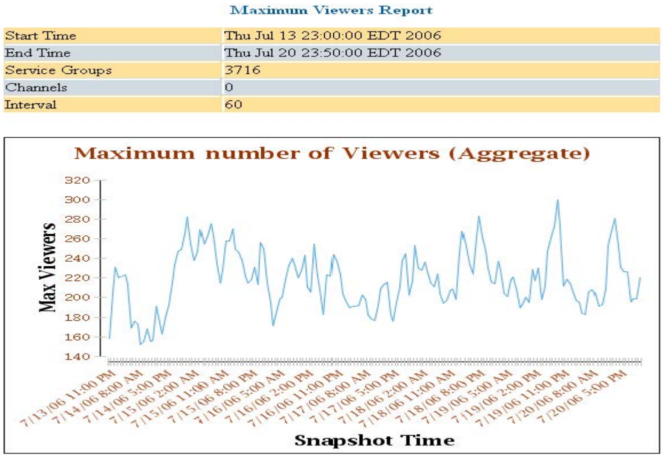
The reports presented in this section should not only be used in a preliminary switched broadcast deployment study but can also be used throughout the life of a switched broadcast deployment to ensure that the conditions are not changing and that the switched infrastructure remains optimized.

## DURING DEPLOYMENT

A cable operator can benefit from leveraging the same monitoring platform used in the prior characterization effort when installing / turning-up a switched broadcast system: the statistics collected previously can facilitate easier troubleshooting of deployment issues such as STB upstream communication issues and QAM mis-configurations.

As new service groups are being migrated to the switched tier one should confirm that STBs are registering with the SBSS and there are no upstream communication issues. An example of such a report, showing the total number of active STBs as a function of time, is provided in figure 7.

A side benefit of this type of report is that it can be used to verify that the number of active tuners in each service group doesn’t exceed expectations. Trending growth over time will indicate when a service group split may be appropiate.

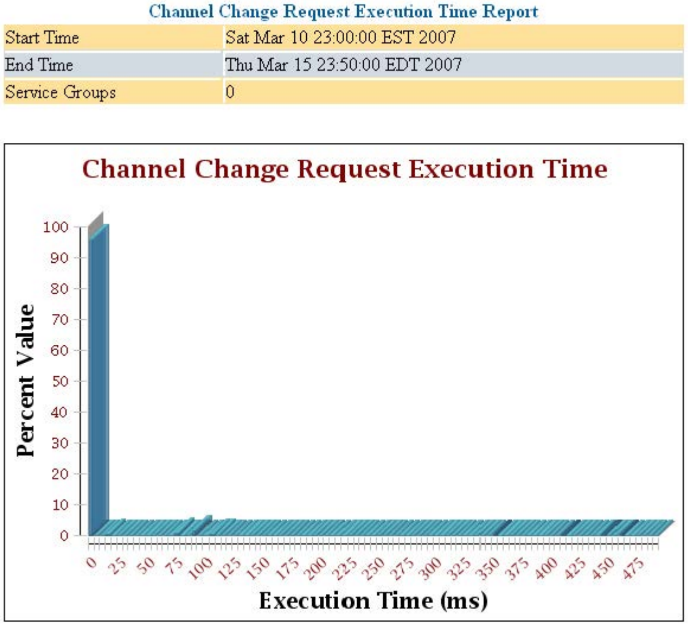


Maximum Viewers Report

| | |
|---|---|
| Start Time | Thu Jul 13 23:00:00 EDT 2006 |
| End Time | Thu Jul 20 23:50:00 EDT 2006 |
| Service Groups | 3716 |
| Channels | 0 |
| Interval | 60 |

Figure 7: Number of active STBs in a specifc service group

Once a channel has been mapped to a QAM group a STB will tune to it without delay, a direct benefit of the protocols used in switched broadcast. Figure 8 shows that over 99.9% of the channel changes (not including the STB decoder actually locking into the stream) occurred in 5msec or less.



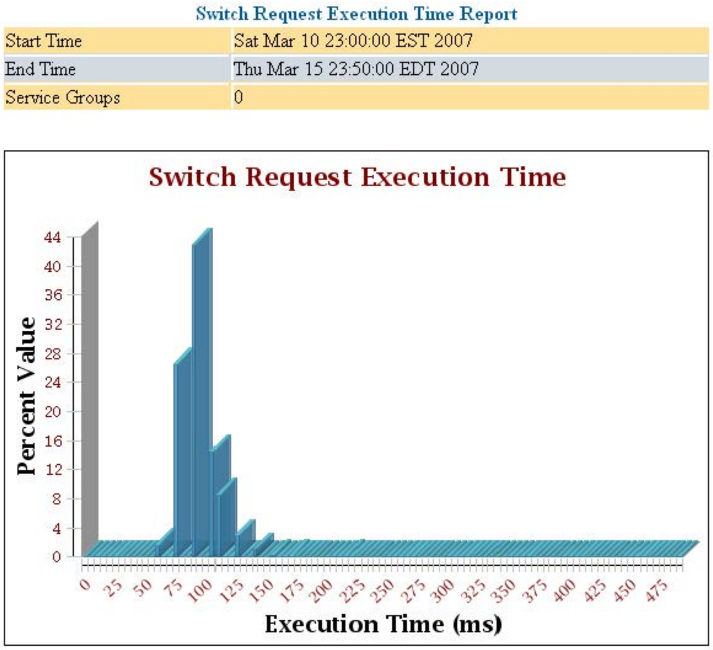
Channel Change Request Execution Time Report

| | |
|---|---|
| Start Time | Sat Mar 10 23:00:00 EST 2007 |
| End Time | Thu Mar 15 23:50:00 EDT 2007 |
| Service Groups | 0 |

Figure 8: Channel change times already mapped on the switched tier

A departure from a 0 to 5msec range would indicate an issue with the system and along with likely impact to the subscriber’s viewing experience.

On the other hand, when a channel is not already mapped to one of the switched QAMs because no one in that service group was previously watching it, additional steps will be required for the STB to tune to it. Figure 9 shows a typical distribution of such delays, which include the time for the request to go from the STB to the SBSS, the SBSS processing of the request and forwarding to the Edge QAM, the Edge QAM joining the multicast group for that stream, adding the session to one of the service group QAMs, and updating of the MCP to send the tuning information to the STB. It typically takes under 120msec to accomplish this, a negligible amount of time in terms of subscriber experience. This metric can be affected by delays in the network or SBSS overloads and is, therefore, important to closely monitor during and, after deployment. Doing so will help address potential issues and avoid subscriber complaints about tuning delays.

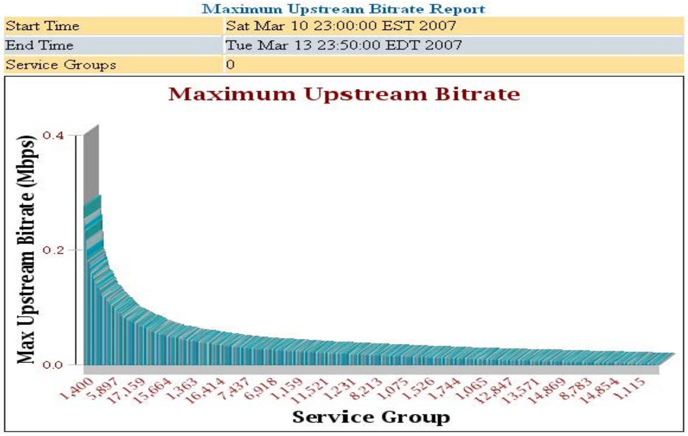


Switch Request Execution Time Report

| | |
|---|---|
| Start Time | Sat Mar 10 23:00:00 EST 2007 |
| End Time | Thu Mar 15 23:50:00 EDT 2007 |
| Service Groups | 0 |

Figure 9: Additional delay for channel changes requiring mapping of an unwatched switched program

Because the OOB upstream bandwidth is limited it is important to keep track of how much is being used by switched traffic flows between the STBs and their respective SBSSs. Typical loads vary between 0 and 200Kbps depending on the number of active STBs per service group, as figure 10 illustrates.



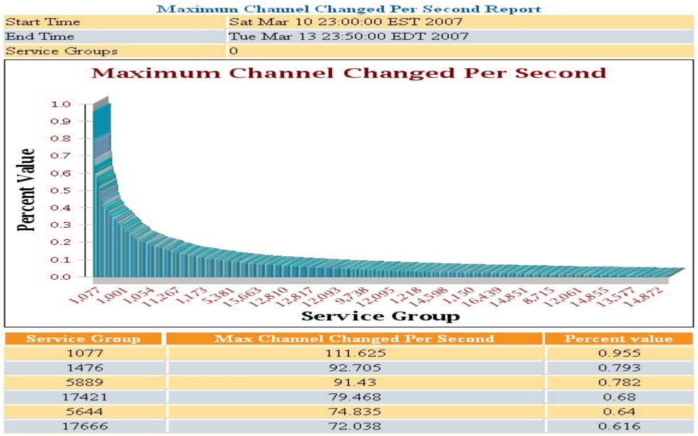
Maximum Upstream Bitrate Report

| | |
|---|---|
| Start Time | Sat Mar 10 23:00:00 EST 2007 |
| End Time | Tue Mar 13 23:50:00 EDT 2007 |
| Service Groups | 0 |

Figure 10: Maximum upstream bit rate for switched broadcast traffic per service group

The SBSS performance (see figures 8 and 9) is a strong function of the number of channel change requests the server has to handle. That metric is presented in Figure 11. Once a baseline has been established one can monitor how this metric changes with the number of active STBs (see figure 7) and measure the resulting impact on SBSS performance. A typical scenario would be an increase in the number of STBs and their respective traffic flows, requiring the addition of an SBSS to split the load and maintain an acceptable level of service.

Because the viewership and performance monitoring tool provides real-time visibility into the system behaviour, there is no need for empirical rules, since action can be taken as soon as the measured subscriber experience threatens to dip below the deemed acceptable thresholds.



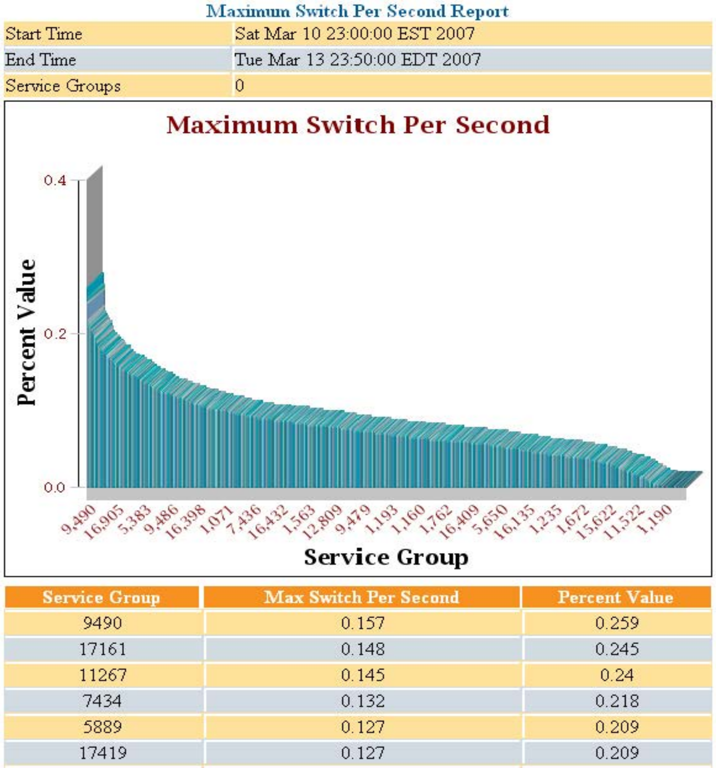
Maximum Channel Changed Per Second Report

| | |
|---|---|
| Start Time | Sat Mar 10 23:00:00 EST 2007 |
| End Time | Tue Mar 13 23:50:00 EDT 2007 |
| Service Groups | 0 |

| Service Group | Max Channel Changed Per Second | Percent value |
|---|---|---|
| 1077 | 111.625 | 0.955 |
| 1476 | 92.705 | 0.793 |
| 5889 | 91.43 | 0.782 |
| 17421 | 79.468 | 0.68 |
| 5644 | 74.835 | 0.64 |
| 17666 | 72.038 | 0.616 |

Figure 11: Maxmum number of channel change requests per second per service group

The number of channel change requests requiring a new program t be mapped to the service group lineup should typically be fairly small unless the switched broadcast channel map is large and the over-subscription ratio high. As can be seen from the data provided in figure 11 and 12, the ratio between the two is about 1000 to 1. This ratio, however, does change significantly from one system to another. A large and sudden departure from the established baseline would typically indicate issues with system configuration or the MCP reaching the STBs.



Maximum Switch Per Second Report

| | |
|---|---|
| Start Time | Sat Mar 10 23:00:00 EST 2007 |
| End Time | Tue Mar 13 23:50:00 EDT 2007 |
| Service Groups | 0 |

| Service Group | Max Switch Per Second | Percent Value |
|---|---|---|
| 9490 | 0.157 | 0.259 |
| 17161 | 0.148 | 0.245 |
| 11267 | 0.145 | 0.24 |
| 7434 | 0.132 | 0.218 |
| 5889 | 0.127 | 0.209 |
| 17419 | 0.127 | 0.209 |

Figure 12: Maxmum number of channel change switch requests per second per service group

## AFTER DEPLOYMENT

After installing a switched broadcast system the monitoring focus can be broadened to allow identification of trends such as un-forecasted growth that may lead to insufficient bandwidth being available to support all requested channels or management servers becoming overloaded. Early identification of such trends can allow pro-active remedies, minimizing negative subscriber impact.

Experience has shown that one of the most relevant metrics to follow is the occupied bit rate concentration ratio. This metric takes into account the number of switched QAMs available within each service group and the number of active streams (actively watched channels). It also encompasses the bandwidth of individual streams, something that is not necessarily constant when using multi-width CBR (constant bit rate) to maximize video quality, or when mixing standard and high definition progams. Trending this metric over time for each service group will characterize the growth rate, and a good rule of thumb is to consider adding more QAMs or splitting service groups when the ratio consistently surpasses 80 to 90%, or spikes to 100%.

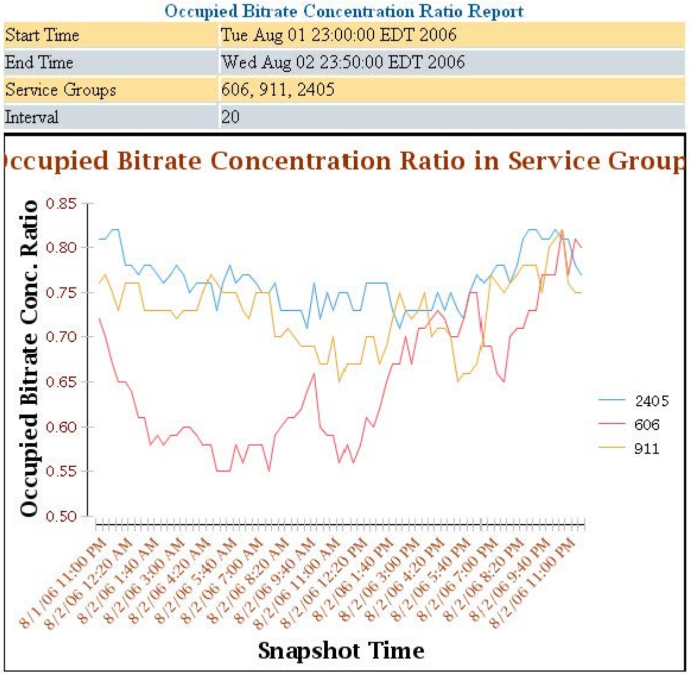
Figure 13 shows one such daily summary of the maximum occupied bitrate concentration ratio for each service group.

Occupied Bitrate Concentration Report

| | |
|---|---|
| Start Time | Tue Aug 01 23:00:00 EDT 2006 |
| End Time | Wed Aug 02 23:50:00 EDT 2006 |
| Service Groups | 0 |

Figure 13: Maximum occupied bit rate concentration ratio per service group

Figure 14 examines the top three service groups in Figure 13 in more detail. One of the key things to note here is that the 80% utilization threshold is only surpassed briefly during the typical 10PM to midnight busy hour, meaning that there is no immediate need for any infrastructure changes.

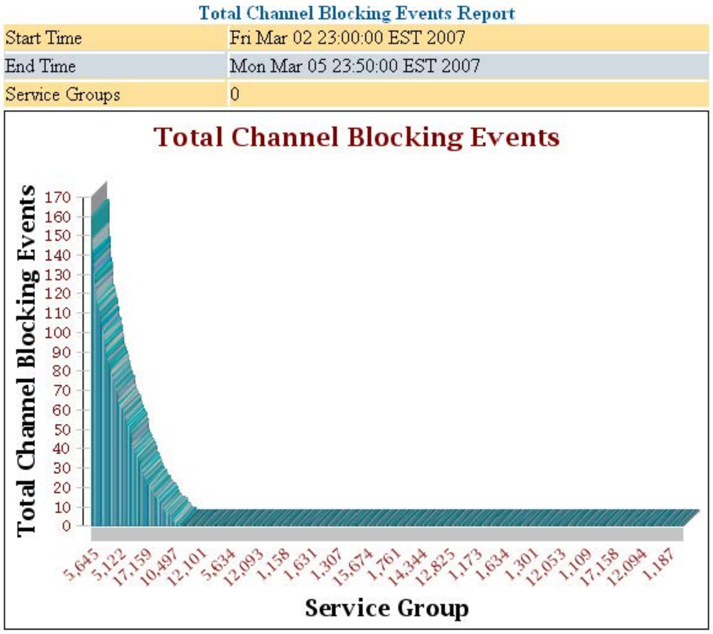


Occupied Bitrate Concentration Ratio Report

| | |
|---|---|
| Start Time | Tue Aug 01 23:00:00 EDT 2006 |
| End Time | Wed Aug 02 23:50:00 EDT 2006 |
| Service Groups | 606, 911, 2405 |
| Interval | 20 |

Figure 14: Daily variation of occupied bit rate concentration ratio for three service groups

Once the occupation ratio is consistently in 100% vicinity and the over-subscription ratio is greater than 1:1, the likelihood is high that no bandwidth will be available on the QAMs for new channels. This is defined as a blocking event and results in a subscriber receving a “Channel not available – Please try later” message, instead of the desired program. Figure 15 presents a summary, for each service group, of the number of blocking events over the duration of a few days. Approximately 15% of the service groups show a significant number of blocking events.

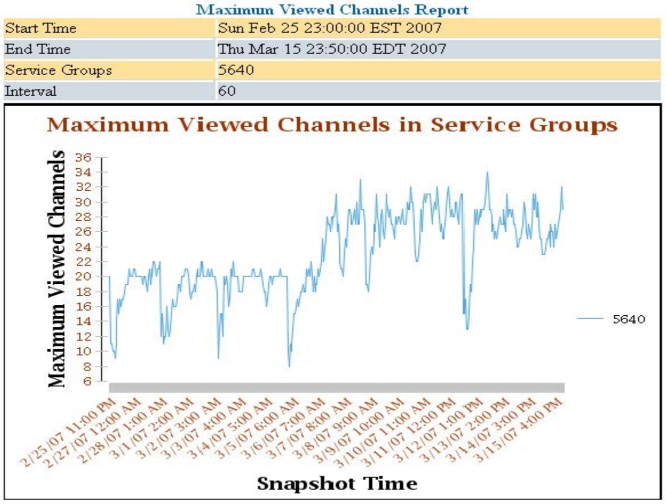


Total Channel Blocking Events Report

| | |
|---|---|
| Start Time | Fri Mar 02 23:00:00 EST 2007 |
| End Time | Mon Mar 05 23:50:00 EST 2007 |
| Service Groups | 0 |

Figure 15: Maximum number of blocking events for all service groups

As a remedial step to bandwidth scarcity additional switched QAMs can be added to the service groups most affected by blocking. The results of doing so ae highlighted in Figure 16. This example shows that, in one of the service groups under consideration, the number of active streams on the switched tier jumped from about 20 to about 30 as soon as the additional QAM became available.

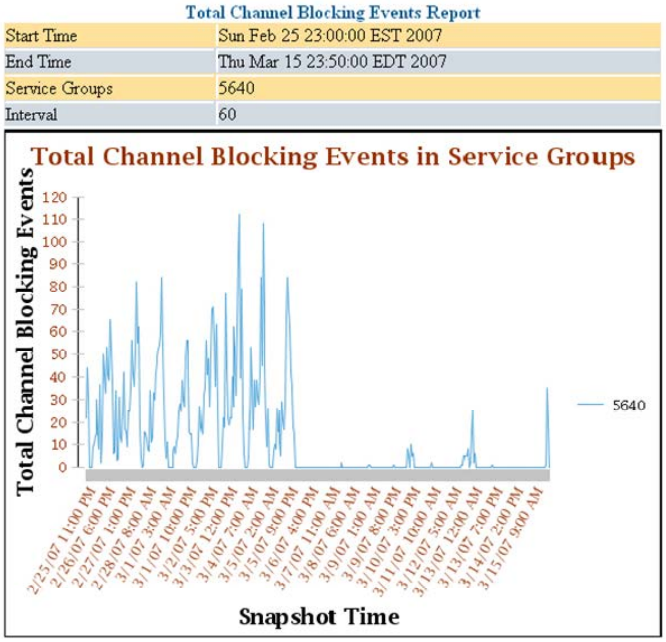


Maximum Viewed Channels Report

| | |
|---|---|
| Start Time | Sun Feb 25 23:00:00 EST 2007 |
| End Time | Thu Mar 15 23:50:00 EDT 2007 |
| Service Groups | 5640 |
| Interval | 60 |

Figure 16: Maximum number of viewed channels for a specific service group across the addition of a switched QAM

Coincident with the increase in available bandwidth the numbers of blocking events virtually disappeared, as figure 17 shows. Additionally, the over-subscription ratio decreased from 2.5 to 1.7 in the affected service groups


Total Channel Blocking Events Report

| | |
|---|---|
| Start Time | Sun Feb 25 23:00:00 EST 2007 |
| End Time | Thu Mar 15 23:50:00 EDT 2007 |
| Service Groups | 5640 |
| Interval | 60 |

Figure 17: Blocking events for a specific service group across an additional switched QAM

This can be explained by the larger size of the service groups, subscriber demography and switched channels lineup. The positive point is that because of the detailed monitoring information available, the capital needed for an additional switched QAM was limited to a very small number of clearly identified service groups. Moreover, the expense could be delayed until the blocking situation became too affecting for the subscribers.

## BUSINESS BENEFITS OF STATISTICS MONITORING

Benefits accrue when statistics monitoring is applied to both linear and switched broadcast networks. In addition to the operations benefits already described, namely capacity planning and diagnostics, the ability to collect accurate viewership data full-time introduces new revenues opportunities. These are examined in the following sections of this paper.

## CONTENT PERSONALIZATION

Personalized news is one example of ways that precise viewership records collected by a statistics monitoring tool could be leveraged for new revenue generation. Since the real-time collection of viewership data allows content and subscriber interests to be accurately correlated, enterprising newrooms, or third-parties, could create news summaries that are more likely to retain the attention of viewers, as compared to traditional broadcast TV news programs. For example, newsrooms could record a series of short news stories on a wide range of topics, allowing cable operators to combine into personalized bulletins that address a subscriber's specific interests, whether it's baseball teams or Internet start-ups.

A personalized version of a music network is another example of how switched unicast provides cable operators the opportunity to offer increasingly customized content.

Customized services like these could be offered as premium services, potentially earning the operator additional revenues. Alternatively, personalized content could be offered at no additional charge, with the expectation that increased customer loyalty results.

## ADDRESSABLE ADVERTISING

Addressable advertising is a methodology for more closely matching advertisements to subscribers' interests. Figure 18 illustrates how three subscribers, all watching the same program on the switched tier receive different ads during the commercial breaks. For example, subscriber #1, an avid teen snowboarder, receives an ad about snowboard sales. During the same commercial break, subscriber #2, a thirties-something bachelor, views an ad about an upcoming motor show. Subscriber #3, an avid traveler in her fifties, receives information about cruises in the Caribbean.

Studies, such as a recent CTAM report, reveal that marketers are willing to pay premium rates if their ads achieve improved response rates among their intended audiences.

By building precise databases about which programs individual subscribers are watching, cable operators can determine which ads to forward to viewers.

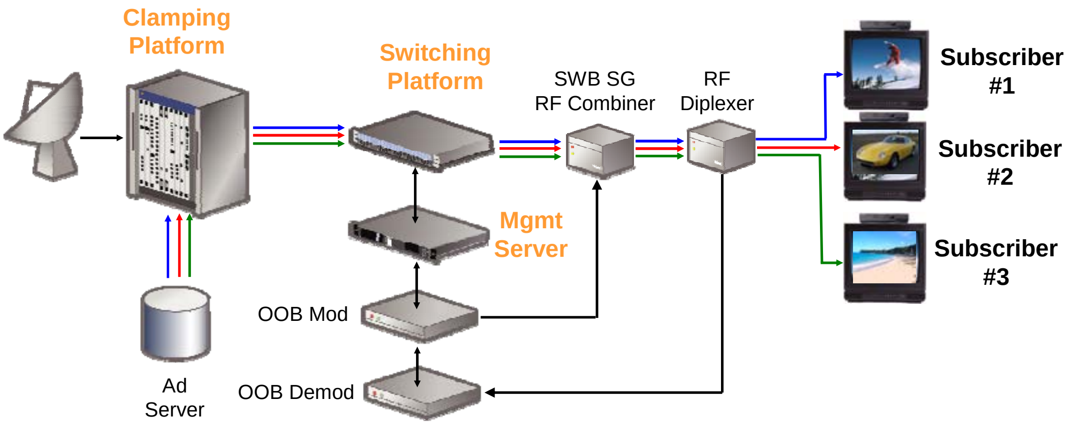


Figure 18: Subscribers receiving different ads though all view the same programming

In contrast to third-party research firms that typically collect broadbrush viewership data, the statistics monitoring platform described in this paper can provide insights that would otherwise be lost. For example, a subscriber, that may have paused to watch a short segment about the fashion industry while channel surfing, may be a viable candidate for an advertisment about a relevant local fashon event. Other methods of recording consumer viewing habits are unlikely to yield this level of detail.

Additonally, a consumer that routinely tunes to a home improvement network may be receptive to an advertisement about a sale at a local hardware store, and could receive such an ad, even when watching a different network.

## GUARDING PRIVACY RIGHTS

The Cable TV Privacy Act of 1984, is intended to protect personal information. In particular, it prohibits cable TV providers from disclosing personally identifiable information, and allows users to view and verify their information. Naturally this means that gathering and reporting on an individual's channel viewership history is likely illegal. In order to satisfy this requirement, the SBSS servers offer the option to scramble the STB MAC addresses in the logs making it impossible to identify any single user.

The gathering of individual subscriber data can be implemented on an "opt-in" basis. For example, the more enterprising cable operators could explicitly ask their subscribers about their ad preferences. In return for providing a cable operator with a list of the subjects and categories they'd be interested in viewing ads on, subscribers could receive a complimentary gift or upgrade to an expanded service package, or some other incentive.

## CONCLUSIONS

The ablity to obtain precise viewership statistics assists cable operators in a variety of ways. These include providing the insights needed to engineer effective switched tiers and speed system deployment and turn-up. The capability to monitor network perfomance allows adjustments to be implemented that ensure channel change times and service quality are optimized.

Although primarily developed for switched broadcast environments, the non-invasive tool described in this paper has applications in linear broadcast networks also. Broader applications can include support for addressable advertising business models and other content personalization oppportunies.

However, maintaining subscribers' privacy is paramount and the viewership statistics needed to empower these benefits can be obtained / stored in ways that enable cable operators to meet their obligations under the Cable Privacy Act.

# CABLE THIN CLIENTS: A NEW REVENUE OPPORTUNITY FOR MSOS

Daniel Howard
Motorola Connected Home Solutions
Technology Office

## *Abstract*

*A thin client computer typically a PC that consists of a motherboard only and boots the operating system off of the network. Thin clients are essentially intelligent display devices that normally run applications on a server located somewhere else on the network. The lack of moving parts and ability to manage them remotely makes them a very cost effective technology for business applications such as point of sale devices, and many corporations are using thin clients for their employees based on the much lower total cost of ownership of this technology. Recently, with the availability of Open Source office applications and educational titles, schools all over the world are switching to thin clients as a more cost effective approach to providing reliable and inexpensive student PCs.*

*In this paper, a thin client computer technology that uses cable modem infrastructure is described. Because of its broadcast architecture, the cable plant is uniquely positioned to take advantage of thin client computing, and can provide the service more efficiently and cost effectively than other broadband wired networks. Thin clients are as robust as a cable modem, and can be rebooted and managed remotely as easily as a cable modem, which means MSOs can offer a managed home PC service without concern for additional truck rolls. Best of all, it opens up new revenue opportunities for MSOs, including pay-as-you-go data services, children's PCs with preloaded educational titles, and elderly residents who would only get a PC if a service provider supported it for them. And since a cable thin client is not much more than a cable modem that you plug a monitor, keyboard, and mouse into, MSOs can enter the market at far less than would be required for conventional computer technology. Finally, since the servers are located in the MSO's facilities, the operating system is completely under the MSO's control, and thus can be easily maintained and provide an additional portal to users.*

## INTRODUCTION

With the growth of competition for cable operators in the triple and quadruple play arenas, there is renewed interest in emphasizing the content and services provided by the operator to subscribers. And this content is not just limited to TV, music, and Internet portals, but many operators are providing free software, such as anti-virus applications, along with their high speed data services. These software application content offerings are generally being used today to enhance the marketability of high speed data service, rather than to specifically generate new revenue streams.

The main drawbacks of offering software content are that it can increase the software support cable operators must provide to subscribers, and the multiplicity of platforms and operating systems to support can make such content offerings undesirable to the operator. The cost of licensing such software can also be a burden to operators' revenue growth.

Businesses have been plagued by these same issues in providing PC platforms and software applications to their employees since the advent of the PC. An approach that is currently gaining ground in the business community is to use a thin client architecture to reduce the total cost of ownership and reliability of employee PCs. Thin clients are essentially a PC that consists of a motherboard only, and boots the operating system off of a server located elsewhere in the network. In a typical thin client deployment, over one hundred clients can be run by a single server, which means the business IT department need only support a single PC, the server, and the thin clients are more like appliances that at most require an occasional rebooting. Thin clients are much cheaper than conventional desktop PCs ($150 vs. $300+), and software additions, updates, patches, etc. are done on a single server, and then apply to hundreds of clients.

To further reduce the total cost of ownership (TCO), many businesses and schools are switching to Free Open Source Software (FOSS), and indeed an increasing number of cable operators are requesting Open Source solutions for their networks and services. As opposed to proprietary software offerings, FOSS is developed by the world community and the source code is freely available to programmers to download and improve. The Linux operating system is probably the most famous FOSS package, but FOSS applications exist for all major operating systems, and in fact there is now an Open Source office suite, OpenOffice, that rivals proprietary offerings in function, power, and compatibility.

So why would a cable operator be interested in thin client computing and Open Source Software? Because these are enabling technologies for cable operators to offer new software content and new high speed data services to subscribers. An example new service that would generate an entirely new revenue stream for operators would be to offer a home PC platform, running Open Source Software, that can be cost-effectively deployed and that is easily supported by the operator. By using thin client technology as the home platform, the cable operator can standardize both the hardware and software for the system, and can manage the clients from the operator's facilities, much as cable modems and settop boxes are managed currently. For the purpose of this paper, this proposed PC platform is termed a cable thin client.

An introduction to thin client technology, an architecture for, and the challenges of delivering thin client computing over cable networks are presented in this paper. Benefits to both consumers and operators are provided, as are new MSO revenue opportunities. Finally, future network architectures that would include cable thin clients are referenced, and a summary of why cable thin clients make sense for cable operators is presented.

## THIN CLIENT TECHNOLOGY

While most PCs can be converted into thin clients merely by altering the BIOS to boot the PC off of the network instead of the hard disk drive, the real benefits of thin clients can be seen in platforms designed to operate only as a thin client. Most thin clients are about the size of a cable modem, but there are also thin client solutions built into LCD monitors and even wall data ports, such as the Jack-PC from Jade Integration as shown below: (www.jadeintegration.com/jackpc.php).

**Figure 1. Example Thin Clients.**

Most are diskless and fanless, which means the moving parts that typically fail in a home PC are completely eliminated, and so thin clients are projected for TCO purposes to have 7-10 year lifespans, or roughly twice that of conventional PCs. During its life, all thin clients in the network may be upgraded by upgrading the servers.

The basic thin client architecture is to have at least one server on the network from which the thin clients boot their operating system and run applications, as shown below.

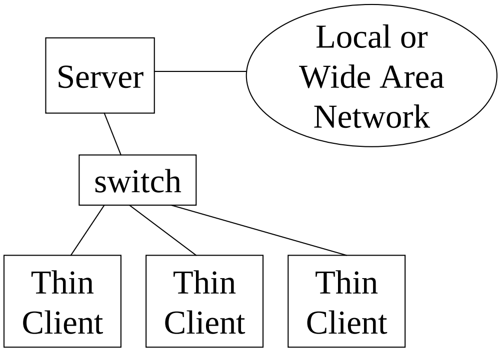


**Figure 2. Basic Thin Client Architecture.**

Thus, the server has two Ethernet interfaces, one to the outside world and one to the thin clients. The server can either be a single server that provides all required functions of boot, authentication, and applications, or these functions can be divided into servers optimized for each function, as is done in enterprise-style deployments of thin clients.

The author’s own initial experience with thin client technology using Free Open Source Software was at his child’s elementary school several years ago, where classroom computers were 7-10 years old and typically unused by teachers and students due to lack of adequate maintenance. After installing classroom servers in each room, the three existing PCs were suddenly brought back to life, working faster than ever before, and maintenance issues all but disappeared, even for PCs that were up to 10 years old. Furthermore, the software suite used was completely sufficient for all of the teachers’ and students’ needs, and the fact that it was free meant that copies could be distributed to as many servers as desired, and indeed install disks could be sent home with students so they could use the software at home as well as at school. The package used, K12LTSP, includes dozens of educational applications, with an emphasis on math and science. After the initial deployment, and with over 100 donated PCs from local businesses, a seachange occurred in how the PCs were used in the classrooms, and test scores even went up dramatically the number of working PCs in each room was tripled. The result is that the public school district of Atlanta is now deploying this technology, using brand new thin clients the size of a cable modem, servers which support over 100 clients each and LCD monitors. The LCD monitors were chosen because when classroom ratios of student to PCs of 2:1 are approached, the electricity available in a room becomes the limiting constraint, and lower power-consuming devices must be used. A a new thin client with an LCD monitor uses about 1/6 the electricity of a conventional desktop PC and CRT monitor.

The low cost of thin client technology, the depth of FOSS software available, and the complete success of this approach for school districts that typically struggle to maintain their information technology, led to the proposed cable thin client solution provided in this paper.

## ADAPTING THIN CLIENT TECHNOLOGY TO CABLE DATA NETWORKS

To implement thin client technology over a cable data network, several modifications to the basic architecture must be made. First, the enterprise version must be implemented, wherein powerful servers in the cable operator’s facilities are used. This is so that truck rolls likely involve using separate servers for boot, authentication, file, and application serving. There are two approaches currently used for application servers: first, use multiple application servers that each have all of the application software. This approach has the benefit of providing load balancing and automatic failover if a server has issues. The second approach is to use application servers that have been optimized for specific applications. This approach often optimizes performance in the system, but unless it is combined with redundancy, can lead to unhappy subscribers.

A proposed architecture for a cable thin client (TC) network is shown below. Note that four types of CPE are shown: standalone cablemodem (CM) and standalone TC; combined CM/TC, and then two similar variants where the CPE also includes a server. These latter two approaches would be applicable to MDUs and small businesses.are not required to manage and maintain the servers, as well as providing additional security for the servers and control over the user interface presented to clients. This would also

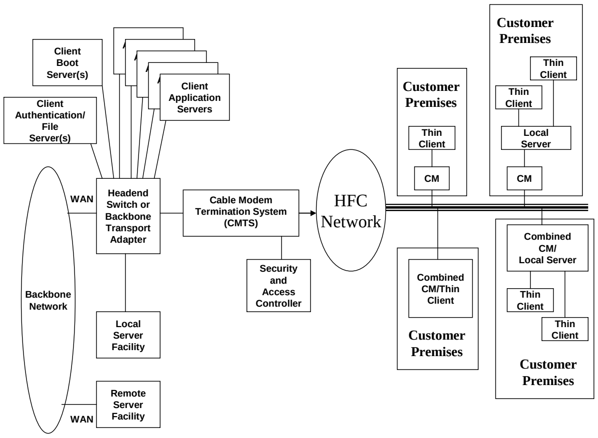


**Figure 3. Example Cable Thin Client Network.**

The bandwidth requirements of thin client networks can be substantial. Without any form of compression of the screen data before sending to the thin clients over the network, a single thin client can require several Mbps downstream data rate, and for particularly animated applications, a client can require over 10 Mbps downsteam data rate. Hence, even in the basic architecture, if more than 6 or 7 clients are connected to a server, then Gigabit Ethernet is required for the link between the switch and the server, which aggregates the traffic of all clients. However, most applications, like OpenOffice and even video streaming from websites like YouTube and CNN took only 1 Mbps or so per client.

Nevertheless, there are already available compression packages such as the FreeNx package (http://www.nomachine.com), which can be combined with Linux terminal server packages, and purports to offer thin client support, even over dialup networks. Experiments with this package indicate that while the latency is noticeable, it is tolerable, and clearly an optimization of such an approach would lead to better user experience, while keeping the bandwidth requirements low.

But cable data networks can take advantage of their broadcast architecture to enact further improvements in delivering thin client screen data. The desktop itself can be standardized such that much of the screen data is common across many thin clients. Further, as an overload protection for events such as popular web video suddenly becoming available, the video streams can be delayed and synched so that multicast delivery efficiencies can be obtained. Also, the thin client platform can be made thicker so that certain especially bandwidth intensive applications can be run locally. Finally, an adaptive system, which trades off client bandwidth requirements vs. latency can be used.

## BENEFITS FOR THE SUBSCRIBER

Why would a customer pay a monthly subscription to have a headache-free PC in the home loaded with software that is maintained by the cable operator and replaced by the operator if it ever fails? First, there are many subscribers, especially the low income and the elderly, for whom the desktop PC is either too expensive to purchase initially, too complicated to decide what to get, or too worrisome to have in the home unless a service provider supported it. Second, many homes with children still do not typically have a PC dedicated to the children's use, which means the children end up using the parent's PCs and often the web sites for kids end up downloading software that slows the PC down. Were the children's PC to be a cable thin client however, since the server is in the cable operator's facilities, the operator can use web content filtering and URL-blocking applications like DansGuardian and Squidgard, to provide a safe web accessto all children's thin clients on the network. Further, the operator could include other content related to the cable operator's business, such as flash-based children's programming, and built in links to cable operator' web sites for children.

The protection of user data and privacy may be effected in several ways. First, the user can take responsibility for their own data protection and use an encrypted USB flash drive plugged into the thin client. Second, the cable operator can maintain files in their facility, keeping them private via built-in mechanisms in Linux, or using other applications which increase file security. Especially for non-computer literate subscribers, the knowledge that files are automatically backed up by the operator could either enhance the service or provide additional monthly revenue. Printing, external CD ROM drives and other USB devices can also be installed in the client from the server, although a support call would typically be required.

User benefits thus include: all typical office applications (web browsing, office applications, email); children's edutainment applications; automatic, secure network backup of user files; automatic software updates; and all at no licensing cost.

Interestingly, one other user benefit of this architecture is that a user's experience is not fixed in space: a user can go to a neighbor's home, Internet café, etc., and can log onto the cable modem network there, perhaps requiring an encrypted USB key, and then their desktop environment, including all services to which they currently subscribe, are available. This kind of seamless mobility is a new wrinkle on PC mobility. In the school environment described earlier, this architecture has proven to be a method of providing ubiquitous PC access that is less costly and more reliable than 1:1 laptop initiatives, for example.

## CABLE OPERATOR OPPORTUNITIES

The main new opportunity for cable operators is a completely new source of revenue from offering a cable thin client subscription service. This addresses both existing high speed data customers as well as a currently untapped base of subscribers. Minimum 2 year contracts with free thin clients are a possibility, similar to the cell phone model. And a cable thin client goes beyond even the quad play arena into the so-called x-play scenarios, where new devices leverage multiple services, applications, and user experiences; a cable thin client is a platform for PC-TV as well as IPTV and interactive TV. Since the thin client is a "12 inch" interface, but can easily include more typical video content, the operator now has a platform for ITV that moves into the next generation of applications without concern for the often limited processing power and capabilities of older settop boxes.

The opportunity for ad revenue enhancement is also significant. With a cable thin client using a common portal and linked to the operator's content, the cable operator has unprecedented influence on the web and video experience of the subscriber. And by integrating data from such a service with other home services provided by the cable operator, the operator has valuable usage data across a variety of platforms and services that can raise ad revenue rates for the operator.

## FUTURE DIRECTIONS AND SUMMARY

Looking to the future, when high definition video, ultra high bandwidth data services, peer-to-peer video sharing and conferencing are all being delivered via cable operator's networks, the cable thin client represents yet another bandwidth consuming service, but one that easily integrates with existing services. In these scenarios, other means of delivering high definition video, which can include thin client screen data, may be required. One example is the DOCSIS IPTV Bypass Architecture (M. Patrick and G. Joyce, "Delivering Economical IP Video over DOCSIS® by Bypassing the M-CMTS with DIBA, SCTE 2007) as shown below, where the thin client servers have been moved deeper into the back office and the thin client uses DOCSIS 3.0 technology to further increase the bandwidth available to the client.

The benefits of thin client computing over cable networks can be summarized as follows:

For the service provider:

- New revenue opportunities
- Better metrics on customers, enhanced content value
- Leverages existing broadcast infrastructure
- Leverages existing Open Source applications

For the consumer:

- Maintenance-free computing, automatic updating of software and features
- Low cost entry to Internet computing
- Security, privacy, and reliability

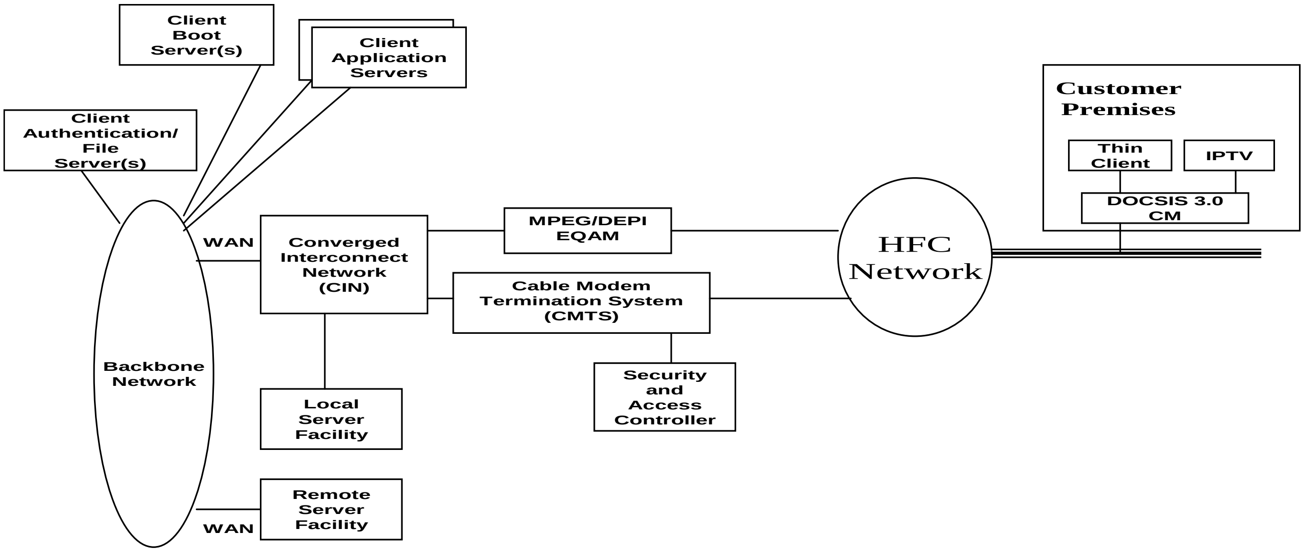


**Figure 4. Thin Client Architecture Using DIBA.**

# CREATING LOGICAL CHANNELS AND IMPLEMENTING ADVANCED SPECTRUM MANAGEMENT

Jack Moran and Michael Cooper
Motorola Connected Home Solutions

## *Abstract*

*DOCSIS® 2.0 support for logical channels enables cable operators to increase the available upstream bandwidth. This paper will highlight the business and service advantages of creating logical channels and implementing advanced spectrum management to monitor performance and efficiently utilize HFC bandwidth.*

## INTRODUCTION

Operators can quickly deliver increased upstream performance to HFC networks consisting of mixed mode DOCSIS 1.x and 2.0 Cable Modems (CMs) as well as for new networks consisting entirely of DOCSIS 2.0 cable modems. But there are issues to address to reap these rewards, and this paper will explore them.

This paper will explain the technology, discuss the challenges involved in implementing logical channels, and provide examples of performance gains realized by deploying logical channels in both mixed-mode and pure DOCSIS 2.0 environments.

It will also discuss how cable operators can utilize a spare receiver on a CMTS to perform spectrum management and non-obtrusively measure spectrum impairments, predict the dominant impairment in any area of the spectrum by analyzing signal-to-noise measurements, and qualify any area of the return path spectrum so they can achieve the maximum data signaling rate for logical channels.

They can leverage logical channels and advanced spectrum management to increase the aggregate bandwidth. Automation is key to the successful implementation of logical channels and advanced spectrum management.

The CMTS, Advanced Spectrum Management and DOCSIS 2.0 Logical Channel Operation can begin the tedious process via a script and automatically compile the various CMs or MTAs into the up to 4 Logical Channel Configurations supported. From the perspective of maximum throughput being maintained, it is quite possible to eventually allow the CMTS to automatically make the decision regarding Logical Channel Assignments and to affect the change, thus allowing the cable operator to adaptively change the maximum throughput possible based solely on return path conditions presented to the CMTS.

Today, this is not performed due to the cable operator having to gain confidence in the Logical Channel estimation process which is being performed by doing the analysis automatically and assigning all DOCSIS 2.0 CMs and MTAs the correct Logical Channel Assignment, but first allowing the cable operator to make the final decision as to the throughput change. Once the cable operator gains confidence in the accuracy of the throughput estimation via Logical Channel Operation, it is logical to assume that the cable operator will take the next step in the process and allow fully automated Logical Channel Operation.

## UNDERSTANDING LOGICAL CHANNELS

The concept of logical channels was introduced as a mechanism to allow legacy DOCSIS 1.0 and 1.1 CMs—which only support the Time Division Multiple Access (TDMA) protocol—to coexist with the Synchronized Code Division Multiple Access (SCDMA) protocol introduced with DOCSIS 2.0. That is, logical channels were created out of necessity to support the large, existing 1.x CM base already fielded. However, this feature supports so much more than just legacy operation.

Before exploring the benefits of logical channels, a clear definition is warranted. DOCSIS defines a logical upstream channel as, "A MAC entity identified by a unique channel ID and for which bandwidth is allocated by an associated MAP message. A physical upstream channel may support multiple logical upstream channels. The associated UCD and MAP messages completely describe the logical channel."

Logical channels are a unique set of transmission characteristics that are dealt with via the use of different modulation (burst) profiles in the upstream direction. All Logical Channels (modulation profiles) are time division multiplexed into the same physical DOCSIS channel. A physical DOCSIS channel is defined by such parameters as: a) symbol rate (six rates from 160 ksym/sec to 5.12 Msym/sec in octave steps), and b) center frequency. Logical channels further define the channel with burst profile attributes.

Prior to the introduction of logical channels in DOCSIS 2.0, each physical channel was required to utilize only one value for each of the burst profile attributes. Thus, the configuration of an upstream channel in DOCSIS 1.1 or 1.0 was driven by the worst performing CM or tap in the plant. For example, if a CM was located at a point in which a significant micro-reflection was present which was beyond the modem's ability to successfully transmit 16QAM using its 8-tap equalizer, then the operator was forced to configure the entire channel (and therefore all modems on that channel) to utilize the QPSK modulation. Now, if that micro-reflection characteristic was only present for 1% or 2% of the CM population on the channel, then we have a scenario in which the operational throughput of the channel is being dramatically limited by a small fraction of the modem population. The logical channel feature allows us to circumvent this problem by defining up to four logical channels on that one physical channel.

For example, logical channel 1 might be configured with a QPSK modulation and the 1-to-2% of poor performing modems would be assigned to that logical channel. Similarly, logical channel 2 could be configured for 16QAM and the remaining population would be assigned to this logical channel. Using this one configuration change, we have now reclaimed nearly a 98% increase in the upstream throughput. Another interesting fact is that we realized a significant increase in throughput by only leveraging DOCSIS 2.0 functionality within the CMTS. That is, we realized this improvement even with only 1.0 CMs present on the plant. It is easy to see how additional logical channels might be created which leverage higher-order modulations (32QAM, 64QAM, and even proprietary 256QAM). For these logical channels, only the corresponding modems which support such capabilities would be assigned to the logical channel.

While changes in modulation type provide for the most dramatic changes in realized throughput, the use of logical channels is not limited to this single parameter. In cable plants where micro-reflections or amplitude roll-off are significant issues, the use of preamble lengths may be exploited to yield improvements in throughput. For example, modems encountering greater linear distortion could be assigned to logical channels which utilize longer preambles and therefore yield better equalizer performance, while modems encountering less distortion would be assigned to logical channels utilizing shorter preambles. A similar technique can be applied to FEC [both codeword length (K) and number bytes corrected (T)], byte interleaving, guard times, or any combination thereof. This provides the needed mechanism to deal with the variances of modem signal quality resulting from such factors as system loss, amplifier cascades, and micro-reflections.

Cable operators can leverage logical channels to improve the throughput of legacy DOCSIS 1.0, 1.1 CMs as well as optimize performance for DOCSIS 2.0 and even 3.0 CMs. They can implement logical channels on pure DOCSIS 2.0 or 3.0 deployments, but the reality is that most cable networks also consist of legacy DOCSIS 1.0 and 1.1 CMs. Logical channels can support mixed-mode operation and they can optimize the performance of diverse CMs deployed throughout the access network.

## MANAGING IMPAIRMENTS

DOCSIS 2.0 has opened opportunities in which increased efficiencies and greater throughput can be achieved within the return path. However, increasing efficiencies and enhancing throughput is not just a simple matter of enabling the new features in DOCSIS 2.0.

One must first understand the dominant characteristics and impairments of a given return path before channels can be reconfigured accordingly. When fairly simple characterizations are performed, dramatic increases in throughput can be achieved.

Channels which would be unusable with DOCSIS 1.0/1.1 can now be reclaimed, and throughput of legacy channels can be increased 50% or more with an optimal configuration. Further, the characterization methodologies presented in this paper will yield relationships between actual plant devices and dominant impairments present within a given return path. By identifying dominant impairments and actual plant devices causing such impairments, this methodology supports targeted maintenance activities for so-called low-hanging fruit improvements that yield major performance benefits.

The higher bit rates achieved with higher modulations come at the expense of greater Modulation Error Ratio (MER) requirements [sometimes incorrectly referred to as Signal-to-Noise Ratio (SNR)] on the upstream channel. These requirements go beyond the traditional first order issues, such as:

- Thermal noise

- Ingress noise
- Impulse noise

This paper focuses on the topic of increased modulation levels and the issues to be overcome to support such levels. Specific issues that will be discussed within this paper are:

- Linear impairments
- Non-linear impairments

## ADVANCED SPECTRUM MANAGEMENT

A spare receiver architected into a CMTS module can allow cable operators to best understand the impairments on the DOCSIS infrastructure. It runs in parallel to the live ports to monitor performance of any one of the upstream ports without materially impacting the subscriber experience.

The receiver is connected in parallel with a selected receiver port so the operator can measure traffic and performance in real time on any given receiver port. The parallel receiver can access all of the mapping information as well as a full list of cable modems available to whichever receiver port is currently being evaluated.

Therefore, while the receiver port being monitored is performing its function at full capacity, the parallel receiver can non-obtrusively gain access to all of the return nodes connected to one of the receiver ports and perform tests on each upstream channel to assess its quality and take the time required to complete detailed, coherent MER measurements.

Cable operators can leverage advanced spectrum management to optimize the performance of cable modems and better understand how to automatically compensate for linear and non-linear impairments.

With the release of DOCSIS 3.0, spectrum management will become even more critical because the CMTS will be faced with maintaining quality of service on multiple bonded upstreams. Multiple upstream channels will require that MSOs reclaim more and more of their upstream frequency spectrum, including regions which have historically been avoided due to their greater susceptibility to various impairments. Maintaining quality of service across many service flows across multiple bonded upstream channels can not be performed manually by an operator and will require advanced spectrum management to make sure proper flows are assigned to physical channels capable of meeting quality of service requirements. With that stated, it is vitally important that the fundamental building blocks for a more detailed analysis that will be required for the future DOCSIS 3.0 be first proved out in all DOCSIS 2.0 services for the cable operators today.

## LINEAR IMPAIRMENTS

Linear impairments refer to a class of impairments that are signal-dependent and largely unique to a given responding CM because its transmission path through the return path network possesses its own micro-reflection (impedance mismatch). Moreover, the number of amplifiers in cascade also impact the amount of amplitude distortion and group delay (phase) distortion that a CM signal will be impacted by, that is to say the more amplifiers in cascade the more diplex filters the CM signal must traverse. This simply means that the effects of a linear impairment can only be observed while in the presence of a signal. The signal required for evaluation of linear impairments can either be very expensive test equipment, or the cable operator can simply opt to use a very inexpensive DOCSIS CM as the source and a spare receiver architected into a CMTS access module as the measurement tool.

In the end, when it comes down to convenience, speed of measurement, and cost, there can be little question that the combination of a DOCSIS CM and DOCSIS CMTS is the most effective characterization tool available to the cable operator. Obviously, the measurement accuracy of a system using a DOCSIS CM and DOCSIS CMTS is considerably less than using a Vector Signal Generator (VSG) such as the Agilent E-4438C or Arbitrary Waveform Generator (AWG) such as the Agilent E5182A and a Vector Signal Analyzer (VSA) such as the Agilent 89640A, 89650S or N9020A or a second generation CATV Analyzer such as the Sunrise Telecom AT-2500RQ, but the speed of the DOCSIS CM/DOCSIS CMTS system more than makes up for the accuracy limitations. Consider the fact that the CM is already installed in the network and that even a detailed DOCSIS 2.0 CMTS – CM measurement takes less than 100 ms and a less detailed measurement takes less than 5 ms. The CATV engineer using any one of the devices mentioned above will have an average measurement time of no less than 10 seconds for a less detailed measurement and the measurement time can easily be over three minutes for a detailed analysis. Obviously, one isn’t even discussing the time for a CATV technician to get to a remote site location and connect the VSG or AWG to the CATV network.

If a cable operator wants a definitive linear characterization of a return node cascade of amps and the optical link, then nothing substitutes for setting the entire circuit in a lab environment and using a classic RF network analyzer such as the Agilent 8753ES (75 Ohm network analyzer) to report the definitive amplitude and group delay response from 2 MHz to 52 MHz. Alternatively the VSG/VSA combination is actually more useful in that one can not only receive a detailed report regarding the amplitude and group delay distortion, but one can also receive a report regarding the definitive MER versus the carrier frequency as well. With lab characterization accuracies set aside, the realities of characterizing a live CATV node in the field poses a variety of issues that lab equipment simply cannot deal with such as:

- Physical distance
- Ingress noise
- Impulse noise
- Live traffic
- Time Synchronization

Therefore, the DOCSIS CM and DOCSIS CMTS system offers the cable operator the only characterization technique that is both non-disruptive to customer traffic and extremely convenient both from availability and from a time to measurement perspective.

It is also important for the cable operator to understand why until recently, linear impairments have not been a concern. This is due to the fact that QPSK modulation has no amplitude modulation associated with it, and it is fairly immune to linear distortion affects. This is fundamentally due to the typical micro-reflection tending to be in the 15 dB to 25 dB range and given that QPSK needs only an MER > 14 dB to be perfectly acceptable from a performance perspective, one can easily see why this linear impairment has not been a show stopper for QPSK modulation. Another more important point regarding all linear impairments is that the impact on MER performance is also a function of DOCSIS channel bandwidth.

That is to say the significance of a micro-reflection or even the effects of multiple diplex filters has a dramatically larger impact on a channel bandwidth of 3.2 MHz than it ever did for the older 1.6 MHz bandwidth services. One can then understand that linear impairments are a major impact on the wider DOCSIS 2.0 channel bandwidth of 6.4 MHz. This is the primary reason why the DOCSIS 2.0 Equalizer was increased from 8 taps to 24 taps in length. As a result, linear impairments have only recently generated attention as more and more operators have moved to a 3.2 MHz bandwidth first and are now moving to 16QAM modulation, which with the combination of wider bandwidth and 16-QAM modulation is more susceptible to these affects. As operators seek to enable more advanced modulations of 32QAM and 64QAM provided by DOCSIS 2.0 and utilize the widest channel bandwidth available of 6.4 MHz, these impairments will become even more critical.

As far as the HFC plant is concerned, linear impairments generally fall into one of three classes:

- Micro-reflections –
  impedance mismatches
- Amplitude and group
  delay distortion –
  diplex filters
- Amplitude tilt or slope
  – coaxial cable

## NON-LINEAR IMPAIRMENTS

As in the case of linear impairments, system non-linearity is also a signal-dependent distortion. Moreover, while linear impairments such as micro-reflections and diplex filter effects can be observed in the way the noise floor is shaped; there is by definition no system non-linearity without the presence of the signal. In essence, if there is no signal there is no system non-linearity occurring. More importantly, the impact on a DOCSIS signal that a system non-linearity presents is also a function of the level of QAM that is being transmitted. Simply stated, since system non-linearity is signal dependent, then it also holds true that the larger the signal power the more severe the system non-linearity impacts the DOCSIS signal.

Traditionally, most users rely almost entirely on MER to understand channel quality and predict performance under various configurations. Because of the fact that non-linearities impact higher levels of QAM more so than lower levels, an operator must be careful when interpreting the MER (SNR) of a communications signal at say QPSK so as not to quickly extrapolate the capabilities of that channel to support higher levels of QAM.

Given that system non-linearity has a much greater impact on a higher power signal, it also follows that system non-linearity impacts the outer points of the transmitted signal constellation more than the inner points of the same transmitted constellation. This is fundamentally due to the fact that the outer points are transmitted at a significantly higher power than the inner points and thereby are impacted significantly more than the inner points. This, too, implies that an operator must be careful when interpreting the meaning of MER values reported for a communications channel.

For example, if a communications channel is configured to transmit 64QAM in a nearly perfect communications channel except for non-linearity, then all but the outer corner points of the constellation will be nearly perfect. However, depending upon the degree of non-linearity, the outer points may be so corrupted that they are compressed to a point of being non-distinguishable from inner points; that is, a very high error rate may still result. For example, if this non-linearity only impacts the 3 outer constellation points at each corner, then only 12 points (3 in each of the 4 corners) of the total 64 are impacted. The impact on MER is 12/64 or 19%, which implies that it will have only a minimal impact on the reported MER.

There is a significant difference in impact to a DOCSIS transmitted constellation when it is subjected to a 2nd Order Inter-modulation Distortion (IMD)—commonly referred to in the CATV Industry as Composite Second Order (CSO)—versus a 3rd Order IMD Distortion—commonly referred to as Composite Triple Beat (CTB). It is in fact the 3rd Order IMD Distortion that is far more damaging to the DOCSIS transmitted constellation since it impacts the outer points more significantly than the inner constellations, while 2nd Order IMD tends to impact all of the constellation points evenly.

The system non-linearity that we have been referring to is created by two completely different circuit areas of the CATV Network. Common Path Distortion (CPD) is the result of dissimilar metals acting as a diode and exhibiting both 2nd and 3rd Order Distortion that occurs on the coaxial cable path that is common to both forward and return path directions.

CPD is well understood by the CATV industry in general. It is the phenomena of a coaxial connector becoming or temporarily acting as a diode. It is easily observed by seeing analog video carriers spaced 6 MHz apart throughout the return path. While CPD is easily detectable, the return laser being either clipped or just becoming marginally non-linear can only be witnessed today by advanced spectrum management on a dedicated receiver on a CMTS card, or by deploying vector signal analyzer test equipment or second-generation CATV analyzers with a DEMO function and a known reference signal installed on the network.

With advanced spectrum management, one can easily observe that the effect of any non-linearity is that the outer constellation points are impacted far greater than the inner constellation points.

## AUTOMATING LOGICAL CHANNEL CONFIGURATIONS

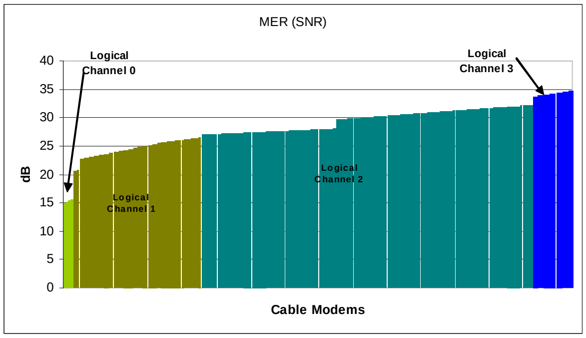
Logical channels allow cable operators to realize higher upstream bandwidth capacity per port by enabling different modulation profiles on a single port.

Proactive configuration tools are required so that cable operators can implement proactive maintenance strategies that enable automated decision making by providing actionable data.

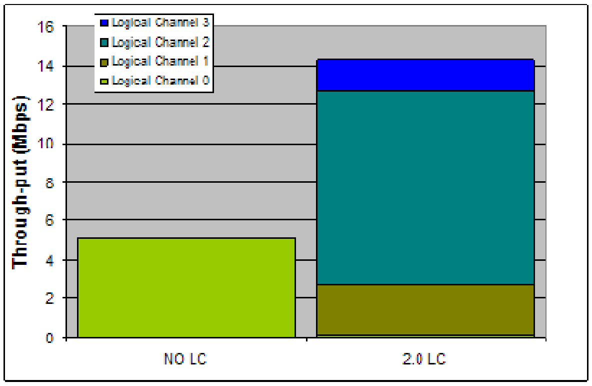
Much of the necessary data required to increase bandwidth and perform plant maintenance is captured in the CMTS; however, automated configuration management tools are needed so that cable operators can proactively allow automated decision-making that configures CMs for maximum performance based on rules and policies determined by the cable operator.

Now that we have quantified the improvements that can be realized using logical channels, we now turn to the issue of automated configuration. That is, how does an operator easily identify what logical channel configurations to define and which modems should be assigned to each logical channel. A software tool which supports the collection and sorting of various CM performance statistics is necessary.

DOCSIS 2.0 introduced requirements for a multitude of new performance statistics that are unique to each modem within the plant. Specifically, per modem statistics were added for MER, micro-reflections, and FEC statistics. The number of statistics in combination with the hundreds of thousands of modems makes for a large and complex data-mining problem that can not be solved manually by a human operator. By extracting these modem statistics, a tool could allow the operator to sort and analyze the distribution of modems on the plant and their associated performance measures. The following figure provides an example of a sample distribution. This distribution could then be used to isolate groupings of CMs with similar characteristics, and logical channels meeting the needs of each grouping could then be created.



This management tool could then be used to predict the throughput performance of a new configuration, illustrated in the following figure.

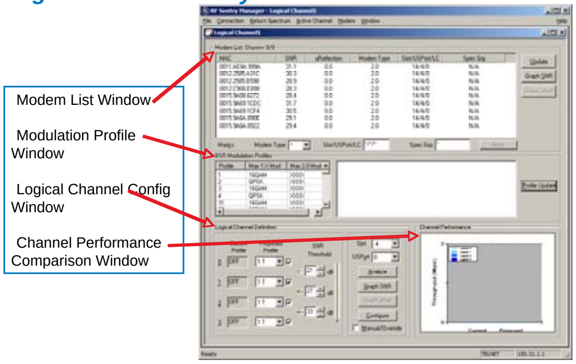


Logical channels offer huge opportunities to reclaim throughput efficiencies on the upstream. While the magnitude of improvement is dependent upon the characteristics of the cable plant, these significant benefits are even available to plant environments which are dominated by 1.0 CMs. The key roadblock to leveraging this technology is the need for automated tools that facilitate the cumbersome task of assigning each individual modem to the appropriate logical channel. However, as the benefits of this capability are realized within the industry, we fully expect vendors to begin offering tools to automate such a process.

Configuration management tools can serve as GUI-based front-ends for CMTS platforms, and they can leverage the insights gained from advanced spectrum management to automate the configuration of CMs to maximize performance.

Operations personnel could then benefit from a logical channel analysis user interface that graphically presents the salient performance information necessary for configuring logical channel operations.

*Logical Channel Analysis User Interface*



They will be able to easily analyze logical channels and monitor the automated configuration of logical channels. Graphing capabilities allow operators to quickly identify problem CMs and the corresponding percentage of the total modem population impacted. Operations personnel will be able to review summary update of CM profiles while gaining the flexibility to drill down into detailed windows that present full modulation profiling information.

Analysis features provide a graphical presentation of the total channel capacity for both the current profile configuration and the proposed profile configuration. Cable operators can allow the configuration manager to automatically initate reconfigurations of profiles and move modems to logical channels while retaining the option to manually accept or reject reconfigurations.

This powerful management tool therefore provides cable operators with the ability to automate configuration of CMs to maximize performance as well as the option to only allow the system to reconfigure CMs after an operator has manually accepted the reconfiguration recommendations.

## SUMMARY

DOCSIS 2.0 support for logical channels allows cable to provide significant amounts of increased upstream bandwidth. Operators can quickly deliver increased upstream performance to HFC networks consisting of mixed mode DOCSIS 1.x and 2.0 cable modems as well as for new networks consisting entirely of DOCSIS 2.0 cable modems.

Cable operators can implement logical channels in both mixed-mode and pure DOCSIS 2.0 environments and they can concurrently deploy advanced spectrum management to utilize a spare receiver to non-obtrusively measure spectrum impairments, predict the dominant impairment in any area of the spectrum by analyzing signal-to-noise measurements, and qualify any area of the return path spectrum so they can achieve the maximum data signaling rate for logical channels.

They can leverage logical channels and advanced spectrum management to increase the aggregate bandwidth. Automatic logical channel configuration complements and adds value to advanced spectrum management, and it allows cable operators to successfully optimize the highest throughput of a given channel and better support the QoS demands of real-time services such as voice and video.

# DELIVERING ECONOMICAL IP-VIDEO OVER DOCSIS BY BYPASSING THE M-CMTS WITH DIBA

Michael Patrick
Motorola Connected Home Solutions

## *Abstract*

*DOCSIS IPTV Bypass Architecture (DIBA) DIBA refers to any of a number of techniques whereby downstream IPTV traffic is directly tunneled from an IPTV source to a downstream Edge QAM, "bypassing" a DOCSIS M-CMTS core. It will allow operators to deliver economic IP video traffic over DOCSIS infrastructure.*

## INTRODUCTION

MSOs are considering IP-video and IP Television (IPTV) to supplement their current digital video delivery. IP-based video enables new video sources (the Internet) and new video destinations (subscriber IPTV playback devices).

Of course, such a transition to IPTV requires significant additional downstream DOCSIS® bandwidth. In a reasonable 5-7 year "endgame" scenario, VOD is expected to require 26 times the bandwidth of High Speed Data (HSD). What's more, attempting such an increase in downstream bandwidth with a conventional Modular-CMTS (M-CMTS) will be expensive.

As an alternative, we propose that IP-video content travel directly to an Edge QAM, bypassing the M-CMTS core altogether and saving local switch and CMTS core capacity. This technique is called the DOCSIS IPTV Bypass Architecture (DIBA). DIBA refers to any of a number of techniques whereby downstream IPTV traffic is directly tunneled from an IPTV source to a downstream Edge QAM, "bypassing" a DOCSIS M-CMTS core. By bypassing the high-cost components of the CMTS core, DIBA is an architecture for delivery of high-bandwidth entertainment video over DOCSIS to the home for a price per program that matches the cost for conventional video over MPEG2 delivery to set-top-boxes.

DIBA is an innovative bypass architecture in which the conventional Modular-CMTS architecture is replaced by a hybrid architecture consisting of an integrated CMTS and DOCSIS External Physical Interface (DEPI) or MPEG Edge QAMs. This paper/presentation will demonstrate the effectiveness of DIBA, discuss how operators can evolve their infrastructure to implement DIBA, discuss alternative bypass encapsulation protocols to tunnel traffic to either DEPI or MPEG Edge QAMs, and outline the economic advantages of DIBA migration. It will explain how DIBA removes the boundaries of delivering video using a M-CMTS, and how operators can accelerate service velocity by efficiently delivering IPTV services that bypass M-CMTS platforms.

## THE CHALLENGE OF DELIVERING IPTV

IP-video and IP Television (IPTV) create opportunities to supplement their current digital video delivery. IP-based video enables new video sources (the Internet) and new video destinations (subscriber IPTV playback devices). It should be considered as a delivery option of Video on Demand (VOD).

The extent to which VOD is delivered as IPTV governs the extent to which additional downstream DOCSIS bandwidth must be provided for IPTV. Consider a hypothetical 5-

7 year “endgame” scenario where each video subscriber can download “What I Want When I Want” (WIWWIW). Assume a fiber node of 750 homes passed has 75% video subscriber penetration, and during the busiest video viewing hour 50% of the video subscribers require an average of 10 Mbps VOD content. The total VOD content required to be delivered to that single fiber node during the busy hour is thus 750 * .75 * .50 * 10 Mbps = 2.8 Gbps, or 73 carriers of 6 MHz QAM 256. What fraction of this “endgame VOD” bandwidth is IPTV over DOCSIS and what fraction is traditional Digital Video to Digital Set-Top Boxes (DSTBs) is anybody’s guess.

VOD bandwidth swamps High-Speed Data (HSD) bandwidth. Last year, most MSO deployments offered only 20 Mbps of HSD per 750-household fiber node. A reasonable endgame scenario for HSD would provide 100 Mbps channel bonded service to 50% of households passed with a 0.25% concurrency, for a total HSD fiber node throughput requirement of 750 * 0.5 * .0025 * 100 Mbps = 94 Mbps. Thus, under reasonable hypothetical endgame scenarios, VOD is expected to require 26 times the bandwidth of HSD.

A key goal of the DOCSIS Modular CMTS (M-CMTS) architecture was to define a common Edge QAM (EQAM) architecture that could support a cost-effective transition from Digital Video (DV) VOD delivery to IPTV VOD delivery over DOCSIS. A straightforward strategy of using M-CMTS to provide IPTV VOD would call for deploying M-CMTS EQAMs for DV-VOD first, and adding M-CMTS core capacity as needed to transition the DV-VOD to IPTV-VOD. As we have seen, though, the M-CMTS core capacity used for IPTV-VOD will quickly exceed the CMTS capacity used for HSD. A straightforward implementation of M-CMTS for IPTV, therefore, will require *many times* the cost of M-CMTS core capacity as that currently used for HSD.

## AN INTRODUCTION TO DIBA

But there is a simple alternative for IPTV support with a M-CMTS: IPTV should be tunneled directly to the EQAM, bypassing the M-CMTS core altogether.

DIBA refers to any of a number of techniques whereby downstream IPTV traffic is directly tunneled from an IPTV source to a downstream EQAM, “bypassing” a DOCSIS M-CMTS core. By bypassing the high-cost components of the CMTS core, DIBA is an architecture for the delivery of high-bandwidth entertainment video over DOCSIS to the home for a price-per-program that matches the cost for conventional video over MPEG2 delivery to set-top-boxes. IPTV is by definition an IP packet with video content from an IP source to an IP destination address.

In traditional IPTV, an IPTV server sends an IP packet to the destination address of a subscriber playback device, which we call generically an “IP Set-Top Box” (IPSTB). With the M-CMTS architecture forwarding, the IPTV packet is routed through the M-CMTS core. In a conventional M-CMTS architecture without DIBA, therefore, the IPTV content is forced to make two transits through the Converged Interconnect Network (CIN) switch that connects regional networks with the core CMTS.

The video/IP traffic travels to the M-CMTS core, then “hairpins” back through the CIN on a DOCSIS External Physical Interface (DEPI) pseudo-wire to the DEPI EQAM. If-and-when IPTV becomes a significant fraction of overall bandwidth delivered to a fiber node, this hairpin forwarding of IPTV content will require significant expenditures by MSOs for M-CMTS core and CIN switching bandwidth.

In DIBA, rather than pass through the M-CMTS core, the high-bandwidth video/IP content is tunneled from the IPTV server, through the MSO’s converged interconnect network directly to the DIBA EQAMs. The IPTV emerges from the DIBA EQAM with full DOCSIS framing, suitable for forwarding through a DOCSIS cable modem to home IP devices. This architecture will deliver high-bandwidth entertainment video/IP/DOCSIS to the home for a price-per-QAM that matches that of the most advanced, cost reduced MPEG2 EQAMs.

## REDUCING DELIVERY COSTS

Operators can deploy independently scalable numbers of downstream channels without changing the MAC domain or the number of upstream DOCSIS channels. These downstream channels are available for VOD/IP and switched-digital-video/IP. They can also lower the cost to deliver video over DOCSIS service to be competitive with today’s MPEG VOD. DIBA accomplishes the goals of the M-CMTS without the unnecessary expense of the M-CMTS. DIBA avoids the expense of the DOCSIS MAC domain technology for the video/IP traffic by using both synchronized and unsynchronized DOCSIS downstream channels. The synchronized channels pass through the integrated CMTS or the CMTS core and provide the many DOCSIS MAC functions, including:

- Conveying the DOCSIS timestamps
- Managing ranging to provide the proper time-base to the cable modem
- Instructing the cable modems when to transmit upstream
- Delivering other MAC layer messages for cable modem registration, maintenance, etc.

In contrast, the unsynchronized DOCSIS channels are generated by the EQAMs (including installed MPEG EQAMs) for the bypass traffic and omit these functions. With an integrated CMTS and no timestamps in the un-synchronized channels, the DOCSIS Timing Interface which is required in the M-CMTS architecture is not necessary in DIBA.

In DIBA, a DOCSIS 3.0 cable modem requires only one synchronized channel from the existing integrated CMTS to provide timing, control the upstream transmissions, and provide the other MAC functions. The DOCSIS 3.0 cable modem will receive video/IP on additional un-synchronized, inexpensive, DIBA-generated channels, which can even be bonded.

## LEVERAGING EXISTING EQAMS

DIBA offers a variety of bypass encapsulations. Using a standard M-CMTS DEPI EQAM, two bypass encapsulations are possible, depending on the video server and MSO network. In either case, the server originates a Layer 2 Tunnel Protocol Version 3 (L2TPv3) tunnel to the DEPI EQAM. The tunnel payload can be:

- The DOCSIS Packet Streaming Protocol (PSP) in which the video/IP is encapsulated into DOCSIS MAC frames. This permits the EQAM to mix both IPTV traffic originated from the IPTV server with non-IPTV (e.g. HSD or VOIP) traffic originated from the M-CMTS core on the same DOCSIS downstream carrier.

- The DOCSIS MPEG Transport (D-MPT) layer which consists of an MPEG-2 Transport Stream of 188 byte packets. All DOCSIS frames, including packet-based frames and MAC management-based frames, are included within the one D-MPT flow. The EQAM searches the D-MPT payload for any DOCSIS SYNC messages and performs SYNC corrections. It then forwards the D-MPT packet to the RF interface. The intent of D-MPT mode is to allow MPEG packets to be received by the EQAM and forwarded directly to the RF interface without having to terminate and regenerate the MPEG framing. The only manipulation of the D-MPT payload is the SYNC correction.

A standard MPEG2-Transport Stream (MPEG2-TS) EQAM can also be used. Here the video server may transmit an IPTV PSP formatted data packet. A PSP/MPT converter, either attached to or embedded within the CIN networking device, then changes the format into an MPEG-2-TS that a conventional MPEG EQAM can process. An alternative is for the VOD server to directly generate IPTV/MPT formatted packets that the MPEG EQAM can process. In the case of non-synchronized DOCSIS channels, a non-DOCSIS Program ID (PID) is used because each D-MPT program would require a separate MPEG-2 PID. A required extension for sending multiple programs streams of D-MPT to the same downstream QAM channel would be for DOCSIS 3.0 cable modems to be programmed to accept D-MPT-formatted packets with other than the standard DOCSIS PID.

## TUNNELING BROADCAST IPTV

There are several tunneling options that can be used. One of these is “Interception”, in which SPTS/UDP/IP encapsulation is intercepted by a CIN multicast router and encapsulated into PSP tunnel. Another is PSP Tunnels directly from the video server to the DEPI EQAMs. For bonded IP multicast, a “DIBA Distributor” component co-located with a CIN IP multicast router distributes or “stripes” the IP packet to a pre-configured DOCSIS 3.0 Downstream Bonding Group. Switched broadcast IPTV multicast is sent without BPI encryption on non-synchronized downstream channels to DOCSIS 3.0 cable modem IPTV devices.

## THE ROLE OF THE CONTROL PLANE

While there is no design yet for the DIBA control plane, there are several other control planes that must be considered in that design. The current MSO switched broadcast architecture is intended to save bandwidth for a moderate-sized optical node, and that feature will have to be retained in any DIBA implementation. The PacketCable Multimedia (PCMM) architecture is intended to provide QoS within a DOCSIS system, and that function will have to be retained. The M-CMTS architecture is intended to provide a scalable number of downstream DOCSIS channels managed by an edge resource manager, and this feature is an essential part of DIBA. IP Multimedia Subsystem (IMS)-based IP services will include entertainment video as well as personalized and on-demand video.

## VOD AND SWITCHED BROADCAST

Many cable operators have existing VOD and switched digital video architectures. These two architectures have quite different purposes, but are similar in one major way to conventional digital broadcast video. In all cases the video is brought over the MSO network to the EQAMs as MPEG-2 SPTS/UDP/IP. The EQAM then strips off the UDP/IP encapsulation, multiplexes the material into multi-program MEPG-2 TSs that are transmitted over QAM carriers. The set-top boxes are able to demodulate the MPEG-2 TSs and decode the video. Since the set-tops cannot receive video over an IP stack, they must be instructed which QAM carrier frequency and which Program ID to demodulate. In the case of switched digital video, the objective is to make more video channels available than there is bandwidth for in a typical RF spectrum. All the available video content is carried as a collection of IP multicast sessions, in most cases with all content available within the hub. Only when a set-top within a fiber node requests a particular video title is that title carried to that node.

Statistically, some number of set-tops will be viewing each of the popular video broadcast titles. Other, less popular titles will likely be viewed by only one set-top at a time. Still other less popular titles will not be viewed at all. Thus, the limited bandwidth available to that node will accommodate a great many more available titles than could be carried simultaneously.

When the control plane receives a request from a set-top for a new title, it searches for an EQAM serving that node which has sufficient available bandwidth. This EQAM is instructed to issue an IGMPv3 join to that IP multicast session. In the case of VOD, the control plane must locate a server with the desired title. This server will then send the title as unicast MPEG2-TS/UDP/IP to an appropriate EQAM serving the set-top making the request. In the case of VOD, there can be thousands of available titles.

## PACKETCABLE MULTIMEDIA

PacketCable Multimedia will be an important element of the control plane for DIBA. The objective of managing QoS for the myriad of services to be provided over the limited bandwidth of DOCSIS 1.0 and 2.0 networks led to the development of PCMM. PCMM uses such elements as the Application Server, Application Manager, Policy Server, and CMTS. These are all in the headend and under control of the MSO. PCMM is only available for IP-based services. There are several scenarios for the operation of PCMM for IPTV. In one, the client (video viewing) device does not itself support the QoS signaling mechanisms and relies on the Application Server as a proxy. The client sends a service request, and the Application Server sends a service request to the Application Manager. Upon receipt of this request, the Application Manager determines the QoS needs of the requested service and sends a Policy Request to the Policy Server. The Policy Server in turn validates the Policy Request against the MSO-defined policy rules and, if the decision is affirmative, sends a Policy Set message to the CMTS. The CMTS performs admission control on the requested QoS envelope (verifying that adequate resources are available to satisfy this request), installs the policy decision, and establishes the service flow(s) with the requested QoS level.

## M-CMTS EDGE RESOURCE MANAGER

The M-CMTS architecture takes the PCMM architecture one step further in terms of granularity. Within the PCMM architecture, it is the CMTS that makes the admission control decision for new flows and assigns then to a suitable DOCSIS channel. In the M-CMTS, the core CMTS must in turn request QAM resources from the edge resource manager.

## IP VIDEO AND IMS

There has been much work within the Internet community to use IMS to control video services over IP to offer roaming and two-way features. This has generally not been for entertainment video. However, the current aim of IMS has become the universal deployment of all IP-based services through both fixed and mobile networks, regardless of location. So it is possible that IMS will become a suitable vehicle for the deployment of entertainment video as well, particularly personalized and on-demand video.

This is an aggressive goal, but not an unreasonable one, given the history of and effort that has gone into IMS. IMS began as an effort by telecom carriers in the Third Generation Partnership Project (3GPP) to converge mobile services with VoIP and IP multimedia by using the Session Initiation Protocol (SIP). SIP is an IP-based peer-to-peer protocol used to establish and control two-way flows carrying voice, video, and gaming. The actual flows use Real-Time Protocol (RTP) and UDP/TCP. True to its heritage, non-IP devices such as analog telephones are supported through gateways.

Gradually the range of services and access technologies increased, resulting in a powerful Next Generation Networking (NGN) architecture. The access technologies include any IP/SIP-enabled devices, including phones, PDAs, computers over DSL, DOCSIS, 802.11, etc. The network control plane is built around a Call Server Control Function (CSCF) that comprises the following:

1. Proxy-CSCF, which is an initial control plane element to interact with an IMS device and acts as a proxy in the SIP processing. Beginning with registration, each IMS terminal is assigned to a Proxy-CSCF. The Proxy-CSCF authenticates the IMS device and establishes security. It also authorizes the IMS device to use network bandwidth, and may apply QoS policies to this use. It is in the path of all signaling.
2. Server-CSCF, which is a SIP server and binds the user IP address to its SIP address. It interfaces to the user database (the Home Subscriber Server or HSS) where user information is kept and sees all signaling messages. It also determines which of the application servers to route particular messages to.
3. Interrogating-CSCF is another SIP proxy, this one serving as the interface by which IMS messages from the outside world access the local network. The I-CSCF is able to inspect the user database to determine, for a particular IMS terminal, the correct Server-CSCF to forward signaling messages.

The network also has application servers that are linked to the control plane via the Server-CSCF. In the past, these applications have been call related, such as call-waiting, call-forwarding, conference calls, voicemail, etc. However, it is possible to have these services include such network-oriented tasks as resource management for VOD sessions.

There is an obvious advantage to using IMS and SIP to provide entertainment video services in that the control plane would naturally extend to any IP video sources in the Internet on the one hand, and to a multitude of fixed and wireless video playback devices on the other hand. The roaming ability and the device presence enable handoff between devices, a necessary component of seamless mobility. There are many issues to deal with in integrating the VOD and switched digital video control and data planes with IMS/SIP.

## BROADCAST IPTV OVER DIBA

In the typical MSO approach to switched broadcast video, the set-tops are non-IP so it is the MPEG EQAM that joins the IP multicast group. The multicast group is then mapped to a particular QAM and PID, and it is this information that is forwarded to the set-top to enable it to receive the MEPG-2 video transport stream.

In fact, DOCSIS is not generally used to deliver broadcast entertainment video, since the bandwidth is too expensive. However, the intention of DIBA is to bring the cost of DOCSIS bandwidth low enough to make IPTV viable. With DIBA, the IP set-top will be able to join the IP multicast group. The switched broadcast control plane determines which programs are multicast to which fiber node, consistent with the IP set-tops receiving which programs.

## BROADCAST IPTV TO DOCSIS 2.0 AND 3.0 MODEMS

Broadcast IPTV is possible with both DOCSIS 2.0 and 3.0 modems. In either case, the IP set-top box sends an IGMP join to an IP multicast session for a particular program. The DOCSIS 2.0 modem will have to use a synchronized DOCSIS downstream channel, while a DOCSIS 3.0 modem can receive on a Bonded Channel Set. It may or may not be necessary to change DOCSIS channels to change to a new IPTV ‘channel’. However, any DOCSIS channel changing requires communication from the CMTS to the cable modem in the form of a Dynamic Channel Change (DOCSIS 2.0) or Dynamic Bonding Change (DOCSIS 3.0) to change a tuner of the cable modem to the new channel.

## SUMMARY

Cable operators will be able to economically supply IPTV capacity of gigabits per fiber node by bypassing the M-CMTS core and tunneling IPTV directly to the M-CMTS EQAM. DIBA allows operators to avoid the cost of additional M-CMTS core capacity for IPTV. DIBA is proposed as a work item for CableLabs after DOCSIS 3.0 draft specifications. If it is approved, CableLabs would standardize:

- Data Plane operation from IPTV server to DIBA PSP/MPT Converter, DIBA Distributor, and DEPI/MPEQ EQAMs

- NGOD Session Manager control signalling to IPTV server

- CMTS control signaling to DIBA Distributor and PSP Converter functions in CIN

- CMTS control signaling to EQAM

DIBA will remove the boundaries of delivering video using a M-CMTS, and it will allow cable operators to leverage service velocity by efficiently delivering IPTV services that bypass the M-CMTS core.

## ABOUT THE AUTHOR

Mike Patrick is a Data Networking Architect, Motorola Connected Home Solutions. In this role, He is responsible for the architecture of Motorola’s Cable Modem Termination System products (CMTS). He is currently leading technical development of the DOCSIS® 3.0 feature set for Motorola’s BSR 64000 CMTS/Edge Router.

Mike joined Motorola in 1994. Over the last 13 years, Patrick was a Data Networking System Engineer and was a major contributor to the Downstream Channel Bonding and IP Multicast features of DOCSIS 3.0. Mike was a major contributor to version 1.1 of the Data Over Cable System Interface Specification (DOCSIS), which standardized QoS operation for VoIP cable modems.

Before joining Motorola, Mike served in various engineering and management positions at Digital Equipment Corporation, General Computer Corporation, Cryptall Corporation, Ztel, and Texas Instruments. He has 14 U.S. patents granted to date.

Mike earned his bachelor's of science and master's of science, computer science degree from the Massachusetts Institute of Technology in 1980, as well as a Masters in the Management of Technology from MIT Sloan School in 1992. He can be reached at michael.patrick@motorola.com.

# Enhancing Switched Digital Video for Support of Next-Generation Advertising through Switched Digital Programming

By Joe Matarese, SVP, Advanced Global Technology
C-COR, Inc.

*Abstract*

*Several years into the development, and a little over a year into the deployment of Switched Digital Video (SDV) technology, is it possible that the cable industry continues to overlook much of SDV's promise? This paper considers a specific aspect of SDV where cable providers have the ability to make significant advances in delivering value to programmers and advertisers, as well as the opportunity to monetize such value. We define this aspect as Switched Digital Programming (SDP).*

*SDP is characterized by a focus on optimal design of switched digital video systems to deliver programming, not simply channels, in the most bandwidth efficient manner. Existing models of SDV typically consider channel popularity in determining the amount of narrowcast bandwidth that one must dedicate to delivery of the SDV channels. While reliance on channel popularity information is useful in obtaining a rough estimate of SDV sizing, to get a more accurate reading one must consider the actual programs. For example, while Network A might have a general popularity rating, averaged over the week, of six points audience share, it is clear that some programs offered by the network will have much higher audience shares while other programs will have lower audience shares. A similar phenomenon is seen with VOD programs. The relative popularity of a specific program, averaged over its entire license window, is often significantly less than the relative popularity of the same program averaged over a shorter period like the prime time viewing window or the new release window.*

*In order for SDV to take advantage of specific program knowledge and for SDP to be deployed, several things must happen:*

*Modeling of SDV narrowcast bandwidth requirements must be redone to make use of program-specific audience measurement data, perhaps from Nielsen Media Research.*

*SDV systems must be instrumented to generate audience measurement data from real time collection of channel change signaling.*

*SDV systems must be enhanced to facilitate explicit program joins and leaves, either by a human viewer or by a DVR, so that program-specific audience measurement does not incorporate artificial "channel loitering".*

*This paper will investigate each of these undertakings in detail and provide statistics from an actual field-trial of SDV in order to build these measurement models.*

*As a consequence of deploying SDP, programmers will benefit in addition to MSOs. Through the MSOs programmers can obtain accurate, "raw" audience measurement data on which to base program lineup decisions. Moreover, programmers have the potential to deliver multiple network variants, possibly targeted at different demographic profiles, rather than deliver a single "one size fits all" network that tries to satisfy all viewers.*

## INTRODUCTION

With new advertising models much on the minds of cable industry strategists, it’s important that strategic planning respecting use of switched digital video technology take into account the great benefits to be derived from SDV for advertising purposes.

Over the past year the case for development of new advertising models has moved to the industry’s front burner amid signs of dissatisfaction from Madison Avenue over ROI performance of television advertising in comparison to advertising on the Internet. Operators recognize cable’s VOD and SDV infrastructure opens new opportunities to extend the power of the TV medium to advertisers at levels of addressability and viewership assurance never seen before in TV advertising.

But to fully exploit the advertising benefits of this interactive infrastructure operators may need to go farther than they have to date to ensure the capabilities are in place that will allow addressable advertising to become a major revenue center for the future. This may add marginal costs to SDV infrastructure costs, but these will be more than offset by the contribution a robust new advertising revenue stream makes to the return on the overall investment in that SDV infrastructure.

Presently, planning surrounding implementation of addressable advertising models in video-on-demand and other digital service applications is pursued as a largely separate endeavor from the planning associated with implementation of switched digital broadcast and unicast capabilities. The latter are primarily tied either to bandwidth savings, in the case of switched broadcast, or to time-shifted delivery of broadcast programming in the case of unicast models. But full realization of the revenue-generating and operations-savings potential of SDV in both modes requires that the goals of next-generation advertising be made a vital component of SDV applications planning.

Several considerations enter into an assessment of what must be done to optimize the SDV environment for enhanced advertising, including:

Adjustments in SDV architecture that allow operators to move away from switching based on channel demand as measured over long time segments to a per-program switching capability based on close tracking of recent individual program performance over short time intervals;

Implementation of data-collecting capabilities to enable this SDP paradigm, which will require generation of audience measurement data from real-time collection of channel-change signaling;

Agreement on a process that facilitates system recognition of explicit program joins and leaves by both viewers and DVRs so that program-specific audience measurement does not record “channel loitering” as actual viewing;

Modeling of advertisement placement and the underlying support architecture around demographic profiles as opposed to geographic zoning.

### Switched Digital Programming

SDP is a method of using switched digital video systems to deliver programming, not simply channels, in a manner that’s optimized for addressable advertising. Implementing switching at this level of granularity gives operators an opportunity to maximize the placement of ads based on the profile of each individual viewer of programs with the lowest viewing levels while making sure the programming with the most appeal to multiple viewers is always in the broadcast tier.

Existing models of SDV typically determine which programming belongs in the SDV channel cluster on the basis of measurements of network popularity over relatively long periods of times. But while reliance on channel popularity information is useful in obtaining a rough estimate of SDV sizing, to get a more accurate reading one must consider the actual programs.

For example, while Network A might have a general popularity rating, averaged over the week, of six points audience share, it is clear that some programs offered by the network will have much higher audience shares while other programs will have lower audience shares. A similar phenomenon is seen with VOD programs. The relative popularity of a specific program, averaged over its entire license window, is often significantly less than the relative popularity of the same program averaged over a shorter period like the primetime viewing window or the new-release window.

A compilation of SDV data from field experiences based on three different criteria for setting switching priorities illustrates the benefits of SDP. These measurements, as shown in Figures 1, 2 and 3, were taken across a service group consisting of 500 tuners.

Each chart reflects how an 85-channel lineup might be delivered digitally to a 500 tuner user group through a combination of broadcast and SDV multicast. In this model ten channels are allocated to broadcast, leaving the remaining 75 networks to be delivered via SDV based on viewer selections. Studies show that at peak hours the percentage of viewers tuned to SDV averages about 14 percent of the viewing population. In the charts shown here the focus is on the number of unique channels being watched via SDV through the course of the day (shown in purple with the highest peak) and the number of unique tuners that are engaged with SDV-delivered programming.

Ideally, in terms of the benefits of using individually targeted advertising, the operator will want the number of SDV programs being watched to sync precisely with the number of tuners that are viewing those programs. When that happens there is just one viewer per SDV-delivered program.

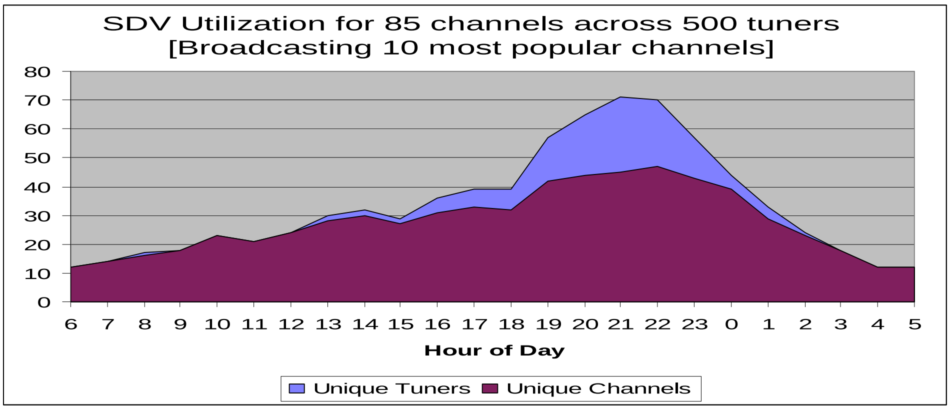


Figure 1 SDV utilization for 85 channels across 500 tuners (broadcasting 10 most popular channels)

Figure 1 assumes the choice of programming to be broadcast over the ten broadcast channels is based on network popularity as reflected in viewership over the course of the entire day. Here, as shown by the blue area of the chart, there are substantially more tuners watching SDV programming than there are programs being delivered in SDV mode, especially at peak hours but also on either side of the peak times. In other words, more than one tuner is tuned to a significant percentage of the SDV programming at these times. In Figure 1 the blue peak in the evening is tallest of the three scenarios, which reflects the fact that there are some popular programs airing in peak hours that are on otherwise less popular networks.

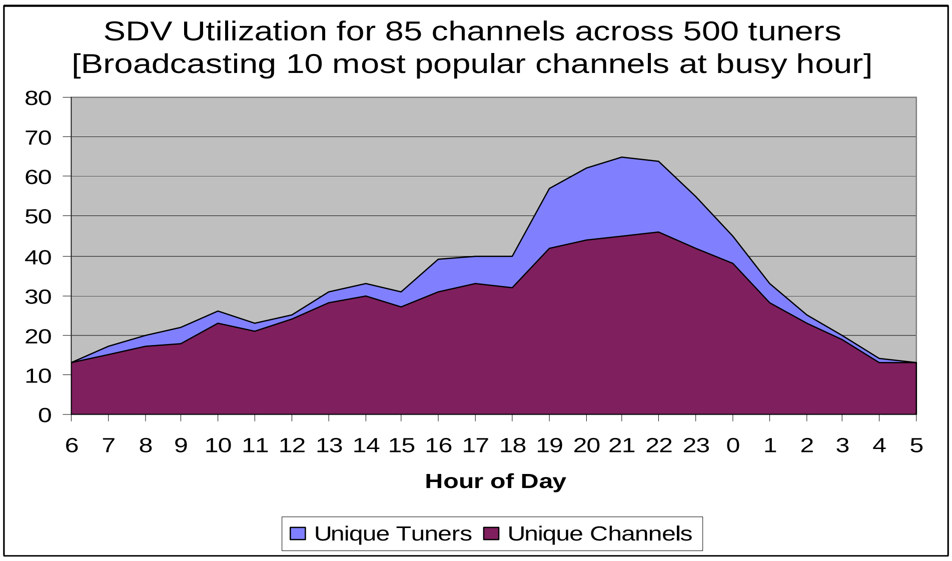


Figure 2 SDV utilization for 85 channels across 500 tuners (broadcasting 10 most popular channels at busy hour)

Figure 2 shows SDV utilization when the ten broadcast channels are chosen based on network popularity measured at the peak busy hour. This lowers the number of multiple tuners per SDV channel at peak but spreads the distribution of multiple SDV turners across the entire day. This pattern results because there are some networks that are very popular over the course of the day but don't necessarily have the most popular programs on air during peak hours.

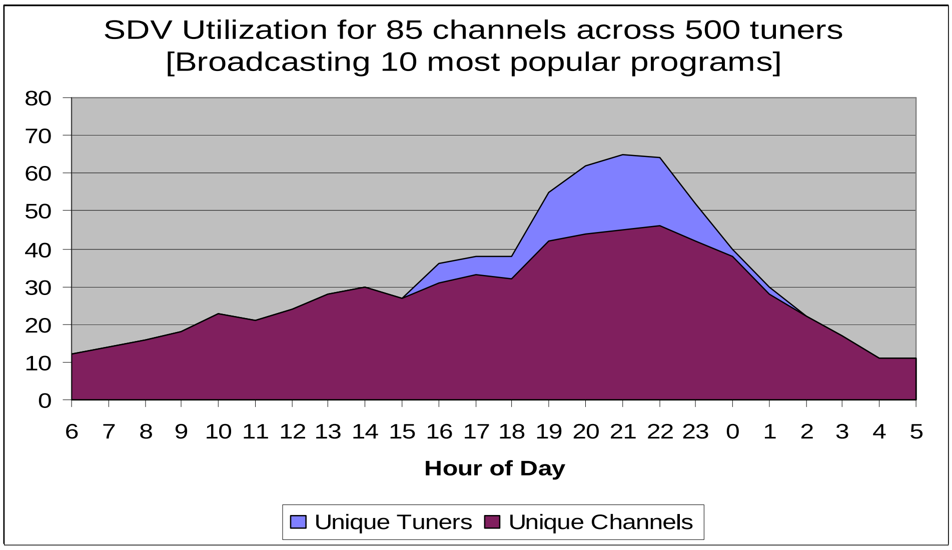


Figure 3 SDV utilization for 85 channels across 500 tuners (broadcasting 10 most popular programs)

Figure 3 shows SDV utilization when the ten most popular programs are broadcast based on program popularity measured at each hour of the day. This results in the most effective use of SDV for advertising purposes.

If the SDV system is agile, always moving the most popular programs at any given time to the broadcast QAM and switching the rest, the blue area is mitigated at peak and throughout the day. This increases the percentage of time that just one viewer is watching each SDV-delivered program. Thus, while SDV multicast is being used in this example, the operator is maximizing the amount of programming that is effectively being delivered as if it were unicast.

SDP and Advertising

As demonstrated by Figure 3, SDP ensures that there's a big chunk of the day where each tuner is on a unique program, which means those unique programs are available for ad insertions based on demographic profiles that provide a more granular focus for advertising than is the case with geographic- or zone-based ad insertion. Niche programming is thus optimized for setting higher CPM (cost per thousand) avails on either side of the two
or three busy hours than would be the case in the other scenarios.

The concept of demographic profile-based advertising can be an important driver behind operators' planning for use of SDV, especially if SDP capabilities are factored in to maximize the addressable advertising opportunities. One illustration of the benefits of using SDV to support an advertising model based on demographic profiling can be found in a large metro interconnect scenario where, today, the niche programming lineup is determined on a zoned geographic basis.

Consider, for example, a hypothetical metro interconnect where there may be 80 zones, each of which is served by 100 niche channels matched to the zone profile. In this scenario, to support addressable advertising just on a zone basis in effect requires the interconnect to be equipped to insert ads individually on what amounts to the equivalent of 8,000 channels.

The use of SDV technology in conjunction with demographically determined ad insertion greatly reduces the insertion scale requirements while creating a much better means of matching ads to viewers, thereby raising the CPM value of such advertising. For example, assume there are eight defined demographic profiles for advertising purposes and that, on average, each of the 80 niche channels (or programs in the case of SDP) on view at any given time might need to be transmitted in four streams to accommodate the demographic profiles of viewers tuned to that programming across the entire interconnect region. (Some programming might have just one demographic viewing profile, while others might have more than four, but this can be assumed to be the average.)

In this case, the number of channels the insertion system must be able to serve goes down drastically. With four channels required per niche channel, the interconnect needs to be able to insert ads into just 400 channels to cover the entire region. Thus, a far more advantageous degree of addressability is achieved in a far more manageable insertion environment.

Improving Accuracy of Viewing Data

Introducing the idea of SDP has important implications for how program viewing is monitored and for the collection of data that can be extremely useful not only to managing bandwidth and determining viewer profiles, but also to fulfilling advertisers' needs for accurate information about ad viewership. To ensure SDP insertion is responsive to the fast-paced changes in audience viewing preferences, the system requires a very agile feedback loop that regularly delivers audience information across all channels.

Currently available data collection and management systems provide operators the means to gather and categorize such information into useful components for operations, ad insertion and advertiser performance analysis applications. But it's essential that the monitoring be applied across broadcast as well as narrowcast QAMs to create the full picture of program performance and ad viewing that a SDP system requires.

Moreover, from an operational perspective, generating reports that look at SDV and broadcast holistically is essential to capacity planning, knowing what switch requirements are, when nodes should be split, etc. With an agile, comprehensive feedback platform, operators will be able to look at programming in 30-minute intervals, take the measure of shifts in popularity and determine which programs to move in and out of the broadcast tier.

The compilation of data for advertisers' purposes as well as for accurate measurement of true viewing patterns on a per-program basis will require that DVR tuners be monitored with an understanding of which recorded programs are actually watched and whether ads are skipped. Existing data management systems allow operators to collect such information and verify the amount of time spent watching recorded programming and the ads.

Supporting SDP also requires a valid measure of when a TV tuned to programming is actually being watched. Such information would also contribute greatly to the advertiser's understanding of program and ad performance. But obtaining this information would require implementation of a mechanism that would require consumers to proactively signal when they are joining or leaving a program.

For example, the user interface could be programmed to put a message on screen after each program asking the viewer to indicate whether the next selected program is going to be watched. If the answer is no or there is no response, the system could switch the screen to a barker channel until the viewer signals a program request.

Does the need for accurate viewing information merit implementing a "lean-forward" requirement that is intrusive to traditional viewing behavior? One can argue that the time has come to recognize the potential of advertising as a major contributor to revenue growth going forward. As competition intensifies downward pressure on subscription prices and as addressable advertising provides cable operators a way to offer much greater value to advertisers, there is reason to expect the balance between ad and subscription revenues will shift toward a greater share on the advertising side.

For example, operators are already considering the possibility that by giving viewers the option to view advertising as an alternative to paying full price for premium on-demand programming they could expand the audience for such programming and thereby derive higher revenues through the CPM rates charged for the ads. The opt-in potential of advertising-support strategies could be applied to the SDP model as well, where consumers who express a willingness to signal their viewing intentions would be given a discount on their subscription price. The percentage of viewers willing to do this would likely be large enough to comprise a valid statistical base for advertisers to use in judging ad performance on a per-program basis.

All of these contributions to an enhanced data-collecting capability for purposes of supporting SDP will provide operators the ability to deliver the improved advertising performance analysis advertisers are looking for. Exploiting these capabilities will require coordination of efforts between technology and advertising departments and a high-level strategic commitment to the potential of next-generation advertising in cable.

Cable operators will need to proactively pursue the support of audience measurement entities such as Nielsen to ensure the records are compiled in a consistent manner endorsed by the advertising industry. Programmers will have to be persuaded that the pain of delivering the "truth" about ad performance, which is to say, valid performance data that accounts for ad skipping and channel loitering, is well worth the benefits to be gained from providing advertisers the ROI they are looking for. Fortunately, this may not be too hard a sell at a moment when the risks to not delivering more accurate performance data are already apparent in the growing share of ad dollars going to the Internet.

## The Advertising-Optimized SDP Architecture

All the functionalities of an advertising-optimized SDP infrastructure are available for implementation by cable operators, should they decide to make advertising revenue acceleration a core driver to SDV and data management planning.

The essential requirements are: traditional silos be eliminated through coordinated management of broadcast and narrowcast QAMs; VOD and SDV resources must be managed from a single platform, and advertising and applications policy management must be applied across all applications.

A unified session resource management system ties all this together. It is responsive to command flows from both the VOD client on the set-top and the program selection process within the electronic programming guide. When the subscriber requests a program, whether from the VOD or the SDV side, the system identifies the specific server or switch interface, instigates the subscriber join controls for SDV multicast, registers the bandwidth requirements for a VOD or SDV session and identifies the path through the access network for delivery to the set-top box.

In instances where core and edge QAMs serve to modulate bit streams onto RF frequencies that have been allocated to a specific server group, the unified bandwidth management system selects the appropriate QAM module for content delivery. And, by tracking bandwidth utilization across the QAM and server arrays, it performs the load balancing essential to preventing bandwidth bottlenecks in the distribution system.

Moreover the unified platform provides the resource management control mechanisms that are essential to executing interactive applications requiring transmission of specialized data content to a particular set-top. For example, in an advertising application where a viewer is requesting transmission of a long-form version of an ad, the SRM sets the bandwidth requirements and duration for the session by coordinating bandwidth allocations through the policy server and appropriate edge QAM to allow the expanded ad to be transmitted to the viewer.

By eliminating the platform silos between advertising, SDV and on-demand content, the unified platform also allows operators to stream content and advertising to any device - mobile, IPTV, PC or television - from a single server. With the switching architecture in place to support SDP, this device-specific, user-specific capability puts operators on course to meet market requirements as they unfold for years to come.

## Summary

The importance of SDV to advertising can’t be overstated at this moment of turmoil in the TV advertising domain. Advertisers are looking for ways to overcome the low returns on TV advertising as audiences fragment around niche programming, avoid advertising through use of DVR technology and spend ever more time away from the TV and on the Web. Advertisers not only seek greater efficiency in reaching desired audience segments; they want better information concerning the performance of their ads.

Cable operators can greatly expand the value of SDV for advertising purposes by developing a Switched Digital Programming capability that ties in with a full accounting of viewer behavior across broadcast, SDV, VOD and DVR segments. Demographic profiling can be used rather than simple zone segmentation to allow advertisers to reach the people they want to reach within each program.

Highly accurate accounting of individual programming and ad viewership enabled by a comprehensive data-collection and management system allows operators to provide advertisers an unprecedented level of assurance as to the return on their ad dollars. The network and back-office components are available to support the new advertising paradigm. Operators who make advertising requirements a key consideration in their use of SDV and on-demand infrastructure will be well positioned to make advertising an ever larger contributor to their revenue stream.

# EXPLOITING MICROCELLS ON CABLE PLANT TO LEVERAGE SPECTRUM ASSETS

Stuart Lipoff
IP Action Partners

## *Abstract*

*This paper will describe how cable operators can exploit their outside plant assets to host wireless microcells to do more than compensate for their smaller spectrum assets relative to mainstream cellcos who employ macrocell architectures. By means of computations and modeling, it is demonstrated that exploiting microcells can not only provide competitive parity to mainstream cellcos with more spectrum, but in fact a hybrid macrocell/microcell architecture with only 20MHz of licensed spectrum can actually enhance the ability of the wireless operator to offer leaf-frog services that incumbent cellcos will not be able to duplicate without adding more much more spectrum.*

## INTRODUCTION

Following the FCC's AWS auction the SpectrumCo consortium of MSOs now owns approximately 20MHz of spectrum with nearly ubiquitous national coverage. The 137 licenses acquired cover 260.5 million pops at a cost of $2.4 billion (equates to $0.46/MHz-pop).

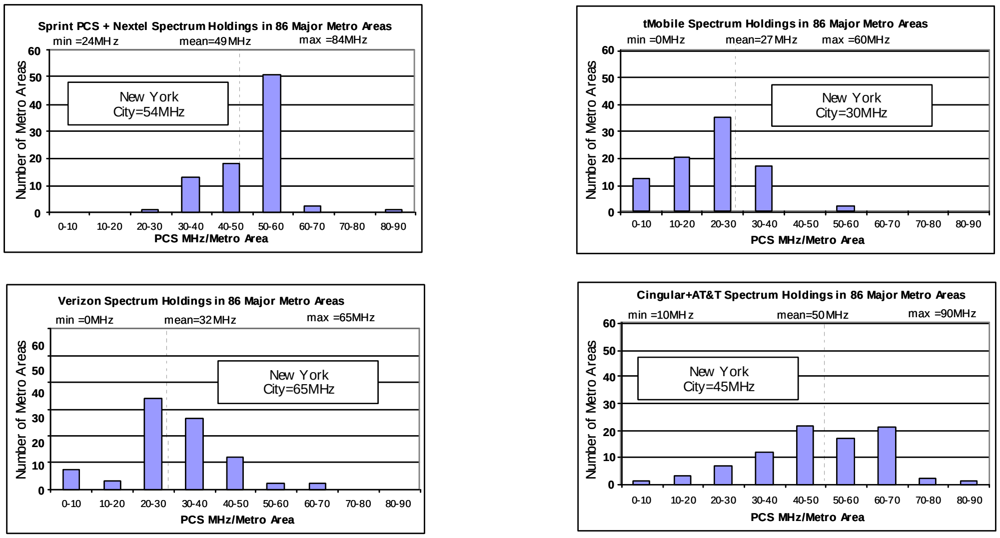
This spectrum is situated between the existing 800MHz and 2GHz incumbent cellular bands and is prime "real estate" to launch wireless services that could be employed to enable a "quad-play" service offering that would be competitive with incumbent cellular operators (cellcos).

A competitive wireless service for cable is not just an upside opportunity, it is a survival imperative as Table 1 demonstrates.

**Table 1 Strategic Imperative for Cable**

| Issue | Implication |
|---|---|
| Verizon and AT&T already lead in wireline and wireless market share, have significant high speed internet adoption, and are moving rapidly to add IPTV to their mix | To defend against churn cable must be able to offer at least a competitive quad play bundle |
| The mainstream cellcos have covered their historical fixed CapEx investment in wireless plant and are largely investing success capital | For cable to build a me-too greenfields wireless network with competitive coverage and quality of service will require massive fixed capital investment with long payback periods |
| Wireless voice adds already exceed wireline adds and over time, significant migration from wireline to wireless can be expected | Although cable is realizing impressive growth and revenue from wireline services, unless the industry has the ability to offer wireless voice it will lose out in the future migration |
| Cellular is so competitive that voice margins are depressed so the mainstream carriers are attempting to develop high value added non-voice services to boast ARPU | Cable has control of, and experience in formulating, highly valued multimedia assets that could allow cable to leap-frog over simple me-too voice service offerings |

**Figure 1 Histograms of Spectrum Holdings of Mainstream Cellcos**



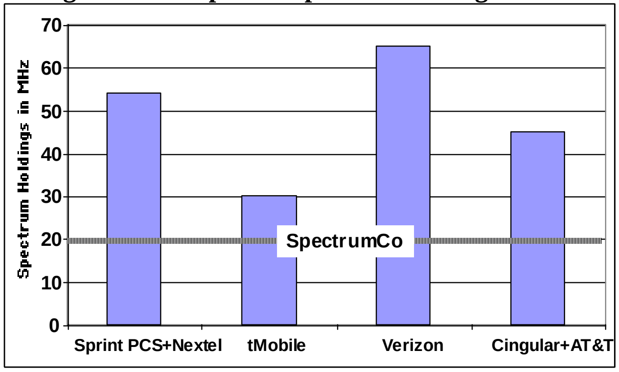
Source: computed from RCR News database prior to AWS auctions

## Spectrum Required

The mainstream cellcos typically hold on the order of 40MHz of nationwide spectrum nearly twice the spectrum won by cable in the AWS auctions. As shown in Figure 1.

Furthermore, as Figure 2 shows, to support the very high population density of New York City (NYC) and their >20% share of market, Verizon is holding 65MHz of spectrum in NYC.

**Figure 2 Example Competitive Holdings in NYC**



In fact as Figure 3 shows, based on today's rapidly growing voice+data traffic load, more than 20MHz of spectrum is required in the top 20 USA cities (as ranked in order of highest population density).

**Figure 3 Spectrum Required to Support 2006 Levels of Cellular Traffic**

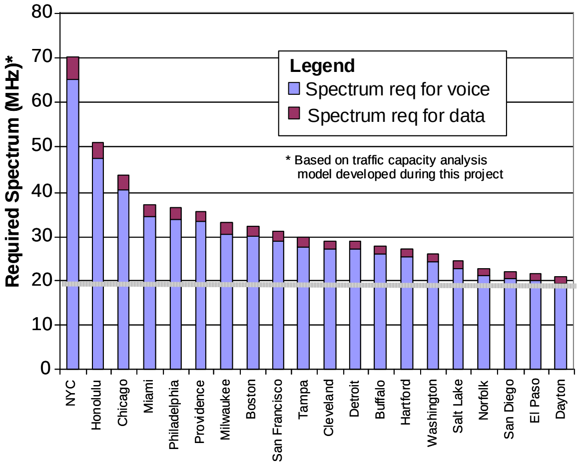


Figure 3 is based on a model to be further described in this paper under the following set of key assumptions:

- CDMA2000 all IP EVDO-Rev A air interface
- A 25% market share typical of Cingular and Verizon market share leaders

- Traffic load of
  - 142 minutes of use per month of voice traffic engineered for 2% grade of service blocking
  - Additional data traffic based on 10% adoption of data by voice subs who consume on the average of 25MBytes/month of downstream data
- Macrocell architecture with the minimum radius of macrocell of 0.5 mile

Figure 3 is actually a best case for incumbent carriers since they need to *waste* some spectrum to support legacy, less spectral efficient, air interfaces and manage dual use while they transition to new 3G and 4G technologies.

Because MSOs could build a greenfields wireless network that fully exploits 3G+ air interfaces that are more spectrally efficient than legacy cellco networks, it would be possible to build out a fully competitive voice-centric cellular network with comparable voice capacity to cellcos even with one half the spectrum. However, it would be desirable to do more than construct a "me-too" wireless voice service offering. In the short term, a "me-too" service offering would make the acquisition of subscribers a slow process that is a highly undesirable position to be in as a late entry competitor in the marketplace. In the longer term, cellco competitors would be likely to use their additional spectrum to launch non-voice services that would make the spectrum starved MSO's cellular network non-competitive.

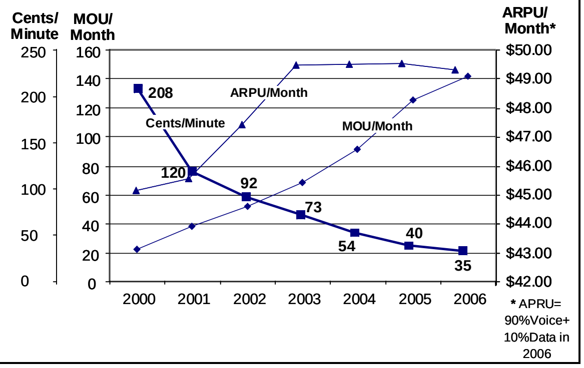
For both short and long term reasons, it would be highly desirable for the cable AWS spectrum owners to leap-frog beyond a me-too voice centric wireless offering and instead offer a multimedia rich voice+non-voice service from the start. However it would also be highly desirable to build this leap-frog network without having to engage in the expensive and uncertain process of acquiring additional spectrum much beyond the 20MHz already in hand.

Market Reality

Formulating a plan that optimizes the cable industry assets requires an understanding of the marketplace realities. In particular what must be understood are the underlying economics and the implications for the current and future traffic load.

As Figure 4 shows, the average revenue per user (ARPU) is flat to slight declining while minutes of use continue to grow as a rapid rate. Data traffic today is minimal since only it accounts for only 10% of ARPU and further because most of the data revenue is from very low bandwidth messaging services.

**Figure 4 Historical Trends in Cellular Voice Traffic Load and Revenue**



The net result is that the ratio of revenue to traffic is rapidly declining. This trend toward a commodity status for voice demonstrates the need to find new non-voice services that will generate high margin ARPU.

**Table 2 Revenue and Bandwidth of Over-the-Air Unicast Non-Voice Wireless Services**

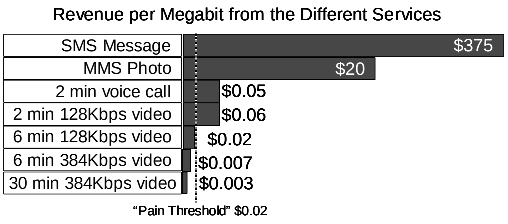
| Service | Bandwidth | Average Revenue | Capacity Usage Mb/s |
|---|---|---|---|
| SMS Message | | $0.07 | 0.0002 |
| MMS Photo | | $0.20 | 0.0098 |
| 2 minute voice call | 12 Kbps | $0.07 | 1.4400 |
| 2 minute low-resolution video | 128 Kbps | $0.99 | 15.3600 |
| 6 minute low-resolution video | 128 Kbps | $0.99 | 46.0800 |
| 6 minute high-resolution video | 384 Kbps | $0.99 | 138.2400 |
| 30 minute high-resolution video | 384 Kbps | $1.99 | 691.2000 |

Source: “The Economics of Mobile Broadcast TV”, Yoram Solomon, Mobile DTV Alliance

The challenge however is that although subscribers will pay substantial fees for highly valued multimedia services, these highly valued services (e.g. video) generate very large traffic loads. Examples of non-voice traffic are shown in Table 2.

These traffic loads are so intense that that their $/megaByte can not be economically supported on today's macrocellular 3G cellular networks at economic levels of CapEx investment in infrastructure as shown in Figure 5.

**Figure 5 Revenue Per Megabit from Example Non-Voice Wireless Services**



Source: "The Economics of Mobile Broadcast TV", Yoram Solomon, Mobile DTV Alliance

## A NATURAL STRATEGY FOR CABLE

Assuming greenfield wireless build-out for cable, a leap-frog wireless strategy is called for given cable's late start and the need to accelerate subscriber take-up to cover the large fixed CapEx investment needed to build a network with competitive coverage. Greenfields spectrum can be built by cable using most advanced spectral efficient technology to gain on the order of a 2X capacity advantage over incumbents who carry legacy technology baggage; however, incumbents will soon catch up as they rebuild to next generation air interfaces.

Cable's natural strategy is to employ multimedia highly valued video and other entertainment assets integrated with wireline and high speed data products to create highly differentiated services and counter declining ARPU/minute trends. Although next generation air interfaces can provide some temporary capacity advantage, it will not be possible to execute a multimedia bandwidth hungry leap-frog strategy employing conventional cellular architectures without spectrum well in excess of that held by competitors. The ideal situation for cable is to execute a strategy that not only builds on content assets but also employs a non-conventional architecture that leverages the installed base of cable plant and technology.

Compared to cellcos, the cable industry has unique access to a broadband low cost wired backbone network that can be leveraged to provide economically attractive high bandwidth multimedia services. Furthermore, we will show that even with 40-60MHz of spectrum, incumbent cellcos can not economically scale their macrocellular networks to support the capacity needed to satisfy consumer's appetite for high bandwidth multimedia services. On the other hand, the cable operators can exploit their extensive HFC network to economically support a microcell underlay network* that can more than compensate for the cable operators limited 20MHz of spectrum.

Not only will we demonstrate the ability for such a microcellular based underlay network to offer a non-voice traffic capacity advantage, we will also note that a hybrid macrocellular network combined with a microcellular underlay will allow additional degrees of freedom in the overall architecture and the associated service offerings that further leverage cable industry multimedia interests and assets.

### Multimedia Non-Voice Traffic Loads

Today's cellcos offer data plan pricing that top out at about 50MB/month, but even modest multimedia traffic in excess of webpage surfing breaks this 50MB budget and places a non-economic load on their >40MHz of macrocellular infrastructure. Some examples of multimedia traffic are shown in Table 3. Note that all the examples

---

* While microcells on aerial plant are straightforward to add, in congested business districts with underground plant the costs will be higher.

beyond webpage downloads, break the 50MB/month threshold.

**Table 3 Illustrative Multimedia Data Traffic Load**

| Media Event | MegaBytes (MB)/Event | Events/Day | 20% Busy Hour kb/s | MB/Month (22 days) |
|---|---|---|---|---|
| 6 min 384 kb/s video QVGA (320x240 pixels) at 15 f/s | 17.28 | 1 | 7.68 | 380 |
| MP3 Song Download | 5 | 5 | 11.11 | 550 |
| eMailFile Attachment Downloads | 0.3 | 20 | 2.67 | 132 |
| WebPage Downloads | 0.05 | 50 | 1.11 | 55 |

Since building a bottom up forecast for consumer's appetite for non-voice data consumption is so dependent on forecasting the adoption of specific services, it is highly speculative to rely upon a bottom up buildup of traffic. A top down approach looking at consumer's appetite for fixed wired data consumption on high speed cablemodem or DSL residential services can be viewed as a high end upper bound for wireless data usage.

Table 4 is based upon today's average data consumption by cablemodem subscribers. It shows that the current fixed wired downstream consumption of 1,209 Mbytes/sub/month is 240 to 24 times the 5-50 Mbytes/sub/month consumed by today's wireless data subscribers. Furthermore if one considers that today's wireless data ARPU is only 10% of today voice+data ARPU, if data were adopted by a larger percentage of voice subscribers well in excess of what 10% ARPU implies; the total data traffic load on the wireless system could well grow on the order of 10 times additional.

**Table 4 Average Data Consumption by Fixed Wireline Cablemodem Residential Subscribers**

| | |
|---|---|
| Average Wired BB Downstream Busy Hour Data Rate/Sub (kb/s) | 25 |
| Busy Hour Downstream Data Consumption/Sub (Mbytes) | 11 |
| Busy Hour Concentration Factor (percent of all daily usage) | 20% |
| Monthly Downstream Data Consumption/Sub/Month (Mbytes) | 1,209 |
| Ratio of Upstream to Downstream Data Rate | 20% |
| Monthly Upstream Data Consumption/Sub/Month (Mbytes) | 242 |

## Study Methodology

Our analysis employed the following steps:

- Employ the central business districts in New York City as a worst case USA example for spectrum needed in high population congestion limited areas.
- Compute the amount of spectrum required to serve today's traffic load for a conventional macro-cellular architecture
- Propose an alternative microcellular architectures for cable
- Re-compute the capacity provided for each microcellular alternative under consideration and compare that to
  - Today's voice+data traffic loads
  - Future possible multimedia intensive data traffic loads

## MACROCELLULAR ANALYSIS

In order to have a baseline to compare with a recommended microcellular approach for cable, we begin by analyzing the capacity of a conventional macrocellular architecture in the example NYC CBD.

### Manhattan Central Business Districts

The amount of spectrum required in any one market area is determined by the traffic capacity of the most density loaded cellsite in that market area. Using the NYC market area as an example, the most density loaded cellsite can be expected to be in either Lower or Midtown Manhattan areas shown in Figure 6 Lower and Midtown Manhattan Central Business Districts.

## Table 5 Macro Cell Analysis Model for CDB in Manhattan

**Manhattan Demographics**

| | | | |
|---|---|---|---|
| 1 | Manhattan land area (miles^2) | 28.4 | input |
| 2 | Nighttime population | 1,487,536 | input |
| 3 | Daytime Population | 3,389,300 | input |
| 4 | Ratio Day/Night | 2.3 | calc |
| 5 | Mid-town Manhattan land area (mile^2) | 1.64 | input |
| 6 | Lower Manhattan land area (mile^2) | 1.72 | input |
| 7 | Nightime PopDensity pops/mile^2 | 52,378 | calc |
| 8 | Nightime Pops in lower Manhattan | 90,090 | calc |
| 9 | Nighttime Pops in midtown Manhattan | 85,900 | calc |
| 10 | Pop Increase btwn Night to Day | 1,901,764 | calc |
| 11 | Pop Increase into lower due to migration into Manhattan | 973,522 | calc |
| 12 | Pop Increase into midtown due to migration into Manhattan | 928,242 | calc |
| 13 | Pop Increase into lower due to 25% migration within Manhattan | 190,369 | calc |
| 14 | Pop Increase into midtown due to 25% migration within Manhattan | 181,515 | calc |
| 15 | Total daytime Pops in lower Manhattan | 1,253,981 | calc |
| 16 | Total daytime Pops in midtown Manhattan | 1,195,657 | calc |
| 17 | Total Daytime PopDensity in lower (mile^2) | 729,059 | calc |
| 18 | Total Daytime PopDensity in midtown (mile^2) | 729,059 | calc |

**CellSite Loading**

| | | | |
|---|---|---|---|
| 19 | CellRadius (miles) | 0.50 | input |
| 20 | CellArea (miles^2) | 0.79 | calc |
| 21 | Daytime Pops in lower Manhattan cellsite | 572,602 | calc |
| 22 | Cellular Adoption | 72% | Input |
| 23 | CellSubs in lower Manhattan cellsite | 412,273 | calc |
| 24 | Market Share | 10% | input |
| 25 | Own subs in lower Manhattan cellsite | 41,227 | calc |

**Voice Traffic Load per Subscriber**

| | | | |
|---|---|---|---|
| 26 | Voice Minutes of Use/year/Sub | 1,700 | input |
| 27 | Minutes of Use/work-day/Sub | 6.54 | calc |
| 28 | Busy hour concentration of minutes | 20% | input |
| 29 | Busy Hour Minutes/Sub during work day | 1.31 | calc |
| 30 | Busy Hour voice traffic load/Sub (mE) | 21.8 | calc |

**Data Traffic Load per Subscriber**

| | | | |
|---|---|---|---|
| 31 | Average Wired BB Downstream Busy Hour Data Rate/Sub (kb/s) | 25 | input |
| 32 | Ratio of Upstream to Downstream Data Rate | 20% | input |
| 33 | Throttle on Wireless Data Rate Ratio to Wired Rate | 4.13% | input |
| 34 | Busy Hour Concentration Factor (percent of all daily usage) | 20% | input |
| 35 | Percent of Subs who Take Data Service | 10% | calc |
| 36 | Equivalent downstream megabytes/sub/month of Data Usage | 50.0 | calc |

**Data Only CellSite Capacity for CDMA2000 EVDO Rev A all IP**

| | | | |
|---|---|---|---|
| 37 | Downstream throughput for 3Sector Site (Mb/s) | 2.7 | input |
| 38 | Upstream throughput for 3Sector Site (Mb/s) | 1.35 | input |
| 39 | Downstream Capacity/sub requirement for data traffic (kb/s) | 1.03 | calc |
| 40 | Upstream Capacity/sub requirement for data traffic (kb/s) | 0.21 | calc |
| 41 | Downstream Capacity/site requirement for data traffic (Mb/s) | 4 | calc |
| 42 | Upstream Capacity/site requirement for data traffic (Mb/s) | 1 | calc |
| 43 | Number of Channels Required for Downstream | 2 | calc |
| 44 | Number of Channels Required for Upstream | 1 | calc |
| 45 | Worst Case Number of Channels Required | 2 | calc |

**Voice Only CellSite Capacity for CDMA2000 EVDO Rev A all IP**

| | | | |
|---|---|---|---|
| 46 | Downstream throughput for 3Sector Site (Mb/s) | 2.7 | input |
| 47 | Upstream throughput for 3Sector Site (Mb/s) | 1.35 | input |
| 48 | Capacity requirement for voice traffic (kb/s) | 10 | input |
| 49 | Voice Circuits Available (upstream limited) | 135 | calc |
| 50 | Traffic Capacity/3Sector Site in Erlangs @2% Blocking | 87.97 | calc |
| 51 | Number of Voice Subs Supported/Site/Channel | 4,036 | calc |

**Computation of Spectrum Required**

| | | | |
|---|---|---|---|
| 52 | Number of 1.25MHz channel pairs required for Voice | 10 | calc |
| 53 | Required Spectrum for Voice (MHz) | 25 | calc |
| 54 | Number of 1.25MHz channel pairs required for Data | 2 | calc |
| 55 | Required Spectrum for Data (MHz) | 5 | calc |
| 56 | Total Spectrum Required for Both Data+Voice (MHz) | 30 | calc |
| 57 | Ratio of Data Channels/Total Channels | 17% | calc |

**Figure 6 Lower and Midtown Manhattan Central Business Districts**

Lines 1 through 18 of the Table 5 model shows the process of arriving at a population density of 729,059 pops/mi$^2$ which is the same in both Lower and Midtown districts.

The approach presented in the table for estimating the worst case NYC population density in the central business districts (CBD)is based upon:

- The base population in the CBD is the nighttime population (lines 7&8)

- One component of increase above the nighttime base population is migration into Manhattan from elsewhere. The computations in the table assume this migration is divided between lower and midtown based on their respective land areas (lines 11&12)

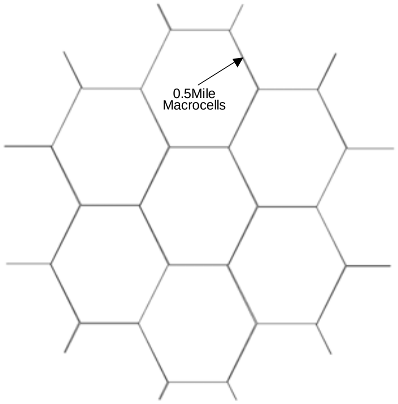
- A second component of increase in the CBD is based upon an estimated 25% migration from other parts of Manhattan into the CBD (lines 13&14)

- The resulting work week daytime population density in the CBD is computed and is the same for both lower and midtown districts (lines 17&18)

## CellSite Loading

Although the standards for modern 3G air interfaces support smaller cell sizes, we assume for this analysis the smallest economic macrocellular architecture is built using 0.5 mile radius 3 sector macrocells as shown in Figure 7.

**Figure 7 Conventional Macrocellular Architecture**



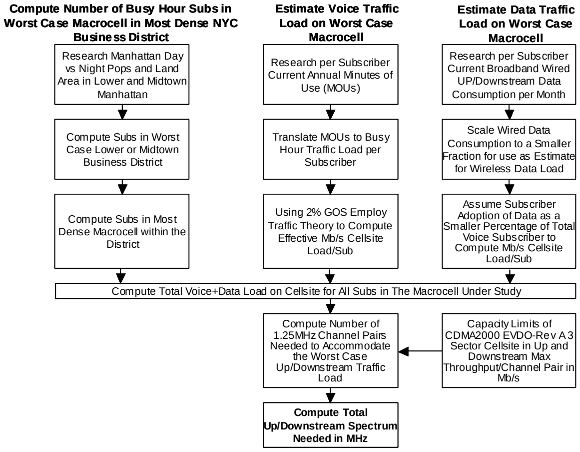
Employing the CDB population density computed above, the number of subscribers in the worst case 0.5 mile cellsite is computed in lines 19 through 25 of the Table 5 model.

The cellsite area in line 20 is multiplied by the CBD population density to arrive at line 21 pops in the macrocell. Using the CTIA 72% figure for cellular adoption results in the cellular subscribers in the macrocell in line 23. On line 24 the market share for the cable company is assumed to be 10% comparable to that enjoyed by t-Mobile but much less than the 25% market share enjoyed by Verizon and AT&T/Cingular. This results in a line 25 estimate of 41,227 of your own subscribers generating traffic in your worst case macrocell.

## Model Flow Chart

The total spectrum required to support today's voice+modest data load is computed to be 30MHz in line 56 of the Table 5 model for a 10% market share carrier but would need to be on the order of 75MHz for the mature carriers who enjoy 25% market share. The process for computing the required spectrum is shown in the Figure 8 flow chart.

**Figure 8 Flow Chart for Macrocellular Analysis**



The analysis assumes the following key inputs to the model:

- Each macrocell employs CDMA2000 EVDO-Rev A all IP technology
- The traffic load is typical of today's cellular system averages of
  - Voice traffic of 142 MOUs/month
  - Data traffic based on 10% fraction of voice subscribers who consume 50 megabytes/month of downstream data
- The market share of the carrier under study is 10%

### Voice Traffic Load

Lines 26 through 30 of the Table 5 model compute the voice traffic load. Starting with the CTIA reported average 1,700 minutes of use per year per subscriber (MOUs) of voice traffic, the number of MOUs per work day is computed on line 27 and based upon an assumption of 20% of all daily minutes being used in the busy hour, the busy hour MOUs is computed on line 29 and shown in mili-Erlangs per subscriber during the busy hour of 21.8 mE on line 30.

### Data Traffic Load

Lines 31 through 36 of the Table 5 model are employed to compute the busy hour non-voice data traffic load. Line 31 represents today's average daily downstream consumption of today's wired cable modem service as reported by a cable operator. Since modern air interfaces generally support higher downstream (i.e. base to mobile) throughput then upstream (i.e. mobile to base), the upstream subscriber data consumption is computed based on a ratio of upstream to downstream of 20%.

Line 33 of the model represents a factor to scale the wireless rate to a 4.13% fraction of the wired rate. The 4.13% number was computed in order to generate a typical data consumption of 50 megabytes per data subscriber per month (MB/mo) as shown on line 36. Further more, line 34 estimates the busy hour usage by assuming that 20% of all daily usage occurs during the busy hour. Finally line 35 assumes that data subscribers represent only 10% of total voice subscribers (i.e. data consumption averaged across all voice subscribers is 5 MB/mo).

### Data Only Cellsite Capacity

Lines 37 though 45 of Table 5 compute the number of channels and spectrum required to support the upstream and downstream data only traffic loads on the CBD cellsite during the busy hour. Line 37 represents the downstream throughput per site for a three sector EVDO-Rev A macrocell while line 38 represents the upstream throughput limit for the site. These throughputs are based on a private communications to me of the results of a simulation conducted by Nortel of various 3G and 4G air interfaces under mobility conditions.

The total traffic load per subscriber is computed by using lines 31 through 36 of the model as input to compute line 39 and line 40 of the model. By using the number of data subscribers the total down and upstream load on the site is computed on lines 41 and 42. Dividing line 31 by line 37 for downstream and line 42 by line 38 for upstream yields the number of 1.25MHz channel pairs needed to support the non-voice data traffic load on lines 43 and 44 and the worst case of down and upstream resulting need for channel pairs is shown on line 45.

### Voice Only Cellsite Capacity

Lines 46 through line 51 of the Table 5 model compute the maximum number of voice subscribers that a cellsite can support in order to maintain a 2% grade of service blocking factor. The throughput on line 47 and the assumption that a header suppressed VoIP channel will require 10 kilobits per second on the channel computes to a line 49 result of 135 voice circuits supported on the site. The Erlang B model is employed to compute the maximum traffic capacity of the site in Erlangs and by dividing the voice load in mE per subscriber from line 30, the line 51 result of a maximum load of 4,036 voice subscribers per site is computed.

Computation of Spectrum Required

The data only capacity from line 45 and the subscriber voice limit from line 51 is employed in lines 52 to 57 of the Table 5 model to compute the line 56 result of 30MHz of total spectrum required.

On line 52 the number of channel pairs for voice is computed by simple division of line 25 voice subs in the site by the capacity per channel on line 51.

On line 54 the number of channel pairs for data is repeated from line 45. By multiplication of the channel pairs needed on line 53 for voice and line 54 for data by 2.5MHz of spectrum per channel pair the line 53 resulting spectrum need is computed for voice and the line 55 corresponding spectrum need for data is computed. Line 53 and line 54 are added together to result in the need for 30MHz of spectrum shown on line 56 of the Table 5 model.

As a perspective on the relative drivers of bandwidth, the ratio between data and total channels is computed as 17% on line 57.

## MICOCELLULAR ANALYSIS WITH LICENSED SPECTRUM

Unlike competitive carriers, cable operators have extensive HFC plant throughout their franchise area that can be used to provide economic backbone interconnection between microcells with coverage range much less than 0.5 mile radius. For the microcell case, it is assumed that the macrocell network continues to exist to support high mobility applications as well as allow for less than 100% coverage by a microcell network that underlies the macrocells. Two microcell underlay options were considered in this study: 1) Borrowing licensed spectrum from the macrocell pool of spectrum to use in microcells and 2) Using unlicensed WiFi spectrum for the microcells. In this section we analyze the licensed spectrum case.

Examination of the technical specification for CDMA2000 handsets reveals that compliant devices have the ability to backoff their transmitter power to support operation cell sizes much smaller than the 0.5 mile macrocells assumed in the previous analysis. In Table 6 we compute the range to be limited to about 350 feet under the case of maximum power backoff. Allowing for some overlap at the edges, this supports microcell sizes as small as a 300 feet coverage radius.

**Table 6 Coverage Range Computation of CDMA2000 Compliant Cellular Handsets**

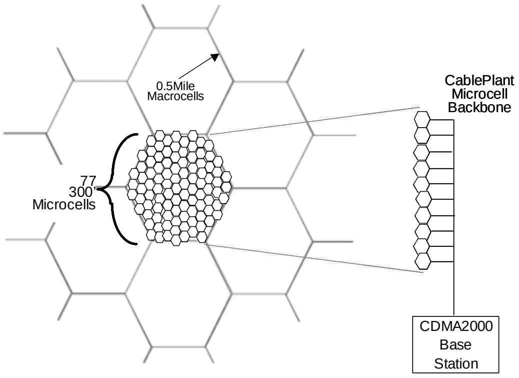
| | |
|---|---|
| **Receiver Sensitivity (dBmW)** | -104 |
| **Transmitter Minimum Backoff Power (dBmW)** | -25 |
| **Link Budget (dB)** | 79 |
| **Free Space Propagation Range @ 2000MHz (feet)** | 350 |

Free Space LinkBudget(dB)=36.56+20Log$_{10}$(Frequency in MHz)+20Log$_{10}$(Distance in Miles)

**Source:** "Recommended Minimum Performance Standards for CDMA2000 Spread Spectrum Mobile Stations", Release B, Version 1, 3GPP2C.S0011-B, December 13, 2002

With reference to Figure 9, it can be seen that a 0.5mile macrocell can be filled with 77 microcells of 300 foot microcells and using the ability of CDMA2000 to allow 1:1 frequency reuse, the same spectrum can be used in each cellsite. Assuming the limited space to mount microcells and the need to keep the cost low, it is assumed that an omnidirectional antenna is used (i.e. single sector) and that a maximum of one channel pair per microcell is supported. The figure also shows the application of cable plant as a backbone for connection of the cells. A variety of backbone technologies are possible including DOCSIS to stand alone microcellular base stations as well as radio over fiber to connect a single base station to remote antenna drivers within each microcell.

**Figure 9 Microcell Architecture**

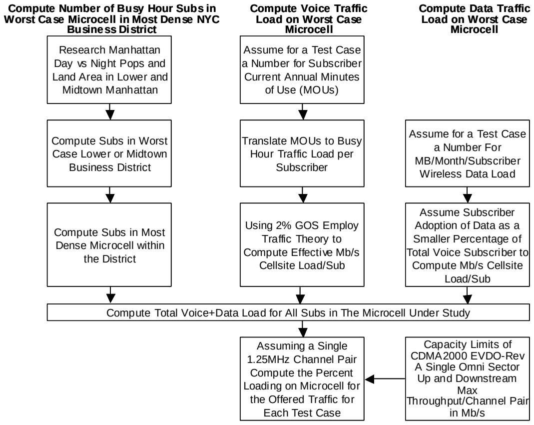


Model Flow Chart

Unlike the approach employed for computing the spectrum required per macrocell based on traffic, in this licensed microcellular option under study it is assumed there is only one channel pair allocated to each microcell. The analysis approach focus is upon understanding the capacity of such an architecture under several traffic load cases and then to compare the traffic capacity of this microcellular+macrocellular hybrid to the previous macrocellular only case.

The process for computing the capacity of the microcellular architecture is shown in the Figure 10 flow chart.

**Figure 10 Flow Chart for Microcellular Analysis**



The analysis considers several cases in which the inputs to the model are varied:

- Each microcell employs CDMA2000 EVDO-Rev A all IP technology with single channel pair and an omni-sector antenna
- The market share of the carrier under study is 10%
- The traffic loads studied consist of two cases
- Today's voice with multimedia data traffic max'ed out to use full capacity of the microcell
  - Voice of 142 MOUs
  - Each data sub consumes 653MB/month downstream traffic
  - Data subs are 10% of voice subs
- No voice traffic with data capacity max'ed out
  - All voice traffic on the macrocell and none on the microcells

  - Each data sub consumes 816MB/month downstream traffic

  - Data subs are 10% of voice subs

capacity to handle today's voice plus non-voice data load. So the analysis objective in this case is not to compute how much spectrum is needed, but instead to compute how much more than today's traffic can the microcell support.

**Table 7 Licensed Spectrum Micro Cell Analysis Model for CDB in Manhattan**

**Manhattan Demographics**

| | | | |
|---|---|---|---|
| 1 | Manhattan land area (miles^2) | 28.4 | input |
| 2 | Nighttime population | 1,487,536 | input |
| 3 | Daytime Population | 3,389,300 | input |
| 4 | Ratio Day/Night | 2.3 | calc |
| 5 | Mid-town Manhattan land area (mile^2) | 1.64 | input |
| 6 | Lower Manhattan land area (mile^2) | 1.72 | input |
| 7 | Nightime PopDensity pops/mile^2 | 52,378 | calc |
| 8 | Nightime Pops in lower Manhattan | 90,090 | calc |
| 9 | Nighttime Pops in midtown Manhattan | 85,900 | calc |
| 10 | Pop Increase btwn Night to Day | 1,901,764 | calc |
| 11 | Pop Increase into lower due to migration into Manhattan | 973,522 | calc |
| 12 | Pop Increase into midtown due to migration into Manhattan | 928,242 | calc |
| 13 | Pop Increase into lower due to 25% migration within Manhattan | 190,369 | calc |
| 14 | Pop Increase into midtown due to 25% migration within Manhattan | 181,515 | calc |
| 15 | Total daytime Pops in lower Manhattan | 1,253,981 | calc |
| 16 | Total daytime Pops in midtown Manhattan | 1,195,657 | calc |
| 17 | Total Daytime PopDensity in lower (mile^2) | 729,059 | calc |
| 18 | Total Daytime PopDensity in midtown (mile^2) | 729,059 | calc |

**CellSite Loading**

| | | | |
|---|---|---|---|
| 19 | CellRadius (miles) | 0.06 | input |
| 20 | CellArea (miles^2) | 0.01 | calc |
| 21 | Daytime Pops in lower Manhattan cellsite | 7,394 | calc |
| 22 | Cellular Adoption | 72% | Input |
| 23 | CellSubs in lower Manhattan cellsite | 5,324 | calc |
| 24 | Market Share | 10% | input |
| 25 | Own subs in lower Manhattan cellsite | 532 | calc |
| 26 | Voice Minutes of Use/year/Sub | 1,700 | input |
| 27 | Minutes of Use/work-day/Sub | 6.54 | calc |
| 28 | Busy hour concentration of minutes | 20% | input |
| 29 | Busy Hour Minutes/Sub during work day | 1.31 | calc |
| 30 | Busy Hour voice traffic load/Sub (mE) | 21.8 | calc |
| 31 | Voice Load on Site from own subs (Erlangs) | 11.6 | calc |
| 32 | Number of Voice Ckts to be Reserved for 2% QoS Blocking | 18 | calc |
| 33 | Capacity requirement for voice traffic (kb/s) | 10 | input |
| 34 | Upstream & Downstream load on Site from Voice Ckts in Mb/s | 0.180 | calc |

**Data Traffic Load per Subscriber**

| | | | |
|---|---|---|---|
| 35 | Average Wired BB Downstream Busy Hour Data Rate/Sub (kb/s) | 25 | input |
| 36 | Ratio of Upstream to Downstream Data Rate | 20% | input |
| 37 | Throttle on Wireless Data Rate Ratio to Wired Rate | 54.00% | input |
| 38 | Busy Hour Concentration Factor (percent of all daily usage) | 20% | input |
| 39 | Percent of Subs who Take Data Service | 10% | input |
| 40 | Equivalent downstream megabytes/sub/month of Data Usage | 653.1 | calc |

**MicroCellSite Capacity for Single Channel CDMA2000 EVDO Rev A all IP**

| | | | |
|---|---|---|---|
| 41 | Downstream throughput for Omni Single Sector Site (Mb/s) | 0.9 | input |
| 42 | Upstream throughput for Omni Single Sector Site (Mb/s) | 0.45 | input |
| 43 | Downstream Site throughput left after voice reserve (Mb/s) | 0.720 | calc |
| 44 | Upstream Site throughput left after voice reserve (Mb/s) | 0.270 | calc |
| 45 | Downstream Capacity/sub requirement for data traffic (kb/s) | 13.50 | calc |
| 46 | Upstream Capacity/sub requirement for data traffic (kb/s) | 2.70 | calc |
| 47 | Total Downstream data traffic Load/Site (Mb/s) | 0.719 | calc |
| 48 | Total Upstream data traffic Load/Site (Mb/s) | 0.144 | calc |
| 49 | Total Downstream Data+Voice Traffic Load on Site (Mb/s) | 0.899 | calc |
| 50 | Total Upstream Data+Voice Traffic Load on Site (Mb/s) | 0.324 | calc |

**Ratio Analysis**

| | | | |
|---|---|---|---|
| 51 | Ratio of DS Voice Load/Site to Total Single Channel Site Capacity | 20% | calc |
| 52 | Ratio of DS Data Load/Channel to Total Single Channel Site Capacity | 80% | calc |
| 53 | Ratio of DS Data Load/Channel to Max Possible Data Load | 100% | calc |
| 54 | Ratio of US Voice Load/Site to Total Single Channel Site Capacity | 40% | calc |
| 55 | Ratio of US Data Load/Channel to Total Single Channel Site Capacity | 32% | calc |
| 56 | Ratio of US Data Load/Channel to Max Possible Data Load | 53% | calc |

## Licensed Microcellular Capacity Computation

In a matter parallel to the macrocellular analysis model, the model of Table 7 was created to compute the capacity of the licensed microcellular alternative based upon a single CDMA2000 channel pair in a 300 foot omni-sector microcell using an EVDO-Rev A air interface.

The Table 7 model tracks the Table 5 macrocell model from lines 1 through line 30. The analysis departs from the macrocell case beyond line 30, because even this single channel pair microcell has more than enough capacity to handle today's voice plus non-voice data load. So the analysis objective in this case is not to compute how much spectrum is needed, but instead to compute how much more than today's traffic can the microcell support.

Starting at line 31 in the Table 7 model the Erlang voice load on the microcell is computed and using Erlang theory the number of voice circuits that must be reserved to support a 2% grade of service blocking is computed on line 32. Using the voice data rate on line 33, the load on the microcell in Mb/s needed to support today's voice traffic is computed on line 34.

Lines 35 through 40 in the Table 7 model directly parallel the computation in macrocell model. The input parameters for the data consumption are adjusted upward so that the capacity of the microcell is max'ed out. Line 37 of the model is set so that line 52 of the model just reaches 80% of site capacity which when added to the 20% needed for voice results in 100% of the site capacity consumed by the offered traffic.

The finding from this analysis is shown on line 40 of the Table 7 model indicating that a licensed microcellular network can support all of today's voice traffic as well as allowing non-voice data usage to grow from the 50MB/month/data-sub to 653MB/month/data-sub. What is more remarkable is that unlike the macrocellular case which required 30MHz of spectrum to support the much smaller non-voice data load, with the microcellular architecture only a single channel pair of 2.5MHz (i.e. 2 x 1.25MHz) is needed.

The licensed microcell model of Table 7 was also run with the voice traffic set to zero such as would be the case if all voice traffic was supported on the macrocell overlay and the microcellular network employed only for data. Under these conditions, the model computes a non-voice data only capacity that supports each data sub consuming 816MB/month of downstream non-voice data traffic.

## MICOCELLULAR ANALYSIS WITH UNLICENSED SPECTRUM

This alternative differs from the just licensed microcell analysis, just above, in that unlicensed WiFi spectrum is employed in the microcells in an architecture similar to today's muni-WiFi systems. As with the previous licensed microcell analysis, the macrocell network continues to overlay the unlicensed microcells to support high mobility and fill any microcell coverage gaps.

Examination of the performance of WiFi network access points in Table 8 shows that such outdoor unlicensed microcells have 300 foot radius of coverage comparable to the same coverage radius assumed for the licensed microcellular option considered above.

**Table 8 Typical Performance of WiFi Network Access Points**

| Protocol | Release Date | Operating Frequency | Data Rate (Typ) | Data Rate (Max) | Range (Indoor) | Range (Outdoor) |
|---|---|---|---|---|---|---|
| 802.11a | 1999 | 5.15-5.35 / 5.47-5.725 / 5.725-5.875 GHz | 25 Mbit/s | 54 Mbit/s | ~25 meters | ~75 meters |
| 802.11b | 1999 | 2.4-2.5 GHz | 6.5 Mbit/s | 11 Mbit/s | ~35 meters | ~100 meters |
| 802.11g | 2003 | 2.4-2.5 GHz | 25 Mbit/s | 54 Mbit/s | ~25 meters | ~75 meters |
| 802.11n | 2007 draft | 2.4 GHz or 5 GHz bands | 200 Mbit/s | 540 Mbit/s | ~50 meters | ~125 meters |

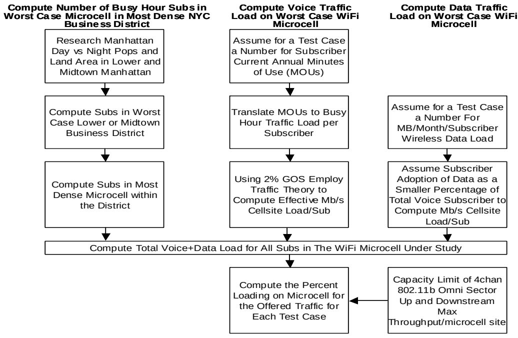
Source: Wikipedia entry for 802.11

Unlike the approach employed for computing the spectrum required per macrocell based on traffic, in this unlicensed microcellular option under study it is assumed there are a fixed number of four WiFi channels allocated to each microcell. As with the licensed microcellular option, this analysis approach focus is upon understanding the capacity of such an architecture under several traffic load cases and then to compare the traffic capacity of this microcellular+macrocellular hybrid to the previous macrocellular only case.

### Model Flow Chart

The process for computing the capacity of the unlicensed microcellular architecture is shown in the Figure 11 flow chart.

**Figure 11 Flow Chart for Unlicensed Microcellular Analysis**



## Table 9 Unlicensed Spectrum Micro Cell Analysis Model for CDB in Manhattan

**Manhattan Demographics**

| | | | |
|---|---|---|---|
| 1 | Manhattan land area (miles^2) | 28.4 | input |
| 2 | Nighttime population | 1,487,536 | input |
| 3 | Daytime Population | 3,389,300 | input |
| 4 | Ratio Day/Night | 2.3 | calc |
| 5 | Mid-town Manhattan land area (mile^2) | 1.64 | input |
| 6 | Lower Manhattan land area (mile^2) | 1.72 | input |
| 7 | Nightime PopDensity pops/mile^2 | 52,378 | calc |
| 8 | Nightime Pops in lower Manhattan | 90,090 | calc |
| 9 | Nighttime Pops in midtown Manhattan | 85,900 | calc |
| 10 | Pop Increase btwn Night to Day | 1,901,764 | calc |
| 11 | Pop Increase into lower due to migration into Manhattan | 973,522 | calc |
| 12 | Pop Increase into midtown due to migration into Manhattan | 928,242 | calc |
| 13 | Pop Increase into lower due to 25% migration within Manhattan | 190,369 | calc |
| 14 | Pop Increase into midtown due to 25% migration within Manhattan | 181,515 | calc |
| 15 | Total daytime Pops in lower Manhattan | 1,253,981 | calc |
| 16 | Total daytime Pops in midtown Manhattan | 1,195,657 | calc |
| 17 | Total Daytime PopDensity in lower (mile^2) | 729,059 | calc |
| 18 | Total Daytime PopDensity in midtown (mile^2) | 729,059 | calc |

**CellSite Loading**

| | | | |
|---|---|---|---|
| 19 | CellRadius (miles) | 0.06 | input |
| 20 | CellArea (miles^2) | 0.01 | calc |
| 21 | Daytime Pops in lower Manhattan cellsite | 7,394 | calc |
| 22 | Cellular Adoption | 72% | Input |
| 23 | CellSubs in lower Manhattan cellsite | 5,324 | calc |
| 24 | Market Share | 10% | input |
| 25 | Own subs in lower Manhattan cellsite | 532 | calc |
| 26 | Voice Minutes of Use/year/Sub | 1,700 | input |
| 27 | Minutes of Use/work-day/Sub | 6.54 | calc |
| 28 | Busy hour concentration of minutes | 20% | input |
| 29 | Busy Hour Minutes/Sub during work day | 1.31 | calc |
| 30 | Busy Hour voice traffic load/Sub (mE) | 21.8 | calc |
| 31 | Voice Load on Site from own subs (Erlangs) | 11.6 | calc |
| 32 | Number of Voice Ckts to be Reserved for 2% QoS Blocking | 18 | calc |
| 33 | Capacity requirement for voice traffic (kb/s) | 10 | input |
| 34 | Upstream & Downstream load on Site from Voice Ckts in Mb/s | 0.180 | calc |

**Data Traffic Load per Subscriber**

| | | | |
|---|---|---|---|
| 35 | Average Wired BB Downstream Busy Hour Data Rate/Sub (kb/s) | 25 | input |
| 36 | Ratio of Upstream to Downstream Data Rate | 20% | input |
| 37 | Throttle on Wireless Data Rate Ratio to Wired Rate | 112.00% | input |
| 38 | Busy Hour Concentration Factor (percent of all daily usage) | 20% | input |
| 39 | Percent of Subs who Take Data Service | 100% | input |
| 40 | Equivalent downstream megabytes/sub/month of Data Usage | 1355 | calc |

**MicroCellSite Capacity for Single Channel CDMA2000 EVDO Rev A all IP**

| | | | |
|---|---|---|---|
| 41 | Typical Down+Upstream throughput per 802.11b channel (Mb/s) | 6.5 | input |
| 42 | Number of active 802.11b channels/access point | 4 | input |
| 43 | Derating percentage based on contention | 70% | input |
| 44 | Down+Upstream throughput for Omni Single Sector Site (Mb/s) | 18.2 | calc |
| 45 | Down+Upstream Site throughput left after voice reserve (Mb/s) | 17.84 | calc |
| 46 | Downstream Capacity/sub requirement for data traffic (kb/s) | 28.00 | calc |
| 47 | Upstream Capacity/sub requirement for data traffic (kb/s) | 5.60 | calc |
| 48 | Total Down+Upstream data traffic Load/Site (Mb/s) | 17.888 | calc |
| 49 | Total Downstream Data+Voice Traffic Load on Site (Mb/s) | 18.068 | calc |

**Ratio Analysis**

| | | | |
|---|---|---|---|
| 50 | Ratio of Voice Load/Site to Total WiFi Site Capacity | 1% | calc |
| 51 | Ratio of Data Load/Channel to Total WiFi Site Capacity | 98% | calc |
| 52 | Ratio of DataLoad Max Possible Data Load | 100% | calc |

The analysis considers several cases in which the inputs to the model are varied:

- Each microcell employs 4 channels (1, 4, 8,11) of 2400MHz 802.11b WiFi network access points.
- The market share of the carrier under study is 10%
- The traffic loads studied consist of two cases however the capacity of the system is now so great, we assume the limit case in which non-voice data adoption grows from the 10% assumption employed in the licensed cases to approach 100% adoption of non-voice data services by basic voice subscribers.
  - Today's voice with multimedia data traffic max'ed out to use full capacity of the microcell
    - Voice of 142 MOUs
    - Each data sub consumes 1,355MB/month downstream traffic
    - Data subs are 100% of voice subs
  - No voice traffic with data capacity max'ed out
    - All voice traffic on the macrocell and none on the microcells
    - Each data sub consumes 1,379MB/month downstream traffic
    - Data subs are 100% of voice subs

<u>Unlicensed Microcellular Capacity Computation</u>

In a matter parallel to the licensed microcellular analysis model, the model of Table 9 was created to compute the capacity of the unlicensed microcellular alternative based on a muni-WiFi type of architecture employing 802.11b technology.

The Table 9 model is the same as the licensed microcellular model from line 1 through line 40 with the exception that the line 37 data load is set to max out the cellsite capacity as shown in line 52. Lines 41 through line 43 have been added to unlicensed model to compute the cellsite (i.e. 4 channel WiFi access point) capacity on line 44. Otherwise this model is the same as that used for the licensed microcellular case.

Line 41 of the Table 9 model shows an assumption for a typical data throughput of 6.5Mb/s per WiFi 802.11b channel which is multiplied by 4 channels on line 42 and further by a 70% derating factor on line 43 to result in the throughput of 18.2Mb/s on line 44. The 70% factor is meant to represent overhead associated with the WiFi contention protocol. Employing more recent 802.11X technology could increase the throughput.

Because the voice load on the site is so small in comparison to the non-voice data load, there is virtually no difference in the maximum non-voice capacity of site with or without voice traffic. The model computes a maximum capacity of 1,355MB/month/sub with voice traffic on the microcell and 1,379MB/month/sub with voice traffic moved to the macrocellular overlay network.

Since today's cable modem subscriber consumes 1,209MB/month/subscriber of datastream high speed internet services, this unlicensed microcellular network has a capacity that exceeds today's wired cablemodem data consumption.

## ALTERNATIVE ARCHITECTURES

One concern that might be raised with regard to implementation of a microcellular network is the difficulty of obtaining reliable 100% area coverage with such small 300 foot radius cells. However, this concern is mitigated when you consider that the underlay need not have 100% coverage of the macrocell in order to support high quality multimedia services because the wide area microcell is always available for fallback when microcell coverage may be lacking or marginal.

Also of interest is designing non-voice applications which take advantage of the increasingly low cost but very capable storage and intelligence in today's cellphones to implement applications in which low latency low bandwidth user interface traffic is delivered over the wide area macrocell and high latency tolerant block file transfers are conducted using the microcells. Such an approach in which the wide area network is employed for user interface and control and the very wide bandwidth delay insensitive traffic can be delivered over the microcellular data network without any noticeable impairments notice by user for many applications of interest, e.g:

- Synchronizing mp3 music files stored locally in the phone with an online server
- Enhancing *Cache and Carry* of video by supporting intervals of a few minutes between reloading the cache
- Automatically delaying large email file uploads or downloads until within range of a microcell

## SUMMARY AND CONCLUSIONS

Although cable operators may currently own 20MHz of spectrum versus competitors holding 2-3 times more, cable can more than compensate by exploiting unique backbone cableplant assets that would be non-economic for competitors to match. Stealing one 1.25MHz channel pair from the macrocell inventory to implement a single channel pair microcell underlay would support today’s voice traffic load as well as allowing data subscribers to scale from today’s 10%

adoption to 100% adoption at the same 50MB/month/sub rate of consumption. Using four unlicensed 802.11b channels in the microcell in a MuniWiFi like architecture would allow cable to offer competitive wide area voice services as well as support 100% adoption by subs who consume data at rates even greater than today's wired cablemodem subscribers.

A summary of the modeling results in Table 10 demonstrates the significant increase of a microcellular network versus conventional macrocellular plant.

an unimpaired level of user satisfaction resulting in significant CapEx cost savings for the microcell underlay.

By intelligent design of the services offered, the coverage quality of the microcell underlay need not be made equal to the macrocell overlay network in order to provide

**Table 10 Summary of Systems Capacity Modeling Results**

| Case Under Study | Architecture | Per Subscriber Traffic Load | | Total Licensed Spectrum Employed (MHz) | Portion of Licensed Spectrum Needed | | Total UnLicensed Spectrum Employed (MHz) | Portion of UnLicensed Spectrum Needed | |
|---|---|---|---|---|---|---|---|---|---|
| | | Voice Traffic Load (Minutes of Use Per Month) | Downstream Data Consumption (MegaBytes per Month) | | For Voice Traffic (%) | For Data Traffic (%) | | For Voice Traffic (%) | For Data Traffic (%) |
| Macrocells with Unlimited Licensed Spectrum | 0.5Mile radius 3-sector EVDO-Rev A all IP CDMA2000 | 142 | 5<br>DataSubs=10%<br>Each Sub 50MB | 30=1.25X2X12 | 83% | 17% | None | 0% | 0% |
| Underlay of Microcells Employing a Single Channel Pair<br>Mixed Voice+Data | 300' radius single channel omni sector EVDO-RevA all IP CDMA2000 | 142 | 5<br>DataSubs=10%<br>Each Sub 50MB | 2.5=1.25X2X1 | 40% | 2% | None | 0% | 0% |
| Underlay of Microcells Employing a Single Channel Pair<br>Voice+Maxed Out Data | 300' radius single channel omni sector EVDO-RevA all IP CDMA2000 | 142 | 65<br>DataSubs=10%<br>Each Sub=653MB | 2.5=1.25X2X1 | 40% | 60% | None | 0% | 0% |
| Underlay of Microcells Employing a Single Channel Pair<br>Voice+Maxed Out Data | 300' radius single channel omni sector EVDO-RevA all IP CDMA2000 | 142 | 65<br>DataSubs=100%<br>Each Sub=65MB | 2.5=1.25X2X1 | 40% | 60% | None | 0% | 0% |
| Underlay of Microcells Employing a Single Channel Pair<br>Maxed Out for Data Only | 300' radius single channel omni sector EVDO-RevA all IP CDMA2000 | 0 | 82<br>DataSubs=10%<br>Each Sub=816MB | 2.5=1.25X2X1 | 0% | 100% | None | 0% | 0% |
| Underlay of WiFi Microcells<br>Today's Wired Data Consumption | 300' radius four channel omni sector 802.11b WiFi | 0 | 1209<br>DataSubs=100%<br>Each Sub=1209MB | None | 0% | 0% | 74MHz<br>4 Chans<br>1, 4, 8, 11 | 0% | 88% |
| Underlay of WiFi Microcells<br>Maxed Out for Data Only | 300' radius four channel omni sector 802.11b WiFi | 0 | 1379<br>DataSubs=100%<br>Each Sub=1379MB | None | 0% | 0% | 74MHz<br>4 Chans<br>1, 4, 8, 11 | 0% | 100% |

# EXTENDING CONDITIONAL ACCESS SYSTEMS TO SUPPORT DRM SYSTEMS

Balu Ambady
Cable Television Laboratories, Inc.
Bill Helms
Time Warner Cable, Inc.

## *Abstract*

*Conventional implementations of content security by Multiple System Operators (MSOs) rely on hardware based broadcast Conditional Access Systems (CAS). However, competitors like IPTV providers, and Internet based movie-on-demand sites have started offering high value content using content protection based on software Digital Rights Management (DRM) systems like Microsoft WMDRM, and Real Helix DRM. It appears that content providers are increasingly becoming comfortable with the reduced level of security provided by DRM systems for certain levels of content.*

*Under currently available solutions, while the content is under CAS protection, opportunities to offer innovative content packages to the customer are preserved.*

*This paper describes architectural options where CAS protected content can be transferred securely from Set Top Box (STB) devices to Portable Media (PMD) or PC devices with DRM protection. This would allow the operator to provide an enhanced home networking and content usage experience to the subscriber.*

## INTRODUCTION

### Content Protection Domains

When user subscribed content is moving from the Operator's head-end device to a STB, and then on to a Home Network device, from a content protection point of view, the content may be considered to be moving between one of the three content protection domains as shown in Figure-1:

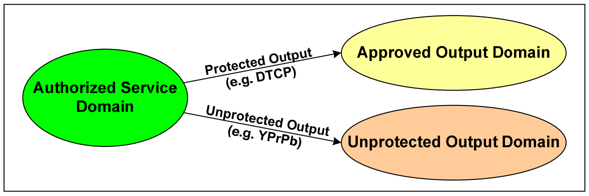


**Figure 1 – Content Protection Domains**

1) Authorized Service Domain (ASD)

Within the ASD, content is secured using the mechanisms provided by the MSO according to usage rules set by the Operator's provisioning system. Content is protected using CAS while it is transferred from the MSO head-end to the subscriber home, and by an ASD specified protection mechanism while it moves between one or more trusted devices that are part of the ASD network.

2) Approved Output Domain (AOD)

As permitted by FCC encoding rules[1], content may also be released to non-operator Content Protection (CP) or DRM systems that have been approved[2] by CableLabs®. When the content is released from the ASD to another CP system, it is considered to exit the ASD and enter the AOD. The content is still securely protected by the CP/DRM system and the encoding rules are enforced by the CP/DRM.

Several Link Protection systems like DTCP and HDCP, and DVD based recording solutions like VCPS fall under the AOD category. A PC based architecture, the OCUR (OpenCable Unidirectional Receiver) using the Microsoft WMDRM and Real Helix DRM systems for content protection also fall under this category.

3) Unprotected Output Domain (UOD)

Lower value content that require only minimal levels of protection, or no protection at all may be released to the UOD, for example to the unprotected Analog video outputs.

High Level Design Goals

While a large number of the devices that are capable of serving MSO delivered content fall under one of the three domains described earlier, many new devices like PMDs and cell phones do not readily fall under these domains. In order for the MSOs to offer the best entertainment experience for the subscriber, it is desirable to bring more of these new devices either under ASD, or AOD. In order to join the ASD or AOD, these devices must be capable of providing sufficient levels of content protection, as well as be capable of supporting the MSO defined usage rules. We offer some architectural options to achieve this goal.

The options provided in the next section provide solutions while striving to attain a balance between the needs of the main stakeholders and their needs:

1) Subscribers

- Ease of use/transparency
- Fair use
- Interoperability with other devices in the home

2) Content Provider/owner

- Sufficient protection of owner rights
- Monetization of content

3) MSO

- Ensuring a link between services provided and payment received
- Support for flexible business models, freedom to innovate
- Leverages existing infrastructure where possible
- Provide consistent MSO branded experience

4) Device manufacturer

- Easy to license technology
- High level of interoperability
- Freedom to innovate

## ACHITECTURAL OPTIONS

We offer four solutions to achieve the goals outlined above. Three of the solutions described below use Operator provided security mechanisms and thus allow the content to stay fully within the ASD, and the fourth option provides a “bridge” mechanism from ASD to a third-party DRM solution.

In the case of the three ASD solutions, encoding rules permit the operator to extend the usage rules beyond the limited usage rules expressed using the 8-bit Copy Control Information (CCI) bits[3]. This would allow for new business models for the MSOs, and more choice for the subscriber. For example, current CCI is not sufficient to express a “rent” model where a subscriber may rent a digital content for 7 days.

In the case of the bridge from ASD to the native DRM space, certain additional encoding rule restrictions may apply.

### 1) Hardware ASD

This architectural approach would require all devices that are interested in joining the ASD to embed the Downloadable Conditional Access System (DCAS) Secure Micro chip and store the associated keys, secrets and Root of trust. An architectural diagram of a typical implementation of home networked ASD Host with a DVR application is shown in Figure-2.

The System on a Chip design would allow the Root, critical security parameters and keys to be securely stored on the Secure Microchip. The software modules that support ASD functionality can be securely downloaded on to the device based on this Root of trust. The ASD modules are developed by an ASD vendor, and all vendors follow a standard set of protocols, content formats and encryption schemes. In addition to allowing multiple ASD vendors to participate in this eco-system, this has the added advantage of allowing the same ASD infrastructure and standards to be utilized by both MSO leased and retail devices.

The CA protected content is converted into ASD protected content. Security Packages are created that store the usage rights to be kept with the encrypted content. The Secure Micro is responsible for key generation during encryption, and for retrieval of keys during decryption. The Host Transport processor is responsible for the actual encryption/decryption of the content.

Devices that belong to the same ASD are managed by the ASD controller by using a “Trusted List” that identifies the devices that are tied to a subscriber billing account.

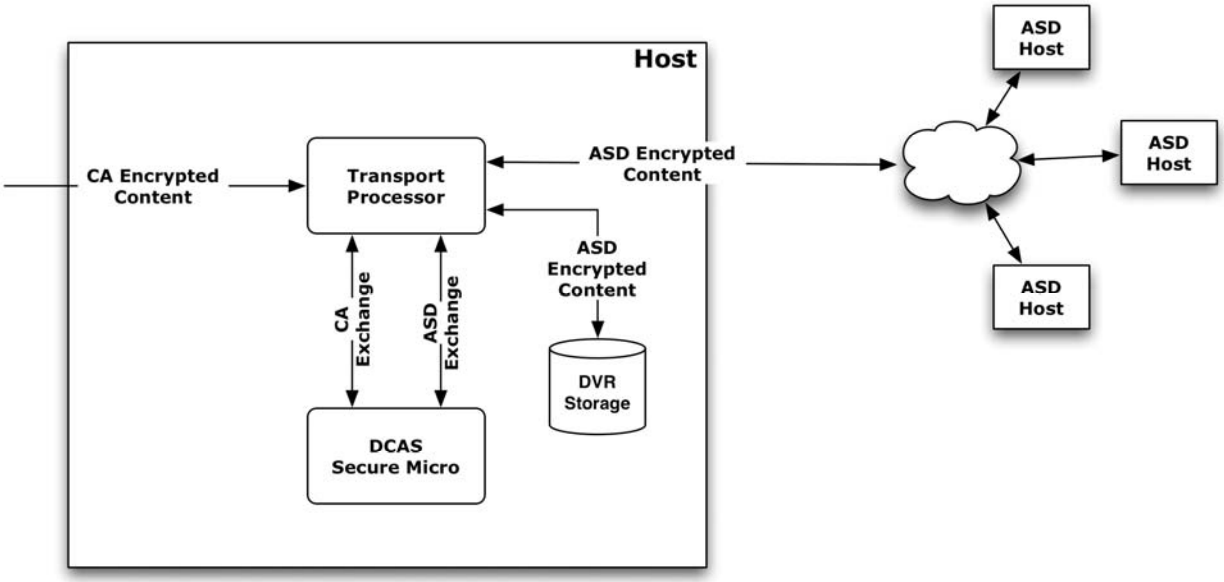


**Figure 2 – Simplified DCAS ASD Host Block Diagram**

2) Cable DRM (CDRM)

An alternate approach would be for the Cable industry to develop its own DRM solution that can run on a variety of devices that do not have the Secure Microchip or other Hardware secure storage for keys and other security parameters. Ideally, this would be a Platform/Operating System (OS) independent technology that is easy to implement on a wide variety of devices like cell phones and PMD made by different manufacturers. Content security will be based on obfuscation of keys and code, and software renewability.

The architecture for such a solution is shown in Figure-3. The components are similar to typical DRM systems, there is a DRM Client on the portable device, a Controller providing Keying, Registration and Individualization for the Client, a License Server that packages the content in DRM format and also enforces Licensing rules, and a Content Server from where the Client can access DRM protected content.

In a simple use case, as the 1st step a personalized copy of the Cable DRM Client that is uniquely tied to the PMD is securely downloaded. The 2nd step, high value content that the subscriber is authorized to receive is moved with Conditional Access protection to the Content Server, in this case the STB. The 3rd step, as authorized by the License server, the STB transcrypts the content in to CDRM format. The 4th step, the PMD can download this content from the STB and consume it as allowed by the CDRM usage rules.

Since the Cable DRM devices use the Operator's protection system, devices implementing this are part of the ASD. These devices support the extended usage rules using standardized protocols as in the case of Hardware ASD solution.

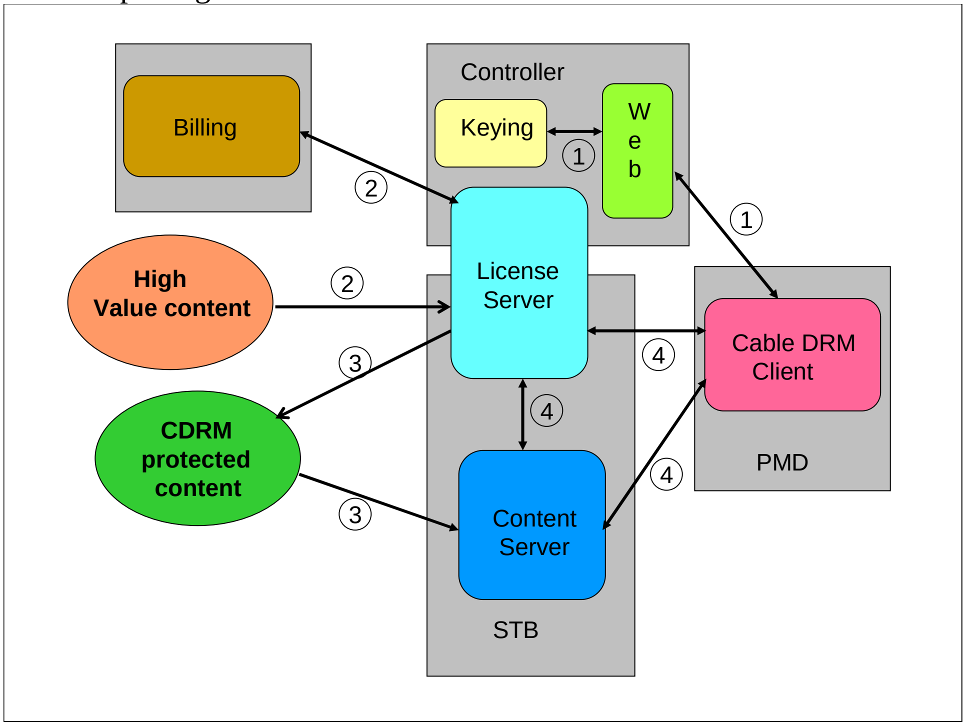


Figure 3 – Cable DRM

Because of the relative ease of circumventing Software DRM systems, the DRM Client should be carefully designed to preserve execution integrity and algorithm secrecy, prevent tampering, impersonation, and input spoofing.

These security goals are usually achieved in conventional software DRM systems by using the following mechanisms:

- Use of information diversity and complexity
- Use of masking techniques to mask: control flow, data/code itself, location, and usage
- Use of one-way transformations, temporal and spatial diversity
- Use of Index computations, aliasing techniques
- Use of "bad programming" practices: e.g., use of pointers, goto
    - Hard to debug (on purpose)
    - Prevent Dynamic analysis (during execution)
- Use of encrypted functions
    - Integrity Verification Kernel
- Use of "white box" techniques
- Resistance to side channel attacks ("grey-box") - however, this may be less of a concern as S/N ratios are significantly degraded in GP
- Software Key hiding
    - Many skews to prevent domino effect (of hacks)
- Use of Protected Media Paths
    - Utilize OS support (e.g., Vista PVP)
    - Disk/CD/DVD drive solutions
- Secure boot up/firmware
- Encrypted/protected memory

Other design attributes supporting security should include:

- Secure clock
    - Anti-rollback time
    - Secure (external) time source
- Secure random number generation
- Support of standard cryptographic algorithms
- Support for secure Proximity checks
- Efficient software implementation
- Robust Revocation, and Renewal mechanisms
    - Breach response readiness
    - Heart-beat checks
- Secure DRM infrastructure
    - Secure build servers
    - Secure (firewalled) download portals
- Ability to add features later
    - Watermarking, fingerprinting

- Ability to use Trusted Platform Module

Similar to the Hardware ASD, the CDRM would allow the MSOs to use a common messaging protocol, encryption scheme, data structures and extended usage rules on all of the devices on which the CDRM Client is available.

### 3) Software ASD

This is a modified approach that would combine the options-1 and 2 above, this option would allow existing CAS/ASD vendors to port their ASD clients into the Portable devices. The architecture and requirements for the Software Client are similar to the CDRM case, except that some of the messaging protocols may be proprietary to the CAS/ASD vendor. Another difference is that instead of a single DRM, there would be multiple CAS/ASD systems, one for each vendor.

Since the main difference between Hardware ASD and the current solution is that the security is based on obfuscated software instead of hardware stored keys, this solution maybe considered as a “Software ASD” implementation.

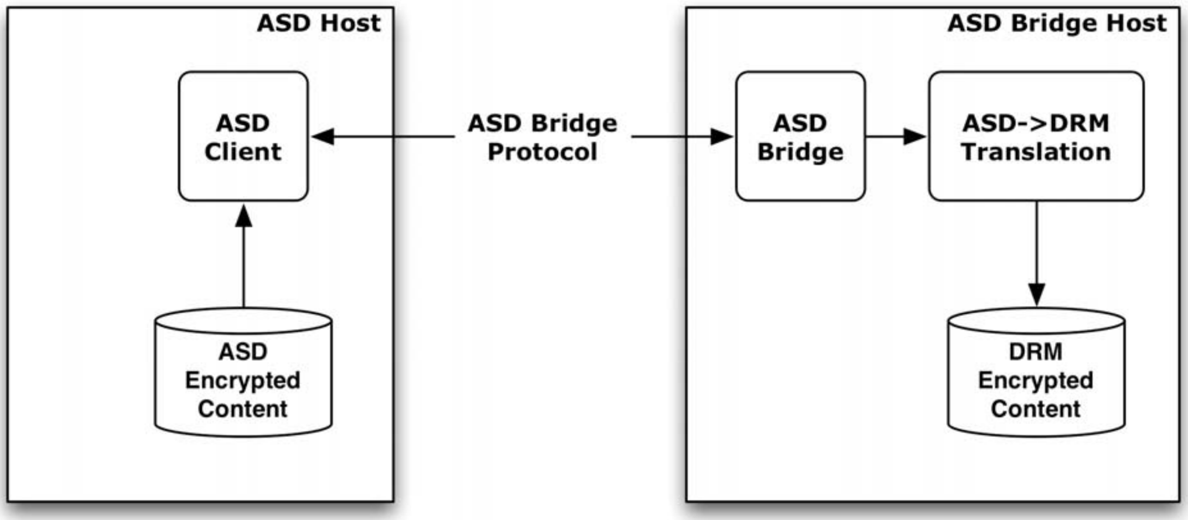


**Figure 4 – ASD/DRM Bridge**

### 4) ASD/DRM Bridge

The ASD/DRM Bridge model shown in Figure-4 is a compromise between the previous three models. It uses ASD protocol to exchange content with the ASD Bridge device (e.g., PMD). However, the native DRM is responsible for securing the content and obeying the MSO usage rules contained in the ASD protocol.

Of course, this requires that the bridge device translate from the ASD encryption and protocol to the native DRM encryption and protocol. Under this model, the implementation of the bridge to DRM is the responsibility of the device vendor, rather than the operator or ASD vendor.

### Comparison of Options

The relative advantages and disadvantages of the four options outlined above are shown in Appendix-A.

## CONCLUSIONS

Depending on the implementation timeline, complexity of implementation, ease of licensing, cost and other factors, the Operator may select one of the four options presented above. An ASD solution that embraces the new generation of Portable devices and other Home Networked devices would help to enhance the subscriber's entertainment experience and thus would help to strengthen brand loyalty.

## References

1 The FCC Second Report and Order (FCC 03-225) Section VI-B. http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-03-225A1.pdf

2 The FCC Second Report and Order (FCC 03-225) Section V. http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-03-225A1.pdf

3 SCTE 41 – POD Copy Protection System http://www.scte.org/documents/pdf/ANSISCTE412004.pdf

## Appendix – Comparison of Options

| | **Hardware ASD** | **Cable DRM** | **ASD Client** | **DRM Bridge** |
|---|---|---|---|---|
| **Complexity of Licensing** | Low | Low | Low | High |
| **Cost to MSO** | High | High | Medium | Low |
| **Complexity of Development** | High | High | Medium | Medium |
| **Interoperability** | High | High | Medium | Medium |
| **Extended usage rules** | Yes | Yes | Yes | Yes with possible restrictions |

# HOUSE CHECK - A NEW PROCESS TO VERIFY VOICE INSTALLS

Mike Gordish, Product Manager, Mobile Workforce Manager, C-COR
Robert F. Cruickshank III, VP OSS Market Strategy, C-COR
Dan Rice, VP Assurance Solutions, C-COR

## *Abstract*

*As MSOs accelerate efforts to satisfy strong market demand for cable digital voice services they have an opportunity to adopt new installation procedures that can contribute significantly to customer retention while reducing ongoing operational costs. Back office-based 'Assurance' software that delivers comprehensive analysis of voice service performance on the MTA (multimedia Terminal Adapter) across all household cable outlets with DOCSIS equipment installed, allows installers to quickly obtain a view of whole-home service quality that simplifies identification and repair of problems at installation and provides a reference base for use in future trouble shooting. By aggregating and interpreting DOCSIS-generated data on RF signal quality, jitter, packet loss and other parameters, these "house checks" provide easy-to-understand readouts on installer handhelds and give Customer Service Representatives, Dispatchers, Engineers and Network Operations Center (NOC) personnel a reliable record of the true state of performance before the installer leaves the premises.*

## INTRODUCTION

Early market response to IP-based cable digital voice service has prompted MSOs to move as quickly as possible to extend service reach within existing markets and into new markets. At the same time, operators recognize that the faster they proceed, the greater the challenge they face in meeting the fundamental requirement that their voice services match or exceed the performance quality of incumbent switched circuit telephone service.

Where, initially, operators could afford to devote whatever amount of time it took for well-trained in-house installation personnel to perform comprehensive testing and repairs on drop connections and home wiring to ensure quality performance, the industry is long past the stage where spending several hours per installation is feasible. In order to scale to millions of customers, MSOs have had to reduce installation times while often turning to outside contractors to handle the volume of orders.

This accelerated pace of deployment and service provisioning raises the risk that problems will arise after installers leave, which in turn can lead to higher levels of customer dissatisfaction. One MSO's recent study[i] of the impact of faulty service on customer retention found that 10 percent of voluntary disconnects are either directly or indirectly caused by service problems. The MSO found that repeat service problems are the largest single cause behind voluntary disconnects.

Another recent study[ii], by an independent research company, found that technicians had to visit 34 percent of new cable VoIP customers within 90 days of installation and that 16 percent required visits two or more times. Twenty-nine percent of the people experiencing problems said their "primary frustration" was with recurring quality deficiencies, while twenty-four percent said the frustration was with recurring reliability problems.

With various reports citing truck roll costs of anywhere from $130 to $200 per average visit, the operational costs combined with the costs of customer dissatisfaction are clearly too high for operators to sustain these levels of service problems. The obvious place to begin in the effort to cut these costs is with the installation process itself. Fortunately remedies are at hand that will help operators achieve a much higher level of assurance that service quality is where it should be when the installer leaves the home without adding significantly to the amount of time installers spend executing the work order.

## THE INSTALLATION CHALLENGE

Two major issues stand in the way of an MSO's ability to ensure digital voice service is performing at required quality levels upon completion of service installation. The first has to do with the amount of time it takes to thoroughly examine all potential problem areas using traditional means of testing and measurement. Getting a good quality read on levels from the primary outlet to be used for connecting the MTA (multimedia terminal adapter) to the cable network is typically all that time allows, which means that if the subscriber subsequently moves the MTA to another outlet, there is no assurance the same performance levels will apply.

Splitter locations, outlet-specific points of ingress, variations in microreflections depending on where lines are capped off can all produce significant variations in performance from one outlet to the next. Knowing where the weak points are gives the operator the option of either fixing them or advising the customer not to use those outlets.

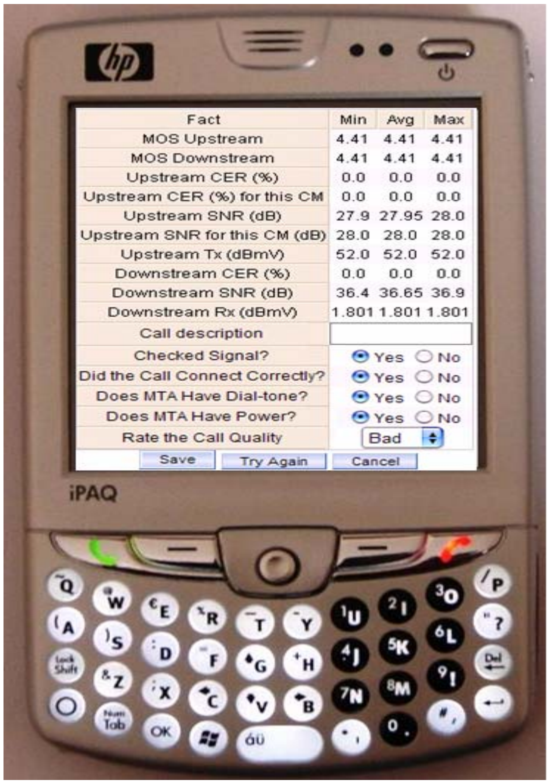
Another problem has to do with the fact that contract workers are typically paid by the number of jobs they complete per day. Every time they encounter a problem that must be fixed, their time on the job is extended. Unfortunately, with manual entries of performance parameters it's easy to 'fudge' the results to produce a performance report that makes it seem like a sub-par service meets minimal threshold levels. If the service is relatively clean and works at the time of installation, the fact that levels are below threshold may not lead to discernable problems until long after the installer has departed.

Measurements taken by installers also provide few clues as to the source of problems, with the result that the installer may go through time-consuming remedial steps that can be innefectivebefore he knows he must turn to network technicians for help. And even then he may not know whether the problem is network related or a problem associated with IP addressing, device identification or other data issues that require IT department attention.

Beyond the immediate impediments to efficient and successful installation, manual measurement and data entry provides an imprecise record of device and link status for use as a reference baseline for troubleshooting problems in the future. Thus, current methodologies not only contribute to post-installation problems; they impede efficient resolution of those problems when they occur.

## THE OPTIMUM HOUSE CHECK

Thoroughness in monitoring service performance is vital to ensuring potential sources of future problems are mitigated while the installer is still on the job. This includes taking stock of key internal and external network performance parameters as well as all metrics that are vital to proper performance of premise devices.



| Fact | Min | Avg | Max |
|---|---|---|---|
| MOS Upstream | 4.41 | 4.41 | 4.41 |
| MOS Downstream | 4.41 | 4.41 | 4.41 |
| Upstream CER (%) | 0.0 | 0.0 | 0.0 |
| Upstream CER (%) for this CM | 0.0 | 0.0 | 0.0 |
| Upstream SNR (dB) | 27.9 | 27.95 | 28.0 |
| Upstream SNR for this CM (dB) | 28.0 | 28.0 | 28.0 |
| Upstream Tx (dBmV) | 52.0 | 52.0 | 52.0 |
| Downstream CER (%) | 0.0 | 0.0 | 0.0 |
| Downstream SNR (dB) | 36.4 | 36.65 | 36.9 |
| Downstream Rx (dBmV) | 1.801 | 1.801 | 1.801 |

Call description
Checked Signal? ● Yes ○ No
Did the Call Connect Correctly? ● Yes ○ No
Does MTA Have Dial-tone? ● Yes ○ No
Does MTA Have Power? ● Yes ○ No
Rate the Call Quality Bad

Figure 1: Handheld with House Check Metrics

These measurements must be accessible to the installer via standard handheld devices with user interfaces that clearly spell out the metrics with respect to each type of measurement. And measurements must be recorded and stored for immediate and ongoing access from Customer Service Representatives, Dispatchers, Engineers and NOC personnel. See Figure 1 for an illustration of handheld User Interface.

The MTA, of course, is an essential source of information about service performance. The House Check performed at the MTA should measure and record:

- Levels and persistence of errors generated from the external network;
- Premises wiring noise, distinguishing as to types such as ingress, device-generated noise and spurious noise from microreflections;
- Transmit and receive power;
- Provisioning status with regard to performance of voice and data interfaces, related service elements such as Soft Switches, NAP (Network Access Point) addresses, security, service levels, etc.

This degree of granularity in measurements not only assures all bases are covered with regard to assessing service performance; it also provides the information an installer needs in order to know whether he can fix the problem or, if he needs to request help, whether the problem should be addressed by network ops or an IT technical unit.

To provide a full perspective on service performance measurements from all VoIP-specific interior outlets the House Check should tabulate:

- MOS (Mean Opinion Scoring) – MOS is a measure of phone call quality and requires use of sophisticated algorithms in a software program that looks at several quality parameters, such as jitter, latency and packet loss, to determine where the voice signal should be placed on the industry standard MOS scale.
- Stability – It’s essential that enough measurements be made over time across enough data points during installation to prevent over-reliance on what could turn out to be an anomalous performance reading taken on one data point at a single point in time.
- Inside wiring performance – The installer must be able to assess power and signal-to-noise levels at every outlet and take remedial steps to adjust levels to compensate for the impact of splitters and spurious noise from microreflections.

In addition, the installer should have at his disposal a ready means of measuring and recording levels at the tap and ground block. These measures not only serve to identify whether there are local outside plant issues; they also create a clear benchmark for gauging contributions of in-home wiring and other elements to signal degradation at all VoIP-specific outlets. And the technician must be able to perform these measurements with respect to transmit and receive parameters across all VoIP-related DOCSIS channels.

## THE ENABLING SOFTWARE INFRASTRUCTURE

Creating an easy-to-use, thorough House Check process as described above requires a well-conceived software infrastructure that fully exploits the data collecting power of DOCSIS with complete compatibility with multiple vendors and support for multiple DOCSIS and PacketCable devices, including CMTSs, call management servers, media gateway controllers, cable modems, DOCSIS set-tops and MTAs. To be useful beyond the installation stage, the software should interface with a complete VoIP lifecycle management solution that optimizes on-going fault and performance management.

The system must be able to isolate and resolve issues at the time of installation and beyond by providing real-time and historical data about premises devices, supporting interfaces and the CMTS. This includes the ability to identify and resolve typical transport issues such as jitter, fragmentation, latency and packet loss. The system must be able to draw together and interpret real-time and historical information for any given premises device, providing detailed views to the NOC on location, configuration, settings and performance while delivering read-outs of pertinent device information on the technician’s handheld at the customer site.

Beyond the initial install, the system must be able to generate real-time and historical reports reflecting performance over time. With this level of comprehension the system should be able to isolate and prioritize problems and provide quick and efficient recommendations to resolves the issues, both at the time of install and subsequently whenever problems arise.

As part of the analytical process, the House Check support system should be able to go beyond generating raw performance metrics on a per-device, per-location basis to calculate a MOS score for the service at all end points. Such MOS summaries measure network health by determining availability in terms of Degraded Modem Hours (DMH) or Severely Degraded Modem Hours (SDMH), and identify network issues that are causing QoS-sensitive applications such as voice to fail.

### The Full Scope of House Check Benefits

There are many immediate and long-term benefits to be derived from implementing a House Check process like the one described here. From the installation perspective, the process reduces the time spent by installers on each job and provides operations support staff a complete picture of post-install performance at the site versus what the performance looked like when the installer started running the analysis.

The process eliminates time-consuming manual measurements and correlation of different types of performance parameters at the tap, ground block and premises outlets. Instant feedback minimizes time spent on troubleshooting. And, if more work is required, the process further cuts installer work load by automatically creating thorough maintenance work orders.

Another enormous benefit is the ability to quantitatively measure and report 'success rate' of installs by individual installer or groups of installers/contractors. The most capable installers can be recognized and given more 'challenging' installs and trouble calls. And the less capable installers can receive training in specific areas that are deficient. Gone are the days of wondering which staff are the most capable and which staff repeatedly have install/trouble call difficulties.

Beyond the installation phase, the information generated by the House Check establishes a comprehensive data base which Customer Service Representatives, Dispatchers, Engineers and NOC personnel can use to trouble shoot problems over time. For example, a shift in MTA readings that matches the readings taken at a secondary or tertiary outlet during installation can indicate the device has been moved, allowing the CSR to suggest it be put back in its original location to avoid a maintenance call. And, if a truck roll is necessary, the comparative readings can tell Customer Service Representatives, Dispatchers, Engineers and NOC personnel whether the problem is plant related or, if it's on premises, what type of expertise and equipment will be required to perform the work.

The operational savings resulting from implementation of a House Check process are evident across several parameters. The process ensures a higher success rate for VoIP installs and virtually eliminates post-install trouble calls. Over time there are fewer complaints per customer, which increases customer satisfaction. And when there are problems, less time is spent trouble shooting and fixing them. Moreover, the accumulated performance record across service areas allows NOC personnel to discern patterns in the problem reports that can track problems geographically.

## CONCLUSION

Now that MSOs have made the business case for digital voice to where they are scaling service to millions of households, the bottom-line focus necessarily shifts to finding ways to reduce costs and raise customer satisfaction.

The opportunity to eliminate the primary cause for voluntary disconnects, namely repeat trouble calls, while significantly lowering installation and ongoing operations costs is one of the great benefits the cable industry derives from operating in the packet-based DOCSIS environment. Now, with the means at hand to support whole-home House Check functionality, operators can look forward to ongoing improvements in the bottom-line performance of digital voice.

## References

---

[i] Undisclosed MSO Customer Data

[ii] MSO-commissioned study

# HYBRID IP-QAM VIDEO SOLUTIONS

Christopher Poli, P.E., Richard Moore, William Weeks
Motorola Connected Home Solutions

*Abstract*

*A Hybrid IP-QAM System Architecture is one of the most effective architectures available today in expanding plant capacity – allowing operators to offer an expanded and wide spectrum of high definition, international and niche programming. The architecture is capable of fully supporting Broadcast, Unicast or Narrowcast services (i.e., VOD) and Switched Digital Video; multiple video encoding formats (e.g., MPEG-2, MPEG-4) and can support a transition to an all IP STB solution. While a complete IP network build-out will likely take years and millions of investment dollars to complete; a hybrid system provides a bridging architecture without stranding existing set-top assets. This paper provides an overview of this architecture including a comparison of the bandwidth delivery provided in hybrid architectures to that of other architectures.*

## INTRODUCTION

The competitive landscape is changing and becoming more challenging. Satellite delivered TV has publicly announced plans to launch 150 HD channels (source: DirecTV website) – nearly ten times the typical HD offering of most cable MSO's. It would take 75 QAM streams of cable plant capacity to carry that many HD services in MPEG-2 format. Competitors are deploying highly competitive video services in many major markets with plans to offer richer content and more bandwidth than is possible in an HFC plant. VOD capacity requirements continue to grow – pushing nodes deeper into the cable plant. Symmetrical applications such as VOIP and internet applications continue to increase pressure on high speed data rates.

Over the past couple years, at least one major successful deployment of a national hybrid QAM IP architecture has already occurred (and continues to expand). Municipal entities are in the planning stages or have already begun overbuilding as well – some are looking at a pure IPTV play and others at hybrid architectures that provide both broadband IP and QAM delivery capabilities.

There are options to increase the effective plant bandwidth capacity beyond splitting nodes to the point of fiber exhaust. Moving to 1 GHz for data services may also provide some relief (video still up to 860 MHz). Going all digital – and recapturing analog spectrum. Switched Digital Video is an emerging technology as another means to reclaim channel bandwidth – and combined with DOCSIS 3.0, could provide significant relief to forward bandwidth demand – both from a video services standpoint as well as data services. All of these activities certainly help a great deal in the forward path, but they neither address the longer term need for significantly more bandwidth nor the need for significant additional symmetrical bandwidth.

This paper will cover an End to End Network and System Architecture culminating in a PON based delivery to the home with hybrid IP and QAM capable set-tops. A hybrid IP QAM set-top has the capability to receive services both on an IP interface and through the more traditional QAM tuner. The set-top applications determine the video source to be displayed or recorded. The IP interface can use either a traditional ethernet interface or a home network coax-based interface such as Multimedia over Cable Architecture (MoCA) or HomePNA. A home network based architecture such as MoCA can provide 70-100 Mb/s IP interface over an RF channel that runs above 860 MHz. A properly deployed, Hybrid architecture offers strategic advantages to the service provider.

The paper will also cover benefits of integrating Video on Demand or other singlecast services onto the IP data path, and briefly describe the use of multicast on the IP pipe to deliver carousel data, tunnel singlecast data, or deliver IP video services. The paper will finally cover how an IP capable architecture can enable seamlessness in communications to a multitude of home and mobile devices – a two way IP pipe to stay connected to the world.

## BANDWIDTH NEEDS

How much bandwidth is enough? With the proliferation of HD content and multiple HD TV sets in a single home, bandwidth required to the home in either a traditional HFC or Switched Digital delivery will drive the need for architectures that reach beyond simply pushing QAM nodes further and further outward toward the edge. As the price of HDTV sets continues to fall, and the obvious quality difference between HDTV and Standard Definition TV continues to rise within the subscriber base, the need to deliver richer content – certainly beyond the typical 10 to 20 HD services offered in an HFC plant - will rise significantly in the near future. Add to that the ever increasing demand for data services bandwidth and it may become apparent that subscriber demand will exceed today's HFC QAM channel plant capacities. The cable provider who is prepared to deliver the bandwidth to enable these services and more will have a significant competitive advantage over the operator who cannot.

In today's MPEG-2 environment, delivering 100 HD streams and their equivalent SD streams will consume on average 60 QAMs. That is nearly half of the QAM channels available on an 860 MHz HFC plant. Take away analog channels and those dedicated to data and there is very little bandwidth left to deliver compelling services such as VOD that are capable of significantly reducing subscriber churn. While data services may likely be moved above the 860 MHz channel frequency to extend available bandwidth, it still requires pushing the nodes further and further out into the plant to provide the needed bandwidth to deliver all the services desired. As nodes are split, fiber exhaust may become a serious consideration in the long term.

How much bandwidth is required in building out a pure IP network that has the ability to deliver enough bandwidth to each home to provide satisfactory service? Even with the promised bandwidth advantages of MPEG-4, high definition services will consume 8 Mbps/service. If we assume a peak requirement of simultaneously delivering two HD services, an SD service and 6 Mbps best effort data service, then the peak BW per home is approximately 26 Mbps to a single home. Although some of the bandwidth is delivered to multiple homes in an IP Switched Digital Video deployment, that network will need to be managed – meaning that decisions to allow new video sessions to traverse a particular path may be dependent on current bandwidth consumption on that path. Depending on specific market assumptions, the 26 Mbps/home number may be adjusted upward or downward. It certainly would be no less than 20 Mbps/home and an argument might be made for as much as 40 Mbps/home or more. While this number may seem absurdly high today, how long ago did 750 Kbps DSL, 1.5 Mbps DSL or 6 Mbps Cable Modem rate seem absurdly high when phone line data connections were only capable of delivering 56kbps? Data customers moved very quickly from slower data service providers to high speed cable modems. The future may hold another migration to competing services that offer more or business-necessary erosion to service margins to keep the current customer base.

Today's HFC services compete for available bandwidth. In a typical HFC architecture, QAM channels are allocated across multiple service types. As bandwidth demands increase, HFC architecture may move toward Switched Digital Video where a number of QAM channels are allocated to switched broadcast (essentially multicast), singlecast (i.e., VOD) or data services. Depending on the allocation and number of subscribers served per node, these nodes will push deeper and deeper into the network requiring more and more fiber to provide the more effective bandwidth per subscriber to remain competitive.

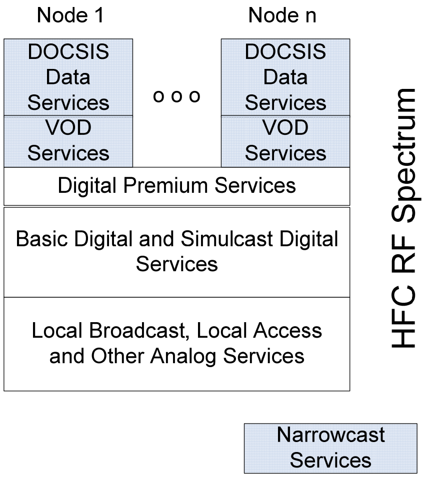


Figure 1: Typical HFC Services

This paper is not aimed at converting from a typical HFC architecture to PON architecture as conversion may have an insurmountably high network infrastructure, headend and customer premise cost and does not adequately leverage the enormous investment in the current QAM plant. Rather, the description above points out what is already available on the QAM side of the hybrid architecture – including the application management layer. This HFC foundation can be expanded by adding in the PON architecture with only minor changes to the existing video infrastructure.

## HYBRID QAM-IP ARCHITECTURE

A QAM-IP architecture enables existing, previously integrated QAM based applications and augments that plant capacity with an IP infrastructure that can grow as demand for bandwidth grows. Deployed applications can easily migrate to utilize the IP back channel as opposed to traditional thin-pipe RF return.

Hybrid architectures provide significantly more forward broadcast and narrowcast capacity than non-hybrid architectures. At the same time, density of homes per GigE served can be increased since the majority of forward video broadcast load can be delivered on the QAM portion of the plant. A hybrid QAM IP architecture using a typical GPON delivery can offer a full 860 MHz QAM plant (or 1 GHz if data rides above 860 MHz). In addition, the OLT acts as a multicasting engine and switch that allows 32 users to share up 2.4 Gbps IP downstream and 1.2 Gbps IP upstream.

A typical Hybrid IP-QAM PON Architecture utilizes a dedicated wavelength for the traditional 860 MHz cable plant services and separate dedicated wavelengths for the IP connection upstream and downstream. These can be carried on the same fiber or on different fibers to a node accumulation point or to the side of a home. Figure 2 illustrates a system deployment to the side of a home.

The Headend or Supernode has typical cable headend functions including derivation of broadcast video services, VOD (may also be delivered directly on IP path), IPPV services, and out of band messaging (may include CAS, EPG, and set-top related configuration services). Services are received, processed (may include ad insertion), encrypted (as required), RF combined and transmitted optically on 1550 nm wavelength. Data services including VOIP and optionally video services such as VOD can be delivered on a separate IP path, transmitted optically on 1490 nm wavelength. 1550 and 1490 wavelengths are wave division multiplexed onto the PON at the output of the OLT. The return path from the ONT to the OLT is IP on 1310 nm and no RF demodulators are required to receive upstream transmission from the set-top. In addition, application communications downstream may also traverse the IP path and an OOB communication channel downstream (such as utilized with some RF return networks) is not required. Specialized headend equipment to enable communication from application server in headend to application client on the Q-IP set-top is eliminated/not required. All connectivity can be provided through an IP network connected through an upstream edge switch/router to the OLT.

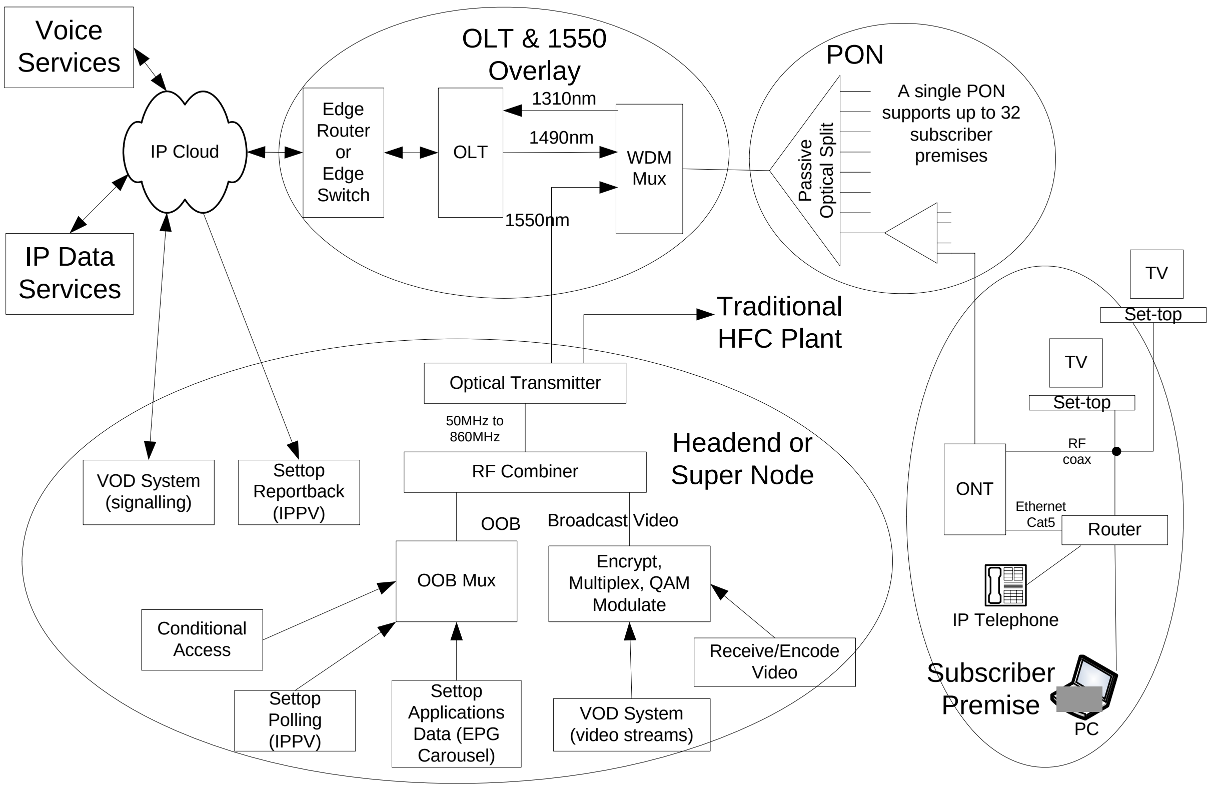


**Figure 2: Functional Hybrid PON Network – IP Return Channel**

In figure 3, the OOB multiplex, set-top application data and VOD system are moved completely over to the IP data path. In addition, new services such as streaming video and even switched video can be provided across the IP path. Lastly, all return channel functions are also moved to the IP path. Carousel data is easily handled with the multicast capability of GPON. Specific requirements for IP multicast data carousel would likely be application dependent.

Application signaling (such as required for VOD) and application service delivery is typically singlecast to the client. By providing for both QAM and IP paths, bandwidth utilization – particularly on the forward path – can be optimized, and flexibility in the manner in which services are delivered is maintained.

The subscriber premise as shown in both figures 2 and 3 assume a home IP network over coax such as Multimedia over Cable Architecture (MoCA) or HomePNA which provides return path channel and enables advanced features like whole home network or multi-room DVR. MR-DVR will allow video to be selected and streamed from a DVR set- top to a non-DVR set-top on the same home network. Set-tops may be provisioned behind a home router or directly from the network / system. Ideally, set-tops sit behind a home router and are provisioned locally on the home LAN. DHCP and DNS functions are used for IP provisioning and application client-server connection. Standard IP protocols govern both provisioning and communications.

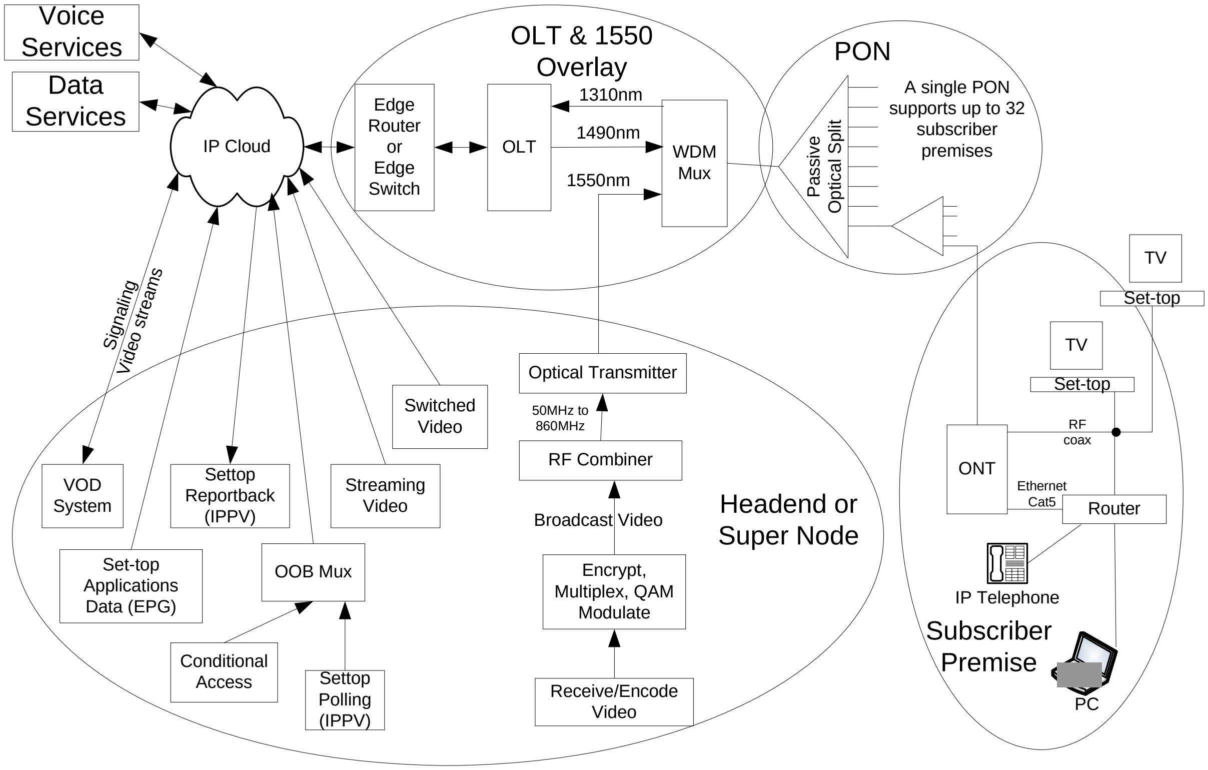


**Figure 3: Functional Hybrid PON Network – IP Services**

Specific deployment requirements may vary based on existing operator infrastructure and longer term desires. Considerations may include existing network and equipment deployments, fiber availability, provisioning requirements, and the like. Benefits of integrating Video on Demand or other singlecast services onto the IP data path, and the use of multicast on the IP path to deliver data or video services provides more bandwidth on the QAM side of the hybrid architecture which will allow high bandwidth HD broadcast services to be delivered without consuming all important bandwidth for future applications on the IP side. Lastly, an IP home network can more readily enable seamlessness in communications using an IP network to a multitude of home and mobile devices – a two way IP pipe to stay connected to the world.

## PROS AND CONS

Reasons for deploying PON based architecture include substantially enhancing bandwidth to the home easing future bandwidth constraints, using it in targeted areas where competitors are focusing, and lower maintenance costs than an active and largely metallic network. FTTP certainly must be considered for greenfield deployments. Challenges include costly rebuilds that may not be popular in the near term to stockholders and potentially stranding some deployed investment assets as new technology is deployed.

## VISION FOR THE FUTURE

An IP enabled cable plant allows for more diverse service offerings as well as provides a framework for the future. Transcoding to formats such as MPEG-4 in the headend will allow for more services in less bandwidth. It is expected that MPEG-4 content would be delivered over IP simply because it is assumed the receiving consumer device will have an IP interface.

Transcoding content in the HE may also enable delivery over other mediums such as wireless/WiFi/cellular/DVB-H networks. In such a seamless network, IP video delivery requires the creation of SPTS, individually transcoded and encrypted and either stored on a headend server or streamed directly to the mobile user.

In a seamless mobility solution, operators collectively deliver a continuous experience of content. Using a network management protocol such as IMS, switching Cable, Wireline and Wireless sessions occurs at the session layer. The transport layer – meaning any transport layer, including a network managed IP or QAM broadcast service – provides transport services only. Applications become transport indifferent. Subscriber access points may be Cable, Data/IP (FTTP), Cellular, WiFi, or WiMax, but are driven by the subscriber and associated applications. In an ethernet centric world, a bandwidth rich IP infrastructure is absolutely necessary to stay competitive and the additional QAM transport layer provides downstream capacity that offloads the IP infrastructure substantially.

Bottom line - Hybrid architectures allow the operator to be ready to fully enable the flexible Connected Home.

# INTRODUCING *Lc*WDM™ – THE NEXT WDM TECHNOLOGY FOR THE CABLE INDUSTRY

Oleh J. Sniezko, Sudhesh Mysore, Charles Barker
Aurora Networks, Inc.

## *Abstract*

*This paper presents a WDM technology for downstream HFC communication. The technology, trademarked as Lc***WDM™***, is a dense wavelength division multiplexing (DWDM) technology based on an extension of the ITU-T Recommendation G694.1*[1] *to the optical O-Band (1260 to 1360 nm).*

*An order of magnitude decrease in wavelength spacing (as compared to CWDM) allows all Lc***WDM™** *wavelengths to be within ±20 nm window about the nominal zero dispersion wavelength ($\lambda_0$) of standard SMF-28 type fiber (ITU-T G.652). This greatly reduces chirp-induced CSO due to fiber dispersion and allows for longer fiber spans.*

*This paper describes the various fiber phenomena – both linear and nonlinear – that limit the fiber distance and the number of wavelengths that can be supported by the two technologies (CWDM and Lc***WDM™***). Descriptions of the detailed testing of fiber nonlinearities and how they affect system performance as well as architectural applications of the Lc***WDM™** *system is also presented.*

*The Lc***WDM™** *system has been carefully engineered to ensure that the fiber dispersion and nonlinearities of a typical (SMF-28) fiber, prevalent in HFC optical node links, do not degrade the link performance below the acceptable levels.*

## INTRODUCTION

### Quest for Bandwidth Never Stops

The demand for bandwidth among the customers of the telecommunications network operators increases continuously. On the other hand, the telecommunications network operators today are looking at adding services and programming to increase revenue potential and to match the demand for HDTV channels. This trend includes addition of bandwidth capacity on fiber to serve residential and commercial customers. HFC broadband network operators are also facing competition from satellite operators in video services and from Telcos in all services. Two of the most leading challengers from this side are Verizon (FiOS) and AT&T (U-Verse).

The continuous demand for increased bandwidth capacity per user is forcing segmentation of the node areas into smaller service groups. In parallel to this trend, the expansion of urban areas force HFC new builds with many nodes deployed every year to serve areas that were not served previously. Both of these trends (demand for higher bandwidth per user and demographical expansion of urban areas) lead to increased demand for fiber capacity through new construction or more efficient fiber capacity utilization. The first choice, new fiber construction, may prove costly in areas already equipped with fiber (at least to the existing serving area boundaries) but which lack dark fiber. Even at modest construction cost per mile and average population density, the cost per household (served or not served) will amount to $150 and more. For downtown areas, this cost can be much higher, even prohibitively so.

The second choice is preferred as it is an order of magnitude lower in cost than the cost of new fiber construction. There are several technologies that improve the efficiency of fiber utilization. These technologies can help to free up fiber on fully occupied optical cable routes for segmentation, serving area expansion and even new service addition (SMB packages).

The most effective and mature technology that has been deployed by many operators is C-Band DWDM (1530-1565 nm) and its distributed DWDM version. This technology allows for practically unlimited narrowcast bandwidth (upper octave in any system, e.g., upper 500 MHz of narrowcast signal in 1002 MHz systems) on 40-plus wavelengths over a single fiber, supplemented by an additional broadcast wavelength on the same fiber. In addition to its highly efficient fiber utilization, this technology allows for practically unlimited distance (100 plus km up to the DOCSIS latency limits).

DWDM technology, aided by TDM in digital reverse links, has been used for increasing the efficiency of fiber utilization for upstream communication in long-distance or on extremely fiber-starved routes.

At the other extreme of the fiber utilization efficiency spectrum are single-wavelength systems: either 1310 nm directly modulated DFB laser links in downstream and upstream transport, which have limited distance of 40 km if no passive loss is present, or 1550 nm eternally modulated laser links in downstream communication.

The most recent WDM technologies that are being introduced by vendors for downstream HFC communication are CWDM and *Lc***WDM**™. These technologies try to close the gap between the two extremes described above.

The coarse wavelength division multiplexing (CWDM) technology was introduced to increase efficiency of fiber utilization in shorter fiber runs and where the lower number of wavelengths per fiber was sufficient. CWDM wavelengths are specified in an ITU-T Recommendation G.694.2[2]. This document specifies up to 18 wavelengths between 1271 nm and 1611 nm with 20 nm wavelength spacing. In most applications on older fiber with higher water-peak loss, only 15 wavelengths can be used.

This technology has been successfully used in the upstream path, both analog (40km links) and digital (where it is not distance limited to 40 km) and for digital downstream, providing Ethernet links (data, T1 over IP and VoIP) for SMB services.

### Downstream CWDM versus *Lc***WDM**™ Distance/Capacity Limits

In downstream communication, where the optical transmitters are used for transporting SCM (sub-carrier multiplexed) analog video and QAM signals, two major challenges significantly limit the use of CWDM technology[3]. One of them is the high level of dispersion (see Figure 1) in SMF-28 fiber at CWDM wavelengths other than 1311 nm. This fiber type has been dominant in access plant deployment in HFC networks. The high dispersion, resulting from the large 20 nm CWDM wavelength spacing, combined with the chirp of directly modulated laser transmitters, results in high levels of CSO. The other technical hurdle is the high level of crosstalk between CWDM wavelengths on the same fiber due to SRS (see Figure 2) and XPM phenomena.

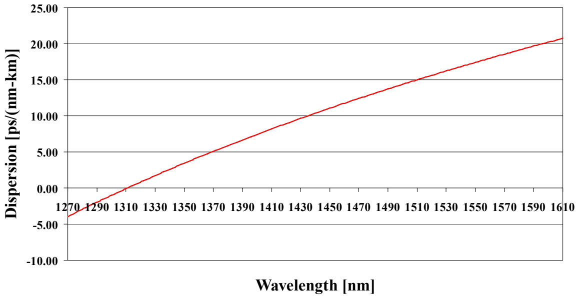
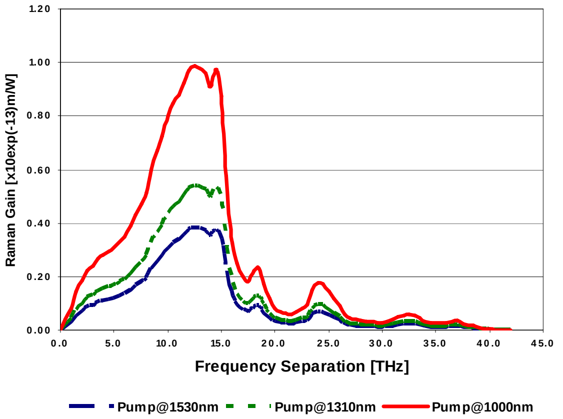


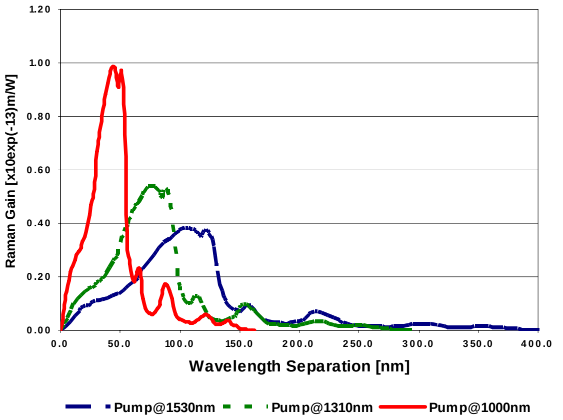
**Figure 1: Dispersion Characteristics of SMF-28 Fiber**

High dispersion limits the fiber span lengths to 15 km even for the two CWDM wavelengths closest to the nominal zero-dispersion wavelength. This limit can be increased only with sophisticated and complex dispersion compensation and/or mitigation technologies, which increase the cost of the transmitter. Use of very low chirp lasers, dispersion pre-distortion techniques (in distance intervals) and passive dispersion compensation techniques are a few examples. However, these are even more complex than techniques used in the DWDM systems described above as they have to be effective with analog channel load across the entire forward RF spectrum, which spans several frequency octaves. Moreover, some of the remedies increase sensitivity to other fiber induced impairments that require additional techniques to neutralize (e.g., low chirp lasers are subject to lower SBS threshold and higher IIN unless SBS suppression and IIN neutralization techniques are deployed).

SRS-induced crosstalk limits the total power allowed into the fiber for wavelength separations larger than 20 nm. Note that in a three-wavelength CWDM system, the separation between the two extreme wavelengths is 40 nm. Since SRS crosstalk increases in severity for larger wavelength separations (up to 13 THz frequency separation – equivalent to about 80 nm wavelength separation at 1310 nm and 110 nm wavelength separation at 1550 nm), this affects CWDM systems and further limits the allowed fiber distance (loss budget) in CWDM systems with more than two wavelengths.



**a) Raman Gain vs. Wavelength Separation**



**b) Raman Gain vs. Frequency Separation**

Based on G.P. Agrawal, *Nonlinear Fiber optics*, 3rd ed. (Wiley, New York, 2001), adapted from R.H. Stolen and E.P. Ippen, *Appl. Phys. Lett.*, 22, 276 (1973).

**Figure 2: Raman Gain for Three Different Pump Wavelengths**

An alternative technology, trademarked as *Lc***WDM**™, is based on extension of the ITU T.G694.1 standard to the optical O-Band (1260-1360nm). All wavelengths used in this system are within a ±20 nm window about the nominal zero dispersion wavelength ($\lambda_0$) of standard SMF-28 type fiber (ITU-T G.652). Each set of wavelengths is engineered for a total difference between the extreme wavelengths not to exceed 15 nm. This significantly lowers the rate of dispersion accumulation with distance and results in lower dispersion-induced CSO in longer fiber spans. Moreover, by placing all wavelengths closer together, SRS crosstalk is minimized even at low RF frequencies.

## IMPAIREMNTS IN OPTICAL LINKS CAUSED BY LINEAR AND NONLINEAR FIBER PHENOMENA

### Nonlinear and Linear Fiber Phenomena

Downstream HFC optical transport links involving analog video and digital video QAM SCM (subcarrier multiplexed) RF channels carry signals that are extremely sensitive to fiber effects that give rise to noise (CNR degradation) and nonlinear signal distortions – primarily second order, which results in CSO degradation. System degradation to varying degrees is observed due to the following linear and nonlinear effects:

1. Fiber Chromatic Dispersion (CD)

2. Interferometric Intensity Noise (IIN)

3. Fiber Nonlinearities
    a. Inelastic Scattering with Phonons
        i. SBS (Stimulated Brillouin Scattering)
        ii. SRS (Stimulated Raman Scattering)

    b. Nonlinear Refractive Index
        i. Cross-Phase Modulation (XPM)
        ii. Self-phase Modulation (SPM)
        iii. Four Wave Mixing (4WM)
        iv. Optical Kerr Effect and Polarization Dependent Loss (OKE/PDL)

4. Higher-order interactions of the above phenomena

The phenomena listed above are well described in the literature. This paper will focus on the impact they have on the performance of different optical link technologies and then concentrate on the *Lc***WDM™** technology based system performance in presence of those phenomena.

The dominant impairments caused by fiber in multiwavelength systems are crosstalk and second order RF distortions (CSO). The challenge in designing the multiwavelength system and its components (active and passive) is to minimize these two impairments while maintaining acceptable CNR, CTB and BER/MER performance of the transported signals. This is achieved by balancing the effects listed above.

### Fiber Phenomena and Multiwavelength Optical Transport Systems

As mentioned above, the three major multiwavelength systems available (some of them widely deployed) for downstream links to optical nodes are:

1. 1550 nm broadcast with externally modulated lasers, together with 1550 DWDM narrowcast overlay on directly modulated lasers

2. CWDM directly modulated lasers

3. *Lc***WDM™** directly modulated lasers

Each of these systems is affected differently by fiber phenomena. Table 1 compares these impacts and lists some remedies implemented in these systems to limit or neutralize the contribution of the fiber phenomena to system performance degradation.

**Table 1: Impact of Linear and Nonlinear Fiber Phenomena on Different Optical Transport Systems to Optical Nodes**

| Fiber Phenomenon | Its Impact on | | | |
|---|---|---|---|---|
| | 1550 BC/1550 DWDM NC Overlay | | CWDM | *Lc***WDM™** |
| | 1550 nm BC | 1550 nm NC (40 wavelengths) | | |
| Chromatic Dispersion | No real impact unless for very long distance (due to SBS suppression system) | Significant impact neutralized by frequency allocation and/or dispersion predistortion (cost impact). Other methods (e.g., selection of low chirp lasers) may require additional remedies. | Significant impact (no simple remedies, low chirp lasers may require SBS suppression circuitry and IIN suppression dithering) | No significant impact if laser chirp is optimized for both chromatic dispersion levels and fiber IIN |
| IIN | No impact (IIN is below the lower forward bandwidth frequency) | The same order of magnitude as for directly modulated 1310 nm DFB lasers | The same order of magnitude as for directly modulated 1310 nm DFB lasers unless low chirp lasers are selected to lower the impact of chromatic dispersion | The same order of magnitude as for directly modulated 1310 nm DFB lasers |
| SBS | Neutralized to the SBS threshold by SBS suppression methods | No impact for standard chirp lasers in the range of optical fiber launch powers | No impact for standard chirp lasers in the range of optical fiber launch powers | No impact for standard chirp lasers in the range of optical fiber launch powers |
| SRS | NA (other wavelengths have different RF frequencies) | No significant impact for QAM channels. | Significant impact for 3 and higher number of wavelengths per fiber | No significant impact on QAM channels and analog channels as long as analog channels are the same |
| XPM | NA (other wavelengths have different RF frequencies) | Not a dominant source of crosstalk | Not a dominant source of crosstalk | Contributes to crosstalk for lower separation wavelengths |
| 4WM | No impact | No impact | Limited impact (under very unlikely scenario) | Measurable impact neutralized by the system design |

## MECHANICS OF DEGRADATION OF OPTICAL LINK PERFORMANCE

### Chromatic Dispersion and Related Chirp Penalties

Chromatic dispersion refers to the variation in group velocity with wavelength. This characteristic is well-described by the Sellmeier equation, and a typical dispersion characteristic of SMF-28 fiber is shown in Figure 1.

Variations in group velocity with wavelength combined with periodic wavelength shifts due to laser chirp in a directly modulated transmitter results in CSO distortion of analog signals[4]. This dispersion/chirp-induced CSO increases very rapidly with dispersion, even for a 20km system, as shown in Figure 3. Several plots are shown for laser chirp (at 100% OMI modulation) ranging from 0.5 GHz to 10 GHz.

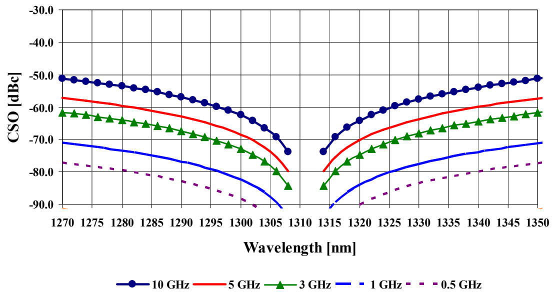


**Figure 3: Chromatic Dispersion Accumulation with Wavelength (20 km of SMF-28 Fiber)**

Note that even with CSO specifications relaxed to −60 dBc, CWDM systems are limited to three wavelengths (1291nm, 1311nm and 1331nm) due to dispersion/chirp-induced CSO for a typical 100% OMI laser chirp of 5 GHz for 8 dBm lasers. This is under the assumption that the average zero dispersion wavelength $\lambda_0$ in the link is equal to the nominal zero dispersion wavelength. The CSO will degrade further for one of the CWDM wavelengths if the $\lambda_0$ in the link is offset from the nominal value. Moreover, this CSO will cascade with optical link CSO generated by other mechanisms (including laser transmitter CSO).

### Crosstalk Due to Stimulated Raman Scattering

SRS refers to optical gain experienced by one wavelength signal at the expense of another shorter-wavelength signal. Although the transfer of optical power is from the short-wavelength signal to the longer-wavelength signal, this transfer is modulated by the product of the optical power of both wavelengths, leading to leakage (crosstalk) of the RF signals in both directions at almost the same level. The coefficient of the product term is denoted by ***g***, the Raman gain coefficient. The Raman gain coefficient increases as the frequency difference between the two interacting signals increases, and reaches a peak at around 13 THz (see Figure 2a). Although the gain peak occurs at a fixed frequency difference of 13 THz when plotted against frequency shift, the position of the peaks vary for different pump wavelengths (i.e., the shorter of the two interacting wavelengths) when plotted against wavelength shift as in Figure 2b. The Raman gain increases for wavelength separation up to 80 nm for a pump near 1310 nm, and up to 110 nm for a pump near 1550 nm. Furthermore, the amplitude of the Raman gain decreases according to well-known scaling rules as the pump wavelength moves towards longer wavelengths as shown by the three curves in Figure 2b.

SRS-induced crosstalk can be extremely high at low RF frequencies as shown in Figure 4 for 55 MHz. For typical system parameters (+10 dBm/ch launch power, 1310 nm pump, 25 km fiber distance) the SRS-induced crosstalk can exceed −40 dBc for wavelength separations larger than 30 nm.

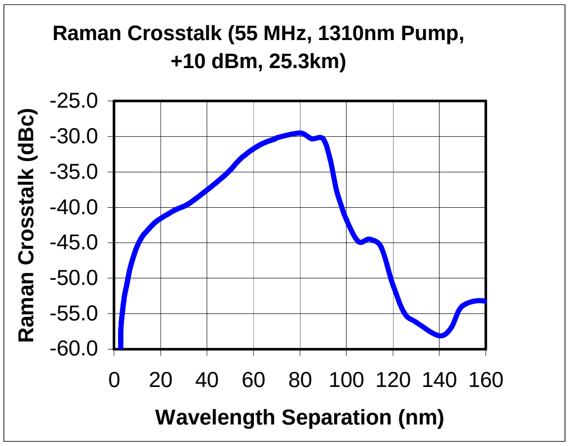


**Figure 4: SRS-Induced Crosstalk**

Cross-Phase Modulation (XPM)

At high RF frequencies (around 500 MHz) and small wavelength separation, the crosstalk does not vanish as predicted by Figure 4 due to the effect of another fiber nonlinearity, namely cross-phase modulation (XPM). This nonlinearity is caused by the nonlinear refractive index – modulation of the fiber refractive index – which causes modulation of one wavelength to induce phase changes in the other wavelengths. This induced phase modulation in conjunction with fiber dispersion results in crosstalk from one wavelength to another.

Fiber dispersion is therefore a necessary ingredient for XPM-induced crosstalk; XPM-induced crosstalk increases with increasing fiber dispersion and decreasing wavelength separations. Like other refractive index nonlinearities, XPM-induced crosstalk depends strongly on the group velocity mismatch (or "walkoff") between the affected wavelengths, which is approximated by the product of the fiber dispersion times the wavelength separation.

The actual crosstalk observed in a multiwavelength system depends on the relative magnitude and phase of the SRS-induced and XPM-induced crosstalk[5]. While optical crosstalk from other wavelengths appear at the same location (under each analog carrier) as CTB distortion (which arise from other RF carriers in the same optical wavelength), the CTB can be differentiated from the crosstalk as it is composed of a large number of discrete RF beat products that are slightly offset from each other.

At this high level of crosstalk (see SRS crosstalk in Figure 4 and XPM crosstalk addition), two major network design limitations must be observed:

1. All analog signals on multiple wavelengths on the same fiber must be the same (the same content at the same frequencies). This limitation allows for lowering the requirements for crosstalk levels.

2. Even for the same signals, multipath effect must be accounted for (from Figure 5, for −35 dBc of crosstalk, only 100 to 250 ns differential delay is allowed for signals traveling on different wavelengths on the same fiber – this translates to 80 to 200 feet of a typical RF cable in the headend combining network).

Four Wave Mixing (4WM)

Four-wave-mixing is another result of the fiber refractive index nonlinearity that results in three optical signals at distinct frequencies $f_i$, $f_j$, and $f_k$ interacting as they propagate along a fiber to give rise to a fourth optical signal at frequency $f_{ijk} = f_i + f_j - f_k$ – similar to a third-order "beat product" in RF systems. The mixing product at frequency $f_{ijk}$ is said to be non-degenerate if $i \neq j \neq k$ (i.e., the mixing product is generated by three distinct signals). It is possible for 4WM to occur with only two optical signals present (at frequencies $f_i$ and $f_j$), giving rise to a so-called "partially-degenerate 4WM" (PD4WM) products at frequencies $f_{iij}$ and $f_{jji}$.

The mixing efficiency of the 4WM process is maximized when the phase-mismatch parameter is zero. The relative 4WM power $P_{ijk}/P_{out}$ generated by three polarization-aligned signals with the same fiber launch power $P_{in}$ and fiber output power $P_{out}$ can be easily calculated from a simple equation[6]. The same slightly modified equation can be used for the case of PD4WM mixing products.

The equations for the mixing products are valid only so long as they are small enough that the "pump power" in the original signals is not depleted. Even so, they predict that very large mixing products can be generated at power levels of +10 dBm/ch if the fiber dispersion is low since the corresponding phase-mismatch parameter is then also small.

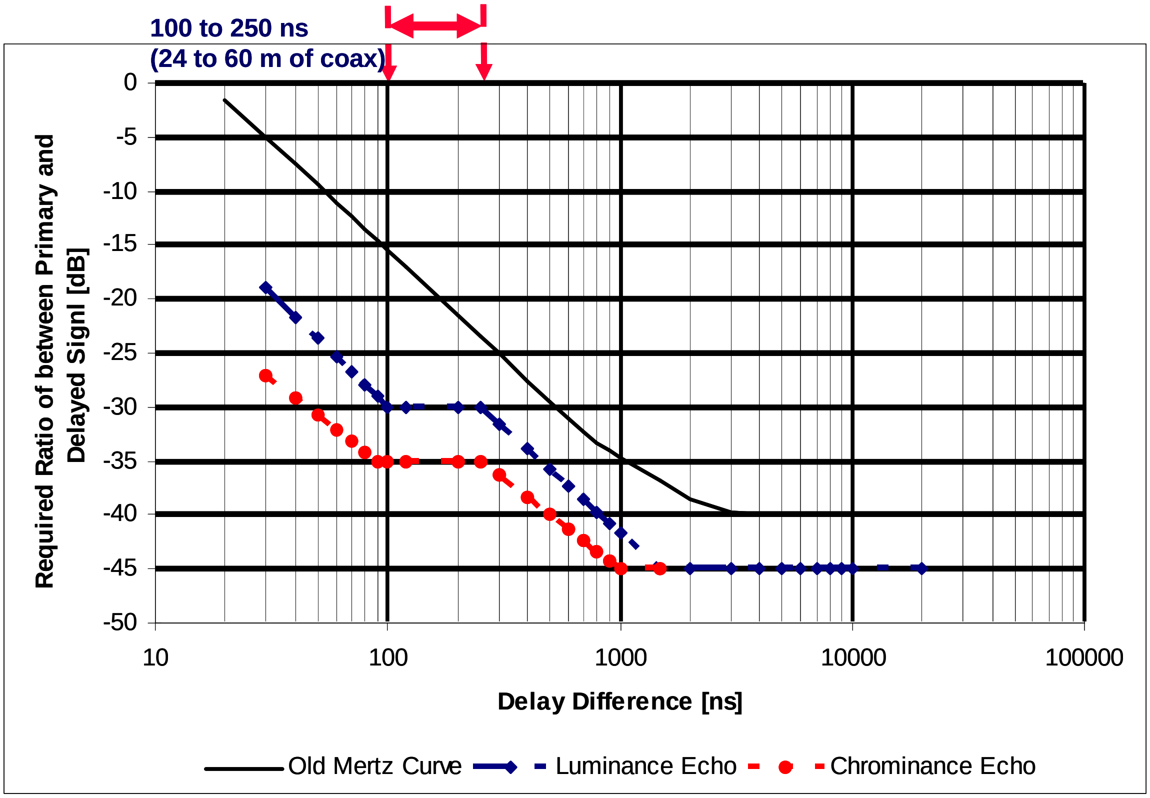


**Figure 5: Echo Tolerance Curves for Color TV Programming Displayed on Large Screen (Based on Rogers Published Test Results – Gary Chan and Nick Hamilton-Pierce and in Line with Dan Pike's[7] Published Test Results)**

Figure 6 shows the calculated power levels of 4WM and PD4WM mixing products as a function of the phase-mismatch parameter for +10 dBm/ch launch power and 20 km fiber distance.

Note that relative mixing products as high as –20 dBc can be generated if the system is not designed properly. Figures 7a and 7b show the input and output spectra, respectively, of an actual 5-wavelength system, demonstrating that such high levels of 4WM and PD4WM can indeed be obtained in not optimized systems.

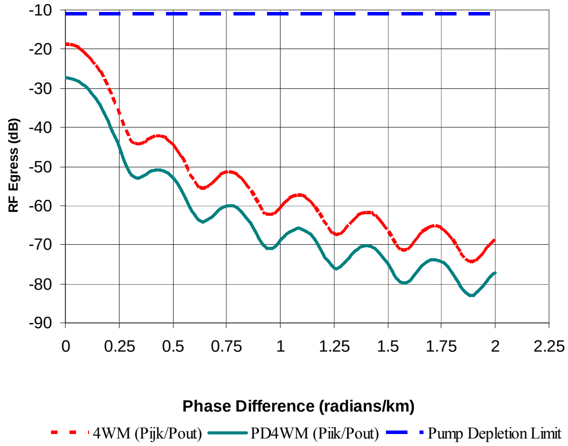
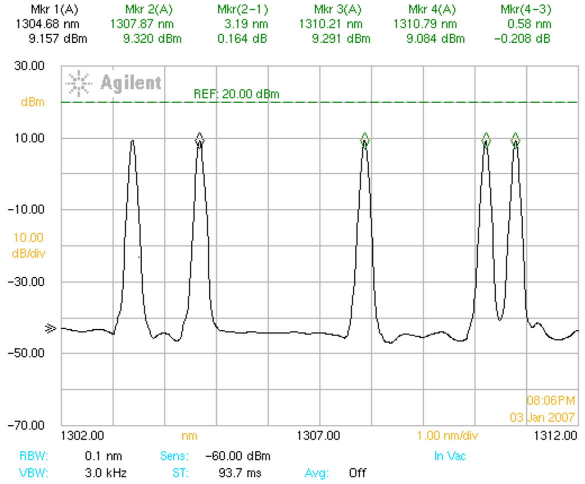
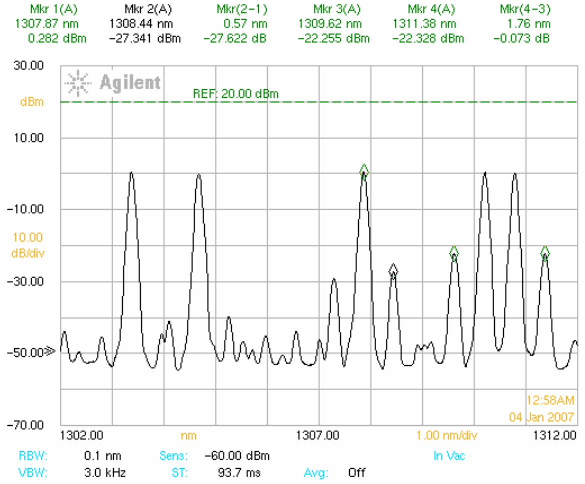


**Figure 6: Dependence of 4WM and PD4WM products on phase-mismatch parameter.**



**a) Input Signals**



**b) Output Signal Spectrum**

**Figure 7: Spectra of 5-Wavelength System with High Levels of 4WM and PD4WM Products**

With N=5 wavelengths, there are a total of N2(N-1)/2 = 50 mixing products generated. Even though it is possible to design the system so that none of the mixing products fall on any transmission wavelength and hence there is no crosstalk caused by ingress of 4WM power into any wavelength, there will still be egress of power out of each wavelength.

In fact, each wavelength is supplying power into a total of (N-1)∗(3N-2)/2 = 26 mixing products. Since this egress power is proportional to $\boldsymbol{P_i P_j P_k}$, where each of the ***P*** term is composed of 79 analog channels, egress of power out of a wavelength generates a multitude of second and third order intermodulation distortion products and degrades the CSO (and to a lesser extent, the CTB) performance of the wavelength.

## *Lc*WDM™ SYSTEM TEST RESULTS

### Test Setup

*Lc*WDM™ systems with two to eight wavelengths were tested in a setup similar to that presented in Figure 8. Each transmitter is loaded with 79 channels of CW carriers from 54 to 552 MHz and 75 channels of 256-QAM signals from 552 1002 MHz. Matrix 1 is used to drive the wavelength under test while Matrix 2 drives the other transmitters with other wavelengths (all wavelengths are tested for performance). This is necessary to measure crosstalk from the other optical wavelengths due to SRS, 4WM and XPM. If it is desired to test only the CTB for the particular wavelength, then the same signal source can be connected to all transmitters.

A polarization controller is attached to each transmitter and the worst-case settings are found prior to each measurement. Testing was performed on several dozen spools of fiber covering a wide range of fiber parameters.

Initial testing started with just two wavelengths in order to test the basic theory. With just two wavelengths present, it is quite easy to model the performance of each wavelength. This allows calculating the strength of each mixing products, in which the particular wavelength is involved and hence the CSO and CTB degradation contribution of each of them. The baseline performance of each single wavelength without fiber is used as a reference.

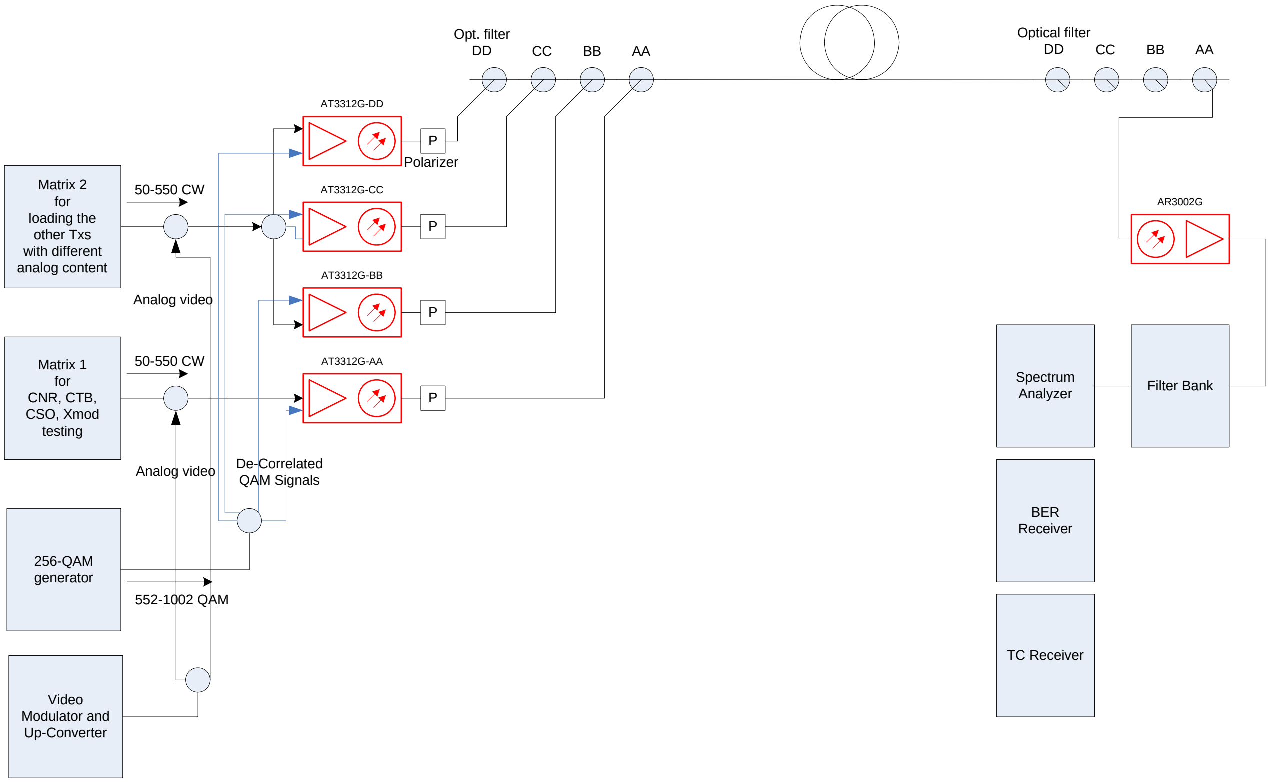


**Figure 8: Test Setup for Multi-Wavelength *Lc*WDM™ System**

Figure 9 shows a 3-D surface that describes the strength of the PD4WM products for a system where one channel is fixed at $\lambda_2$=1308 nm. The x-axis is the zero-dispersion-wavelength $\lambda_0$ and the y-axis is the wavelength $\lambda_1$ of the second channel. The fiber distance has been assumed to be 25 km and the power level to be +10 dBm/ch.

Note that as the phase-mismatch approaches zero, PD4WM product power rises sharply. Also, when wavelength separation is small, the PD4WM power remains high over a much larger range.

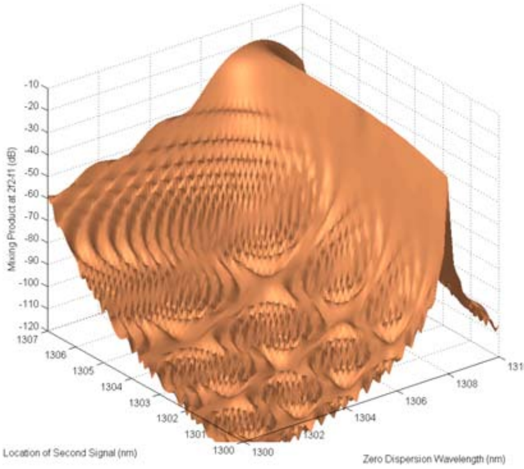


**Figure 9: Modeling PD4WM Power**

Figure 10 shows the results of testing with a fixed wavelength at $\lambda_2$=1311.5 nm. The x-axis is the wavelength $\lambda_1$ of the second channel, which is varied from 1307 nm to 1314 nm. Good correlation is seen between the observed PD4WM power and the calculated values.

Figure 11 shows the same system but over a different fiber. Measurements are made as $\lambda_1$ is varied from 1307 nm to 1314 nm. Again, there is very good correlation between theory and measurements. Note also that the PD4WM power level remains high over a much wider range of $\lambda_1$ values.

Tests on 3-, 4-, 5- and 6-channel systems also showed very good agreement with theoretical models. The tests will be also expanded to 8 wavelengths.

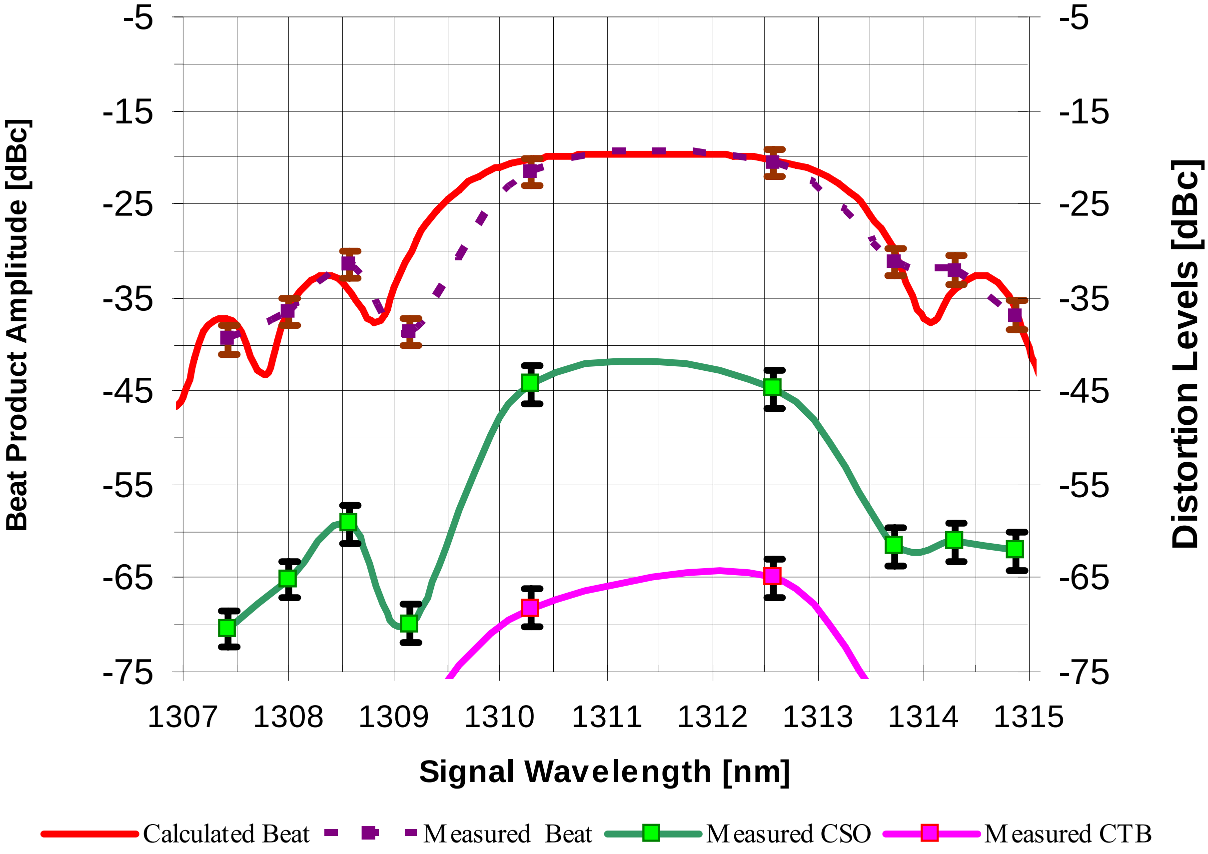


**Figure 10: PD4WM Power versus Wavelength Separation on Fiber with Distant $\lambda_0$**

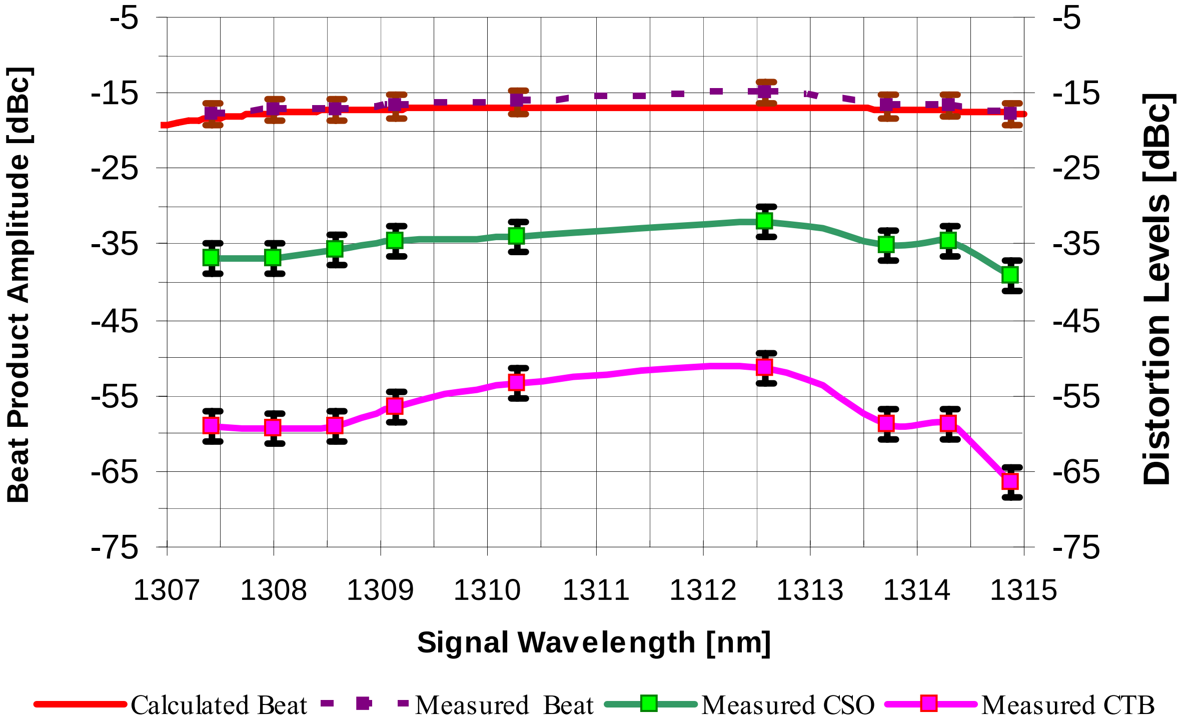


**Figure 11: PD4WM Power versus Wavelength Separation on Fiber with Close $\lambda_0$**

## *Lc*WDM™ SYSTEM DESIGN CHALLENGES

The *Lc***WDM™** design took into account the theoretical modeling results and extensive testing results of the nonlinear and linear fiber phenomena and their effect on analog and digital QAM signal performance transmitted on multiple wavelengths. Two major parameters that posed the design challenges were crosstalk and second order (CSO) distortions. The *Lc***WDM™** system was optimized to meet the requirements for these two parameters while securing acceptable performance for CNR, CTB and MER/BER. Achieving this goal required balancing the fiber effects described previously[8].

### Crosstalk

Three major fiber phenomena cause inter-wavelength crosstalk:

1. SRS
2. XPM
3. 4WM ingress

The optimization of the first two is difficult as the requirements to minimize SRS crosstalk most of the time conflict with the requirements to minimize XPM crosstalk. Fortunately, SRS crosstalk is highest at lower RF frequencies and XPM crosstalk is highest at highest RF frequencies. Moreover, *Lc***WDM™** system has a significant advantage as the separation between wavelengths are relatively small (in comparison to CWDM system) and hence SRS crosstalk is low while XPM crosstalk affects mostly QAM channels. As long as the two assumptions:

1. All analog signals on multiple wavelengths on the same fiber are the same, and
2. Multipath effect is accounted for;

(as explained above) are met, these two effects do not limit the application of the system. Moreover, the third effect is completely eliminated by *Lc***WDM™** system design. Additional contributors to crosstalk (e.g., OKE-PDL interaction) are being controlled by system component selection and specification.

Second Order Distortions (CSO)

Two major phenomena contribute to degradation of the CSO in analog multiwavelength links:

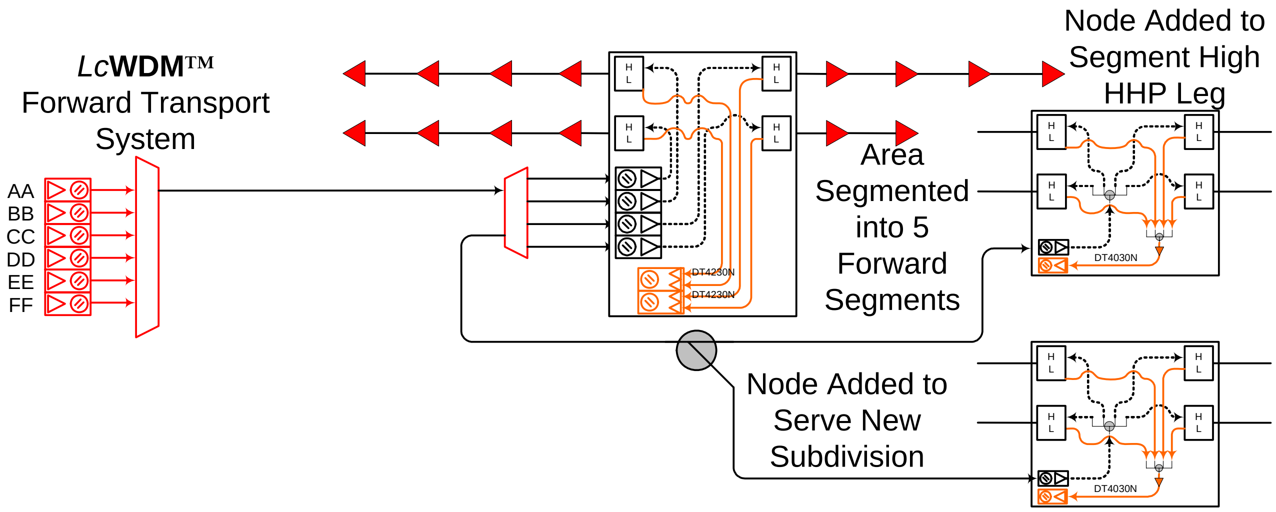
1. Linear effect of chromatic dispersion resulting in chirp-related penalty, and
2. 4WM egress.

The first effect is minimized in *Lc***WDM™** system by the proximity of the wavelengths to the zero-dispersion-wavelength $\lambda_0$ of SMF-28 fiber that is predominant in optical links to the nodes. Moreover, the laser chirp is managed to optimize both CSO and IIN contribution. The second effect is minimized by the system design. The 4WM egress in extreme cases causes CTB degradation but if the CSO degradation is under control, CTB degradation is negligent (see Figures 10 and 11).

## APPLICATION OF *Lc***WDM™** SYSTEM FOR NODE SEGMENTATION AND SERVICE AREA GROWTH

Figures 12 illustrates the use of *Lc***WDM™** technology for downstream transport in conjunction with CWDM technology and digital technology for upstream transport to segment an initial node into 5 forward and reverse segments using only two fibers. Additional node is added to split a node leg that served a high number of households. Sixth node is added (in future) for the new growth area. The upstream fiber is not shown here but at the discretion of the operator, the same fiber that is used for downstream transport can be used for upstream transport (see Figure 13). Alternatively, the second fiber can be used for upstream transport (this will yield longer reach as combining and de-combining filter loss is eliminated from the downstream path).

**a) Initial Node Service Area**



**b) Node Service Area Segmented into Fiver Areas with 6th Node Added to Serve New Subdivision**

**Figure 12: *Lc*WDM™ System Applied for Area Segmentation and Service Area Expansion**

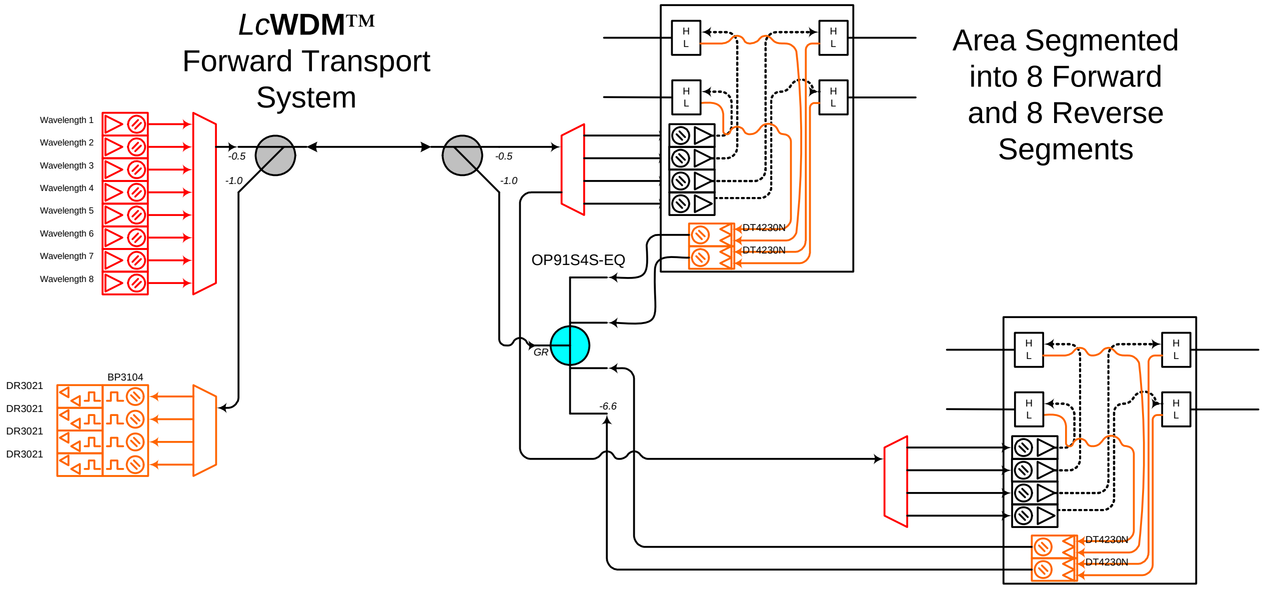


**Figure 13: *Lc*WDM™ Downstream System in Combination with CWDM/Digital Upstream System to Serve 8 Forward and 8 Reverse Segments on Single Fiber**

## CONCLUSIONS

*Lc*WDM™ systems close the performance and cost gap between single wavelength optical links based on 1310 nm directly modulated DFB laser technology and DWDM 1550 nm systems based on externally modulated technology for broadcast links and directly modulated DFB technology for narrowcast overlay. They are much more immune to both SRS-induced crosstalk and dispersion-induced CSO than CWDM systems. This allows for extending their reach and increasing the number of wavelengths per fiber. The fiber length dispersion limits are doubled. The *Lc*WDM™ systems allow for cost-efficient node area segmentation on a single fiber. In combination with CWDM (15 wavelengths after exclusion of water peak and 12 wavelengths after exclusion of CWDM windows occupied by *Lc***WDM™** wavelengths traveling on the same fiber) and digital TDM technology in upstream, *Lc***WDM™** systems can support 6 to 8 forward segments and 24 upstream segments on a single fiber. This allows not only for segmentation of the area but also for node addition in the growth areas without the need for fiber construction on the existing fiber routes. In fact, it allows for fiber recovery to support commercial and other services and generate incremental revenue.

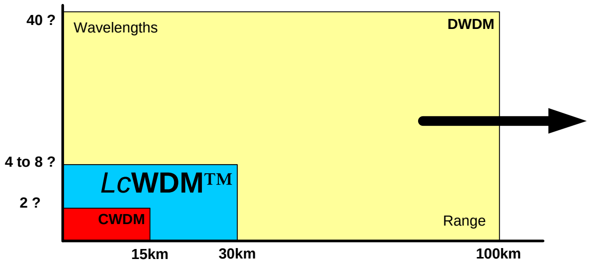


**Figure 14: Capacity/Distance Limits of Various WDM Technologies for Downstream HFC Optical Transport**

A critical advantage of the technology described in this paper is its simplicity. The transmitters carry entire forward load from 54 to 1002 MHz on the same wavelength and in this sense are equivalent in operation to 1310 nm directly modulated DFB transmitters. The *Lc***WDM™** system design follows the same rules as the design of a system with 1310 nm DFB lasers, taking into account the required input level to the node and power budget of the link (including fiber and passive component losses). After accounting for these design rules, the 30 km fiber length limit is rarely reached if more than two wavelengths are used for downstream transport.

## ACKNOWLEDGMENTS

The authors wish to express deep gratitude to Shamino Wang, Samuel Chang and Ricardo Villa of Aurora Networks, who spent hundreds of hours testing all aspects and components of the *Lc***WDM™** systems. Without their dedication, the systems would have taken much longer to develop since its conception in May of 2006. The authors also appreciate the support of the Aurora operations team in building countless iterations of the system components. Extreme gratitude also goes to Wim Mostert who first coined the idea in casual conversation on the way between the Aurora headquarters and the company rental apartment. After peeling the layers and analyzing all possible pitfalls, we realized we have a jewel on our hands, technology that benefits the entire industry.

---

[1] ITU-T G.694.1 (2002), *Spectral Grids for WDM Applications: DWDM Frequency Grid*

[2] ITU-T G.694.2 (2002), *Spectral Grids for WDM Applications: CWDM Wavelength Grid*

[3] T. Werner and O. J. Sniezko, *Exploiting HFC Bandwidth Capacity To Compete With FTTH*, 2006 NCTA Technical Papers

[4] 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 Photonics Technology Letters, Vol. 3, No. 5, May 1991

[5] M. R. Phillips and D. M. Ott, *Crosstalk Due to Optical Fiber Nonlinearities*, Journal of Lightwave Technology, Vol. 17, No. 10, October 1999
in WDM CATV Lightwave Systems

[6] R. W. Tkach, A. R. Chraplyvy, Fabrizio Forghieri, A. H. Gnauck, and R. M. Derosier, *Four-Photon Mixing and High-speed WDM Systems*, Journal of Lightwave Technology, Vol. 13, No. 5, May 1995

[7] D. Pike, *The Effects of Reflections*, 1989 NCTA Technical Papers

[8] H. A. Blauvelt, P. C. Chen, N. Kwong, I. Ury, *Optimum Laser Chirp Range for AM Video Transmission*, 1992 NCTA Technical Papers

# Making FTTH Compatible with HFC

James O. “Jim” Farmer
Wave7 Optics, Inc.

## *Abstract*

*The evolution from Hybrid Fiber-Coax (HFC) to Fiber-to-the-Home is explored. The similarities between the technologies are noted, and we show how an operator who has been deploying HFC can start to deploy FTTH where it makes sense to do so, without excessive investment.*

## INTRODUCTION

Cable operators are starting to deploy fiber-to-the-home (FTTH) in Greenfields, as well as for business applications. The preferred deployment strategies employ one of two standards for FTTH: EPON or GPON. This paper applies to both. Since the FTTH deployments are small compared with the HFC deployments, it is important that minimal or no changes to the operator’s processes or headend technology are necessary in order to integrate the new with the existing.

This paper will review recent experience in integrating FTTH into HFC systems. We’ll cover such topics as data management, video in the two systems, and how voice is handled. There are certain architectural features of FTTH systems that can ease the initial cost of deployment. These features will be explored, showing how to make as much of the deployment cost as possible be based on success.

Cable TV HFC networks are much better equipped to provide all modern residential telecommunications services than are the telephone company’s twisted pair networks. Therefore, we expect that cable operators will continue to use their HFC plants for some time where they exist now. However, HFC is not competitive with FTTH networks that are being installed now, so we recommend that cable operators seriously consider installing FTTH for Greenfield situations, where they must install something anyway. We shall show that it is easy to operate systems that are partially HFC and partially FTTH.

## WHAT IS FTTH?

You can think of FTTH as the logical extension of HFC networks. Cable operators for years have been pushing fiber closer and closer to the home, to enjoy the benefits of greater quality and reliability, and lower maintenance costs. FTTH represents the end game, where fiber has been pushed down to a node size of one home. When fiber is pushed this far, some interesting changes to the way the fiber is used are possible. In FTTH systems:

-One fiber carries downstream and upstream signals, minimizing splicing/connection costs.

-Data is carried on separate wavelengths from RF, resulting in incredible data bandwidths without having to sacrifice any video spectrum. Since data is carried on separate wavelengths from RF, RF spectrum is freed up for more video. There is no tradeoff between data bandwidth and the bandwidth available for video. The entire RF bandwidth from 54 to 1,000 MHz is available for video.

-Reliability is enhanced due to the complete lack of coax in the plant. There are no corrosion issues, no mechanical pull-out issues, no cracked shields, no sheath currents, etc.

-While we are not going to advise you on legal matters, we note that the maximum RF signal level occurs at the side of the home, and is typically less than 20 dBmV. Thus, we would not expect composite leakage index requirements to exist.

-There are no amplifiers to balance: the response you get at the headend is the response you get at the home. Always.

-As shown below, the manner in which upstream signals are handled results in no upstream noise funneling.

-Since the plant is all-dielectric, there is no ingress into the plant.

-Typically there are no actives in the field, and no power connections. Reliability is increased and maintenance is reduced.

-Service disconnect is an integral part of many FTTH systems, removing the need for truck rolls when a subscriber disconnects or connects.

-Expensive CMTSs are not needed in most cases. Data connections are usually gigabit Ethernet to your headend switch.

-Systems come with integral element management – it is not an extra-cost add-on.

-With the exception of the set top box, all equipment is normally located outside the home, where the operator has access. The outside equipment is normally powered from the home (with battery backup for lifeline voice), minimizing the operator's electric bill.

-The FTTH network is architected to carry IPTV should the need arise. IPTV and broadcast video may be carried at the same time.

## TYPICAL FTTH ARCHITECTURE

Figure 1 illustrates a typical FTTH physical architecture. The video headend is just as it is for HFC. The voice facility is not changed from HFC – FTTH networks can support the same protocols specified by CableLabs. The switch(s) used to interconnect data in the headend are the same: Gigabit Ethernet links are usually used to connect to the FTTH equipment. No CMTS is needed for the FTTH portion. A new unit, called an Optical Line Terminal (OLT) is used as the data interface with the FTTH plant. Logically, you can think of it taking the place of the CMTS, though the analogy is not precise. A typical OLT is pictured, and serves up to 2300

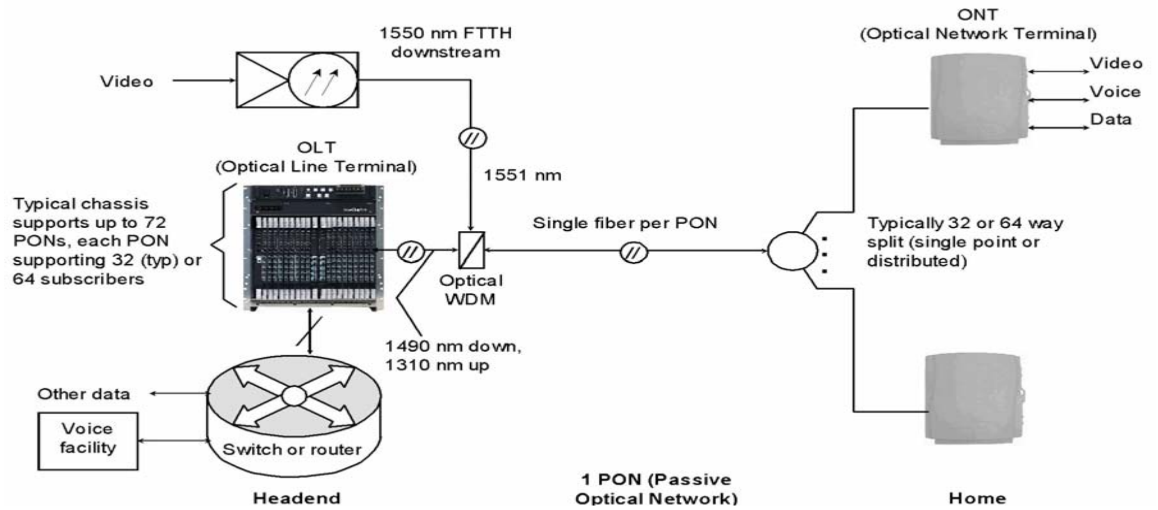


**Figure 1. Typical FTTH Architecture**

subscribers in a typical configuration, in under 1/3 of a standard rack. That gives you up to 6,900 subscribers served per rack.

The video headend drives an externally modulated 1550 nm optical transmitter, whose output is split and amplified (not shown) as necessary to supply signals to all subscribers. Only one transmitter is normally needed. The output of the 1550 nm transmitter is wave division multiplexed (WDM'ed) with the data carriers handled by the OLT. Standard wavelengths used in FTTH systems are downstream data on 1490 nm, upstream data on 1310 nm, and video on 1550 nm (any suitable standard ITU wavelength).

Normally the network to the home is all passive, consisting only of single mode fiber optic cable of the type cable systems normally use, and passive splitting. Distributed splitting (taps) can be used, but it is more common to use concentrated splitting, where all splitting is done in one or two locations. The network between the headend and the home is called a passive optical network, or PON. While it is technically feasible to serve 64 subscribers per PON, serving 32 subscribers per PON is more common. The cost differential, while not zero, is not great.

The termination at the home is called an optical network termination, or ONT. A number of different ONTs are available, but the workhorse configuration is a box that is mounted on the side of the home. It has an RF output for video, one or more data ports (10/100Base-T), and one or more voice ports (RJ-11 with terminal block). RF return support for set top upstream communications is available from some vendors. All ports are controlled through the element management system (EMS), and may be remotely turned on and off. The ONT is usually powered from a power supply mounted in the home. Configurations are available that allow mounting the power supply outside of the home, and taking power from a special ring installed on the power meter. Power can be taken either before or after the meter. If lifeline voice service is provided, a battery is used to provide a minimum of 8 hours standby/2 hour talk time. The battery condition may be monitored through the EMS.

Other configurations of ONTs may be available. For example, a video-only ONT may be available for customers who take basic video only. For business applications, there may be small, indoor-only ONTs that provide data ports only. Some ONTs provide only one data port and one voice port. Other ONTs provide multiple ports. Finally, ONTs designed for multiple dwelling units (MDUs) may be available, with ports to serve several apartments from one ONT. All ports are controlled through the EMS.

## THE ONT RF PORTION

Figure 2 illustrates the ONT with some optional features included. The single fiber from the PON enters the ONT at an optical wave division multiplexer (WDM). The top port of the WDM passes the 1550 nm broadcast signal to a broadcast RF receiver not unlike the receiver portion of an HFC node. The output level is usually set to allow several TVs to be connected. Levels of +12-18 dBmV are common. Recall that the ONT is mounted on the house, so this level exists at the side of the home, not at the pole. AGC is often used to compensate for optical level variation. The AGC may sense the RF signals level, but in this application it is more satisfactory to sense the optical level and correct the RF level.

If RF return to support set tops is used, the RF diplexer separates the upstream signal, digitizes it, and sends it to the headend, where a special device is used to reconstruct the RF to supply it to the set top control system. There are some variations in which the signal is demodulated at the ONT and the demodulated packet is transmitted to the control system. No data is transmitted except when there is a transmission from a set top in the home. Typically the threshold for determining when data is present is set toward the high end of the range of which the set top is capable (accounting for passive splitting/combining in the house), to minimize the effect of any noise in the home. When a set top transmits, usually one or a very few data packets are sent through the data portion of the FTTH system. You can see that this method of transmitting the upstream RF precludes noise funneling in the upstream direction. If a home has a noise problem it would be easy to identify based on the IP address of the transmission.

## DATA PORTION

The bottom port of the WDM (Figure 2) passes the 1490 nm downstream data signal to an optical transceiver, and accepts from the transceiver the 1310 nm upstream optical signal. Data modulation is done at baseband (on-off keying) of the optical transmitter. The data rate at the transceiver is 1 Gb/s or higher, spread over only 32 subscribers. Compare this to DOCSIS, which on a per-RF-channel provides about 37 Mb/s downstream and usually about 8 Mb/s upstream, spread over however many subscribers, maybe 100 or so, if there is a typical application.[i] The upstream data rate is 1 Gb/s, so you can offer outstanding upstream bandwidth as well as downstream. This works especially well for gaming and peer to-peer applications. Following the data transceiver is the digital processing section. The front-end is a protocol-specific chip that manages data transmission according to either the EPON or GPON standard. The digital processing back-end provides one or more 10/100BaseT connections for data, and also includes media conversion to support standard analog phone lines.

In Greenfield applications many homes are now being built with category 5 cable so that you can connect the cable directly to the ONT. If desired, a standard cable modem gateway is connected to provide a firewall and

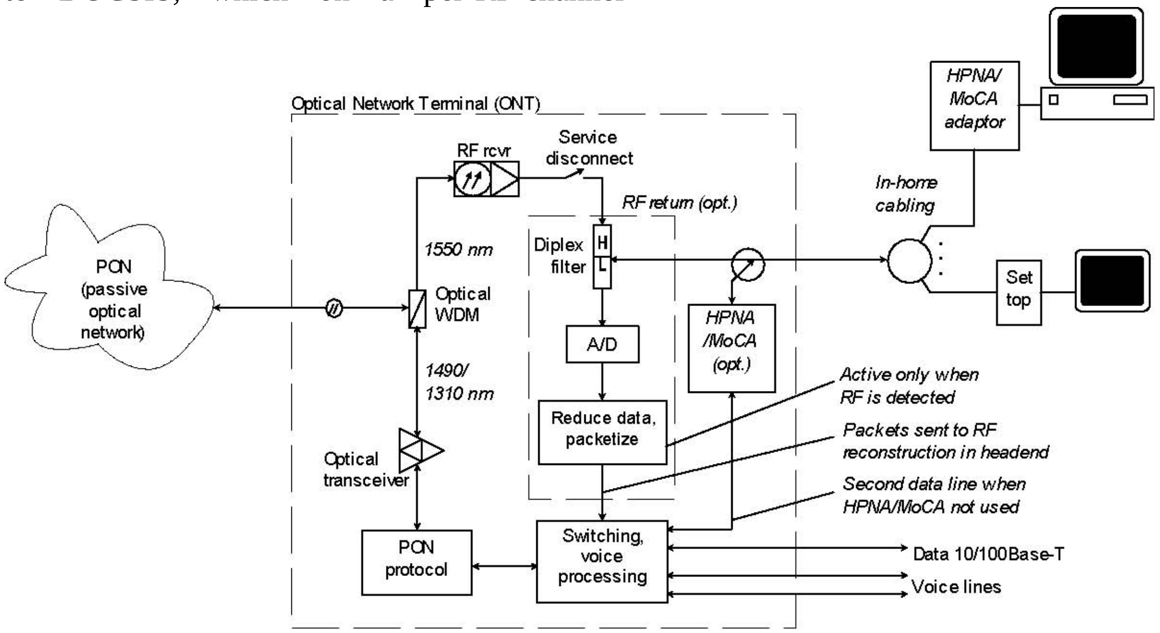
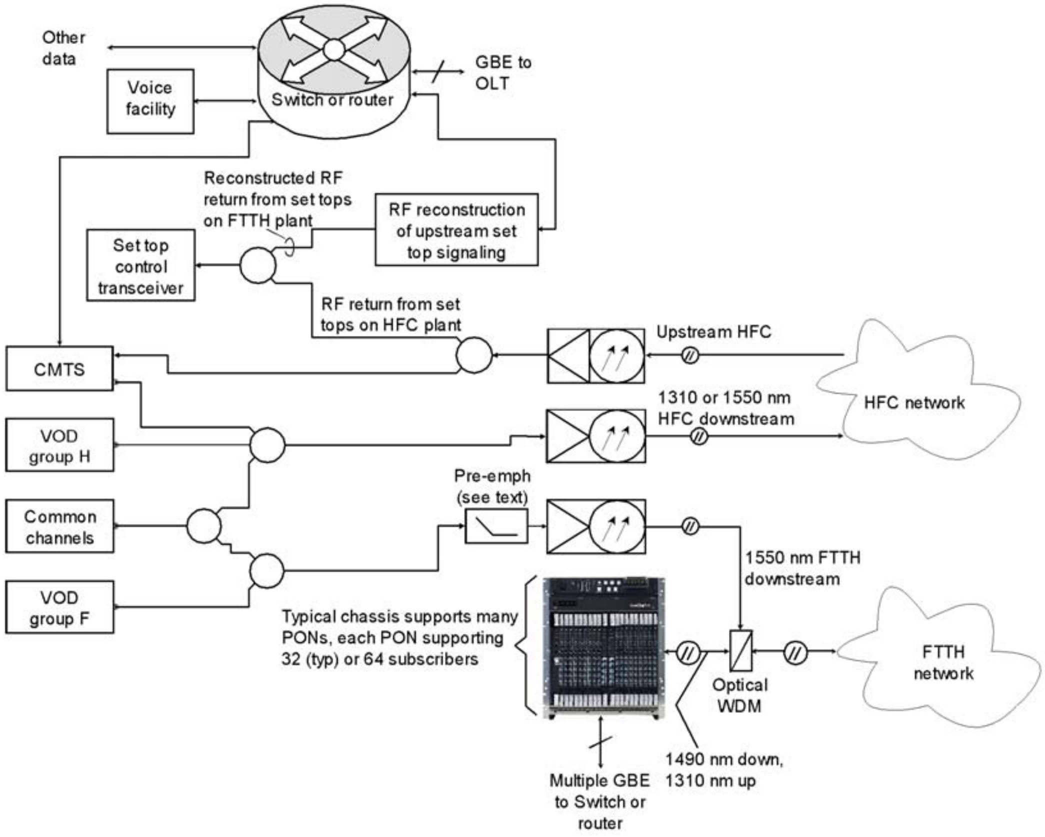


**Figure 2. Optical Network Terminal (ONT)**

DHCP services for the home. For existing construction, there are several options. Many operators are installing cat5 cabling, but some are opting for one of the standards available for putting data over coax. These standards are HPNA over coax, and MoCA. More on those options later.

## MORE ON THE HEADEND

Figure 3 illustrates the headend in more detail. The headend shown is feeding both HFC plant (toward the center right) and FTTH at the bottom. For illustrative purposes, we have shown different VOD zones for HFC and FTTH, but of course this will vary depending on local needs. At the top is the headend switching/routing facility connected to the voice facility and other data services. It is connected to the CMTS for the purpose of data transmission on the cable plant, including voice. Following standard practice, the downstream signals to the HFC plant and the upstream signals from that plant are transported to and from the node on separate fibers. lower frequency channel amplitude by about 4 dB at channel 2, decreasing the boost at higher frequencies, to the point of little boost above about 200 MHz. The reason for this boost is that fiber optic cable exhibits a nonlinearity called Stimulated Raman Scattering (SRS) that causes the downstream data, carried at 1490 nm (in accord with all modern standards), to crosstalk into the video channel, reducing carrier-to-noise ratio (C/N).



Typical chassis supports many PONs, each PON supporting 32 (typ) or 64 subscribers

**Figure 3. Headend serving both HFC and FTTH Networks**

The effect is frequency-dependent, affecting lower frequencies the most. By boosting the low-frequency channels, the C/N can be maintained equal to or better than that delivered by short HFC amplifier cascades.

## DATA FEED IN THE FTTH PLANT

The CMTS does not act as the data interface for the FTTH network, because data is carried as on-off modulation of the data transceiver at different wavelengths than that used for video. Thus, data doesn't take spectrum away from video. The interface taking the place of the CMTS is called an Optical Line Terminal (OLT). It may or may not include front-end data concentration (switching). It interfaces to the headend switch, usually using either optical or electrical Gigabit Ethernet (GBE) connections.

## RF RETURN IN THE FTTH PLANT

In HFC plant, we handle RF return by providing an upstream path between 5 and about 42 MHZ, using diplex filters at the input and output of each amplifier and at the coax side of a node. We normally use a second fiber to return RF signals to the headend from an upstream RF optical transmitter in the node. FTTH doesn't have an RF return path for signals needing it. Some have proposed adding a modulated laser at each home to accommodate RF return, but this is less than satisfactory from cost and maintenance standpoints.

The normal applications for the upstream plant are upstream data from DOCSIS modems, RF return from set tops, and perhaps status monitoring. In some instances there will be contribution video from the field being returned to the headend to be turned around. In the FTTH plant, we don't need the upstream for DOCSIS (except for DSG – more on that later), and status monitoring is replaced with a complete element management system (EMS) that uses the upstream data on 1310 nm for upstream transmission. That leaves the RF return from set tops that we need to handle.

Figure 2 illustrates how RF return is handled at the ONT, and figure 3 shows RF return handled at the headend. Referring to Figure 2, RF upstream signals from set tops are separated in the diplex filter in the ONT. The RF return signal is digitized in the analog-to-digital converter (A/D). Processing consists of doing some fast data compression and putting the resulting digitized data into an IP packet. The packet includes the IP address of the reconstruction hardware at the headend. This packet is transmitted to the headend along with all other data. Note that the only time bandwidth is consumed for RF upstream transmission is when a packet is actually available for transmission. This only happens when there has been a burst of RF from the set top converter.

Figure 3 includes the RF reconstruction. Upstream data from the FTTH plant appears at the GBE connections at the ONT, which are connected to the headend switch or router. The headend switch routes the RF return packets to a special-purpose headend product that reconstructs the RF signal. The reconstructed RF return signal is combined with RF upstream signals from the HFC plant if desired, and the combined signals are supplied to the receiver of the set top control system. Since the destination for the RF return signals is determined by the IP address at the ONT, the RF reconstruction can be located anywhere in the plant. It is only necessary to assign an IP address to the reconstruction unit at the headend. Then part of the provisioning of each ONT is to tell it the IP address of the RF return packets. They will arrive properly at the reconstruction unit.

The RF return system as shown can handle any of the standards in widespread use today, including SCTE 55-1, SCTE 552, and DSG (DOCSIS Set Top Gateway). Somewhat different provisions must be made for each, due to the properties of the different standards. Consult with the manufacturer of your ONTs for specifics.

Because of the way RF return works in FTTH plant, there is no noise funneling. Referring again to Figure 2, the only time there is a transmission from the RF return section is when an upstream transmission is detected. The threshold for detection is set as high as possible, consistent with the need to make sure all set tops can reach that threshold within their specified maximum output level, taking into account in-home splitting and cable loss. The packet can only be transmitted when nothing else in that PON is being transmitted, and though the RF reconstruction hardware in the headend (Figure 3) services many PONs, only one signal can be passed to it at a time. Thus, the system inherently offers protection against noise funneling. In the unlikely event that you have one home that is generating enough noise to make the RF return circuitry think a signal is present, then that noise cannot affect reception from any other home. Furthermore, the offending home is easy to locate: you simply sniff upstream packets and note the source IP address the illegal packets are coming from. Also note that for SCTE 55-1 systems using contention signaling, the system can actually improve efficiency by queuing multiple transmissions.

## PON Vs. NODE

The OLT shown is typical in that it includes a number of optical interfaces, each serving an individual set of subscribers. That set is referred to as a PON, or passive optical network. A PON in this context may be considered as analogous to a node in HFC plant. In FTTH, a PON is effectively a node serving typically 32 subscribers per node. The node analogy breaks down though, when we consider what is in the field. In a node, the node (conversion between optical and electrical) is in the plant, and is followed by a short cascade of amplifiers. In the PON, the 32 (typical) subscriber PON consists of an all-passive network, with the conversion to electrical being done on the side of the consumers' homes. All that is in the field is optical fiber and passive splitters, until we get to the home.

Note that each PON has a WDM after it, which combines the broadcast video at 1550 nm with the data at 1490 nm (downstream) and 1310 nm (upstream). Since each PON has a WDM, it is obviously possible to further divide the RF signals to get more frequency reuse in the FTTH plant, just as you do in the HFC plant. However, if you wish to take advantage of it (your option), you have additional spectrum available in the FTTH plant that you don‘t have in the HFC plant. You are not loosing spectrum for data and voice, since they are carried on separate wavelengths. And you probably have more spectrum available. The RF FTTH spectrum extends to at least 870 MHz, and 1,000 MHz really works (it is often not specified that way for detailed technical reasons, but it works). So you may avoid as much frequency reuse as youi need in some HFC situations.

## CURRENT VIDEO PRACTICE

Entities deploying PONs today have a number of philosophies regarding how to deploy video. While some are deploying IPTV exclusively, we find that most (including Verizon) are deploying broadcast video, while reserving the option to deploy IPTV in the future. The reasons for the continuing popularity of broadcast video, especially in competitive situations, are several. Broadcast video is more mature, as the cable TV industry knows well. There are many more features available on broadcast set tops than there are on IPTV set tops, and the user is accustomed to the broadcast experience. Furthermore, IPTV does necessitate certain in-home wiring (or use of the devices discussed in the next section) not needed for broadcast video. Finally, the ability to supply analog-only services without a set top where that is the only service desired, is a subscriber convenience and an operator cost savings.

A likely future scenario is to provide for a broadcast basic service (analog-only or analog and digital) on broadcast, while offering video-on-demand (VOD) and similar services on IP. This gives you the best of both worlds: broadcast where it is most efficient, and IPTV where it is most efficient. While we are not aware of any operators who have implemented this scenario yet, we do know of operators who are expecting to implement it in the next few years.

## IN-HOME NETWORKING

Figure 4 illustrates some in-home networking options that are available for use with FTTH and, with suitable modification, with HFC networks. To the left are shown the ONT RF and data connections. Since most homes in North America are wired for coax but not data, there has been a lot of work on putting data on the coax network in the home. The cable industry has been following this work closely. Two standards have emerged, and the selection between them may be coming down to whether or not cable TV set tops are being used. Measured throughput for both standards is on the order of 100 Mb/s. Both standards support data delivery to all computers in the home over coax, networking of the computers, and delivery of IPTV if and when desired.

HPNA 3.0/3.1 can be used where RF return from set tops is not needed, such as where IPTV is being employed for premium services. Originally HPNA was an acronym for Home Phone Networking Alliance, but the organization changed its name to the HPNA Alliance.[ii] This standard was developed for use on in-home phone networks, but it was quickly realized that the same standard would work well over coax. The frequency band occupied is 4-21 MHz, so it overlaps spectrum often used for RF return. Equipment supporting HPNA is readily available from a number of vendors.

The other standard, MoCA,[iii] occupies spectrum above 870 MHz, so it is compatible with RF return. Use of that higher spectrum raises concern about it’s viability in older home wiring, but the MoCA association reports good throughput in nearly all homes tested.

Both standards are used in essentially the same way. A unit at the ONT, usually designated as the master, transmits data presented to it on one of the 10/100Base-T connections at the ONT. An adaptor (slave or client) at each computer converts the RF signal back to a 10/100 connection. These adaptors are often called dongles.

The two data-over-coax standards can carry IPTV data to set tops and can also be used for data delivery and in-home networking. Figure 4 shows client-to-client data flow. The standards rely on limited isolation between splitter ports, plus reflections, to achieve client-to-client data communications. This usually works well, as the standards have adequate dynamic range to work over a wide range of attenuation. If a problem were to arise, you can artificially worsen the return loss in the band of interest by installing a filter with a stop band at the device operating range, at the RF output of the ONT.

The use of HPNA or MoCA for in-home networking as well as for delivery of IPTV (if both are being done) will demand that quality of service (QoS) be applied. Both standards have the ability to do basic QoS functions, at least as far as prioritizing one type of traffic over another based on 802.3 quality markings.

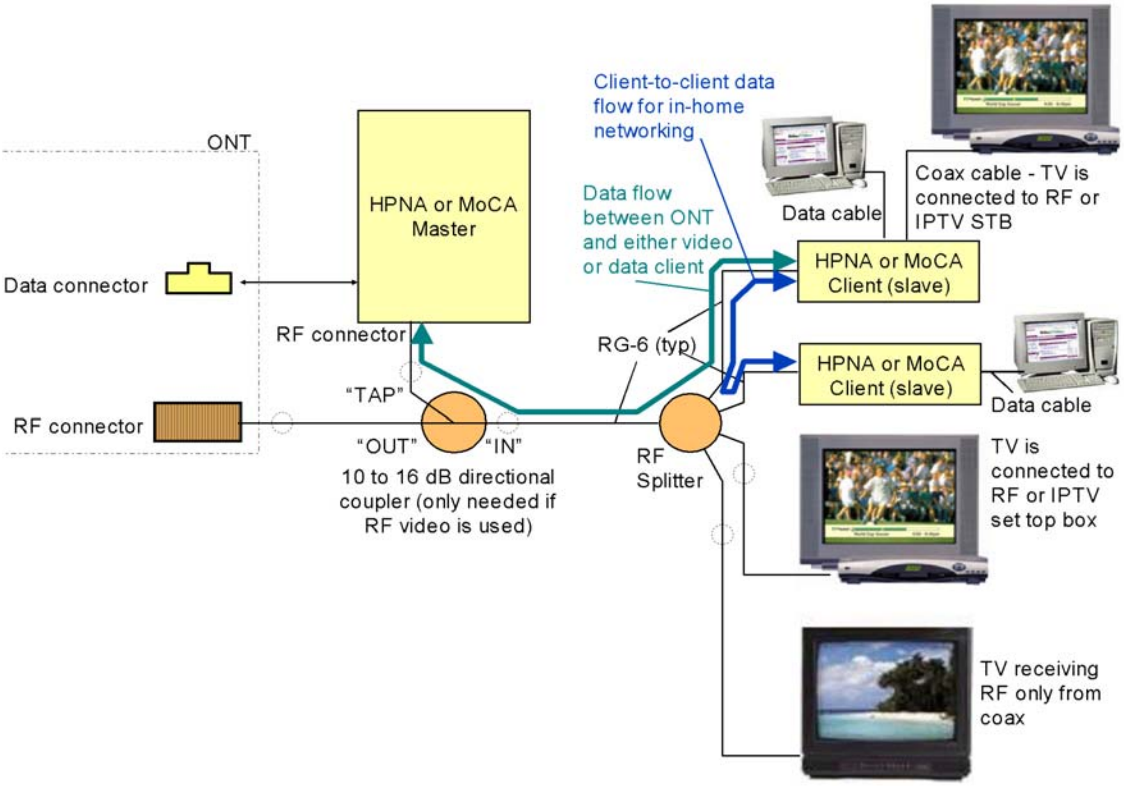


**Figure 4. In-Home Networking**

## DIGITAL TURN-AROUND

Digital turn-around, where you use the upstream data path to send video back to the headend to turn it around and send it to subscribers, is easy in FTTH. This is being done today, where systems are using their facilities to carry local sports. They convert the video to IPTV at the venue, give the packets the IP address of the turn-around device at the headend (or any other point in the network), add quality of service (QoS) parameters, and the turn-around happens. No worries about ingress, no issues of finding spectrum for the upstream video, no plant qualification, no headend re-cabling. Anywhere you have an ONT in the plant is an injection point for entertainment-quality video you can turn around to your subscribers.

## VOICE COMPATIBILITY

Voice compatibility between FTTH and HFC is easy, and there are systems using the same soft switch for both today. Many FTTH systems support MGCP/NCS protocol, and they can support SIP. The CableLabs specified NCS is a profile of MGCP (some people say it the other way around), so it is common to support both signaling standards. SIP is starting to gain a lot of traction in the industry, both in HFC networks and in FTTH networks.

A difference between the two networks is the way QoS is implemented. DOCSIS uses a special QoS paradigm in which a communications path is set up at the beginning of a call and torn down at the end of the call. The soft switch (or other voice facility) is responsible for initiating set-up and tear-down of the circuit. FTTH systems, on the other hand, use IETF- and Ethernet-standardized prioritization. The advantage is that there is no need to set up and tear down the connection. Voice packets are identified and are transported with high priority. Soft switches don't have any problems distinguishing between calls placed on FTTH and on HFC, and there are people today who are operating mixed systems, using the same switch for both HFC and FTTH.

## CONCLUSION

It is very easy to add FTTH to an HFC network. This is advisable where you have greenfield opportunities close to HFC plant. Many land developers are demanding FTTH in their new subdivisions, because they have learned that FTTH adds value to the homes they build. Cable TV operators are in a unique position to service this business, and at the same time, equip themselves for a future in which they will need FTTH in order to compete.

## END NOTES

[i] DOCSIS 3.0 provides more bandwidth by bonding channels, but this takes channels that cannot then be used for video.

[ii] http://www.homepna.org/

[iii] http://www.mocalliance.org/

# MULTI-WAVELENGTH ACCESS NETWORKS: A PRACTICAL GUIDE TO IMPLEMENTATION

Ray Thomas, Time Warner Cable
Venk Mutalik, C-COR

## *Abstract*

*Continued growth in residential data, voice and video traffic is a success story that has driven MSOs to increase throughput capability in all parts of their networks. When increasing throughput capacity, different parts of the end-to-end network require different approaches. In the access (or "distribution") part of the network there have been advances in optical technology that lend themselves to being a rapidly deployable, robust, cost effective solution. Multi-Wavelength Access Networks are an additional tool in the cable operator's toolbox that will allow MSOs to increase throughput capacity or provide additional services on their existing infrastructure at significantly less cost than new fiber construction.*

*This paper will lay out a practical guide to implementation of multi-wavelength access networks. It will first present a model that explains the various benefits and trade-offs of implementing Multi-Wavelength Access Networks. The paper will then document all the optical impairments that accrue within access networks, and conclude with recommendations for the practical implementation of multi-wavelength systems.*

## BUSINESS MODEL FOR IMPLEMENTING MULTI-WAVELENGTH ACCESS NETWORKS

One of the driving forces for the use of multi-wavelength optics in the access network is the requirement to increase throughput capacity. Optoelectronics equipment, installed over the past 15 years, employed a design in which one fiber carried the forward path and a second fiber was used for the return path. Some cable operators built out networks using fiber sheaths containing four or six strands of glass to each optical node, while other cable operators used as few as two or occasionally more than six. In the cases where six fibers inside a sheath are going to a specific area of the system, 3 optical nodes could be served. If a fourth node was needed to serve increased data traffic in that area, another pair of fibers would be required. A typical solution is to install additional fiber to serve that area. The costs for constructing fiber add up on a per-foot or per-mile basis. The process of obtaining construction permits and working during bad weather conditions can add delays to construction projects and the attending service disruptions may also have additional negative consequences.

However, the development of multi-wavelength optical technology presents the opportunity to make increased use of existing fiber and delay the need for building out additional fiber. Additional forward and return paths can be added to the fiber already in place. Construction delays along busy highways and streets or in back easements with difficult access can be postponed. The use of multi-wavelength technology allows for additional throughput capacity to be turned up quickly.

A comparison of the relative costs of multi-wavelength optics versus construction provides a compelling justification. Each mile of aerial fiber construction costs in the range of $12,000-$16,000 (depending on a handful of factors). Constructing 10 miles of aerial fiber would therefore cost $120,000 or more. Underground fiber construction can cost up to twice as much per mile as aerial construction. On the other hand, the total cost of multi-wavelength optical equipment for both ends of the fiber run is less than the cost of constructing two miles of new fiber. Chart 1 shows the relationship between the costs of construction versus multi-wavelength implementation. This is a powerful financial incentive for the use of multi-wavelength optics.

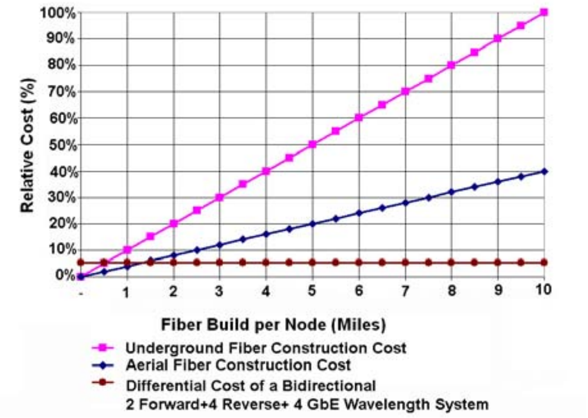


**Chart 1 – Multi-wavelength Access networks are much more cost-effective than construction**

| | **Existing 1310 design** | **Multi-wavelength design** |
|---|---|---|
| Construction of additional fiber | If fiber exhausted must build additional fiber (overlash aerial or trench in u/g) | Only required if extending fiber a short distance to serve additional node(s) |
| Time required for implementation | Requires weeks to design, plan, obtain any required permits, and construct fiber | Can be installed in under a week or even in one night |
| Costs | Approx $12k/mile aerial or $30k/mile u/g | Cumulative differential cost for additional wavelengths is much less than 10% of fiber construction cost |

**Table 1 – Multi-wavelength Access networks can be implemented more quickly than construction**

Continued growth in residential services drives the need to continue increasing throughput capacity. The commercial services segment is a market in which cable operators have staked out a solid business. Businesses represent a market with strong growth potential for cable operators. Cable operators have a range of options for providing service to small, medium, and large businesses. A few of these options include:

- a direct fiber feed into the business
- cable modem business class service
- wireless DOCSIS

A widely preferred approach is to use fiber for serving businesses. Although other options such as wireless DOCSIS can be used as a temporary solution to quickly establish service by reaching across obstacles (railroad tracks, large parking lots, rivers), fiber has long been established as a reliable solution.

Most networks were built with spare fiber capacity in the access part of the network. As fiber is used to connect new business customers, the spare capacity in the fiber sheaths is reduced. Multi-wavelength technology is an alternative path to providing additional services on the networks without incurring the cost of new fiber construction.

## THE FIBER SPECTRUM

The CWDM wavelength plan is detailed in ITU Recommendation G.695, which was ratified in January 2005. The plan provides for 18 wavelengths spaced 20 nm apart over a range from 1271 to 1611 nm. Typically the 1371 and 1391 nm wavelengths are the designated water peak wavelengths of deployed optical fibers (Figure 1 provides insertion loss for a 20 km fiber link for fiber with and without the water peak).

Cable operators can choose to reserve the 1531 and 1551 nm bands for possible deployment of services utilizing the DWDM spectrum. Long haul 1550nm optics, EDFA, and QAM overlay architectures utilize this portion of the optical spectrum. The wide CWDM channel spacing allows the use of lower cost uncooled DFB laser technology for reverse path transmitters.

Additionally, CWDM channel spacing enables the use of cost effective, environmentally hardened optical passives for field deployment. These passives can typically be obtained off-the-shelf from suppliers and do not require unique specification considerations. The CWDM specifications for active devices promote GbE SFPs, reverse path analog transmitters and forward transmitters. The CWDM spec therefore provides for a significant increase in fiber capacity and promotes a level of plug and play capability.

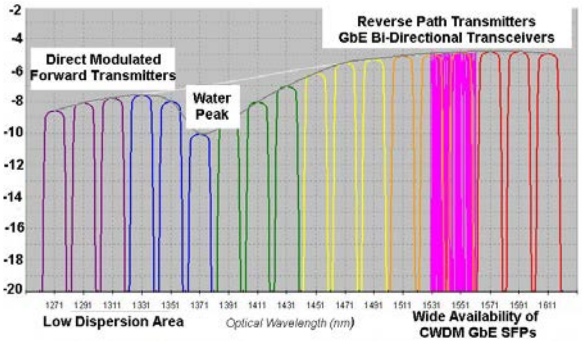


**Figure 1 – Some wavelengths are more suitable for specific services than others.**

From Figure 1, it is also seen that the familiar DWDM wavelengths are in the 1525 to 1560 nm region and are covered by the spectrum of the 1531 to 1551 nm CWDM bands.

## OPTICAL IMPAIRMENTS WITHIN ACCESS NETWORKS

Although distances involved in access networks are quite modest - around 20 km as compared to long haul transport networks of around 500 to 1000 km - there still are numerous optical impairments that could cause measurable degradation of the RF spectrum. The ability to identify all such impairments and manage their impact on the RF spectrum is central to making multi-wavelength access systems work.

Optical impairments are artifacts in fiber networks and the fiber itself that impacts how well the RF spectrum is carried in the network. There are two broad classes of impairment: linear and non-linear. It is generally the case that non-linear impairments are dependent on optical intensity whereas linear impairments are not. The two classes of impairments can be further divided into single and multiple wavelength non-linearities for the optical non-linear impairments, and fiber linear effects and optical passive effects for the optical linear impairments.

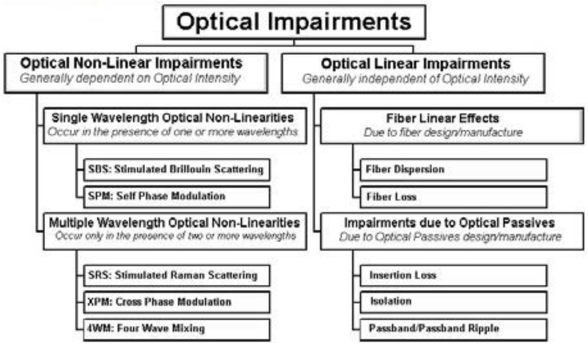


**Figure 2 – Identifying Optical Impairments**

Linear impairments, like dispersion and passband ripple, impact RF spectrum figures of merit, such as CSO and CTB. Most often, linear impairments are the performance manifestations of optical passives/fiber design and manufacture. Careful attention to (and verification of) product specifications can limit the effects of this class of impairment.

Single-wavelength optical non-linearities include the well known Stimulated Brillouin Scattering (SBS) and the lesser known Self-Phase Modulation (SPM). The SBS is often compensated for by manipulating the optical spectrum and/or limiting optical launch power in the network.

Multiple wavelength non-linear impairments, such as Stimulated Raman Scattering (SRS) and Cross Phase Modulation (XPM) induce crosstalk between two or more wavelengths. Crosstalk could be between wavelengths carrying similar services, such as two wavelengths carrying signals in the reverse direction; two wavelengths carrying GbE; or two wavelengths carrying forward signals. In these cases, the RF spectra coincide and the crosstalk results in a direct impact on the RF performance.

Crosstalk may also crop up between wavelengths carrying dissimilar services, such as a forward wavelength and a reverse wavelength, or a forward wavelength and a GbE wavelength. In these cases, if the frequency spectra overlap, there could be a direct impact on the RF performance. For example, since the GbE and forward RF spectra overlap, GbE signals (such as spurious spikes when the GbE is unloaded) could bleed through from the GbE RF spectrum into the forward spectrum due to fiber non-linearities, causing measurable performance degradation.

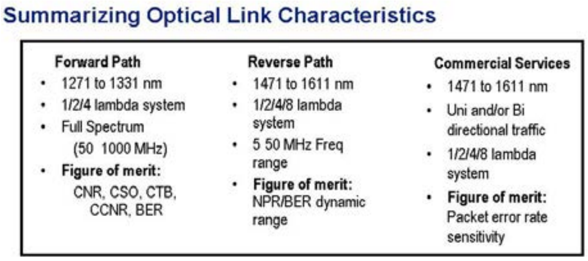
For access network design, SRS and 4 Wave Mixing (4WM) are the dominant non-linearities. Both of these are sensitive to polarization and become progressively worse with higher launch power. Both these phenomena depend upon the wavelength spacing and fiber dispersion as well. While SRS becomes worse with increasing spacing up to approximately 100 nm and then decreases becoming essentially extinct after about 200 nm, the 4WM potentially flares up as the wavelength spacing decreases and one of the signal wavelengths approaches the fiber dispersion zero point.

Typical CWDM multi-wavelength systems employ two forward wavelengths spaced 20 nm apart. Here, the SRS is the dominant non-linearity and the system reach is limited by the total power launched into the optical network. It is possible to reduce the SRS effects by reducing the wavelength spacing below the standard 20 nm used in CWDM and thereby increase the power launched into the network and/or increase the number of forward wavelengths. However, this approach can also substantially increase the 4WM potential when additional variables such as dispersion are introduced in the system, sometimes leading to a less robust system. A nonstandard wavelength spacing plan could also increase the system cost due to lower volumes.

The ability to adequately model and test all aspects of fiber impairments with particular emphasis on the number of wavelengths, polarization and fiber dispersion, along with the earlier mentioned variables such as maximum power launch, wavelength spacing and optimally selected optical filters, is critical to promoting a robust cost-effective solution that also satisfies capacity needs.

## DESIGN TRADE-OFFS

The analog realm features familiar RF trade-offs, such as the fact that CNR can be traded off to achieve better CSO and CTB and vice-versa. Similarly, multiple wavelength access networks will have trade-offs to make sure that each wavelength passes through the network without impacting the other wavelengths on that network.



Summarizing Optical Link Characteristics

| Forward Path | Reverse Path | Commercial Services |
|---|---|---|
| • 1271 to 1331 nm<br>• 1/2/4 lambda system<br>• Full Spectrum (50 1000 MHz)<br>• **Figure of merit:** CNR, CSO, CTB, CCNR, BER | • 1471 to 1611 nm<br>• 1/2/4/8 lambda system<br>• 5 50 MHz Freq range<br>• **Figure of merit:** NPR/BER dynamic range | • 1471 to 1611 nm<br>• Uni and/or Bi directional traffic<br>• 1/2/4/8 lambda system<br>• **Figure of merit:** Packet error rate sensitivity |

**Figure 3 – Multi-wavelength network system performance requires adequate single wavelength performance and limited optical interference**

To employ these trade-offs effectively, the inherent optical linear and non-linear impairments should be studied to identify and quantify their impact on the overall system.

A good deployment strategy would include comprehensive testing and analysis of all optical parameters of the system so that design rules for field deployment can be devised. These rules may govern the locations of optical wavelengths such that the overall optical crosstalk is minimized.

Another rule would consider appropriate intermixing of wavelengths of diverse RF spectra to ensure that the optical level of the composite signal being launched into the fiber remains below specified limits.

Another useful strategy consists of identifying and investing in optical passives that support the selected wavelengths and have adequate isolation and loss specifications. These will often be unique to a specific application. For optimal economic efficiency, it is good practice to set a standard usage plan for wavelengths in order to drive higher volumes of identical optical passive configurations.

Since the launched power of a multi-wavelength system is limited by optical non-linearities, the fiber reach of a multi-wavelength network can still be enhanced by supporting lower optical node receive power. It is often the case that adding a node results in a shorter RF cascade with inherently less CNR degradation, so lower optical receive power can be used without degrading the end-of-line performance. For this reason, segmentable nodes, placed deeper into the network, are particularly well suited for multi-wavelength access networks.

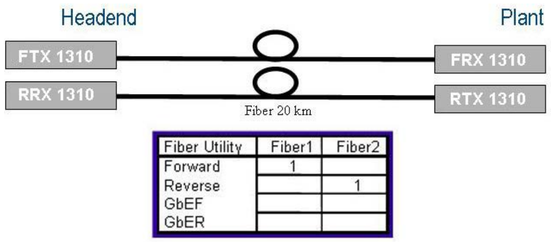
## WAVELENGTH PLAN

The service disruptions that plague new fiber construction can be minimized for multi-wavelength access networks if a wavelength plan is considered in advance of the design. This plan should ideally proceed sequentially from an examination of fiber link lengths and fibers available for deployment (fiber link description) to the services intended for each fiber (deployment package description) and further into future plans for services and fiber usage. Important aspects of future use planning include consideration for 10 GbE usage, preservation of the DWDM band and the desire for route redundancy.

## RECOMMENDATIONS FOR PRACTICAL IMPLEMENTATION

The following well tested design examples present architectures ranging from simple to complex that progressively allow higher and more effective utilization of installed optical fiber. The fiber utility table at the bottom center of the figures will keep a running tally of the number of wavelengths used in each fiber. Although the architectures are presented in the context of the CWDM standard, similar architectures can be conceived for DWDM or other multiple wavelength allocation plans.

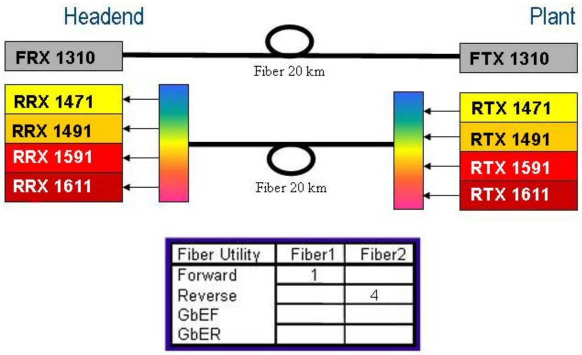
The classic architecture in Figure 4 is essentially characterized by a fiber pair from the hub or the headend terminated into a node. Each fiber then carries only one wavelength, generally at 1310 nm.



| Fiber Utility | Fiber1 | Fiber2 |
|---|---|---|
| Forward | 1 | |
| Reverse | | 1 |
| GbEF | | |
| GbER | | |

**Figure 4 – Basic Architecture**

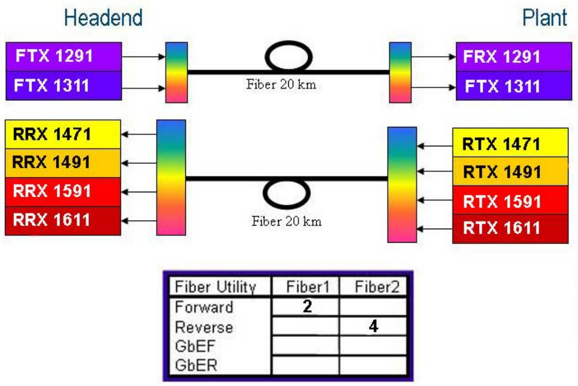
Figure 5 illustrates a very cost effective way of increasing fiber utility and providing reverse segmentation capability.



| Fiber Utility | Fiber1 | Fiber2 |
|---|---|---|
| Forward | 1 | |
| Reverse | | 4 |
| GbEF | | |
| GbER | | |

**Figure 5 – CWDM in the Reverse**

Figure 6 represents an improvement over the previous architecture in that the node can now be segmented in the forward *and* reverse path. This architecture enables the operator to have specific fibers designated for forward or reverse purposes and is least disruptive in providing service augmentation. Please note however that the two wavelengths should have the same analog broadcast signals, but could have different QAM 256 narrowcast signals.

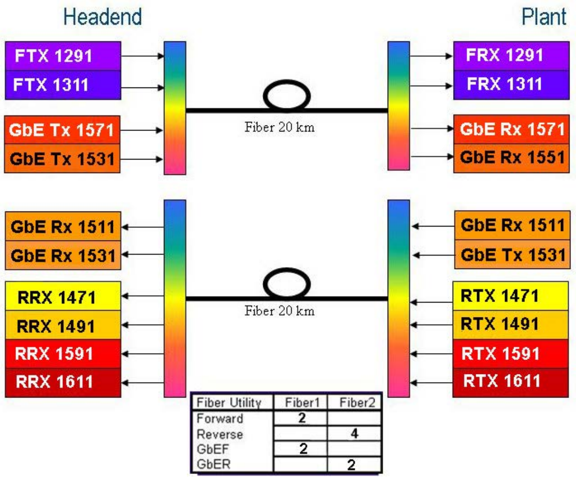


| Fiber Utility | Fiber1 | Fiber2 |
|---|---|---|
| Forward | 2 | |
| Reverse | | 4 |
| GbEF | | |
| GbER | | |

**Figure 6 – CWDM in the Forward and Reverse**

Figure 7 carries the previous architecture further. Here the operator may add the Gigabit Ethernet (GbE) traffic to the previously analog/QAM HFC plant.

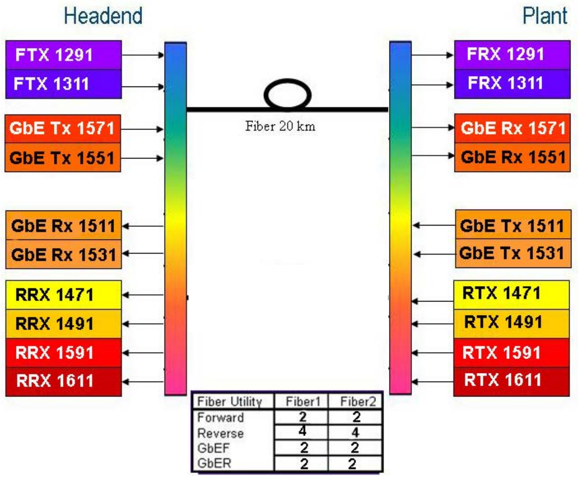
This architecture maintains the forward and reverse designations on the available fibers and is minimally invasive. Higher utilization of the fiber is possible when the two fibers are collapsed into one. That strategy is the subject of the next architecture.



| Fiber Utility | Fiber1 | Fiber2 |
|---|---|---|
| Forward | 2 | |
| Reverse | | 4 |
| GbEF | 2 | |
| GbER | | 2 |

**Figure 7 – GbE in the Forward and Reverse Fibers**

Figure 8 shows the three different services propagated on a single fiber. An architecture of this type provides the most effective usage of the optical fiber. A single fiber is therefore able to provide 2-way segmentation of the forward narrowcast QAM 256 signals, 4-way segmentation of the reverse signals and 2 bi-directional GbE business service links.



| Fiber Utility | Fiber1 | Fiber2 |
|---|---|---|
| Forward | 2 | 2 |
| Reverse | 4 | 4 |
| GbEF | 2 | 2 |
| GbER | 2 | 2 |

**Figure 8 – Typical CWDM Capacity per Fiber**

As indicated earlier, the RF input to each forward transmitter is independent; however, the analog broadcast content of the two transmitters should be the same. The QAM 256 narrowcast content for each of the transmitters can differ. The actual capacity of the optical fiber is much higher than represented here however. The higher capacity is obtained by employing additional wavelengths that are carefully chosen to be compatible with the system needs after considering the optical impairment mechanisms discussed earlier.

## CONCLUSION

Multi-wavelength systems in the metro, long haul and transport arena have been designed for many years now in the form of Multi-Wavelength optical networks. Multi-wavelength systems are increasingly being considered for use in the access (distribution) part of the network. The technology behind multi-wavelength access networks has been evolving over the past year to offer significant new capability.

Traffic continues to grow in cable operator networks. In some metro areas the rate of growth has been astonishing. Development of multi-wavelength optics technology has progressed to the point where it offers a very attractive alternative to construction of additional fiber for increasing throughput capacity. This provides an opportunity for operators to save large amounts of capital by taking advantage of multi-wavelength optics.

## REFERENCES

1. Gigabit IP and Analog/QAM Traffic Compatibility in HFC Networks - A case for Modern physical layer convergence, Shamim Akhtar, Doug Weiss, Comcast Corp., Venk Mutalik, Zoran Maricevic, Marcel Schemmann, David Heisler, Don Wesson, Liyan Zhang, C-COR Inc., SCTE Conference on Emerging Technologies, 2006

2. Cable Television Technology, Second Edition, Chapter 13, Wavelength-Division Multiplexing, The Morgan Kaufmann Series in Networking 2004, by Walter Ciciora, James Farmer, David Large, Michael Adams

3. Nonlinear fiber optics, Govind P. Agrawal, Academic Press, Second Edition, 1995

4. Enhancing cable architectures with WDM - Grafting WDM onto existing cable systems, Venkatesh G. Mutalik, Communication Engineering and Design (CED Magazine), February 1997, Page 54, WDM Series

5. ITU-T Recommendation G.695 - Optical interfaces for coarse wavelength division multiplexing applications: Series G Transmission systems and media, digital systems and networks (01/2005)

6. DWDM: Matching Technology Advancements with Business Requirements, Venkatesh G. Mutalik, National Cable Telecommunications Association (NCTA) Proceedings 1999

7. Maximizing Today's Network: Increasing HFC Network Capacity without Additional Fiber. Jeff Sauter and Bill Dawson. Communications Technology magazine, July 1, 2006.

## CONTACTS

Ray Thomas, Principal Engineer, ATG
Time Warner Cable
Tel: 720-279-2729
Email: ray.thomas@twcable.com

Venkatesh Mutalik, Chief Technologist
C-COR, Inc.
Tel: (203) 630-5763
Email: vmutalik@c-cor.com

# NARROWCAST SERVICES – UNIFYING ARCHITECTURES

Glen Hardin, Time Warner Cable

## *Abstract*

*Cable advertising is a primary source of revenue for the cable industry. It is a massive successful industry by its own right. The ability to target an advertisement to a subset of the cable plant or "ad zone" is one of the key attributes that differentiates cable's advertising capability from that of the national broadcasters. These uniquely zoned advertisements are now more valuable as they are targeted to their intended recipients and narrowly broadcast to a subset of the cable plant to reach the recipients.*

*High Speed Data (HSD) is another highly successful cable service that cable deployed over 10 years ago and is still experiencing strong growth. On a wide scale, Cable's HSD offering is leading the competition with 8Mbps and 10 Mbps HSD offering.*

*Closely tied to the HSD service is the Voice Over Internet Protocol (VOIP) phone service. This relatively new service has quickly become another mission critical offering for the cable operator representing new sources of growth and income.*

*There is little doubt that Video On Demand (VOD) is a success. According to the industry press, across the industry, VOD revenues have grown to over a billion dollars of revenue for the cable industry per year. VOD is widely deployed across all major markets and is the cornerstone of the digital offering.*

*One of the next generation key technologies to unlock Cable's bandwidth potential is the delivery of television through the Switched Digital Video (SDV) infrastructure. Broadcasting the select programs from the digital tier through SDV is projected to save 50 percent of the bandwidth required when compared to broadcasting through normal mechanisms.*

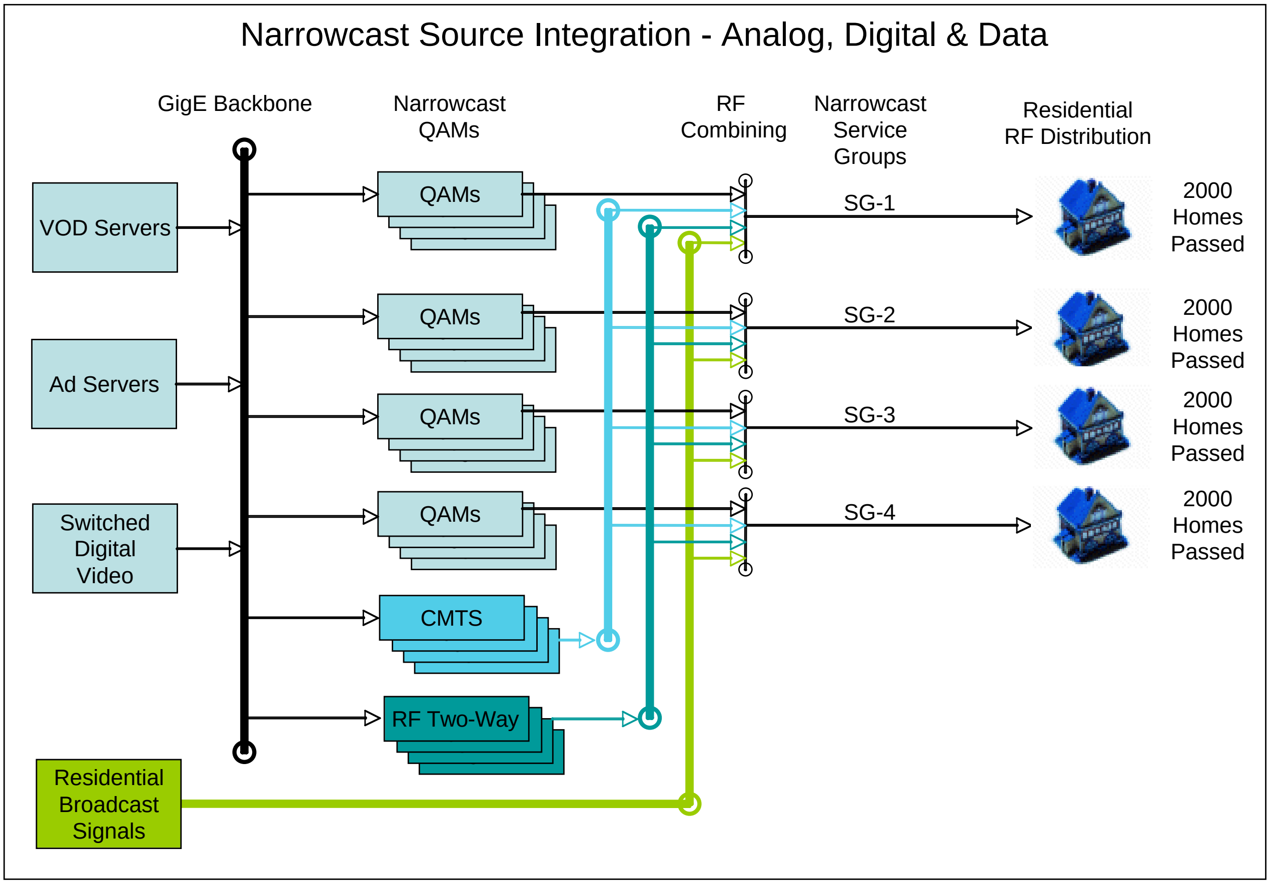
*Narrowcast Bandwidth is the primary technology infrastructure that forms the foundation to allow cable to deliver the aforementioned services of the triple play; Voice, Video and Data. The Narrowcast Bandwidth is native to the cable platform, is one of cable's primary differentiators and serves as a springboard on which next generation services are launched. Cable's current successes can be traced back to investing in the Hybrid Fiber Coax (HFC) architectures. Cable's future success will be built upon the extension of the HFC network, the Narrowcast. It is key to building cable's sustainable network and is cable's primary tool to address competition.*

*This paper seeks to provide context for the Narrowcast - its origins, growth, future and importance to cable's future.*

## Defining "Narrowcast"

First, a few terms need to be defined to clarify their use in this paper.

**Broadcasting** is the transmission of programs or signals to the entire cable plant. That is to say, source signals (analog video, digital video or data) are modulated onto a frequency where it is broadcast across the cable plant unmolested or uninterrupted to the entirety of end users. Cable was built on that capability to transmit signals to the entire plant and it will always continue to broadcast.



**Narrowcasting** is the narrow "broadcasting" of source signals (analog video, digital video or data) to a subset of the cable plant and not indiscriminately to the plant on the whole.

In this manner, end users will receive unique source material based on their geographic location within the plant. This is accomplished by inserting unique or targeted source and modulating it onto a RF frequency that only serves a small subset of the cable plant. These Narrowcast modulators deliver unique RF channels (bandwidth) and thus unique sources to a subset of the cable homes.

It needs to be acknowledged that there is, in fact, both a Forward Narrowcast and a Return Narrowcast, but to avoid too much complexity in the discussion, this paper will focus on the forward Narrowcast.

**A Narrowcast Service Group** is the subset of subscribers within a cable plant that is served by a unique set of source signals (analog video, digital video or data).

A **Channel** is defined as a single 6 MHz frequency slot, whereas a **Program** is defined a single video source. To put it into perspective, an analog source would occupy an entire 6 MHz channel as a single program stream, whereas a digital program stream may be only one of ten programs on a channel. A **Stream** is the transport of a program across a channel at defined bandwidth. In the context of this paper, streams can be analog, VOD and SDV.

## Understanding the Architecture

Cable systems were originally constructed as purely broadcast networks using cascading amplifiers in series to distribute the broadcast signals to the entire cable footprint. This architecture was interdependent and was as much an art as a science to keep the distribution network in “balance” across all the amplifiers within the network.

The figure above represents the basic constructs of the Narrowcast architecture. Modern Hybrid Fiber Coax (HFC) architectures were introduced to compensate for the limitation and complexity of managing series of cascading amplifiers. The reference architecture was based on roughly 500 homes per node and 4 nodes or 2000 homes per laser.

As part of this transformation, real-time two-way Radio Frequency (RF) communication was introduced to lay in the foundation for advanced services. Narrowcasting of the two-way RF communications addressed the problem of trying to communicate to a very large population of Set-Top-Boxes (STBs) by breaking the down the cable plant into subsections. The two-way RF communication equipment would narrowcast on the same frequency, but due to plant isolation they would not interfere with each other. In that way, the same narrowcast frequency could be “re-used” by each sub-plant. To put this into well-known context, the Narrowcast communication technologies that provide the two-way RF communications are Scientific Atlanta’s Quadrature Phase Shift Key (QPSK) modulators and demodulators and the Motorola’s Out of Band Modulator (OM) and Return Path Demodulator (RPD).

This basic technique of plant segmentation and spectrum re-use for creation of Narrowcast Bandwidth serves as the foundation for all of all the following services:

- Two-way RF Communications
- Advertising Zoning
- High Speed Data
- Voice Over Internet Protocol
- Video On Demand/Start Over
- Switched Digital Video

Each service type may serve a unique subset of the cable plant. Thus, within a given cable plant there could be an almost infinite number of different service groups based on service types. This is complex enough in the abstract but in the plant, this represents countless numbers of physical devices, wires, combiners and splitters located across a large geographic area (and this is just the forward path).

## Traffic Modeling Narrowcast Services

Traffic modeling is the exercise in estimating the simultaneous use of limited resources by a given number of users. In context of cable’s Narrowcast services it is the estimation of the amount of bandwidth at peak required per service type to deliver that service.

In cable’s context:

*Two-way RF Communications* – it is the number of STBs that can be attached to a given QPSK Mod/Demod without having too many collisions that would make communication impossible.

*Advertising Zoning* – is the exception. Although it is a Narrowcast service it is multicast to all users within a zone with guaranteed non-contested bandwidth and, therefore, is not bound by a traffic model constraint.

*High Speed Data* – it is the modeling of the number of cable modems that can attach to a
CMTS port and provide the required bits/second service performance.

*Voice Over Internet Protocol Phone Service* – it is the number of simultaneous phone calls that can be made at a given moment in time.

*Video On Demand/StartOver* - it is the number of simultaneous VOD sessions that can support at a given moment in time.

*Switched Digital Video* – it is the number of simultaneous switched television programs that users can be watching at a given moment in time.

When analyzing the various traffic models of the different narrowcast services, it is important to note the key differentiator of service type, and that is the Quality of Service (QoS) requirement. All video services such a VOD and SDV require near perfect QoS delivery, where as with data services like two-way RF communication and data DOCSIS services they have built-in mechanisms to overcome delivery issues such as collisions, dropped packets, burst or sporadic packet delivery and out-of-order packet delivery. Streaming video delivery does not have any of these mechanisms. If a video packet is dropped, arrives out-of-order, is improperly spliced into the main stream or is not properly paced in its delivery, a video artifact will be seen by the end user. Video delivery is all about QoS.

The exception to that statement is VOIP telephone service. With the introduction of VOIP, the data side of the network had to address this new QoS requirement. VOIP is the first DOCSIS service to really demand QoS. Not to overly understate the importance of VOIP QoS but it is by magnitudes less demanding than video QoS. A phone call requires 128Kbps/sec versus 3.75Mbps for Standard Definition (SD) VOD sessions. Network architectures are being adjusted to compensate for this new QoS requirement on the data-networking infrastructure.

Besides QoS, there is another important concept to take into account when performing traffic modeling on Narrowcast services; blocking. Blocking is the term used to describe denial of service due to contention of resource. The familiar analogy is the busy signal associated with trying to make a call on Mother’s Day when everybody else is trying to do the same. There are not enough resources to fulfill everyone’s request at the same time.

Either due to the QoS requirement for both VOD and VOIP each service is guaranteed its full bandwidth requirement for that session or the session/call is denied or blocked in its entirety. This is unlike “best effort” delivery of data services.

For each service type, there may be an acceptable level of blocking. The Service Level Agreement for VOIP phones service and SDV services may require them to be non-blocking where as VOD may be considered a blocking service.

The basic traffic model analysis example will be performed on video service due to the QoS requirement of video.

## Video Narrowcast Modeling

Video Narrowcast utilization is based on three key factors, the number of subscribers, the available bandwidth (i.e. streams both SD and HD) and peak simultaneous utilization. In the examples only SD content is used to model for simplicity.

VOD & Narrowcast

Originally, the narrowcast for VOD was built to a 6% peak streaming utilization of digital households. Four RF channels were allocated with each channel supporting 10 VOD streams per QAM 256 for 40 streams per 666 Digital Households.

The basic math:
500 Homes per node
4 nodes per VOD Service Group
2000 Homes Passed per VOD Service Group
Digital Penetration 33% (Digital Households)
4 QAMs (256) @ Payload of 37.5 Mbps
10 VOD streams per QAM at 3.75Mbps

2000 HP * 0.33 Digital = 666 Digital Homes
4 QAMs/SG * 10 streams=40 streams/SG

40 streams/SG / 666 Digital Homes = .06
or 6%

Therefore, the 40 available streams are to be shared across the 666 Digital Homes within the VOD Narrowcast Service Group.

VOD is considered a blocking service where at peak utilization some requests are accepted to be blocked denying service. Additionally, VOD is a uni-cast narrowcast service. This means there is a one to one relationship between narrowcast bandwidth use and user.

As mentioned at the beginning of this paper VOD is a huge success and with that success, careful management of the Narrowcast Bandwidth is critical to avoid block and denial of services. There are few major factors contributing to high use of VOD.

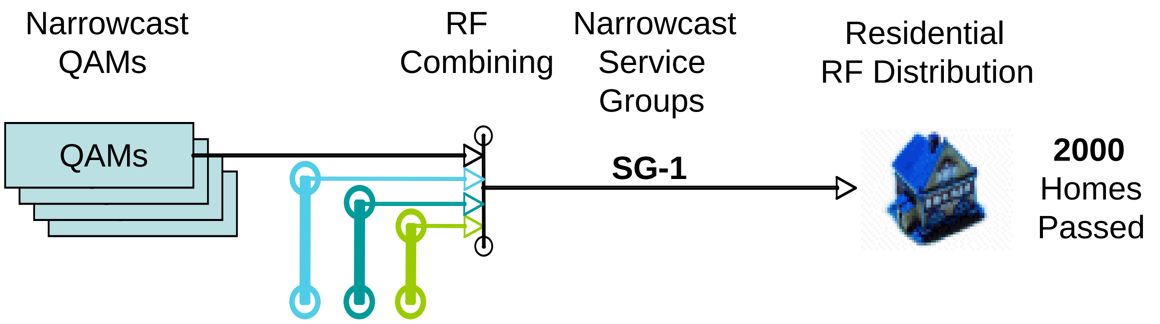
On the system today there is more compelling content, the digital penetration has increased by ten percentage points to around 45%, and markets have launched High Definition VOD (HDVOD) that uses 15 Mbps instead of the SD rate of 3.75 Mbps. However, the biggest contributor to increase usage of the VOD Narrowcast is the StartOver service.

StartOver is an advanced VOD technology that utilizes the existing VOD and Narrowcast infrastructure. Start Over enables the time shifting of live television without the need for upgrades to customer premises equipment. Time-shifted television allows subscribers to re-start and begin watching their favorite broadcast TV program during any point in the broadcast window. Unlike traditional on-demand services that have license windows measured in days or weeks, the Start Over content is only available to start a session within the actual broadcast window of the particular content.

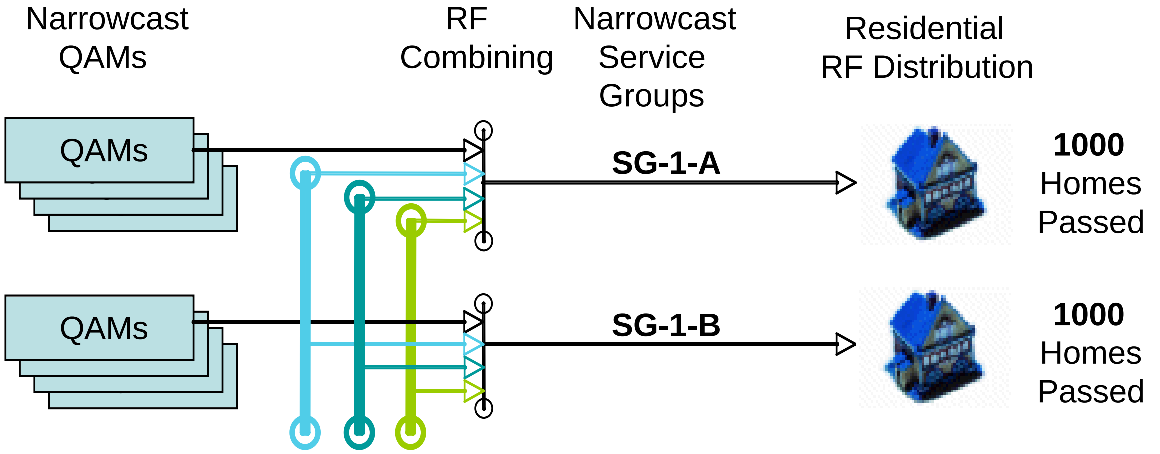
With the launch of the StartOver service stream utilization within the narrowcast of service group has seen a 50% increase in utilization of the narrowcast bandwidth.

Success is a high-class problem to solve and cable is well equipped to address contention within the VOD Narrowcast. The optimum technique to address contention within the narrowcast is to increase the number of channels per narrowcast service group. But this requires additional bandwidth that is typically not available. Therefore the primary tool at their disposal is what is termed “splitting” the VOD Service to “manufacture” enough bandwidth to support the ever-increasing demand.

Splitting of a Service Group - Before



Splitting of a Service Group - After



<u>Manufacturing Bandwidth</u>

In the previous example where the VOD service group consisted of:

500 Homes per node
4 nodes per VOD Service Group
2000 Homes Passed per VOD Service Group

To split a service group the number of nodes defined in the service group are halved thus a split service group now consists of:

500 Homes per node
2 nodes per VOD Service Group
1000 Homes Passed per VOD Service Group
Digital Penetration 33% (Digital Households)

The VOD math - holding Digital Penetration constant:
1000 HP * 0.33 Digital = 333 Digital Homes
4 QAMs/SG * 10 streams=40 streams/SG

40 streams/SG / 333 Digital Homes = .12 or 12%

The VOD math - adjusting Digital Penetration to 45%:
1000 HP * 0.45 Digital = 450 Digital Homes
4 QAMs/SG * 10 streams=40 streams/SG

40 streams/SG / 450 Digital Homes = .089 or 8.9%

The drawing above illustrates the basic concept of splitting of the service group.
By splitting the SG-1 service group into SG-1-A and SG-1-B the affect was to increase the peak simultaneous streaming capacity from 6 % to 12% and even when including the increased Digital Penetration of 45% the peak streaming capacity worked out to roughly 9%, an increase of 3% per service group from before the split of the original service group.

Real work must be performed in a coordinated manner in the RF plant to make a Service Group split occur without impacting the customer. On the physical level within the cable plant to accomplish a Service Group split, new QAMs are require to be:

- Installed
- Configured IP & Controller
- Wired into the combining and distribution network (forward and reverse)
- RF balanced into the plant
- All the STBs on that original service group must be notified that belong to the "new" service group
- Tested

This work represents a manageable amount of work for the cable system and they have become very adept at managing the service group split to a science.

This is the same basic technique used to address contention across any of the Narrowcast Services. Regardless of the service VOD, HSD or VOIP service group splitting has become an organic response to address the changing demographics encompassed within the cable footprint. Diligent monitoring of the various contention thresholds per service and an appropriate response to provide additional narrowcast bandwidth is all that is required to ensure optimum performance. This is why the term "manufacturing" bandwidth is a truism within the cable industry.

Switched Digital Video & Narrowcast

SDV works off the basic premise that not all digital programs are being watched in a given service area at the same time, so why broadcast them, why not narrowcast them just like a VOD stream.

SDV can be thought as of moving the STB tuner out of the STB and placing it just in front of the narrowcast SDV QAMs. To provision these SDV Narrowcast QAMs, the digital programs are multicast onto the transport ring and when a client "tunes" a multicast join occurs on the edge device and the signal is routed through the narrowcast QAM for distribution to the STB.

When studying the narrowcast modeling for SDV the basis for modeling was changed from Digital Households to STB tuners. This was done to accommodate devices that can have multiple tuners per device like DVRs. Where each tuner could be tuned to a different signal source at the same time. For every 250 to 500 STB tuners (DVR STBs have 2 tuners), a SDV narrowcast service group is created across the entire cable footprint.

The SDV math (more factors):
500 Homes per node
2 nodes per SDV Service Group
Digital Penetration 28.6%
Target size of 500 tuners per SDV SG
Boxes Per Home =1.4 average
Tuners Per Box = 1.3 average
8 QAMs (256) @ Payload of 37.5 Mbps
10 SDV streams per QAM at 3.75Mbps

1000HP *0.268 D*1.4STBs/H*1.3 T/B = 520 Tuners per SDV SG
8 QAMs/SG * 10 streams=80 streams/SG

80 streams/SG / 526 Tuners per SDV SG = 0.154 or 15.4%

Therefore, the 80 unique programs are available to be viewed simultaneous across the 520 digital tuners within the SDV Narrowcast Service Group.

The SDV service is considered a non-blocking service where contention is not allowed and must always be available. To borrow an adage from the telcos, POTs (Plain Old Telephone) now represents Plain Old Television and SDV must provide that same seamless service that the customers expect. SDV differs from StarOver/VOD service, as it is a multicast narrowcast service, meaning that there is a one-to-many relationship between narrow bandwidth use and user. That is to say, within the narrowcast more than one user can tune into the SDV narrowcast stream and watch the same program.

SDV is a unique service as it highly dependent on content placed into the SDV tier. Content is the primary factor in determining the utilization of the service and the required bandwidth. For SDV, the number of programs put into the SDV tier will determine the exact amount of bandwidth required to support the number of streams. If there are only 80 services in the SDV tier then there is no possible contention, it is just an expensive broadcast/narrowcast network. SDV becomes much more interesting when it is oversubscribed with content.

If SDV contention is based on tuner math and not Digital Homes, then VOD contention should be normalized to that same basis.

Re-calculating VOD Math based on tuners:
1000 Homes per node
2 nodes per VOD Service Group
Digital Penetration 45% (Digital Households)
Boxes Per Home = 1.8
Tuners Per Box = 1 (DVRs can only order one movie at a time)
4 QAMs (256) @ Payload of 37.5 Mbps
10 SDV streams per QAM at 3.75Mbps

1000 HP *0.268 D*1.7STBs/H*1T/B = 400 Tuners per VOD SG
4 QAMs/SG * 10 streams=40 streams/SG

40 streams/SG / 400 Tuners per VOD SG = 0.099 or about 10%

It is interesting to note that when basing VOD math on Digital Homes we have a contention rate of 8.9% but when it is normalized to tuner math is around 10% for exactly the same service contention level. So it seems that SDV and VOD should be able so share the same physical Narrowcast downstream infrastructure even though running in separate Narrowcast Bandwidth

Going forward, traffic modeling should be calculated using tuner math to normalize the analysis across narrowcast services.

## Unifying the Architecture

A couple of questions need to be asked when architecting a unified narrowcast architecture Voice, Video and Data.

Does it make sense to have one Narrowcast for all services sharing the same downstream infrastructure physical layer or are there many?

Can the narrowcast services share the narrowcast bandwidth?

When all services are normalized to tuner math, either QAM tuners or DOCISIS tuners, the math and traffic analysis becomes an even more interesting exercise in “And & Or “ math and understand the peak trending of services.

A single QAM tuner can either be tuned to a broadcast stream OR to a VOD stream OR to a SDV stream OR turned off. A QAM tuner cannot tune to multiple services at the same time. Therefore, it is easily conceived that for services targeted at QAM tuners the services can coexist and interoperate quite well within the same shared Narrowcast. OR math dictates that is just a series of tradeoffs. Therefore, when the bandwidth for QAM tuner Narrowcast services are shared or pooled together there are economies of sharing.

Current DOCSIS tuners only tune to one frequency at time. With DOCSIS 3.0 and channel bonding the DOCSIS tuners can tune wideband frequencies but as they are tuning discrete frequencies at any one point in time they can be considered bound by OR math when operating in a shared bandwidth pool just like they are today. Typically a single 6 MHz narrowcast channel is allocated for HSD and VOIP phone service across the cable plant.

Today in the residential market, the coexistence complexity arises when looking at the peaks of QAM tuner services and DOCSIS tuner services. While each service type may be bound by OR math when viewed together the two services actually peak at or very close to the same time and are thus actually bound by AND math. Thus, the total load between a QAM tuner and a DOCSIS tuner is cumulative. There are not economies of scale to share the bandwidth between the two tuner service types.

In the near future, when cable has more greatly penetrated into the commercial market with its HSD offering there will be some advantages in sharing the bandwidth between video and HSD narrowcast services. This will be because video narrowcast services under-utilize narrowcast bandwidth during the daytime and that excess capacity could be switched over for commercial HSD use during the day and then back to residential video narrowcast services in the evening.

Combined Narrowcast Services

There are two key technologies that are missing to truly unify the Narrowcast Services, the Global Session Resource Manager (GSRM) and the Business Rules Engine (BRE).

The Global Session Resource Manager (GSRM) is the unifying manager of all the source signal and bandwidth resources. The GSRM negotiates and arbitrates between all services and all requests. It is the key bandwidth allocation mechanism; employing bandwidth optimization algorithms to ensure that efficiencies are realized across the utilization of bandwidth across all narrowcast services.

The fact that the GSRM will be able to share the narrowcast bandwidth will allow for greater bandwidth usage efficiencies across the combined services and is predicted to require less total bandwidth for the same blocking factor for any given service. Efficiency is the key to performance.

The Business Rules Engine is the “uber” Policy Manger for all narrowcast services.

It is the tool that determines how to “sell” the narrowcast bandwidth for how much, to whom and prioritizes services and customers. It “plugs” into the GSRM and is not so much an engineering tool as a business tool. It will allow the cable business to optimize its service, its services and its revenues.

Along with the adoption of the GSRM and BRE a holistic approach must be taken into account when architecting the Narrowcast design. The physical layer of the Narrowcast cannot just be thought of in an abstract way without really identifying with the physical infrastructure of the headend, hubsite, laser, node and customer's home. This is purely a practical operational model concern. If the wiring of the various narrowcast services becomes too complicated to manage and the sheer number of service groups, types of service groups, QAMs combining and distribution networks the field personnel will not be able to support it. Although not detailed in this paper, remember to not forget the complexity of the reverse path traffic modeling and its physical combining and splitting network which is almost equal to the forward path.

## Case Study – The Exponential Growth

The growth of services dependent on the video Narrowcast has dynamically increased the total number of  streams that are "cast"  into the cable plant; either broadcast or narrowcast.

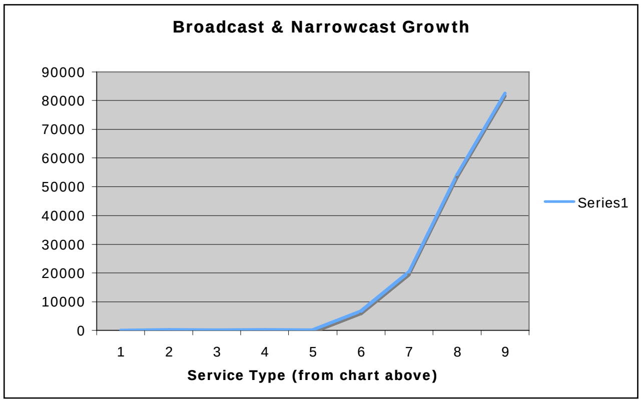
As cable technology advanced in steps from 300MHz, 450MHz, 550MHz, 750MHz, 860MHz and now to 1 GHz and at the same time transitioned from an analog broadcast network, digital broadcast network, HSD narrowcast, VOD narrowcast, VOIP narrowcast and finally to SDV narrowcast networks the number of channels, programs and streams offered in cable system has grown tremendously.

For example in a medium size market:

| | **Service** | **Service Groups** | **Ōcast 6-MHz Channels** | **Streams Per Channel** | **Total ŅcastÓ Streams** | **Cumulative Streams** |
|---|---|---|---|---|---|---|
| **1** | **Analog** | 1 | 78 | 1 | 78 | **78** |
| **2** | **Digital** | 1 | 25 | 12 | 300 | **378** |
| **3** | **Ad Zones** | 4 | Same as above | 68 | 272 | **650** |
| **4** | **HSD** | 338 | 1 | Best Effort | 338 | **988** |
| **5** | **VOIP** | 200 | Same as above | QoS | 200 | **1,188** |
| **6** | **VOD** | 170 | 4 | 10 | 6,800 | **7,988** |
| **7** | **StartOver** | 508 | Same as above | 10 | 20,320 | **28,308** |
| **8** | **SDV** | 678 | 8 | 10 | 52,400 | **82,548** |

The way to interpret the chart above is by culminating the growth of each service stream load and adding it to the next tier of service.

Graphing this culminating growth across services creates a graph that looks like:



Therefore, for the medium size market running all services, the cable system went from broadcasting of a few dozen analog streams to managing over 80,000 streams across their cable footprint within the last 15 years.

In the case study, the Narrowcasts for SDV and VOD have an equal number of service groups at 675 and can therefore share the same physical downstream wiring, combining and distribution path. However, the number narrowcast service groups for HSD/VOIP has a total of 538 services groups and is not in parity with SDV/VOD and thus cannot share the same physical downstream narrowcast wiring. The questions arise in this real-life example:

To simplify the implementation and operational model, does it make sense to increase the number of HDS service groups by 137 to bring it into parity with SDV and VOD so that all services can share the same downstream narrowcast wiring?

Is the peak demand for HSD going to naturally to follow the peak demand for VOD and SDV?

Is it cost justifiable, practical?

## Conclusion – It is One Network

Narrowcast services are pervasive within the cable infrastructure and they form some of the most revenue generating services offered by the cable company. Voice, Video and Data all share the same common cable network but unique Narrowcast Bandwidth.

To fully license this advantage of the Narrowcast Bandwidth, each service should not be looked at discreetly as a unique isolated and separate narrowcast service but in the whole across all service and the cable plant infrastructure.

The emerging and maturing Global Session Resource and Business Rules Engine will unlock the true economies of sharing narrowcast bandwidth within and between HSD and Video services.

The Narrowcast Services will continue to evolve and new yet to be developed and deployed services will guarantee cable's future success.

# OPENNESS AND SECRECY IN SECURITY SYSTEMS: POLYCIPHER $^{SM}$ DOWNLOADABLE CONDITIONAL ACCESS

Tom Lookabaugh, PolyCipher
James Fahrny, Comcast Cable

## *Abstract*

*The PolyCipher Downloadable Conditional Access System provides a new approach to security in cable: a small hardware footprint is used to enable both high security and the flexibility of software downloadable security clients for set-top boxes and other cable-ready devices. An important and often asked question is how much of this system specification should be disclosed publicly? Here we explain the tradeoffs involved in disclosure and motivate the ultimate choice: a combination of public cryptographic primitives embedded in a private defense-in-depth system.*

## INTRODUCTION

PolyCipher is developing with Cable Television Laboratories a proposed foundation for downloadable conditional access (DCAS) for the U.S. cable industry. This system represents a major departure in cable security system design, opening up the possibilities of lower cost and increased flexibility in managing access to cable content, while maintaining backwards compatibility with the industry's installed base of security equipment.

In designing a major security system, there are important questions on what kinds of information are made public and what is kept secret. The choices made affect both the basic security of the system and the industry and the range and simplicity of implementations.

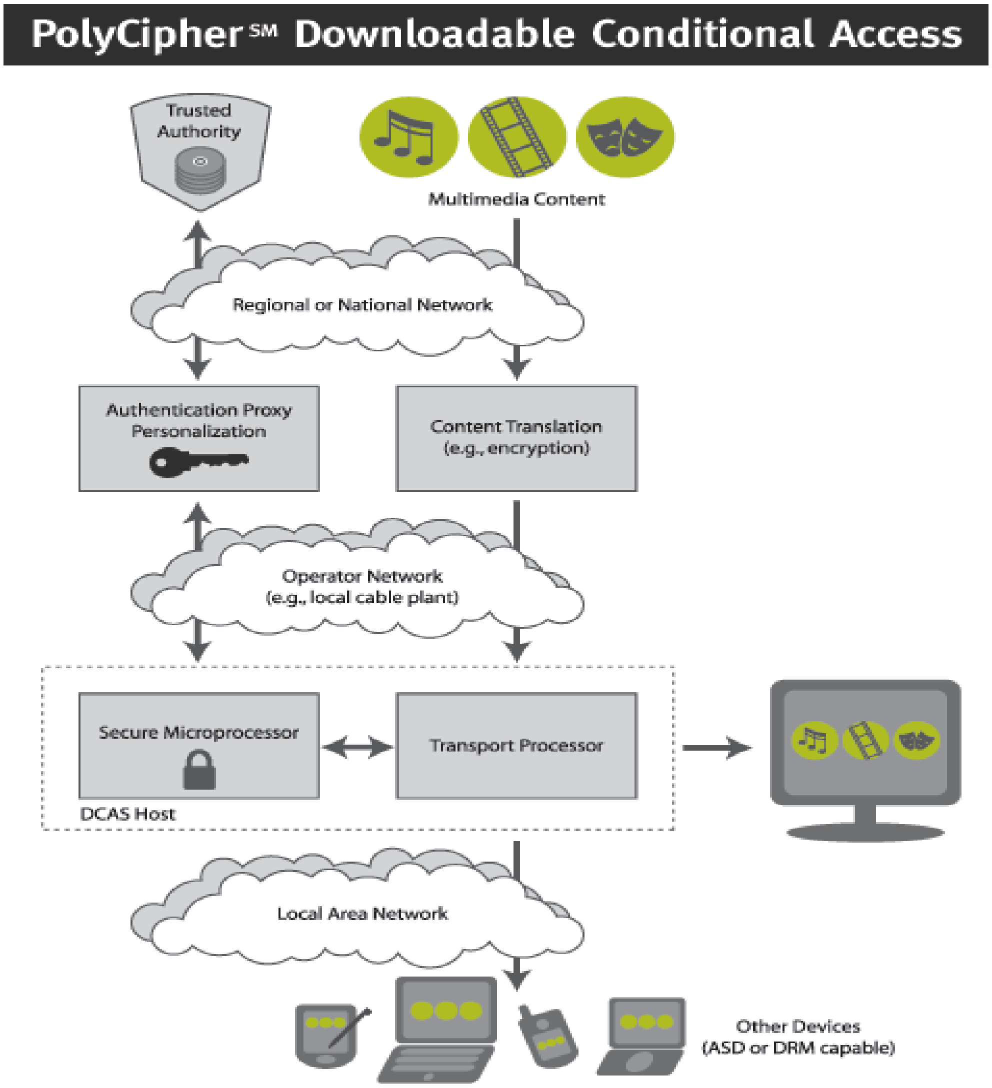
The principle security paradigm employed in DCAS is "defense-in-depth." In the paper we explain how this contrasts and interacts with other important concepts in security design, including "security by design," "security by obscurity," Kerckhoffs' principle, the economics of attack and defense, effective approaches to security system review and qualification, and the role of open source, cryptographic primitives, and modes of standardization for security systems.

The resulting choices are intended to produce a robust security system approach that can achieve both lower costs for the industry and appropriate security to maintain the industry's unparalleled access to high quality content over the coming decades.

## DOWNLOADABLE CONDITIONAL ACCESS

The PolyCipher Downloadable Conditional Access System (DCAS) is an emerging architecture designed to bring increased power and flexibility to the cable industry's effort to combat the piracy of its video, audio and other content.

Specifically, the PolyCipher DCAS architecture is focused on delivering security-related software clients to



A DCAS Host (a set-top box, TV set, or other compliant device) establishes its bona fides by contacting a trusted authority (a database of all authorized devices) via an authentication proxy.

Once authenticated, the DCAS host is personalized with the appropriate security client that will allow it to decrypt and display content transmitted over the cable network (typically, video & audio). Content can also be translated for use by ASD- or DRM-capable devices elsewhere on a local network.

compliant cable-ready hosts, including:

- Set-top boxes and devices
- Cable-ready televisions
- Home entertainment systems
- Cable-ready mobile and portable devices
- Other emerging products

Content security for these devices has traditionally been handled via hardware, through some combination of set-top devices or the installation of CableCARD$^{TM}$s. Unfortunately, this hardware module-driven approach requires significant manual effort to upgrade or change security systems at the cable customer level. Furthermore, hardware modules must be shipped, inventoried and repaired, all of which drives up operating expenses and limits the flexibility of the cable operator.

The PolyCipher DCAS architecture eliminates all this hardware module shuffling because it allows security systems to be automatically downloaded to compliant devices, using the existing cable infrastructure. Furthermore, the PolyCipher DCAS specification provides control over a broad range of security-related functions, including traditional conditional access systems (CAS: control at the host device of access to content), authorized service domain (ASD: control of other local devices all secured by the cable operator's security system), and digital rights management (DRM: bridging to other security systems on other local devices).

Hardware Architecture

The PolyCipher DCAS hardware architecture includes a Secure Micro (SM) and a Transport Processor (TP). The SM is a hardened and limited-capability microprocessor that primarily enables the decryption of multiple video streams, under direction of the installed CAS client. It does this by providing the necessary key management services for the TP and enabling a secure bootstrap of the software system [1].

The download of clients (CAS, ASD or DRM) to the SM is securely managed in the network operator's headend via the interaction of the SM software and a DCAS authentication proxy. The TP is primarily used for encrypting and decrypting the video and media protected by the SM clients.

Software Architecture

The PolyCipher DCAS specification defines many key elements:

- Messages between the SM and the DCAS servers in the headend environment
- Requirements for SM and DCAS Hosts to support DCAS
- Requirements of the headend server
- A new key management infrastructure

The new key management infrastructure supports the DCAS architecture by providing custom protocols, performance and security requirements, and by defining the necessary levels of interoperability, accountability and security.

## OPENNESS AND SECRECY IN DCAS

In simple terms, a security system can be thought of as an algorithm and a key. The algorithm explains what happens to accomplish the functions desired if the key is known. The key is a piece of data that is kept secret except to those who need to use it to operate the system.

While there is no debate about the importance of keeping the key secret, there is an ongoing debate on when to keep the algorithm secret.

Making the algorithm public allows for a broad community with diverse tools and perspectives to carefully evaluate it. Many an assistant professor or graduate student can win glory (and maybe tenure) by finding a critical flaw in a proposed algorithm. The notion that a security system should only rely on the secrecy of the key for its security goes by the name Kerckhoffs' principle [2, 3]. A frequent complement to Kerckhoffs' principle is the notion that reliance on the principle can only be reasonably assured by submitting a security system to broad public scrutiny.

Many important security primitives (algorithms focused on a narrow security function) are publicly vetted in exactly this way, examples include AES, RSA, and a variety of others that are used commonly in security systems of any scale. Some, like DES, were originally created in secret but later subjected to intense public scrutiny.

More generally, software systems such as operating systems and browsers have been the subject of a debate on the relative merits for security of open and closed source. These systems can be quite complex but are also of intrinsic interest to a large community of developers. There does not seem to be a definitive answer as to whether such systems (complex but of broad interest) are more secure in open or closed source form, although a number of authors give the edge to recommending open source for security [4, 5, 6, 7].

However, many large security systems do not fully subscribe to public scrutiny but choose, instead, to keep substantial parts of the security system algorithms private. A relevant alternative paradigm, especially common in pay-TV systems, is called "defense in depth" [8]. This approach views security as an economic competition between the security operator and his opponents (pirates and hackers); the goal is not perfect security – which is deemed to be unachievable – but rather a situation in which the economic gain to the pirates and economic losses of the security operator are both sufficiently low to maintain the viability of the operator's business model. Defense in depth envisions a series of counter measures, all kept secret, which are sequentially deployed when and if previous deployed counter measures are breached. The idea is that with each deployment, there is an extension of the period of low gain for pirates and low loss for the operator.

So, what should reasonably decide which kind or which parts of a security system algorithm should be public versus secret?

A useful answer is: when the system is simple enough that the cost to evaluators to test it is more than offset by their expected gains (in money or, more typically, fame and reputation), the public scrutiny that is frequently equated with Kerckhoffs' principle will be valuable. Conversely, when the system is complex enough and of limited enough interest that public scrutiny will only result in incomplete vetting, public disclosure may backfire, since an attacker needs find only one flaw in a public algorithm to breach it (while the evaluators have the much harder task of attempting to test and eliminate all possible flaws). As Schneier points out, there are limits to the amount of gratis work one can expect of the security community: "Security researchers are fickle and busy people. They do not have the time, nor the inclination, to examine every piece of source code that is published." [9, pp. 343-346]. Or, in Anderson's words: "Arguments against open source center on the fact that once software becomes large and complex, there may be few or no capable motivated people studying it, hence major vulnerabilities may take years to be discovered….the important questions are how much effort was expended by capable people in checking and testing the code – and whether they tell you everything they find" [8, pp. 296-207].

The Data Encryption Standard (DES) encryption algorithm, for example, can be described in about 100 lines of source code. It is straightforward for academics, hobbyists, and professionals to understand

and analyze it in detail from many different perspectives. The same is true of many cryptographic primitives.

The protocols, procedures, and algorithms involved in a full scale conditional access system sit near the other end of a continuum. Describing these fully could easily run to thousands of pages of documentation. It is so expensive to fully comprehend these that the amount of evaluation that can be expected (other from those explicitly paid to do so) is quite limited. And the intrinsic motivation for the security community to protect a particular instance of a conditional access system could reasonably be expected to be much less than that for a widely used application like an operating system or browser. Moreover, systems this complex are simply never bug free; the combinatorics of analysis and testing make this infeasible. The result is that such systems are rarely if ever made public. This does not mean there is an intention that the secrecy of the algorithm is its sole defense; indeed Kerckhoffs' principle is as much an objective here as it is in a cryptographic primitive. But both the reality of large scale system creation and system test and the particular economic incentives of system creators, pirates, and potential reviewers mean that striving for the goal of Kerckhoffs' principle is supplemented by the use of defense in depth to manage the economics that are the fundamental driver in protecting a commercial conditional access.

Swire has developed a thoughtful analysis of the economic, legal, and regulatory implications of tradeoffs in security system disclosure in a pair of papers [10, 11]. There he provides a useful comparison with military cryptography (remember, Kerckhoffs was in fact addressing military uses): why is it that militaries consistently find it valuable to keep cryptographic algorithms secret, even while adhering to Kerckhoffs' idea that they shouldn't design with a dependence on algorithms' secrecy for their success? Ultimately, the analysis there is economically motivated – as in this paper: if defenders profit more from exposure than attackers, then *disclosure* is valuable; if not, then not. Note again, though that Kerchoffs' principle is always valuable in *design*.

The implications for open standardization of security systems follow. Security primitives certainly benefit from the evaluation possible in an open standards setting (although this is not the only way to obtain public scrutiny – for example, an alternative is to publish a patented algorithm and provide a prize to those who breach it). Large scale security systems benefit if they proportionally scale in their interest to the security community – so experts find it worth their while to provide free scrutiny. Less interesting systems, such as a particular conditional access system implementation, do not benefit from open standardization – in the sense that security is weakened rather than improved in the particular and deciding context of system economics. This leaves open the exact boundary between a security primitive and a large scale system, but the principle is clear.

## PolyCipher Approach

The PolyCipher approach attempts to find the best combination of open, public algorithms and closed, defense-in-depth system design. Cryptographic primitives used in the system design are drawn from those that are widely used and well established, such as the DES and AES encryption algorithms.

The overall system is not public, though, and is divided into tiers of system design and documentation that are increasingly access restricted. Although the overall system is too complex to expect that any substantial gratis vetting would be applied if the system were made public, *paid* vetting by experts is extensively applied. Additionally, stakeholders who will depend on the security of the system and who are themselves expert in this type of security are invited to audit the system design. Motivations for stakeholders vary and in many cases they can be expected to be quite critical of the system. The resulting review environment is contentious but thorough.

## CONCLUSION

The PolyCipher DCAS system offers a new strategy for providing hardware rooted downloadable security for the cable industry. The question of how much of the system to publicly disclose is an important one and requires careful thought. But after consideration, the complexity of the system and narrowness of application suggests a hybrid approach: the use of well vetted publicly known cryptographic primitives embedded in a "defense in depth' system subject to extensive but private review.

## REFERENCES

[1] William A. Arbaugh, David J. Farber, Jonathan M. Smith, "A Secure and Reliable Bootstrap Architecutre," in *Proceedings of the IEEE Symposium on Security and Privacy*, May 4-7, 1997.

[2] August Kerckhoffs, "La cryptographie militaire," *Journal des sciences militaries*, vol. IX, pp. 5-38, Jan. 1883, pp. 161-181, Fev. 1883.

[3] Bruce Schneier, "Secrecy, Security, and Obscurity," *Crypto-Gram Newsletter*, May 15, 2002, available at http://www.schneier.com/crypto-gram.html.

[4] Brian Whitten, Carl Landwehr, and Michael Caloyannides, "Does Open Source Improve System Security," *IEEE Software*, September/October 2001, pp. 57-61.

[5] Crispin Cowan, "Software Security for Open-Source Systems," *IEEE Security and Privacy*, January/February 2003, pp. 39-46.

[6] Rebecca T. Mercuri, "Trusting in Transparency," *Communications of the ACM*, May 2005, pp. 15-19.

[7] Jaap-Henk Hoepman and Bart Jacobs, "Increased Security Through Open Source," *Communications of the ACM*, January 2007, pp. 79-83.

[8] Ross Anderson. *Security Engineering: A Guide to Building Dependable Distributed Systems*, New York: John Wiley & Sons, Inc. 2001.

[9] Bruce Schneier, *Secrets and Lies: Digital Security in a Networked World*. New York: John Wiley and Sons, Inc. 2000.

[10] Peter Swire, "A Model for When Disclosure Helps Security: What is Different about Computer and Network Security," *Journal of Telecommunications and High Technology Law*, vol.2, 2004.

[11] Peter Swire, "A Theory of Disclosure for Security and Competitive Reasons: Open Source, Proprietary Software, and Government Agencies," *Houston Law Review*, vol. 42, No. 5, January 2006.

## AUTHORS

Tom Lookabaugh is with PolyCipher, 999 18th St., Su. 1925, Denver, CO 80202, and by email at tom.lookabaugh@polycipher.com.

James Fahrny is with Comcast Cable, 1500 Market St., Philadelphia, PA 19102 and with PolyCipher, 999 18th St., Su. 1925, Denver, CO 80202. He can be reached by email at Jim_Fahrny@cable.comcast.com.

# PRACTICAL VIDEO OVER DOCSIS® IMPLEMENTATIONS WITHOUT FORKLIFTS

Doug Jones, Chief Architect
Roger Slyk, Director, Product Marketing
BigBand Networks

## *Abstract*

*This paper describes a scalable IPTV (video over IP over DOCSIS) solution for providing video content to a wide array of consumer devices that span TVs and PCs, and other device capable of receiving and processing video streams over IP. The solution is extensible to other last-mile edge networks including 3GPP, Wi-Fi, and WiMAX wireless networks, as well as going "over the top" on IP access networks managed by third-party providers. It defines and describes a distribution network that sources live broadcast and stored content from VOD (Video on Demand), PPV (Pay Per View), and Internet sources, with elements of centralized and distributed processing and management. The proposed distribution network utilizes existing network infrastructures and CMTS (cable modem termination system) platforms as the basis for the introduction of advanced services delivery, thus avoiding significant equipment change out, capital expense and operational disruption.*

*The authors describe key business and technology requirements that are intended to be a reference when making decisions about the IPTV system design in order to maintain the business viability of the solution. These include:*

- *Ensuring the solution is agnostic of the last mile network;*
- *Enabling scalable, flexible and cost effective systems for wide scale deployment;*
- *Maximizing the seamlessness of the customer experience;*
- *Placing the intelligence in the network, and minimize the intelligence required in set-top boxes, to protect CPE (customer premise equipment) investment and maximize flexibility.*

*The authors advocate that, while there may be many paths to an IPTV reality, the best evolution plan utilizes M-CMTS (Modular CMTS) designs that leverage existing DOCSIS 2.0 infrastructures while accelerating towards DOCSIS 3.0 broadband speeds and capabilities. The authors describe an evolutionary strategy based on an M-CMTS approach that can drive the availability of downstream DOCSIS 3.0 channel bonding, enable more flexible allocations of upstream to downstream traffic flows, and provide additional performance benefits. Additionally, other parts of the solution are already specified by the cable industry, including PacketCable™ Multimedia and Embedded DOCSIS.*

## INTRODUCTION

Consumers are rapidly demanding an experience that blurs the line between "lean-back" consumption of TV-based entertainment and "lean-forward" multimedia activities on their personal computers. It is not a single killer application, nor an intentional move towards new technology that is driving this rapid change in consumer behavior, but a merging of technology, content and consumer acceptance. The ability to watch similar content in a home theater and on a portable device is increasingly being demanded; the capability to access a broad array of content from many service providers is rapidly becoming expected as well. Cable operators must plan for the flexible delivery of any content from any source to any consumer device in order to remain competitive in the long term. An open architecture based on IP (Internet Protocol) is the preferable long-term solution.

Cable operators need the capability to deliver their services to more than just traditional STBs (set-top boxes) over more than just the HFC (Hybrid Fiber-Coax) network. Cable services are evolving to include to new services that will be delivered to a new generation of devices such as mobile phones, wireless devices and personal computers. The network infrastructure will need to be capable of obtaining content from broad range of sources including broadcast programming, VOD, PPV and emerging "new media" outlets such as YouTube, MetaCafe and Ziddio.

Figure 1: Delivery of video content over DOCSIS infrastructures is increasingly popular

But there is a great incentive to have as many of the new services as possible delivered to the existing base of MPEG-2 STBs as well as to a new generation of CPE with advanced encoding technology and high-speed interactive communications channels based on DOCSIS and DSG (DOCSIS Set-Top Gateway).

The proposed network architecture to deliver these services is based on a core network of video and communications services which provide IP-in, IP-out transport and processing and any number of edge networks which provide IP transport to the customer. The delivery of IPTV services over the existing HFC edge network will be based on high capacity, cost-efficient M-CMTS components, as well as low cost, scaleable MPEG-2 edge QAMs. Other edge networks could include 3GPP (3rd Generation Party Partnership) wireless networks, or any kind of high-speed IP access network which the operator does not even own.

To incent customers, this solution will offer broader ranges of content, time-shifting and place-shifting and greater opportunities for personalization

## IPTV SERVICES

An IP service is a cable service delivered to an end user over an edge network capable of transporting IP. This makes no statement to the origination of the content; that is, it could be a traditional video service that an operator offers today, for example linear television service, or it could be a video clip from a so-called new media outlet such as a website or some other Internet-based source. These are both video services and they can be packaged into a service based on an operator business plan or objective, but either can be delivered to the consumer using IP.

IPTV services generally exhibit several characteristics including:

- content personalization through time-shifting, place-shifting and addressable advertising;
- expansion of programming choices by broadening the universe of content to include new media services available over an Internet connection;
- ubiquity of services due to the prevalence of networks which support IP transport.

Today, HSD and Voice are offered over the HFC network as IP services. The only cable service not offered over IP on the HFC is video although video services are distributed around cable networks (for example backbones and regional area networks) using IP, but delivered to the end user over the HFC edge network using MPEG-2 transport.

Following the trends, video will also go over IP on the HFC and one reason to do this is to prepare to offer cable services on other edge networks which could include wireless networks or other wired networks, with the common denominator that they support IP transport of cable services. It is also expected that services will be delivered over networks which are not necessarily owned by the cable operators themselves (for example, these services would be provided "over the top" on a network owned by third-parties).

Getting back to HFC, with Modular CMTS, DOCSIS 3.0 and other specifications such as PacketCable Multimedia and Embedded DOCSIS, the toolset is available to begin laying the groundwork to deliver all cable services, including video, over the HFC edge network using IP.

## NEXT GENERATION IPTV INFRASTRUCTURES: A VISION

The figure below provides a reference model for Internet TV that includes a core network and one or more last mile edge networks which has the goal of delivering cable services to a variety of end-user devices over a variety of edge networks. This network can deliver linear, switched and On Demand services. The diagram does not imply a one-to-one relationship between components.

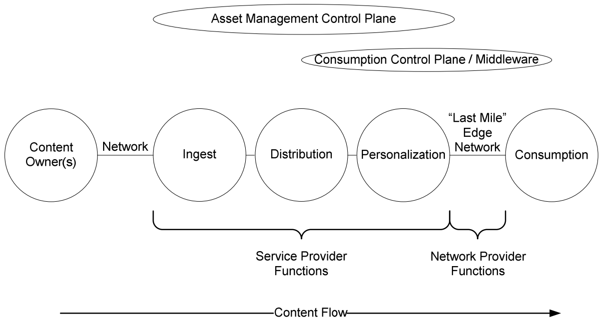


Figure 2: Internet TV reference network

The Ingest, Distribution and Personalization functions are considered part of a core network which can provide a variety of content and services to a variety of IP-capable edge networks. The system uses IP networking and transport as both the method for propagating content and as the method for staging content along the various steps of the system.

The paper first describes the HFC last mile edge network for supporting IPTV services, including M-CMTS and DOCSIS 3.0, to illustrate how these solutions work together to support high capacity, high bandwidth IP services to consumers. This section also describes how IPTV services could be delivered to existing MPEG-2 STBs.

The core network provides functions which are similar to existing cable networks. Later the paper posits how these core network functions are used to deliver cable services in an all-IP environment.

## EDGE NETWORKS

As alluded to earlier, various edge networks are possible and several are shown in the diagrams that follow.

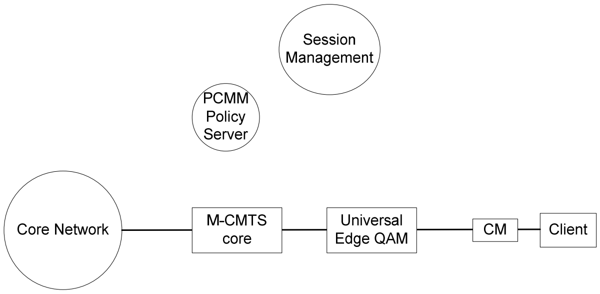


Figure 3: DOCSIS edge network

The figure above shows the DOCSIS edge network at a high level of detail, and includes the PCMM (PacketCable Multimedia) policy server and a session management function. The DOCSIS Edge network will be described in more detail in a later section.

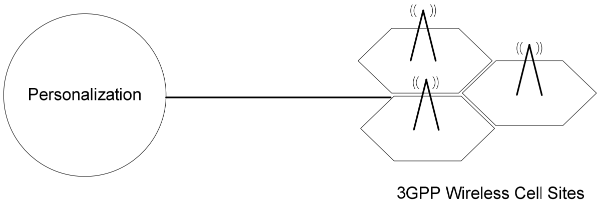


Figure 4: 3 GPP wireless edge network

The 3GPP wireless edge supports third generation mobile network services which include the ability to deliver higher data rates and advanced media features. The 3GPP wireless edge supports native IP services to customer premise equipment. 3GPP wireless technology delivers broadband-like data speeds to mobile devices while allowing consumers to send and receive services including data, voice, digital images, web pages, photographs and video three times faster than possible with a traditional GSM/GPRS network.

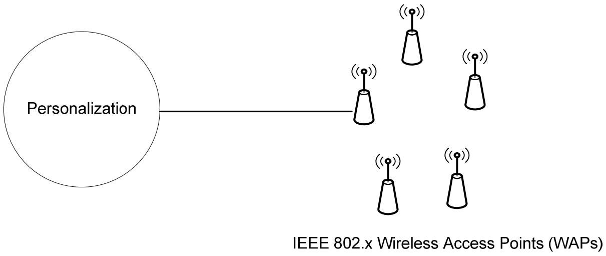


Figure 5: IEEE 802.x edge network

IEEE 802.x refers to a family of specifications developed by the Institute of Electrical and Electronics Engineers for wireless LAN technology. IEEE 802.11 (also known as Wi-Fi) specifies an over-the-air interface between a wireless client and a base station which supports data rates up to 54 Mbps per AP (Access Point) over a distance of up to 100 meters. IEEE 802.11 generally requires line of sight access and is suitable for hotspot access in areas of a city.

IEEE 802.16 (also know as WiMax) refers to a family of specifications developed by the IEEE for Broadband wireless access that provides secure, full-duplex, fixed wireless MAN (Metropolitan Area Network) service. Throughput can reach 75 Mbps and does not require line-of-sight to operate and reach can extend from one mile at full speed to thirty miles at reduced throughput.

IEEE 802.20 (also know as Mobile-Fi) is similar to WiMax but includes optimization for roaming and hand-off at speeds up to 150 miles per hour.

### DOCSIS 3.0 / M-CMTS Edge Network

The advantages provided by DOCSIS 3.0 enable an operator to provide IP video services over their HFC networks. DOCSIS in general provides IP transport over HFC networks and DOCSIS 3.0 in conjunction with PCMM provides the functions needed to deliver high-speed video services over IP over an HFC network.

DOCSIS 1.1 provided the necessary QoS (Quality of Service) for IP services and DOCSIS 2.0 provided higher capacity return path channels. DOCSIS 3.0 provides channel bonding services which dramatically increase both upstream and downstream speeds and capacities as well as the IP Multicast services needed to deliver video. The DOCSIS 3.0 specifications also describe how existing DOCSIS 2.0 and DOCSIS 1.x cable modems operate on DOCSIS 3.0 upstream and downstream channels. The combination of DOCSIS 3.0 and M-CMTS provide both the technical and economic solutions to providing video services over DOCSIS.

The M-CMTS specification describes an economical method for implementing downstream channel bonding which is a key enabler for offering entertainment quality video over DOCSIS.

In the M-CMTS architecture shown below, a device referred to as the M-CMTS Core contains the DOCSIS MAC (Media Access Control) functions which include all DOCSIS signaling functions, downstream bandwidth scheduling, and DOCSIS framing. A UEQ (Universal Edge QAM) contains mainly physical layer related circuitry, such as QAM modulators, and packet tunneling logic to connect to the M-CMTS Core.

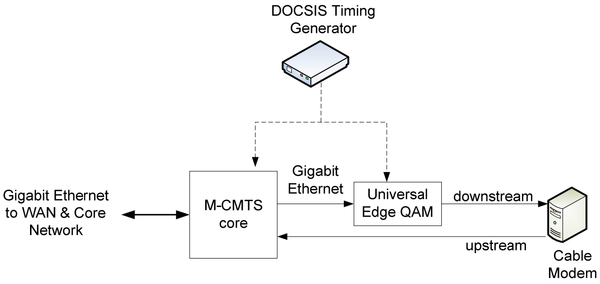


Figure 6: Conceptual view of M-CMTS architecture

Original CMTS implementations placed the CMTS core and QAM functions in an integrated assembly, fed with a common internal clock. The M-CMTS architecture creates interfaces to separate the downstream edge QAM from the M-CMTS core. It is this separation which allows massive, scaleable downstream capacity through the use of universal edge QAMs instead of downstream QAMs integrated into the CMTS as is the case today.

Since M-CMTS enables the separation of these components, the DOCSIS Timing Generator was created to supply the accurate and robust transport of a master clock signal between the M-CMTS components. The DOCSIS Timing protocol is structured to support all SCDMA and TDMA timing requirements.

An added benefit of an M-CMTS architecture is its ability to support flexible allocations of upstream and downstream channels to meet changing traffic pattern demands. That is, M-CMTS can offer more capacity in the downstream direction for video services during the evening hours, but symmetric upstream and downstream capacity during business hours when commercial data services are at peak usage.

## ACHIEVING SCALEABLE DOWNSTREAM DOCSIS CAPACITY

Delivering traditional cable video services over DOCSIS requires a lot of downstream capacity. M-CMTS allows downstream capacity to be added to the system by adding scaleable modular UEQs to the M-CMTS core.

The original integrated CMTS implementations included cards that had both upstream and downstream capabilities. These implementations were suitable for the relatively low bit rate services of high speed Internet and voice. But video services changed the requirements of the system by requiring significantly more downstream capacity than before. Integrated CMTS implementations cannot provide massive economic downstream capacity because each time a card is added, the upstream capacity of those cards is left stranded. Even a downstream-only card is not optimal because of the opportunity cost associated with sparing the additional downstream-only card in the CMTS chassis to provide redundancy for services.

The M-CMTS specification was designed to address these two issues and bring a third solution to bear by supporting the use of dense, scaleable UEQs to provide the downstream capacity while offering the opportunity to share the UEQs with other services such as traditional digital video, switched video and video On Demand.

DOCSIS 3.0 describes channel bonding services which supports the concurrent scheduling of DOCSIS data over multiple channels.

As shown in the following figure for downstream channel bonding, the CMTS distributes individual IP packets over multiple QAM channels in such a way that permits the CM (cable modem) to resequence out-of-order packets. By bonding together multiple downstream channels, the aggregate downstream speeds to a single CM can be 100 Mbps, or more, and have potential to increase broadband subscribers' downstream access speeds by factors ranging from 2 to 20, to support services ranging up to HDTV (High Definition Television).

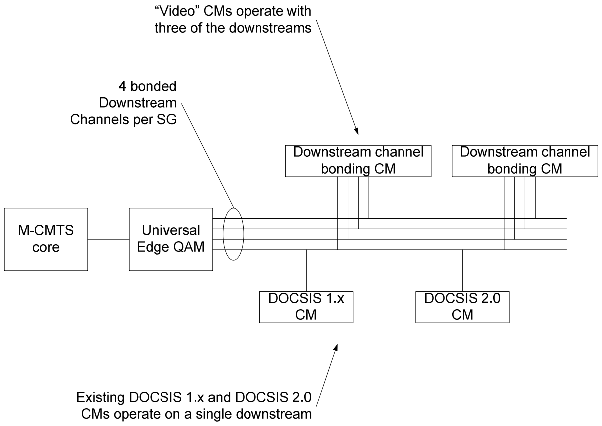


Figure 7: Downstream M-CMTS traffic flows

A key element of downstream channel bonding as described in DOCSIS 3.0 is the ability of existing DOCSIS 1.x and DOCSIS 2.0 CMs to operate transparently on channels which are bonded together for DOCSIS 3.0 operation.

Additionally, DOCSIS 3.0 allows the M-CMTS to assign individual downstream services to particular downstream channels. For example, an M-CMTS may assign a video-over-IP service flow to a downstream channel with deeper interleaving for higher reliability, while also assigning a voice service destined for the same modem to a different downstream channel with shallower interleaving for low latency.

## DELIVERING IP VIDEO OVER HFC NETWORKS

An M-CMTS architecture can be combined with DOCSIS 3.0 to economically deliver cable video services over a DOCSIS network, including linear, switched digital and On Demand services.

## Delivering Linear and Switched Digital IPTV Services

The DOCSIS 3.0 specifications describe a method for delivering linear IPTV services; however, it is the addition of PCMM which allows service guarantees to be offered through policy-based QoS management on the HFC network.

Policy-based QoS management ensures the bandwidth needed for the video service is allocated and managed on the HFC network. Without this management, it would be possible to put more services on the HFC than can be carried by the available bit rate, that is, video services could interfere with data or voice services, and visa versa. It is the PCMM function which allows the operator to manage the HFC resources and make intelligent business decisions in times of resource contention as to which services should be allowed on the HFC network.

There are a couple of models being investigated which use PCMM for delivering linear services and the intent is to provide the best customer experience by minimizing channel change times while providing the operator with the necessary control points to manage the service.

Both methods exhibit a key feature of DOCSIS 3.0 which is the ability to natively provide switched digital video as a by-product of the IPTV implementation using IP Multicast.

The diagram below shows the first method which follows a traditional IPTV model adapted to use PCMM where the IPTV client sends an IGMP Join to the network to "change channels" to a broadcast service.

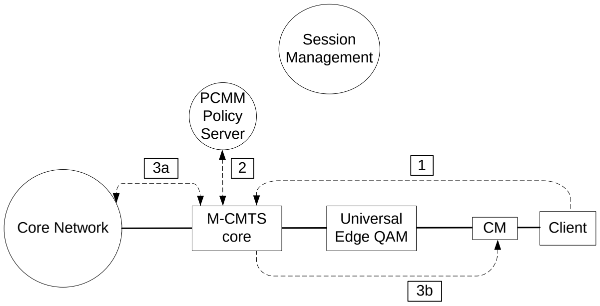


Figure 8: Synergy between PCMM and IPTV

To change to a linear channel, the IPTV client sends an IGMP Join (1) which is received by the M-CMTS core. If that channel is not already on a downstream accessible by that cable modem, but is already flowing into the M-CMTS core, upon receipt of the IGMP Join the M-CMTS will initiate a query (2) to the PCMM policy server to allocate capacity on the appropriate downstream for that service, then configure the CM for the IP multicast flow carrying that service (3b) and then forward the service to that cable modem.

If that channel is already on a downstream accessible by that cable modem, the M-CMTS core will configure the CM for the IP multicast flow which is already carrying that service (3b) on the appropriate downstream and the client will be able to view the content. This is an example of the inherent video switching capability of DOCSIS 3.0. Once a linear service is on a downstream channel, any client which can access that downstream channel can access that service; hence, only one copy of the service is needed on that downstream even if there are multiple viewers of that service. To complete the transaction, the M-CMTS core will signal the PCMM policy server of the channel change request.

If that channel is not already on a downstream accessible by that cable modem, and is also not yet flowing into the M-CMTS core, upon receipt of the IGMP Join the M-CMTS will first complete the query (2) to the

PCMM policy server to allocate capacity on the downstream for that service and then issue an IGMP Join to the core network (3a) to get the service to flow to the M-CMTS core, then configure the CM for the IP multicast flow carrying that service (3b) and then forward the service to that cable modem.

Notice that there is no explicit interaction with the session management function. Since all channel change requests are signaled to the PCMM policy server, session information can be queried from there. This is in contrast to the second way of implementing linear IPTV services with PCMM which explicitly uses the session management function.

The diagram below shows a linear IPTV model with a session-based controller adapted to use PCMM. Interaction with session management is very similar to switched broadcast. In this model, when a channel change occurs the IPTV client issues a session request to the session management function.

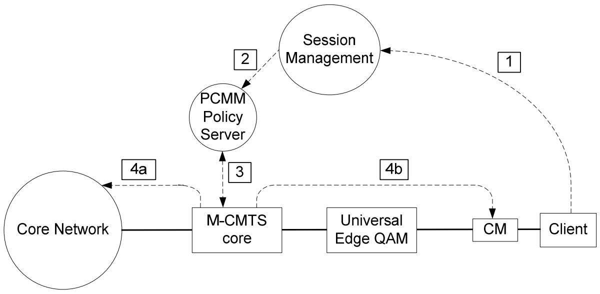


Figure 9: Linear TV model with session-based control

Upon receipt of the channel change request, the session management function interacts with the PCMM policy server to request bandwidth on the M-CMTS core for the service. If that channel is already on a downstream accessible by that cable modem, the M-CMTS core will configure the CM for the IP multicast flow which is already carrying that service (4b) on the appropriate downstream and the client will be able to view the content. This also is an example of the inherent video switching capability of DOCSIS 3.0. Once a linear service is on a downstream channel, any client which can access that downstream channel can access that service; hence, only one copy of the service is needed on that downstream.

If that channel is not already on a downstream accessible by that cable modem, but is already flowing into the M-CMTS core, the M-CMTS will configure the CM for the IP multicast flow carrying that service (4b) and then forward the service to the downstreams which services that cable modem.

If that channel is not already flowing into the M-CMTS core, the M-CMTS will first issue an IGMP Join to the core network (4a) to get the service to flow to the M-CMTS core, then configure the CM for the IP multicast flow carrying that service (4b) and then forward the service to that cable modem.

Note that in this case there is always an interaction with a session server, as well as interaction with the PCMM policy server.

## Delivering On Demand IPTV Services

The DOCSIS 3.0 specifications support the delivery of On Demand IPTV services as unicast streams. As with linear services, the addition of PCMM allows service guarantees to be offered through policy QoS management on the HFC network.

The model for On Demand IPTV services is very similar to existing On Demand services, except rather than send the stream through a video QAM it is sent through a DOCSIS QAM.

The diagram below shows an On Demand IPTV model with a session-based controller adapted to use PCMM. In this model, the IPTV client issues a session request to the session manager when requesting On Demand content.

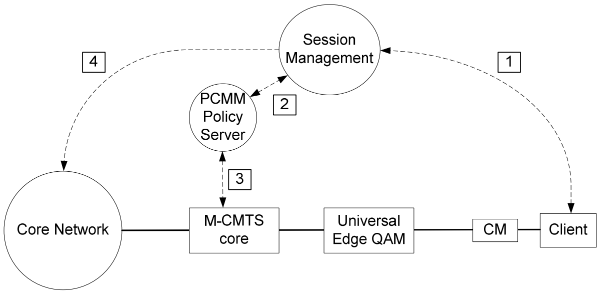


Figure 10: On Demand IPTV with session-based control

The session management then instantiates QoS with the M-CMTS core through the PCMM policy server (2, 3). If the bandwidth needed to deliver that service is available, the session manager will instruct the core network to begin streaming the service to the client.

Considerations for a Cable IPTV STB

An IPTV STB needs a connection to a DOCSIS cable modem. It also needs to be able to decode video from the CM to the decoder.

The following figure shows an example of an eDOCSIS (Embedded DOCSIS) STB solution where the DOCSIS CM follows the eDOCSIS specifications and the Client is considered an Embedded DOCSIS Embedded eSAFE (Service Application Functional Entity).

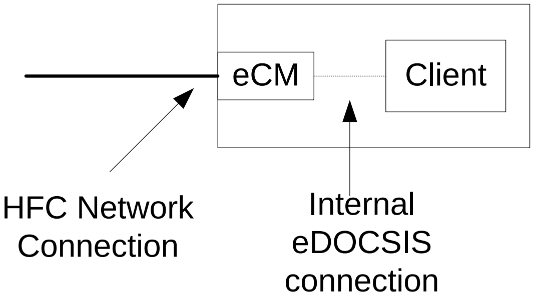


Figure 11: Embedded DOCSIS STB solution

The following figure shows a solution utilizing a stand-alone DOCSIS CM connected to the Client over a home network.

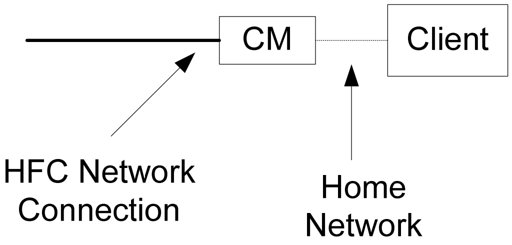


Figure 12: Stand-alone cable modem and client

In this case, the home network supports video streaming in order not to have a negative effect on the service.

With all DOCSIS CM solutions, an M-CMTS solution with downstream channel bonding would provide the most scaleable and economic solution. Downstream channel bonding provides additional capacity which will benefit video services. Additionally, DOCSIS/DSG is used for the out-of-band communication path thereby obviating need for legacy return path components in the STB.

The following figure shows an Off-Net CPE solution where the last mile access IP network is provided by a third party. In this case, a third party manages the network and the relationship with the NID (Network Interface Device). The cable operator is only a provider of content services to the client.

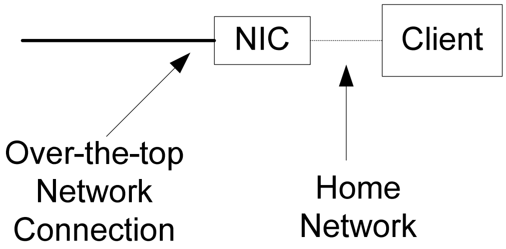


Figure 13: Off-Net CPE in over-the-top application

The client will need the ability to decode video in the format that is sourced from the cable operator. IPTV content can be encoded and delivered in several formats. The traditional format for cable digital video is MPEG-2 transport streams over UDP/IP, specifically packing seven 188-Byte MPEG-2 cells in one IP packet. This is the same format used to deliver On Demand or digital simulcast content over Gigabit Ethernet to a QAM. This same format can be used by an M-CMTS to deliver digital content to a DOCSIS QAM for consumption by the client.

An alternative method to deliver IPTV may be supported by “new media” sourced from Internet content providers. In this case, the content is coded as elementary streams over RTP/UDP/IP with no MPEG-2 framing. In this case, the timing for the content is recovered from the RTP layer instead of from the MPEG-2 transport stream layer. This method is how content is delivered to clients such as Windows Media and Real Player and there are standards available for delivering content using MPEG-2, H.264 and AC-3 codecs as well; hence, this may be an emerging trend for IPTV content.

## Delivering IPTV Services to Existing STBs

The IPTV architecture could be considered incomplete if it did not address delivering IPTV services to legacy STBs.

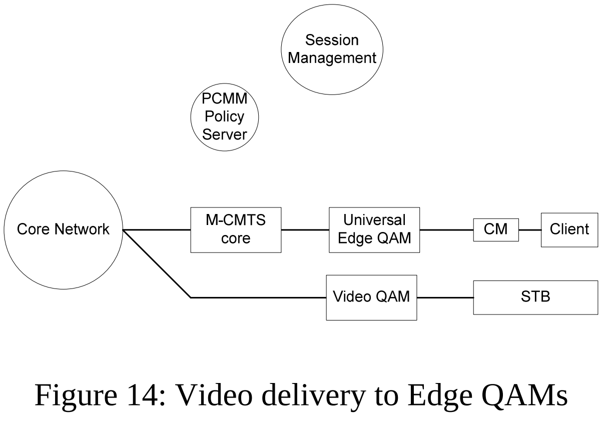


Figure 14: Video delivery to Edge QAMs

The above figure indicates how content available in the core network can be distributed to an existing video QAM just as easily as a DOCSIS QAM through an M-CMTS.

The session management function will know the capabilities of the entity requesting the service, whether it is an IP client or an existing STB application. When the entity places the request, the core network will stream the content over IP to the HFC network. If the content is destined to an IP client then the system will route the content over the M-CMTS QAM. If the content is destined for an existing STB, the system will route the content over IP to a video QAM just as happens today for digital simulcast and On Demand. The video QAM will take the content over IP as input and then strip off the IP and deliver the stream as native MPEG just as it does today to digital STBs.

Existing digital STBs can only render content which is encoded as MPEG-2 transport streams. Some IPTV content can be encoded as elementary streams over IP, with no MPEG-2 framing. For delivery of codec elementary streams to an existing STB, there will need to be a transcoding step to change the codec elementary streams over IP to an MPEG-2 transport stream over IP. This function is currently defined to happen as part of the personalization function.

With this transcoding function, it will be possible to deliver IPTV content, regardless of the encoding and transport, to both IPTV devices and to existing STBs.

## CORE NETWORK

The functions of the core network show many similarities to existing cable networks, including linear and On Demand services. The differences are in the areas of making the existing and emerging services available to a wider array of end devices. In all cases it is assumed the transport of content to these end devices uses IP, but these devices will have characteristics which vary from those of the traditional STB.

The core network can be characterized by three functions including Ingest, Distribution and Personalization.

Ingest Functions

The Ingest functions of the core network are shown in the following figure.

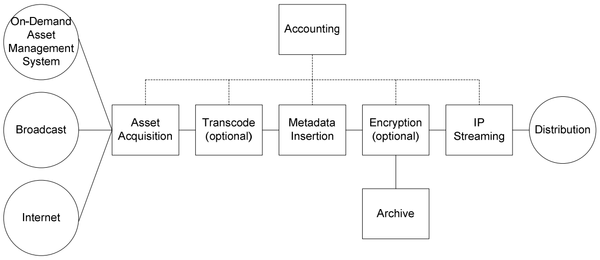


Figure 15: Ingest functionality in network model

The Ingest functions are designed to acquire content from a variety of sources including linear broadcast networks, On Demand storage and content from sources that might otherwise be associated with the Internet.

A key functions of the Ingest network is asset acquisition which includes the ability to connect and retrieve real-time and non-real-time content feeds including the reception and distribution of content from multiple independent sources to multiple distribution servers and applications.

The asset acquisition system streamlines and automates the process of receiving, staging, storing and propagating content including creating index files (for example trick mode files, for On Demand content.)

Another important function of the Ingest network is transcoding. The transcoding system processes MPEG-2 content and generates content in the proper format/profile and with the proper metadata for the media devices supported in the network.

Minimally the transcoding system must support the transcoding of MPEG-2 video and audio to H.264 and MP3 or AAC standards, respectively. In addition, the transcoding system may support conversion from H.264 to MPEG-2 for providing internet-sourced content to existing STBs. Both SDTV and HDTV formats should be supported.

For use with mobile networks, the transcoding system should support the media formats that are described in the PacketCable Release 2.0 Codec and Media specification, available through CableLabs. This specification requires the H.263 codec as well as recommends various profiles of H.264 as optional.

Depending on operator plans, additional codecs could be considered including the transcoding of MPEG-2 video into the following formats:

- Flash
- Windows Media
- Real
- Quicktime

Though these formats are available with CDN systems, there are many system implication to supporting these formats which need to be investigated.

Video content carried over IP can take one of two basic forms. The first encapsulates an MPEG-2 transport stream over IP as is used today to transport video streams over backbone and regional area networks. The second encapsulates codec elementary streams directly into RTP over IP with no MPEG-2 framing. These two methods are fundamentally different and are not necessarily interchangeable at end devices such as STBs and mobile video players. The core network should include the capability to transcode from one to the other while either maintaining the same codecs or when switching between codecs.

The next function includes the insertion of metadata which should comply with the CableLabs metadata specification, which may need an update to support the new types of content which can be ingested including obtaining content and metadata from various internal and external (for example the Internet) sources. This function should support both automatic and manual insertion and include the capability to translate from proprietary metadata formats to the format specified in the CableLabs metadata specification.

Finally, an encryption function should be present to protect content on the core network and an archive function could be present to maintain copies of content.

## Distribution Functions

The distribution functions are shown in the following figure.

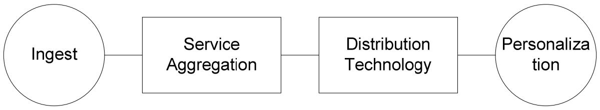


Figure 16: Service aggregation and distribution

Service Aggregation defines the set of service packages and bundles that are ultimately sold to the subscriber. Packages can be composed of real-time transcoded programs, multicast time shifted programs, unicast On Demand programs or a combination of all of the above.

The Service Aggregation system understands the current promotional offers and provides intelligent up-sell opportunities by interfacing with the billing system. The Service Aggregation system makes it easy to create new services, push personalized offers, and recommend add-on services and up-sell.

The Service Aggregation function interacts with the Ingest system and obtains information necessary for the On Demand application from other components, such as the asset metadata from the Asset Management System. This function of Service Aggregation interacts with the Host to present navigation menus and related application features to the On Demand Client and exchanges messages with the On Demand Client to enable the navigation functions, including parental controls.

The distribution technology will be chosen to provide a dynamic system for managing content in the network.

The most common method by which files are transferred on an IP network is the client-server model where a central server sends the entire file to each client that requests it. In this model the clients only speak to the server, and not to each other. The main advantages of this method are that it's simple to set up, and the files are usually always available since the servers tend to be dedicated to the task of serving, and are always on and connected to the Internet. However, this model has a significant problem with files that are large or very popular, or both. Namely, it takes a great deal of bandwidth and server resources to distribute such a file since the server must transmit the entire file to each client. The concept of mirrors partially addresses this shortcoming by distributing the load across multiple servers but it requires a lot of coordination and effort to set up an efficient network of mirrors and it's usually only feasible for the busiest of sites.

On the newer end of the spectrum are Torrent-based technologies which work by taking a large file and breaking it into pieces which are spread out over a network of servers. When the file is downloaded, it travels over the network in pieces from multiple sources which takes less time than downloading the file from a single source. The issue of scale with popular downloads is somewhat mitigated because there's a greater chance that a popular file will be offered by a number of servers. Overall, distributing the download in this fashion reduces the aggregate load on the servers and makes for a more efficient network.

## Personalization Functions

The personalization functions are shown in the following figure.

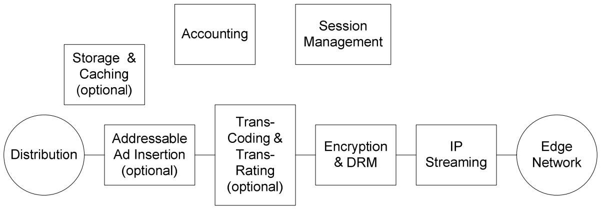


Figure 17: Personalization functions

## Session Management

The Session Manager should support the ability for components to scale to large numbers of active sessions and many session setups / teardowns per second. The Session Manager implements the service and business logic, including policy decisions, associated with Internet TV. The number of session may grow very large and the architecture needs to support optimized resource allocation when many potential resources for a particular session exist.

The architecture should be designed to result in minimal session setup latency by minimizing the number of messages used to set up sessions and allowing as much parallelism as possible between the components involved in setting up the session.

The architecture allows service and business logic to be incorporated into the resource selection process. For example, it may be more important to fulfill requests for sessions associated with a pay per view movie service than for sessions associated with a free On Demand Service. The architecture should be able to translate service and business logic into a set of rules to be incorporated in the resource selection process.

The architecture should make it possible to incorporate new classes of resources along with new resource managers without having to make major modifications to other components of the system. The architecture should allow the ability to modify or add resources to an existing session.

The architecture and resulting protocols should enable the components to be implemented using simple state machines. The architecture should also minimize the amount of state that needs to be maintained across components.

## Addressable Ad Insertion

Addressable advertising provides the capability of addressing specific ads to specific customers. An AA (addressable advertising) architecture is designed to make near real-time decisions on addressability that can differentiate between individual STBs and different users watching that STB. For example, if a home has just one STB the system should be able to discern if it is a parent or a child using that STB and address the appropriate ads. For example, the system could use near real-time analysis of remote control key presses to determine if the user is watching sports, dramas or cartoons and make the appropriate near real-time decisions on which ads to insert. The system can also make use of outside information such as billing profile, demographic and economic data.

Addressable advertising is available for all services provided over the network.

### Instant Channel Change

Instant Channel Change enables channel changing times that are less than traditional cable network tuning of digital channels. ICC (Instant Channel Change) effectively eliminates the delay associated with tuning channels in a digital cable system by exploiting the large video buffers in the STB available with an IPTV system.

### Trans-Coding & Trans-Rating

The trans-coding function is a final opportunity to modify the codecs used to deliver the service.

The trans-rating function is available to "down rate" the service from the bit rate used for transport on the network to a bit rate that is suitable for delivery to the client. For example, a 1.5 Mbps H.264 encoded service could be trans-rated to 300 kbps for delivery over an edge network which cannot support higher data rates.

### Encryption

The encryption function is used to make the service unreadable by users without knowledge of the keys to decrypt the service.

The encryption may be specific to the client and possibilities include Windows Media, Real Helix or any other scheme used by the cable operator.

### Digital Rights Management

Digital rights management is a comprehensive and flexible platform that enables rights holders to create business models for delivered content through the secure distribution of rights to trusted players.

The DRM (Digital Right Management) technology consists of two components including metadata associated with a piece of content which describes the business rules associated with that content. These rules generally indicate how the content can be consumed, by whom, and for how long. Services are built using the metadata. Additionally DRM includes cryptographic algorithms used to protect the metadata and the content from unauthorized use.

The DRM secures both the content and the business rules associated with that content. The devices enforcing the DRM must be trusted, e.g., STBs, PCs, portable devices, etc. and must be manufactured in such a way that the DRM is not easily defeated.

The DRM metadata is used to describe business models and should be flexible enough to allow an operator to implement a wide array of services. For instance, DRM metadata could require a piece of content to be deleted after 5 plays could disallow trick-mode playback during a certain window of time (for example a theatrical release).

DRM could be designed as an extension of the operators existing conditional access systems, and this should include multiroom capabilities, i.e., sharing content over a home network and possibly even over large geographic distances.

### IP Streaming

The IP Streaming function will support outputting content in either IP multicast or IP unicast streams. The IP Streaming function must support IPv4 and must provide an upgrade path to IPv6.

### Quality of Service

The QoS function must include support for DSCP (Differentiated Services Code Point) as well as possibly other IP-based QoS mechanisms as specified by the operator.

### Bandwidth Management

The Bandwidth Management function will monitor and police services and bandwidth usage.

### Storage and Caching

The Storage and Caching functions are used to keep select assets close to the edge. Examples can be assets which are stored as an extension to the Distribution System including movies or advertising content.

Storage at the network edge is also used for remote storage DVR (Digital Video Recording) in conjunction with the Consumption control plane.

## CONCLUSIONS

The paper presents a Video over DOCSIS solution using several available industry specifications including:

- Modular CMTS
- DOCSIS 3.0
- PacketCable Multimedia
- Embedded DOCSIS

Modular CMTS in conjunction with DOCSIS 3.0 provide the most economical means to deliver linear, switched and On Demand digital programming over an HFC IP network. The proposed solution maintains backward compatibility with both existing DOCSIS 1.1 and DOCSIS 2.0 infrastructure and describes a method for delivering IPTV content to existing STBs.

The solution is applicable to not only DOCSIS networks, but also other edge networks including 3GPP, Wi-Fi, and WiMAX wireless networks as well as going over the top on networks owned by third party service providers.

Functions for a core network are proposed that are not that different from cable networks today. Content ingest remains essentially the same with the added ability to transcode content to different codecs. Content distribution capabilities become more sophisticated to replicate traffic around the network where it is needed. Lastly, a group of personalization functions are described which allow cable services to be delivered to customers in a way which matches their method used to attach to the network, whether that be with a STB, mobile device, or some other customer premise equipment.

# SWITCHED UNICAST VIA EDGE STATISTICAL MULTIPLEXING

Ron Gutman, CTO & Co-Founder
Imagine Communications

*Abstract*

*Competitive HDTV offerings compel Cable Operators to make plans for 100+ HD Broadcast channels and 10% HD-VOD service (peak capacity), as well as devise innovative in on-demand and Time Shifted television services. Unfortunately, the current networks are unable to carry the load. The various bandwidth enhancement options are capital intensive, disruptive to existing infrastructure and, for the most part, don't provide sufficient bandwidth. This study analyzes SDV Unicast and Edge Statistical Multiplexing (StatMux), concluding that it is the most cost-effective and least disruptive method for handling the HDTV bandwidth explosion.*

## HDTV-driven Bandwidth Explosion

HDTV is beginning to permeate mainstream America. Currently in over 19 million US homes, i.e. at 17% current penetration, HDTV is projected to climb to 81% by 2010, according to Kagan Research. Unlike the historic migration from analog to digital TV, whereby some percentage of the subscriber base continued to access both formats, it has been recently observed that a significant portion of HD tuners are tuned almost exclusively to HD content, indicating that once exposed to HDTV, some consumers become disenchanted with SDTV.

Concurrently, HD-DVD and Blu-ray are making their debut in the retail market, increasing the demand for HD content and raising the video quality bar. HD-DVDs offer a stunning viewing experience at average bit rates (VBR) of up to 20Mbps using H.264 compression.

Video providers are unfolding their HDTV deployment plans: EchoStar carries 30 National HDTV channels, plus locals, planning on rapid growth. DirecTV plans to carry 60-100 HD by year’s end, with new satellites providing capacity for 150 national and over 1,000 local HD channels. Verizon FiOS carry 20 HD channels, claiming “best video quality”. AT&T U-Verse plans to offer 25 HDTV channels.

It is apparent that the new competitive landscape for acquiring and retaining premium subscribers is centered on HDTV content. Cable operators must make plans for 100 HDTV broadcast channels and improved HD-VOD service.

VOD has established itself as a useful “weapon” in this contest. There are 28 million VOD-capable homes in U.S, with over 3 billion streams served in 2006, and a steady increase in adoption, usage and ARPU. TWC's “Start Over” service has been observed to increase the on-demand usage by 25%, and additional Time-Shifted TV services are emerging. Playlist and dynamic ad insertion techniques are being tested to enable a further increase in on-demand content and profitability.

However, the spectral occupancy of the combined SD-VOD, HD-VOD and HD Broadcast services is identified as a key limiting factor.

## Bandwidth Expansion Options

How much bandwidth will be actually needed? Over the next couple of years, to keep up with competition, cable operators might be required to provide 100 HD Broadcast channels, 8-10% peak capacity HD-

VOD service and up to 25% SD-VOD service, to account for the increase in VOD library content, better movie release windows, ad-supported content and Time-Shifted TV offerings. Advanced high-quality encoders and new post-encoding techniques will enable multiplexing of three HD streams per QAM channel (3:1) at good quality. A straightforward estimate indicates that the operators will require 33 6MHz slots or about 200MHz to provide 100 HD channels. On-demand usage will call for minimizing the Service Group (SG) size at the fiber node, down to about 500 digital tuners per SG. It is estimated that at 20% HD penetration and 10% HD-VOD service, each SG must supply 10 HD-VOD concurrent streams. SD-VOD usage of up to 25% implies 125 SD-VOD stream capacity. Together with HD-VOD, a total of (10 x 4 + 125) / 10 = 17 QAM channels per SG. This is 13 more VOD QAM channels relative to the currently deployed 4 channels per SG! Altogether, HDTV digital broadcast and VOD expansion will further require about 40-50 additional 6MHz slots, or 240-300 MHz.

The primary question is whether it is even physically feasible to provision the required bandwidth. The second question in line concerns the economics: cost per subscriber per 6MHz slot.

In principle there are several bandwidth expansion options. We first consider a costly plant upgrade from 750MHz to 860MHz plant – unfortunately such an upgrade barely provides an additional 100MHz bandwidth, hardly addressing the requisite capacity enhancement. Node split is again expensive and only doubles the Switched tier but not the Broadcast tier, freeing at most two to six (2-6) 6MHz slots.

Other bandwidth expansion options involve massive CPE upgrades. However, these options are too disruptive, too expensive, and entail a certain amount of churn. In particular, analog tier reclamation requires many additional expensive STBs and high installation costs. Moreover, even when the STBs are burdened on the cable operator's balance sheets, it has been observed that such upgrades encounter significant resistance from certain subscribers. Unlike the HDTV experience recounted earlier, some people do prefer analog (and no set-top) over digital SDTV. Similar problems arise upon considering upgrades to H.264, denser QAM, and 1GHz, all of which require STB replacement and expensive installation at the subscribers' premises. There are other options such as spectrum overlay, FTTH and dense home base decoder, which do not entail STB upgrades, however those require expensive plant upgrades and/or expensive CPE home gateways, and even higher subscribers premises installation expenses.

## Switched Digital Video

A promising bandwidth expansion option is Switched Digital Video. The primary architectural alternatives for SDV are multicast and unicast. While multiple subscribers can share a multicast stream, a unicast stream is unique to each subscriber. Unicast SDV enables more personalized media services, included virtually limitless content offerings and highly targeted ads, closely resembling Internet content delivery in its ability to target specific content to specific consumers. Most importantly, Standard Definition SDV Unicast can eventually include the entire SDTV Broadcast tier, freeing up a lot of spectral space for HDTV Broadcast and HD-VOD.

## SD Multicast

A typical deployment of Standard Definition SDV Multicast involves an SDV pool of 120-150 SDTV channels over 8 QAM channels. In a Broadcast tier, assuming good-quality encoding, post-encoding and

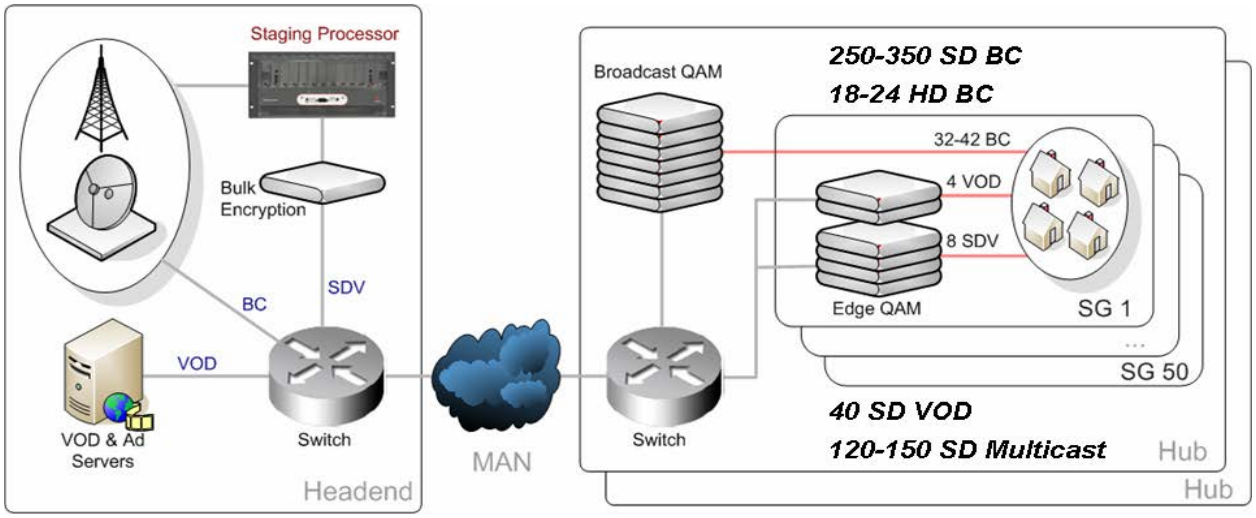


*Figure 1: Today's typical network architecture with SD Multicast*

StatMux technologies, it would be possible to multiplex 15 SDTV streams per QAM channel (15:1) at today's typical SDTV quality. This means that 120-150 SDTV channels in the Broadcast tier would occupy 8-10 QAM channels. Equivalently stated, Standard Definition SDV Multicast savings, in the way it is deployed today, would only amount to, at best, 2 QAM channels. By adopting a more aggressive content-to-capacity ratio (sometimes referred as concentration ratio) it might be possible to save a few more QAM channels, but this is hardly sufficient. Following is typical network architecture with Standard Definition SDV Multicast.

The network architecture illustrated in Figure 1 provides 4 VOD QAM channels, 8 SDV Multicast channels and 32-42 Broadcast QAM channels depending on analog tier size (72-80 channels typically) and the plant bandwidth (750MHz or 860MHz). Assume typical Service Groups (SG) of 500-1000 homes-passed and 500 tuners, 400 of which are SD, and 100 HD at 20% HD penetration. Under these assumptions, the architecture of Figure 1 is able to provision 400-450 SDTV Broadcast channels, around 20 HDTV Broadcast Channels, only 8% SD-VOD service and no HD-VOD service.

## SD Unicast & Edge StatMux

Standard Definition SDV Unicast (SD Unicast), combined with advanced Edge StatMux technology can address the bandwidth issue in a cost-effective way and provide a non-disruptive migration. SD Unicast, powered with Edge StatMux capability, provides: (i) limitless SDTV content, 100+ HD Broadcast channels and 8-10% HD-VOD service without adding a single QAM device while laying the groundwork for a robust future bandwidth optimization roadmap; (ii) significantly improved video quality enabling more effective competition with less constrained media such as HD-DVD/Blu-Ray and Verizon FiOS; (iii) a technology path to truly personalized ad insertion; and (iv) an improved overall user experience, in particular instantaneous channel change support.

In the proposed network diagram illustrated in Figure 2 below, a dense Edge StatMux either replaces the IP Switch for VOD and SDV or is positioned immediately downstream from the IP Switch. The Edge StatMux module receives SPTS/UDP/IP signaling from the VOD Servers and over Multi

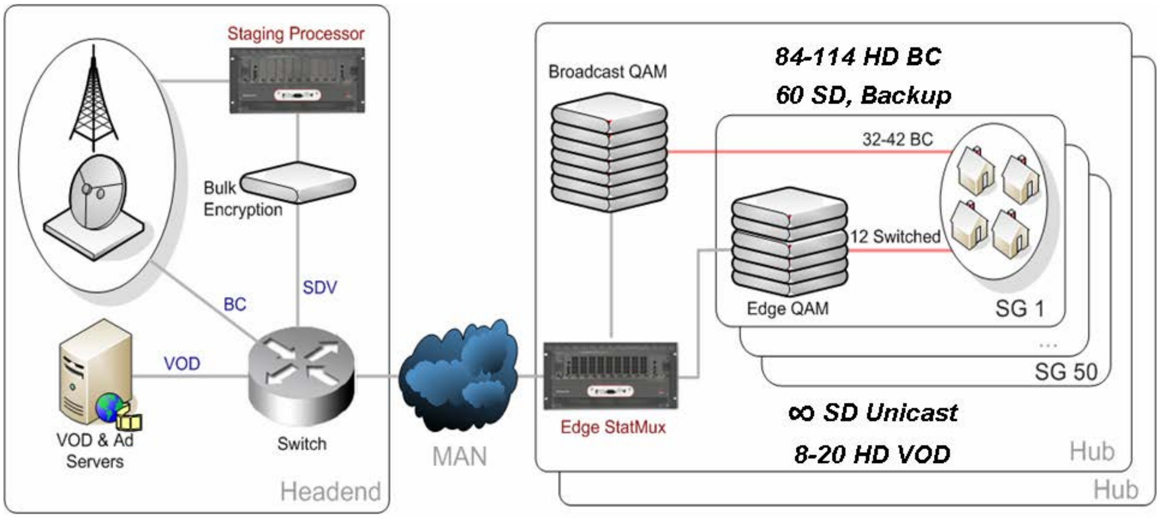


*Figure 2: Network architecture with SD Unicast and Edge StatMux*

cast IP (for the Switched and Broadcast tiers). The Edge StatMux outputs MPTS/UDP/IP to the QAM devices, and optionally locally multiplexes the Broadcast tier as well as providing local ad insertion and the ability to move content dynamically between Switched and Broadcast tiers. The Edge StatMux is connected in a one-to-one correspondence with the QAM channels at the QAM device, together forming the so-called "Edge QAM." An Edge Resource Manager (ERM) or a Global Session Resource Manager (GSRM) communicates directly with the Edge StatMux in such a deployment, instead of communicating with the QAM device, for the purpose of signaling channel changes.

<u>Edge StatMux Requirements</u>

The following list describes the essential set of performance requirements and features to be supported by an Edge StatMux, under the proposed solution:

1. **Price**. The benefits of an Edge StatMux are multiple: substantial savings in QAM device capital expenditures; freeing up of 6MHz spectrum slots; significant improvement in video quality (to be further discussed below); enabling personalized/targeted ad insertion; instantaneous channel change; easy migration to Switched Unicast; and enabling content mobilization between Broadcast and Switched tiers. Notwithstanding these benefits, in order to actually provide a scalable solution, the bottom line should be that the Edge StatMux price be below that of the conventional QAM evolving price, on a per stream basis.

2. **Video Quality**. Under peak usage, the Edge StatMux must provision at least 15 SD streams or 3 HD streams per QAM channel, at the common video quality acceptable for these services today. However, in VOD and Standard Definition SDV Unicast, most of the QAM channels are under-utilized most of the time. At least 99% of time, the video quality threshold can be significantly exceeded. During the brief times of peak usage, the quality should nevertheless be preserved at a minimum threshold. The system should then comprise QoS mechanisms ensuring uniform video quality under all conditions (detailed below).

3. **Density**. A single platform should support an entire hub of 50 or more Service Groups, i.e., at least 600-1000 QAM channels. The real estate utilization of hub's rack should not be increased as a result of deploying the new devices. The size of the platform should approximate the size of the QAM devices being replaced, and not to exceed 5-6 rack units.

4. **IP Switching**. The Edge StatMux must support standard 1GbE fiber or copper interfaces, and 10GbE in the future. It must perform non-blocking IP Multicast duplications, and supply a 2:1 IP unicast input to output ratio for ad insertion and VOD StatMux compression.

5. **Reliability**. The Edge StatMux platform must be Carrier-Grade with “5 Nines” availability, entailing No-Single-Point-of-Failure (No SPOF), redundant hot-swap fans, redundant load-sharing and hot-swap power-supply units, redundant hot-swap platform management card, and hot-swap N+1 processing/multiplexing cards.

6. **Interoperability**. The Edge StatMux must plug seamlessly into the existing infrastructure. It has to be able to re-use the current installed base of broadcast, VOD and SDV QAM devices, as well as off-load advanced functionality from the QAM devices. It should also support the installed base of VOD servers and possibly off-load dynamic ad insertion and slow-motion trick mode functionalities.

7. **Very Low Latency**. The delay is not to exceed an additional 100msec in VOD trick-mode transitions or SDV channel change. The Edge StatMux should also support instantaneous channel change to further reduce the delay by another 250msec on average.

8. **Encryption**. It is usually preferable in SDV deployments to use the more cost-effective upstream bulk-encryption technology rather than session-based encryption. Some amount of VOD content is also being pre-encrypted at content origination. Therefore, an Edge StatMux must support the variety of encryption schemes: SDV bulk-encryption, VOD pre-encryption as well as session-based encryption.

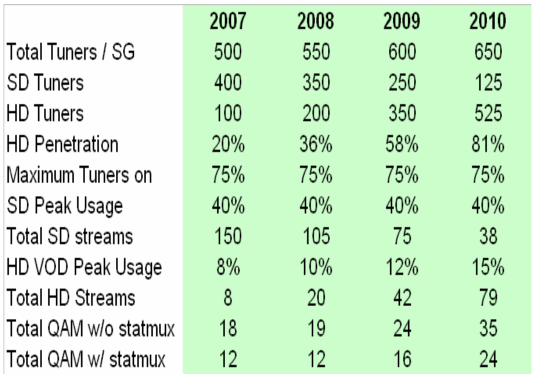
9. **Management**. Must support standard NGOD and ISA element management and session control protocols, managing by proxy the QAM device and the 1:1 StatMux to QAM connection, as well as aggregate alarms and events.

10. **Splicing**. Last but not least, an Edge StatMux should support the insertion of unique ads per stream with no capacity or quality degradation.

Content

Standard Definition SDV Unicast, combined with Edge StatMux capability, can provide 100+ HD Broadcast channels, 8-10% HD VOD service and unlimited SD content for Broadcast, VOD and Time Shifted TV, without adding a single QAM device (!). It also enables controlled content migration of SD Broadcast to SD Unicast and service migration, SG to SG and Hub to Hub.

Let's look at some usage assumptions and results:



| | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|
| Total Tuners / SG | 500 | 550 | 600 | 650 |
| SD Tuners | 400 | 350 | 250 | 125 |
| HD Tuners | 100 | 200 | 350 | 525 |
| HD Penetration | 20% | 36% | 58% | 81% |
| Maximum Tuners on | 75% | 75% | 75% | 75% |
| SD Peak Usage | 40% | 40% | 40% | 40% |
| Total SD streams | 150 | 105 | 75 | 38 |
| HD VOD Peak Usage | 8% | 10% | 12% | 15% |
| Total HD Streams | 8 | 20 | 42 | 79 |
| Total QAM w/o statmux | 18 | 19 | 24 | 35 |
| Total QAM w/ statmux | 12 | 12 | 16 | 24 |

Assuming:

- Typical Service Groups (SG) of 500-1000 home passed and 500 digital tuners out of which 400 are SD tuners and 100 are HD tuners (20% HD penetration in 2007).

- Aggressive digital penetration of 10% more digital STB per year.

- HD penetration growth projected to climb to 81% in 2010 according to Kagan Research.

Regarding the next two numbers to be assumed, it is important note that network design should be based on peak usage. The peak video consumption occurs Thursday evening Prime Time, 8-11PM. The number of tuners that are on in the Sun-Wed window is typically around 20%, however, on Thursday evening it might peak to 60-70%. Therefore:

- Maximum number tuners on is 75% per SG, SD or HD.

- Maximum tuners watching digital SD at peak usage is 40%. On Thursday night, it is estimates that 60-80% of subscribers are tuned to 20-30 channels in the analog/simulcast tier (e.g., ABC, CBS, NBC, Fox, CW, USA, FX, Fox News, ESPN, CNN, MTV, Disney, HBO, Discovery, PBS, etc.). Dynamically leaving 60 SD streams at the Broadcast tier or letting tuners tune to analog under peak usage will readily ensure the 40% peak usage assumption of SD Unicast.

- Although it has been observed that HD tuners predominantly “want” to tune to HD content, since the HD content offering is still insufficient in 2007, it is assumed this year that users will still tune to substantial SD content over prime time and, and this is accounted for in the SD Unicast capacity calculation. At the point where 100+ HD Broadcast channels and 10% HD VOD service are deployed, HD tuners are excluded from the SD Unicast calculation.

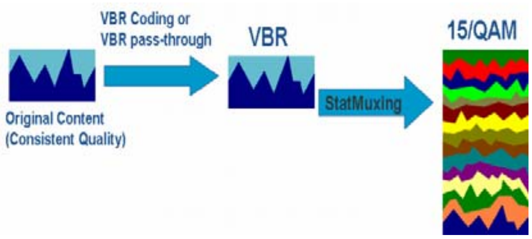
- SD VOD, Time Shifted TV, nPVR services are in high demand. In SD Unicast, these services are unlimited. Over time, similar services will be required for HD and therefore it is assumed that HD-VOD peak usage is to grow at a rate even faster than the SD-VOD growth curve.

The results of this analysis are quite promising. With just 12 QAM channels per SG it would be possible to deliver unlimited SD content and 8-10% HD- VOD service, while freeing up the spectral resource for about 100 HD Broadcast channels. In the 2009-2010 timeframe, if HD penetration increases as projected, a SG split of 2:1 will be further required to maintain 12 QAM channels per SG, in order to keep up with the increased HD-VOD services.

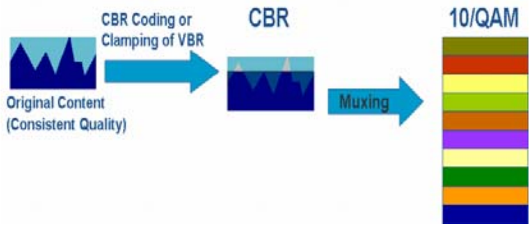
In the future, assuming SG/node splits will eventually reduce the number of HD tuners per SG to 250-300 HD tuners, this will enable, in conjunction with the Edge Stat-Mux architecture further expansion to unlimited HDTV SDV Unicast, completely removing the bandwidth constraint. At this point, the age-old adage “never enough bandwidth: may finally be inapplicable to cable’s robust architecture.

Video Quality Management

Video compression standards such as MPEG-2 and H.264 take advantage of temporal redundancies. MPEG-2 encoding of an SDTV signal, which is temporarily rich in motion and scene changes takes 4-4.5Mbps, providing the same quality as a "talking-heads" scene at 2Mbps. This is why a consistent quality coding technique must use the Variable Bit Rate (VBR) format, which provides the optimal video compression scheme. This is also why DVDs use VBR to optimize video quality and storage efficiency. Moreover, virtually all digital broadcast multi-channels content in North America uses VBR and Statistical Multiplexing.



Significantly, CBR encoding or the "clamping" process of conversion from VBR to CBR, limit the video quality by chopping off the peaks. Moreover CBR multiplexing reduces the QAM and bandwidth efficiency by another 33% down to 10 streams per QAM channel.



Current advanced SDV systems plan to overcome the CBR quality degradation problem by using a method called "Multi-rate CBR", assigning higher bit rates for selected high-complexity services such as sports channels, therefore reducing the QAM efficiency even more. With VBR coding, a good quality sports channel peaks at about 5Mbps but may averages approximately 2.5Mbps. Applying multi-rate CBR coding, it is only possible to carry 8 such “difficult” services over a 38.8Mbps 256QAM channel. In contrast, with VBR/StatMux at the same quality, it would be possible to carry 15 channels, i.e. almost twice as much.

In fact the process of allocating bit rate by the ERM/GSRM in a VBR/StatMux is akin to allocating bit rates in a multi-rate CBR system: every stream is assigned a different effective bit rate which resembles its VBR average, e.g. a sports channel may have a higher average bit rate than a news channel. The Staging Processor continuously monitors the VBR average and publishes it to the ERM/GSRM. The average bit rate is measured statistically per channel and may attain multiple values according to the time of day/week. A video quality monitoring system analyzes the VOD assets in advance in a similar manner, providing a deterministic VBR average to the ERM/GSRM over ADI interface bit rate field.

Until now it has not been economical to extend VBR usage to the video distribution network edge for advanced service architectures such as VOD and SDV. However, the Edge-StatMux architecture changes the equation. The new VBR StatMux solution optimizes video quality over each QAM channel. Advanced Edge StatMux technology can multiplex 15 SDTV streams at broadcast quality, without sacrificing the economics. In VOD and switched video environments, since the QAM channels are not fully utilized 99% of the time, it is possible to further raise the quality bar without sacrificing the capacity gains attainable with the Edge StatMux architecture.

## Conclusions

In this paper we addressed the various options to mitigate the bandwidth demand induced by the accelerated HDTV and VOD penetration trends. It is concluded that a promising avenue to meet this challenging spectral demand is a Switched Digital Video variant involving Standard Definition SDV Unicast and Edge Statistical Multiplexing (Edge StatMux). The benefits of the architecture in terms of bandwidth efficiency, deployment migration, video-quality, and overall user experience (including personalized ad insertion) are coupled with an upward leap in video quality enabled by porting the VBR video format to the network edge, with increased video quality and bandwidth efficiency. Our deployment projection model accounting for content and usage patterns indicates that the new architecture provides an important new tool in the arsenal of the cable television operators, to economically and competitively manage the HDTV and VOD penetration trends, while meeting enhanced quality of service requirements.

ron@imagine-com.com

# THE ‘BRIGHT SIDE’ OF DRM: NEW BUSINESS OPPORTUNITIES

Robert M. Zitter and Craig D. Cuttner
Home Box Office, Inc.

Disclaimer. The potential business scenarios described herein are outlined for technical feasibility and do not represent any proposed or actual business offerings.

## *Abstract*

*Digital Rights Management (DRM) has been, unfortunately, characterized as an “evil” method of restricting user’s rights and abilities to consume content. Quite the contrary is true – the use of DRM, providing usage rules and advanced security, enables new opportunities and new business offerings. This paper describes some of the potential business offerings that might be provided if an operator were to use DRM to distribute content outside the normal scenario of linear content and traditional VOD.*

## INTRODUCTION

Digital Rights Management (DRM) represents a comprehensive system of cryptographic controls tied to business rules that govern the use of content. Basically, DRM is a combination of Conditional Access (CA) that is used as an on/off switch for payment, and Copy Protection (CP) the coarse control over the duplication of content.

DRM adds the component of usage and business rules to the consumption of content – examples of content rules:

- Defined consumption (playback rules) such as number of permitted viewings, time limit before viewing window ends;
- Rules for permitted copies that can be made on certain devices under controlled circumstances
- Rules for content movement and requirements for protected outputs or security of devices including downstream (subsequent) other DRM systems.

Different DRM systems may have additional features or slightly different implementations but for the purposes of this discussion, these core features will be used to build business models.

As an example of how varied DRM system capabilities can be, here is a summary of a few key features of major DRM systems from public sources:

| **DRM Feature** | **Microsoft DRM** | **Apple Fair Play** | **Cable CA & CCI** |
|---|---|---|---|
| Restrict Access to “content for payment” | Yes | Yes | Yes |
| Limit copying to Free / Once / No Copies | Yes | Yes | Yes |
| Can require protected outputs (e.g. Macrovision) | Yes | No | Yes |
| Limit unprotected outputs | Yes | Yes | Yes |
| Limit viewing time and copy retention | Yes | No | No |
| Permit / Restrict movement specifically to portable devices | Yes | Yes | No |
| | | | |

## NEW OPPORTUNITIES

Existing “sales” of content generally reflect either a single narrow use (e.g. VOD / PPV) or almost a perpetual right to view content (such as physical ownership of a DVD). Intermediate variations are also possible such as limited-use subscription, or transaction-based rental.

Most opportunities that use DRM require that all “devices” in the chain of content and/or transaction have an implementation of the same DRM system and/or a compatible downstream DRM system via an approved output – including the “client” device used to consume the content.

### Scenario One – “Burn to Own” DVDs

Conventional (standard definition) DVD’s use a copy-prevention system called Content Scramble System (CSS). Other than the prevention of copies (and, in some cases, enabling output controls such as Macrovision), there are no other proscribed DRM attributes such as limited “permissions” of copying.

Newer disc formats (e.g. HD-DVD) have an advanced system of copy management (Advanced Access Content System (AACS)) that is more of a full DRM system to address this shortcoming. However, business plans that wish to target the millions of existing (CSS-only) DVD players are obliged to “regress” to that technology.

Since DVD content is “authored” (both MPEG formatting as well as menus, chapters and other non-cable formatted data), it is impossible to “capture” a DVD from a live streamed program and burn to a disc. Content to be burned to disc has to be distributed to the burning device (e.g. specialized STB) using DRM to encapsulate the CSS-encoded DVD image file.

### Scenario Two – “Portable Devices”

Currently, conventional “linear broadcasts” are “fat” – i.e. they are distributed for pretty pictures on big screens – this is a poor choice for portable devices with small streams and small disk drives. Thus, a content “move” from DVR to portable device is inefficient and expectation of transcoding (bit-rate reduction) is a burden for existing DVR STBs..

An improved method to provide content to portable devices is to encode specifically for that target device and efficiently move content securely to the portable device.

Secure DRM can be used to authorize the delivery of content to portable devices.

## Scenario Three – "Electronic Sell Through"

Without advanced DRM, a file transfer of content to a PC would be considered "lost" to the Internet, thus any "retail" transactions that enable consumers to have and use content require a sophisticated DRM.

EST is the process where a "locked" file is distributed to a user with a proscribed rights-enabled DRM license to permit viewing and/or movement to portable devices and inside the home (e.g. on a home network).

## Scenario Four – "Early Window High Value Content"

Two variants of potentially-interesting new business opportunities are possible with a highly-secure and usage-restricted environment.

First, "Day and Date" refers to the deployment of on-demand assets that are available to consumers coincident with their home video consumer release. This represents an interesting opportunity since the consumer interest and revenue model of the existing home video marketplace are well characterized. Such distribution would probably require both copy protection (copy never, Macrovision protection for analog outputs) and potentially, restrictions on other outputs – such as component analog outputs that do not support Macrovision, limiting, perhaps this offering to standard-definition content as copy protection is not implemented on high-definition analog outputs. In these cases, either digital outputs with HDCP or image constraint would be required. Unprotected outputs would severely increase the risk that analog-sourced copies could be made and undermine the very significant home video-window revenue stream.

Another "Early Window" content idea represents a potentially new revenue stream that may afford electronic distributors with a very visible (and promotable) offering to consumers. With extremely robust DRM and content controls, content could be offered much closer to the theatrical "first-release" than was previously possible. Movie titles (for example), could be offered on a very restricted and limited basis to HDTV Home Theater households very near their release to theaters. Of course, copying or other content leaks could undermine several subsequent revenue streams and would not be tolerated. Thus, an early window release could include some or all of the following restrictions (which, due to the nature of a new business offering like this, may require review of the FCC's rules that govern output controls, etc.):

- Mandatory highly-secure distribution environment

- Copy Never on protected digital outputs, perhaps restricted to HDCP and not Firewire

- Analog unprotected outputs disabled (Selectable Output Controls (SOC))

- Limited viewing window (e.g. no pause, no 24-hour bookmark / retention)

- HD Only distribution (since this offering may be potentially targeted to home theater owners)

## CONCLUSION

The proper end-to-end implementation of advanced DRM involves cooperation of distributors, manufacturers (CE and others) as well as assurance that the overall robustness of the distribution architecture is without question.

While the form of new product offerings and revenue opportunities will be determined by those who create and provide them, their existence will all rely upon one common denominator: The use of a robust Digital Rights Management System capable of attaching and enforcing the specific rights granted to consumers who use these new products. Our ability to manage the digital rights associated with the video programming will open the door to new choices for consumers and new revenue for programming distributors.

# THE COST OF FAIR BANDWIDTH: BUSINESS CASE ON THE RISE OF ONLINE VIDEO, P2P, AND OVER-THE-TOP; PROACTIVE STEPS OPERATORS CAN TAKE TODAY TO MEET BANDWIDTH DEMANDS IN THE NEAR FUTURE

Michael Eagles, UPC Broadband
Robert F. Cruickshank III, C-COR

## *Abstract*

*The growth of popular video sharing sites such as YouTube, P2P file sharing applications such as Bit Torrent, and over-the-top broadband services has placed unprecedented strain upon the HFC network. Demand for the broadband pipe is growing daily and in order to ensure that all services and applications are delivered reliably, operators must plan for the traffic increase and take proactive steps to mitigate the rising demands.*

*This paper will take a multi-pronged view of the demands placed on today's network, including a recent history traffic analysis of sample systems and projections of where networks are headed tomorrow. Additionally, the paper will investigate both hardware and software options for proactively working to meet bandwidth demands today and over the next three to five years.*

## INTRODUCTION

### Increasing Bandwidth Needs

With operators launching many bandwidth intensive services including High Definition Television (HDTV), Video on Demand (VOD) and ever-increasing High Speed Data rates, there is an ongoing need for greater capacity. Here we take a close look at High Speed Data and the rise of online multimedia devices and applications to better understand current and future demand trends.

In this paper we use the term “bandwidth” to mean “RF bandwidth” and “information rates” to mean advertised product speeds. We discuss “consumption” in the context of usage trends and subscriber “bit-rates” or average speed in Kilobits per second (Kbps).

### Why Look at High Speed Data?

High Speed data is one of the few services where forces, largely external to the cable operator, shape the demand for capacity. In the recent rise of online video, P2P and over the top multimedia services, we notice a trend from the use of High Speed Data for “access” toward its use for “entertainment”. In this paper we seek to understand the impact of this shift on bandwidth needs.

### Methodology

We considered the level of congestion on today’s network by looking at a snapshot of global consumption patterns in addition to a specific per system example.

In assessing the recent and projected growth in demand and bandwidth needs we considered both historical trends in per subscriber average capacity, per subscriber profiles and device and application trends.

Finally we considered a tool-kit of alternatives for the cable operator, grouped into “manufactured bandwidth” or capacity expansion and “technology bandwidth” or capacity management.

## AN ANALYSIS OF TODAY'S NETWORK DEMANDS

In this first section of the paper we explore a snapshot analysis of the demands placed upon the HFC network by identifying global trends and considering data services in a major metropolitan city in the United States.

### Average Usage per Subscriber

We started with a review of global usage trends using a base of 2.75 million DOCSIS® Cable Modem (CM) devices. As shown in Table 1, the average high speed subscriber sends/receives about 300 Megabytes (MB) of IP traffic per day. As one might expect, average daily usage varies among different regions of the world, but not drastically.

|  | Average Daily Directional Usage [MB] | | |
|---|---|---|---|
| Region | Avg. Up | Avg. Down | Avg. Up+Down |
| USA | 108 | 210 | 318 |
| W Europe | 123 | 192 | 315 |
| E Europe | 95 | 183 | 278 |
| S America | 138 | 176 | 314 |
| Asia | 98 | 190 | 288 |

Table 1. Typical Average Daily CM Usage by Region during November 2006[i]

In order to identify sources of congestion on today's network we take a closer look at how this average usage is distributed across the user base and what attributes contribute to this usage.

### Usage by "Power Users"

A common belief is that some subscribers send/receive more or less traffic than other subscribers. A logical follow-on question is "What percentage of subscribers send/receive more than others?" Figure 1 shows the cumulative percentage of traffic sent/received by ~500,000 CMs during June 1995 in a major metropolitan city in the United States.

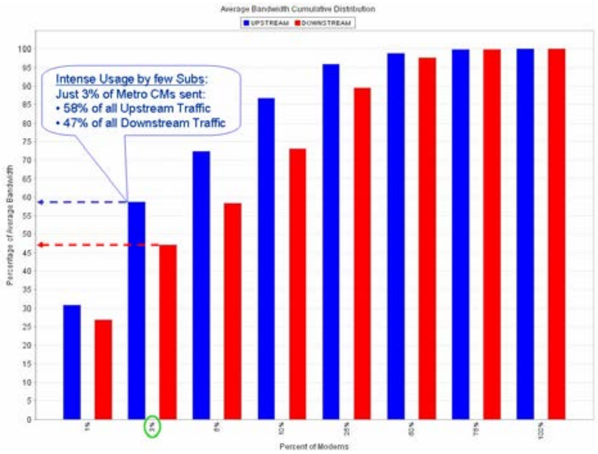


Figure 1. Average Bandwidth Cumulative Distribution in June 2005

Referencing Figure 1, one can see that just 3% of the CMs (on horizontal axis) sent 58% of upstream traffic (on vertical axis in blue) and received 47% of the downstream traffic (on vertical axis in red).!

Fast-forwarding almost two years to today, we see similar results in the upstream—just 3% of the CMs sent 56% of upstream traffic. But look, over the same time period the downstream number has dropped from 47% to 37% as shown in Figure 2.

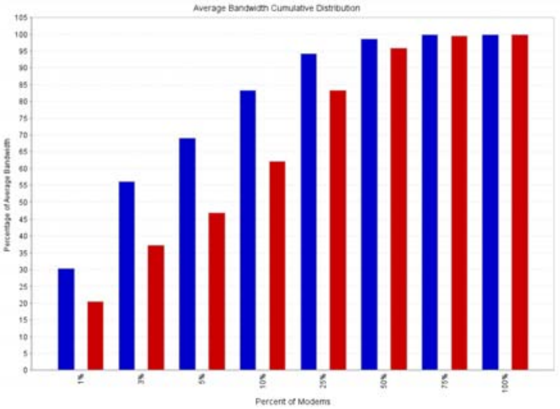


Figure 2. Average Bandwidth Cumulative Distribution March 2007

The takeaway messages from Figures 1 & 2 are: a) half to three-quarters of all IP traffic is sent/received by just 5% or less subscribers, and the remaining 95% of subscribers send/receive the balance of IP traffic on MSO DOCSIS® networks; and, b) the distribution of volume has changed due to an increase in downstream consumption by the "average" user.

We believe a possible reason for the distribution change is that bandwidth intensive, multi-media applications are becoming widely used by mainstream subscribers rather than only early adopters.

## Active vs. Total Cable Modems

Another question is "What percentage of CMs are active throughout the day and night?" Figure 3 shows the number of active and number of total DOCSIS devices (CMs) on a typical Metro upstream hour by hour over 1 week. The total number of devices (across the top) is relatively consistent and ranges from 216 to 220 (the number varying as CMs become unreachable, for example, when subscribers shut down/power up their CMs). The active number of devices (lower periodic trace) varies throughout day from 63 to 148. Dividing the number of active devices by the number of total devices at any hour yields the % of active devices (upper periodic trace) which is read on the right-hand vertical axis.

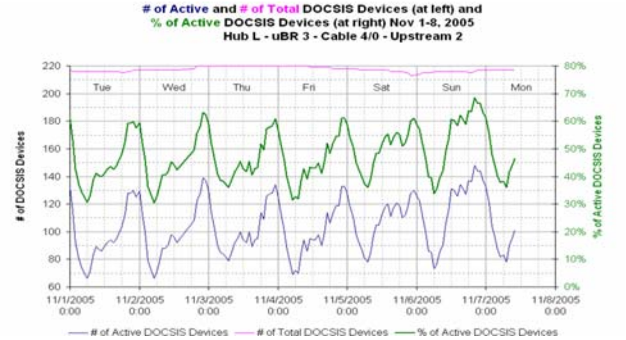


Figure 3. Active & Total Devices on a Typical Metro Upstream in 1 week

Notice in Figure 3 that the percentage of active devices never falls below 31%; put another way, about one-third of all CMs actively send/receive traffic at all times.

Zooming out to a whole year in Figure 4, one can see the percentage of active devices grew more than 10% in 12 months. The takeaway message is that over time a greater percentage of the overall CM population actively send/receive traffic (are active) around-the-clock. We discuss possible reasons for this in a later section.

Fast-forwarding almost two years to today, we see results quite similar to the trends shown in figures 3 and 4.

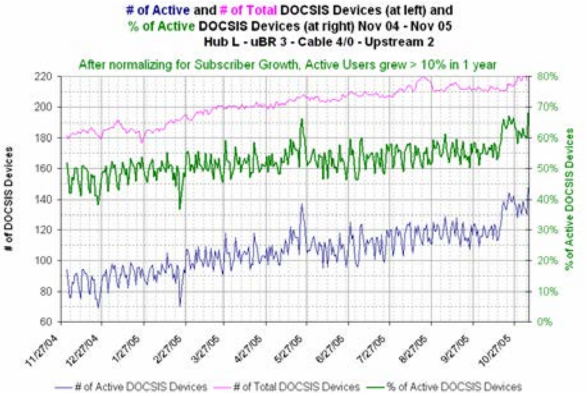


Figure 4. Active & Total Devices on a Typical Metro Upstream in 1 year

## Rise in Measured Congestion

A greater percentage of CMs send/receive more IP traffic more of the time, resulting in a rise in HFC congestion. The top half of Figure 5 shows the hour-by-hour rise and fall of congestion across all of Metro's 500,000 CMs over one week. Notice that peak congestion occurs daily at 8-10 PM. The highest congestion level was late Sunday night when about 40,000 of the 500,000 Metro CMs experienced slowdowns due to congestion.

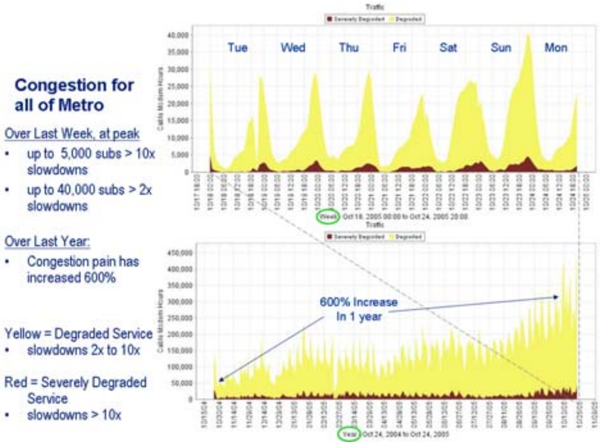


Figure 5. Weekly & Yearly effects of Congestion across all of Metro

The bottom half of Figure 5 shows the day-by-day rise and fall of congestion across all of Metro's 500,000 CMs over an entire year. The vertical axis now represents the number of modem congestion hours (the sum of the number of CMs congested per hour for all 24 hours in the day). Over the course of the year congestion levels rose nearly 600% from an average of ~50,000 modem hours per day to ~300,000 modem hours per day.

## What is the Relationship between Congestion and Bandwidth?

Even when it is assumed that the traffic characteristics do not change, an increase in the available bandwidth to the subscriber - the "information rate" - results in peak rate requirement increases. As a result, the network can serve fewer subscribers at the same level of congestion or the same subscribers with increased congestion.

In considering the recent and projected growth in demand, we consider both historical and projected information rate increases that contribute to congestion as the speed of the sources (CM's) increases. In addition, we consider the impact on traffic characteristics caused by new applications and devices that create/consume IP traffic.

## RECENT AND PROJECTED GROWTH OF THESE DEMANDS

In this section of the paper we explore a historical view of the increase in information rates and the impact of new multimedia online services such as YouTube. By exploring recent growth, we will extrapolate the magnitude of traffic that can be expected within the next three to five years for residential broadband data.

### The Digital Household

We believe there are two primary factors at work contributing to the increase in measured congestion on HFC networks.

(1) Growth in Information Rates: Using historical trends as a base-line we note continued increase in information rates that lead to an increase in measured congestion. This is fueled by the competitive nature of the broadband access market.

(2) Growth in Multimedia Applications and Connected Devices: New multimedia applications, including embedded multimedia, P2P television, and TV place-shifting, result in more bits being transported. In addition, wireless home gateways are entering the mainstream, and we see growth in the number of IP-connected, traffic-generating devices behind each CM.

We see these two factors working together to create significant additional demand for network capacity.

## Predicted Information Rates

To predict the recent and projected growth of demands on the HFC network we also considered an analysis by Bob Scott [ii] that extrapolates historical growth in data product advertised speeds since 1982. This analysis notes a close fit to Moore's Law and provides the following high speed data product information rate scenarios shown in Figure 6.

Using this model we start in 1982 with dial-up telephone Modem speeds of 300 bits per second (bps) and climb to nearly ~20 megabits per second (Mbps) today.

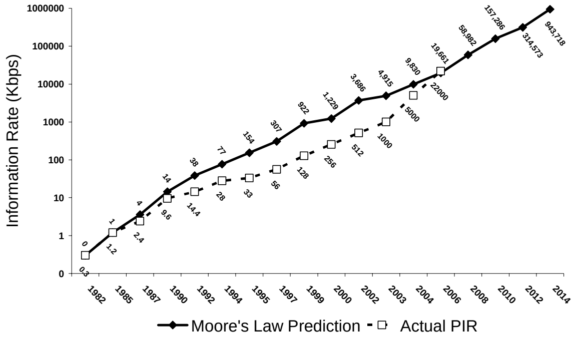


Figure 6. Moore's Law Growth in Historical and Predicted Information Rates

A sensitivity analysis from this work indicates a range of information rate or "speed" possibilities as shown in Table 2. We assume a subscriber take-rate of the different speeds in Table 2 - High Case of 5%, Probable Case of 10%, and Low Case of 85% - reflecting a mix of subscribers using different broadband product speeds to create a weighted information rate each year.

| | **2006** (Kbps) | **2008** (Kbps) | **2010** (Kbps) |
|---|---|---|---|
| High Case | 20,000 | 60,000 | 160,000 |
| Probable Case | 10,000 | 20,000 | 60,000 |
| Low Case | 7,000 | 10,000 | 20,000 |
| Weighted Average | 8,000 | 14,000 | 30,000 |

Table 2: Predicted Information Rates in Kbps

We expect this increase in CM information rates to increase congestion and required bandwidth even if the bits transported and the number of CMs stays the same between 2006 and 2010; and we expect the impact on required bandwidth and congestion to be higher in the case that the bits transported increases through new applications and additional devices per CM.

## Demand Model and Usage Profile

An analysis of traffic profiles described by using John T. Chapman's Multimedia Traffic Engineering Model published in 2002 indicates 1 to 2 traffic generating devices per CM [iii]. Recent research suggests that this has increased to over 3 devices [iv].

We assume an increase from 1.5 devices per CM in 2006 to 3 devices per CM by 2010 explained by increased market usage for new broadband connected applications.

In addition to considering historical per subscriber Kbps data we considered traffic contributions from existing applications on 1.5 connected devices per CM plus contributions from an additional 1.5 connected devices supporting several new applications including Embedded Multimedia (i.e. YouTube); P2P Television (i.e. Joost); VOIP Applications (i.e. Skype); TV

Placeshifting (i.e. Slingbox), and CE Devices (i.e. Connected Handycams).

Using this traffic engineering model in conjunction with recent new applications and devices we are able to estimate the future application traffic in average Kbps per data subscriber from approximately 80 Kbps in 2006 to 350Kbps in 2010 as shown in Figure 7.

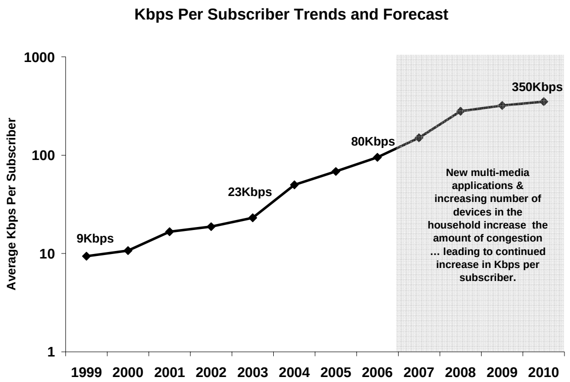


Figure 7. Kbps Per Subscriber Trends [v]

We look in more detail below at changes in the application mix over time to support this.

Changes in Application Mix over Time

We asked the question “What application changes are likely considering historical trends?” Based on information illustrated in Figure 8 we believe that P2P in its current ‘download’ form may have reached its peak and that a ‘TV centric’ and new multimedia web applications may provide the next wave of traffic.

Considering the rapid rise of P2P traffic, we note that the primary traffic generating application may not even exist at the time the network is being planned! The overall mix and contribution of various traffic-generating applications rises and falls over time as shown in Figure 8 and Table 3.

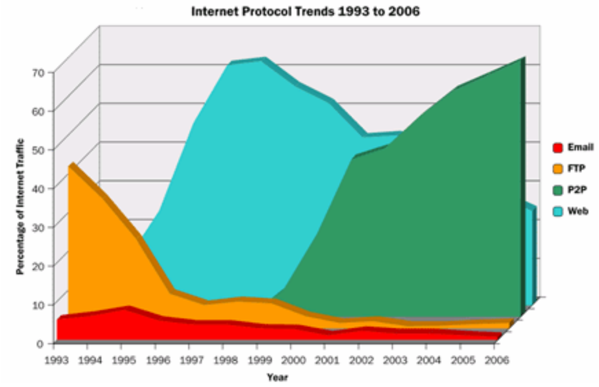


Figure 8. Internet Protocol Trends [vi]

| **1993** | **1999** | **2006** | **2010 ?** |
|---|---|---|---|
| FTP | Web | P2P | P2P-TV |
| Email | FTP | Web | Rich-Web |
| Web | Email | News [vii] | P2P-Other |
| Other | Other | Other | Other |

Table 3. Internet Protocol Trends [viii]

To understand two possible dominant applications we explore the growth of Embedded Multimedia Traffic and PC-Based Television and P2P Multimedia which in our model contribute a significant amount of the new traffic.

Embedded Multimedia Traffic:

Recent developments in web applications include the addition of video-based embedded multimedia. Examples include Brightcove, Coull, Google Video, MediaZone, MetaCafe, PermissionTV, Revver, VideoJug, YouTube, Yahoo! Video, and Ziddio. We explore YouTube in more detail for this category.

Google's YouTube founded in Feb 2005 [ix] provides flash-based video clips embedded in a web browser. Measurement of Google’s YouTube traffic in Europe indicates that YouTube represented 2.5% of all downstream traffic year end 2006 [x], and growth of YouTube traffic was 400% for the 6 months ending 2006.
Normalized for subscriber growth this

translates into 1 Kbps per subscriber today growing to over 100 Kbps per subscriber in just 12 months. Is growth as shown in Figure 9 realistic?

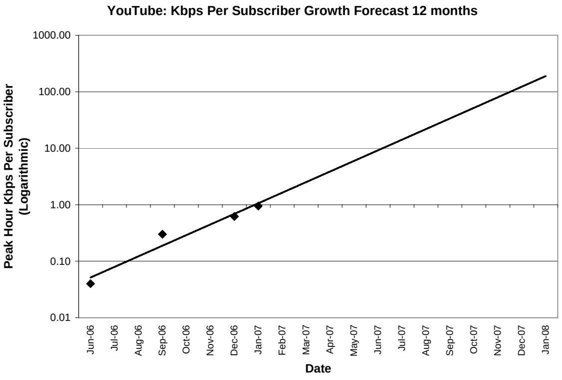


Figure 9. Recent You Tube Traffic Growth

We believe that, more than likely, embedded multimedia traffic growth per subscriber will slow in the future based on reaching maximum market usage for a given broadband subscriber population.

Using the Chapman Multimedia Traffic Engineering single user profile, we assume that by 2010 an increase in embedded multimedia web content including adoption of Google's Click-to-play[xi] advertising model, results in an embedded multimedia user base reaching market usage of 20% of broadband subscribers on average; and that the average bit rate is 250Kbps. This equates to 50Kbps of downstream traffic at maximum market usage using our assumptions.

PC-Based Television and P2P Multimedia

Another category of growth in multimedia applications is PC-based Television and P2P Multimedia. Recent developments include the introduction of several PC based multimedia online platforms including, Arvato, AOL's In2TV, Babelgum, BitTorrent, Grid Networks, Kontiki, Itiva, Joost, JumpTV, MediaZone, Network Foundation Technologies (NFT), Peer Impact, Red Swoosh, and Rawflow [xii].

We note that many P2P applications generate traffic from a series of machine queued requests. Machine-to-machine traffic, as the name suggest, indicates that the application usage continues even without human activity (i.e. mouse-click) present as shown in Figure 10. Perhaps this helps explain the rise in active modems shown in Figures 3 & 4.

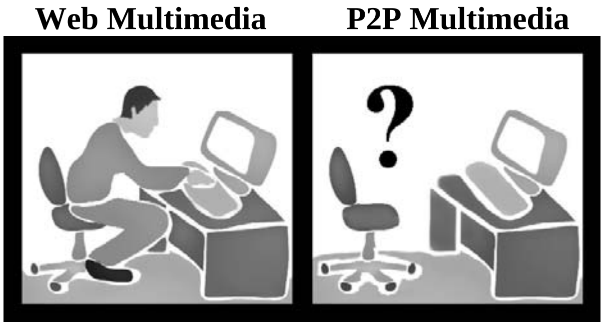


Figure 10. Web/P2P Multimedia Usage

Using the Chapman Multimedia Traffic Engineering single user profile we considered that users watching television have a peak time (on-peak) market usage of 60% and an off-peak value of 33% of broadband subscribers by 2010 based on the Nielsen "Households Using Television Number" [xiii]. We assume the average for P2P video is 40% and a blended average bit rate of ~550Kbps in the downstream and ~125 Kbps in the upstream. This equates to 220Kbps per subscriber in the downstream and 50Kbps in the upstream.

Today P2P video applications are PC-based. However, in the future such an application may find a low cost hardware "host" vehicle such as a set-top, thus significantly increasing market usage levels.

We believe a future statistical projection of Joost[xiv] traffic would assist in identifying growth patterns and a likely market usage peak.

Comparing Each Approach

We considered both Moore's Law historical information rates, keeping congestion constant, and Chapman's Multi-Media Engineering model to create a single CM profile.

| | Historical Information Rates and Moore's Law | Single CM Profile using Chapman's Model |
|---|---|---|
| % Increase in Bit-rate 2006 to 2010 | 300%<br>(320 Kbps) | 341%<br>(353 Kbps) |

Table 4: Comparing Each Approach

For the purpose of determining capacity requirements we noted that Chapman's model generated a greater increase in the per user bit-rate and was therefore used as the basis for our 2010 assumptions.

Predicted Capacity Requirements

As a result of the application analysis above we are able to estimate the amount of additional future downstream and upstream capacity that will be required.

Assuming the information rates offered by the Moore's Law projection and the Kbps per subscriber generated by new applications using the Multimedia Engineering model, this suggests the following over-booking factors as shown in Table 5.

| | **2006** | **2010** |
|---|---|---|
| Avg. Speed | 8,000 | 30,000 |
| Kbps | 80 | 350 |
| Over-booking | 100 | 85 |
| Subs/8MHz | 517 (~500) | 116 (~100) |

Table 5: Speed and Kbps per Subscriber

We assume that 20% of the 50 Mbps (8 MHz) downstream port capacity in 2006 and 2010 is reserved for: a) other services (i.e. eMTA based VoIP and STB return path), and b) additional growth. This results in 40Mbps available for use per 8MHz channel. We also assume statistical multiplexing gains of 1.1 from bonding channels.

In 2006 we support 500 subscribers per 8MHz downstream at 80Kbps per sub compared to 2010 where we are able to support 100 subscribers per 8MHz downstream due to the additional of new applications and new devices. To support the equivalent density per service group we require 32MHz of downstream capacity and 25.6MHz of upstream capacity as shown in Table 6.

| | **2006** | **2010** |
|---|---|---|
| Number of DS | 1 x 8MHz | 4 x 8MHz |
| Number of US | 2 x 3.2MHz | 4 x 6.4MHz |

Table 6. Additional Capacity Required

That's 4 channels x 50 Mbps/ channel = 200 Mbps or 32MHz of total capacity in the downstream as outlined in Figure 11 below.

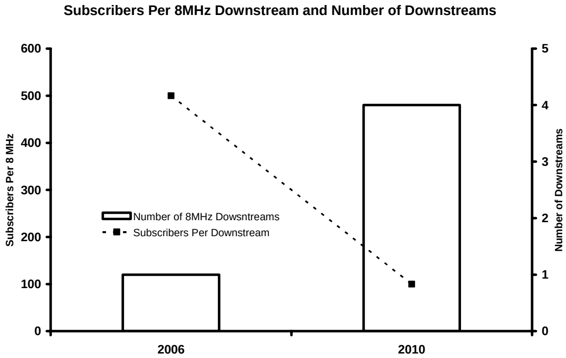


Figure 11. Additional Downstream Capacity

Additionally we would require 60 Mbps or 25MHz of upstream capacity as outlined in Figure 14 below.

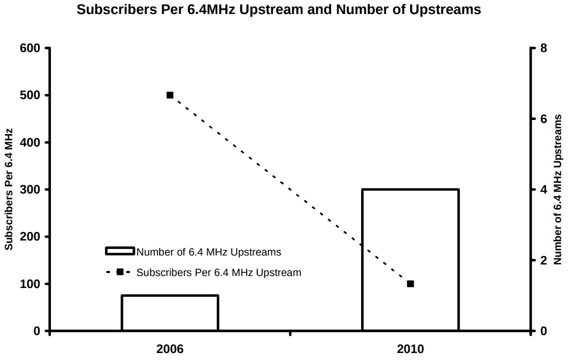


Figure 12. Additional Upstream Capacity

In summary, the growth in usage per subscriber and changes in applications together reduces the net number of subscribers per 8MHz downstream from 500 to something more like 100 as shown in Figure 12.

## MANUFACTURED BANDWIDTH - EXPAND CAPACITY TO MEET DEMAND

Given the usage expected today and within the next few years, how can operators prepare from an HFC perspective? This section outlines the cost of various Capacity Management tactics including:

(a) Reducing Serving Groups

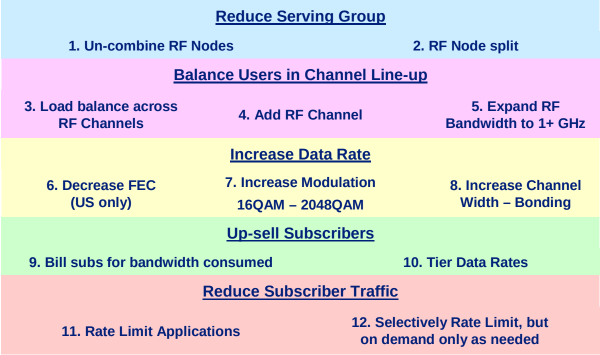
(b) Balancing Users across channel line-up

(c) Increasing Data rates

(d) Up-Selling Subscribers

(e) Managing Subscriber Traffic

Each of the above tactics may have one or more options as shown in Figure 13.



Reduce Serving Group

1. Un-combine RF Nodes
2. RF Node split

Balance Users in Channel Line-up

3. Load balance across RF Channels
4. Add RF Channel
5. Expand RF Bandwidth to 1+ GHz

Increase Data Rate

6. Decrease FEC (US only)
7. Increase Modulation 16QAM – 2048QAM
8. Increase Channel Width – Bonding

Up-sell Subscribers

9. Bill subs for bandwidth consumed
10. Tier Data Rates

Reduce Subscriber Traffic

11. Rate Limit Applications
12. Selectively Rate Limit, but on demand only as needed

Figure 13. Capacity Management Tactics & Options

### Reduce Serving Group

We observe there are two ways to reduce the size of a serving group: Option 1 is to Un-combine RF nodes so that fewer nodes (and subscribers) are connected to an upstream or downstream port. Option 2 is to split a RF node into two or more serving groups with fewer subscribers. In the limit, Option 2 would be considered a “Fiber Deep” approach to reducing serving group size.

### Balance Users across Channel Line-up

As penetration grows it is safe to assume the existence of multiple channels. Option 3 involves spreading users across multiple existing channels to evenly distribute traffic loads. Option 4 involves adding additional channels (assuming spectrum is available), and Option 5 requires expanding the HFC delivery infrastructure to accommodate the transmission of higher frequencies (e.g., 1GHz).

### Increase Data Rates

In DOCSIS systems, upstream traffic ‘packets’ may be ‘protected’ from packet errors by different levels of Forward Error Correction (FEC). In fact, it is possible to not use any FEC at all, though field experience dictates that some level of FEC should always be used. To successfully transport un-errored packets, clean portions of the HFC require less FEC and noisy portions require more FEC. While ensuring critical packets (such as voice) are delivered unerrored, FEC comes at a price; greater FEC (overhead) results in less User Data (payload).

Clean HFC plants can also support higher order modulation which means more bits (a larger alphabet) can be sent at a given symbol rate. The take-home message in Options 6 & 7 is that cleaner HFC carries more User Data than noisy HFC.

Option 8 is the much talked about solution known as channel bonding wherein multiple channels are 'bonded' into a wider (super) channel with higher data rates and higher statistical multiplexing efficiencies

Up-Sell Subscribers

Options 9 & 10 imply that 'heavy' users who send/receive large amounts of traffic could/should pay more for service than 'light' users who don't send/receive much traffic. The thought is that pricing can/will discourage heavy use. Option 9 could be a unit measure that can be thought of as 'Bill by the Byte Transferred' which is akin to how many utilities/consumables are billed today (e.g., milk, water, gasoline, etc., each cost so much per litre.) Option 10 is similar to Option 9, but instead of billing per byte, bills on chunks of bytes (e.g., 0 ≤ 10 GB per month is $40, 10 ≤ 20 GB per month is $50, etc.).

Reduce Subscriber Traffic

Options 11 & 12 refer to methods used to limit certain or all users' traffic and/or traffic types. Before getting alarmed at either Option, take a moment to consider a few practical realities. All highways slow down when they get full; Internet pipes slow down when they get full, too. Option 11 involves rate limiting specific applications (such as P2P) around-the-clock, whereas an example of Option 11 is an on-demand as-needed rate limiting algorithm that enforces application or individual subscriber rate limits ONLY as needed (i.e., at the onset of congestion or when congestion persists).

The Cost of Capacity Expansion

We list the relative capital cost of some of the different capacity expansion tactics & options in Table 7. The important factor is that costs can vary dramatically between expansion options, and that a combination of different options can be considered (i.e. Switched Digital Video & Bonded Channels).

| | $ Per Node Per MHz of Capacity | $ Per HP at 2,000 HP Per Node | $ Per Data Sub (25% pen.) |
|---|---|---|---|
| Upgrading to 1GHz [xv] | $714 | $50 | $200 |
| Fiber Deep [xvi] | $107 | $46 | $184 |
| Node Split [xvii] | $30 | $13 | $52 |
| Switched Digital [xviii] | $312 | $5 | $20 |
| Bond Channels [xix] | $632 | $18 | $70 |

**Table 7. The Cost of Manufactured Bandwidth**

## TECHNOLOGY BANDWIDTH - MANAGE SCARCE RESOURCES WITH NEW TECHNOLOGY

Given the projections, expansion approaches in isolation will not be enough to maintain near-term demand. In addition, a software management approach to directing and managing traffic on the network will stretch the increases gained with HFC equipment upgrades, node splits, etc. This section outlines sample traffic management applications that ensure bandwidth is fairly provisioned so that high priority traffic such as VoIP calls are given top priority, while minimizing the impact of P2P and over-the-top applications.

We see four possible instruments to allocate scarce bandwidth:

(1) Rationing

(2) Delay

(3) Pricing

(4) Access

(5) Supply

Rationing

Operators are able to limit volume and access to certain applications. This can be hard rationing or soft (e.g. first warn heavy users and if behavior doesn't change, then deliver a penalty).

For example, dynamically re-provisioning a subscriber at a lower speed when consumption reaches a pre-set limit is a possible way of making additional bandwidth available to remaining subscribers.

Rationing has the disadvantage of penalizing all applications even those provided by the operators such as email, and also penalizes the subscriber for heavy off-peak consumption that may not impact operator capacity at peak time.

Delay

It is possible to delay the delivery of non-revenue generating consumption though the use of traffic shaping which may be time-based, user-based, protocol or application-based—and therefore 'fair' in the sense that it addresses specific areas of bandwidth consumption.

It is critical that the delayed application can be accurately identified and that a mechanism exists to detect a move by users to alternative applications for the same use. (i.e. a move from a P2P protocol to Network News (NNTP) protocol when wanting to download multimedia content).

Pricing

Pricing may be used to impact consumption and may be based on volume, nature of the service and across the day (as with the electricity industry).

Application specific charges can also impact traffic consumption. For example, it may be possible to offer certain P2P services within the monthly broadband service fee and to charge for access to others.

Alternatively it may be possible to charge a carriage fee for premium P2P content providers in return for a fair allocation of bandwidth.

We believe that pricing is one of the most effective mechanisms for managing heavy bandwidth consumption.

Access

With the growth of wireless home access points, theft-of-service can lead to increased traffic levels associated with a subscriber, even without the knowledge of the subscriber. Offering a managed home network enables the operator to reduce the risk of theft and improves the security offered to the end user.

Supply

Packet Cable Multimedia (PCMM) enables bandwidth supply control at the CMTS when applied to specific applications or service tiers. This enables the operators to monetize Quality of Service. In fact, it would be possible to package Quality of Service as a product.

Additionally as noted in Figure 14 it is possible to manage the supply of bandwidth through the use of a virtual or real two-tier backbone. Using this structure content providers and subscribers that pay for Quality of Service are able to access additional bandwidth.

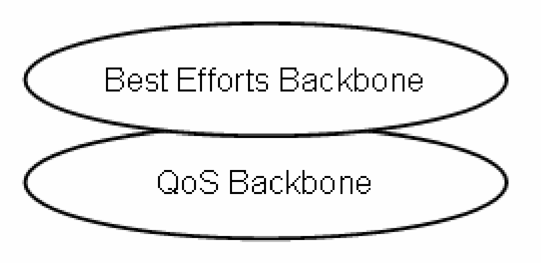


Figure 14. Two Tier Backbone Helps to Manage Supply

## The Cost of Capacity Management

Reviewing the capital cost of capacity management in more detail we see that the cost of effectively managing capacity is far lower than capacity expansion. It follows logically then that options for managing this scare resource should be explored before/as expansion is considered. Table 8 provides examples for consideration.

| | $ Per Node Per MHz of Capacity | $ Per HP at 2,000 HP Per Node | $ Per Data Sub (25% pen.) |
|---|---|---|---|
| PCMM [xx] | $166 | $1.25 | $5 |
| Traffic Shaping [xxi] | $20 | $0.15 | $0.60 |

Table 8. The Cost of Technology Bandwidth – Managing a Scarce Resource.

“You wouldn’t believe how many people try to drive away without paying.”

Figure 15. Bandwidth Expansion and Management is Not Cost-Free to the Cable Operator.

## CONCLUSION

We outlined a trend in the rise of measured congestion and provided an overview of the historical and future application demands causing this congestion, and offer suggestions and relationships (a toolkit) that the cable operator may want to consider. We note that while HFC networks offer great flexibility in expanding bandwidth, not all alternatives are created equal; some are easier and more cost effective than others. A systematic and thoughtful approach to addressing congestion should include evaluating all possible options—considering both the subscriber and the content provider in view of both manufactured bandwidth and technology bandwidth alternatives.

REFERENCES

---

[i] UPC Broadband & C-COR internal reports based on CableEdge® measurements.

[ii] Bob McIntyre, "The QAM Before The Storm", SCTE Emerging Technologies Conference, January 2007

[iii] John T. Chapman, "Multimedia Traffic Engineering: The Bursty Data Model" , Cisco, January 8, 2002

[iv] Ian Fogg, "Home Networks in Europe: Leverage Strong Growth Potential in Congested Digital Device Landscape", Jupiter Research, Jan 2007

[v] Based on historical Kbps per subscriber growth in Europe. Forecast Kbps numbers based on High Speed data applications only and ignores VOIP and Digital Set-top return path traffic which are assumed to be included in the remaining 20% port capacity.

[vi] Andrew Parker, "The True Picture of Peer to Peer File Sharing", CacheLogic, 2004. http://www.cachelogic.com/home/pages/studies/2004_01.php

[vii] A revival in NNTP 'news' traffic has been observed recently in Europe primarily as a mechanism to share multi-part binaries.

[viii] CacheLogic study: The True Picture of Peer to Peer File Sharing. Andrew Parker, 2004. http://www.cachelogic.com/home/pages/studies/2004_01.php

[ix] YouTube website http://www.youtube.com/t/about

[x] Based on UPC Broadband Netflow measurements in Europe using YouTube's AS number.

[xi] Google press release: "Lights, Camera, Action!! - Google Introduces Click-to-Play Video Ads on the Google Content Network" May 23, 2006. http://www.google.com/intl/en/press/annc/clicktoplay.html

[xii] Wired Magazine Feb 2007 and industry sources.

[xiii] James Hibberd, "Cable Late-Night Growth Outpacing Prime - Competition From Other Platforms Creates Opportunity for Daypart", TelevisionWeek, August 7, 2006. Chart notes Nielsen Media Research 2005 Prime-time households using television rates at 60.2% in 2005 and Late-night households using television levels at 33.9%. http://www.tvweek.com/article.cms?articleId=30364

[xiv] Joost is a newly announced television sharing application that utilizes P2P communication to deliver high-quality television across the broadband network.

[xv] We assume the average cost to upgrade to 1GHz is around $45-$55 per HP. We assume the 1 GHz amplifier gain is higher than those at 870 MHz and permit a direct drop-in at existing amp locations.

[xvi] Fiber Deep calculation based on an assumed a Labour of 41,000 and materials of 50,000 per 2,000 Home Passed or a total of 91,000 per 2,000 Home Passed or $46 per Home Passed. Assumes Greenfield … cost would be lower if fiber is already in place.

[xvii] Assumes the worst case costs for a new node, including HE electronics, and new fiber if needed, would be $30,000. Numbers may vary depending on specific labor costs, electronics, and market.

---

[xviii] Switched digital assumes 80% digital penetration, 2,000 homes passed per service group, 4 QAM channels per service group, 100 switched channels, 10 channels re-encoded. Components include QAM modulators, SDV servers, Re-encoding, Multiplexing, Routing and transport. Assumes 4 x 8MHz of capacity can be released for other services.

[xix] Bonding Channel to Increase Bandwidth based on an assumed $10,000 per 8MHz DOCSIS channel today and assuming $10,000 per 4 x 8MHz DOCSIS channels in the future. We assumed 500 subscribers per downstream or $20 per subscriber. We also assumed $50 per CPE device.

[xx] PacketCable Multimedia: Assuming $5 per subscriber for Turbo Button and Gaming Application. Includes cost of Record Keeping Server.

[xxi] Adding Per-Protocol Traffic Shaping: Assuming 25 CMTS port per shaping device based on 900 Mbps per device and 50Mbps per downstream at 70% utilization or 35Mbps. Each shaping device costs $30,000 / 25 ports = $1,200 per CMTS / 4 nodes per downstream serving segment = $300 per node / 2,000 Homes Passed per node = $0.15 per home passed. We assumed that 25% of the total capacity could be reclaimed though shaping.

# THE FUTURE OF TRANSCODING – THE NEED FOR MPEG-2 AND MPEG-4 TO COEXIST

John Hartung, Santhana Krishnamachari
EGT

## *Abstract*

*Virtually all commercial digital video content is stored and distributed using the MPEG-2 encoding standard introduced about 12 years ago. Although this standard enabled a large increase in the number of programs that could be carried in access networks, the capacity of those networks has not kept up with the explosion of content and the increased bandwidth requirement for high definition. The introduction of switched access networks, both HFC and IP, will help alleviate that bottleneck and allow the introduction of new clients that make use of the more efficient MPEG-4 part 10 (a.k.a. H.264) encoding standard. The challenge in realizing these gains, however, will depend on transcoding technology that is both cost effective and maintains the quality of the originally distributed program.*

## THE LIMITS OF MPEG-2 ENCODING

The MPEG-2 standard, like all media encoding standards, is defined by the encoded bit stream syntax, and the semantics, or operations, signified at the decoder by these bit stream elements. The main semantic elements include block based motion compensated prediction, quantized transform based encoding of prediction residuals and reference blocks, and entropy coding of encoding parameters. MPEG-2 encoder operation is not specified by the standard and encoders are not required to make use of all the operations available at the decoder. The standard was designed with complexity in mind so that real time encoders could be economically deployed. Early MPEG-2 encoders were limited in performance due to the required computational complexity needed to generate optimal bit streams. In general, this optimality requires a global search over a large number of encoding parameters, and processors capable of this computation would have made their cost prohibitive. Since the introduction of this standard the performance of integrated circuits has dramatically increased, and algorithms have been developed that achieve near optimal performance at greatly reduced complexity.

Today’s MPEG-2 encoders achieve near optimal performance in terms of objective performance measures such as the peak signal to noise ratio (PSNR) as illustrated in Figure 1. This plot shows the distortion, relative to the original content, averaged over a representative set of 18 video sequences including 24 fps film content, 30 fps interlaced content, and clean and noisy sequences. PSNR of about 34 dB results in high quality encoding that is nearly indistinguishable from the original content. As seen in the plot, the MPEG-2 encoding algorithm can achieve this result, over a broad range of sequences, at about 3.5 Mbps for standard definition video.

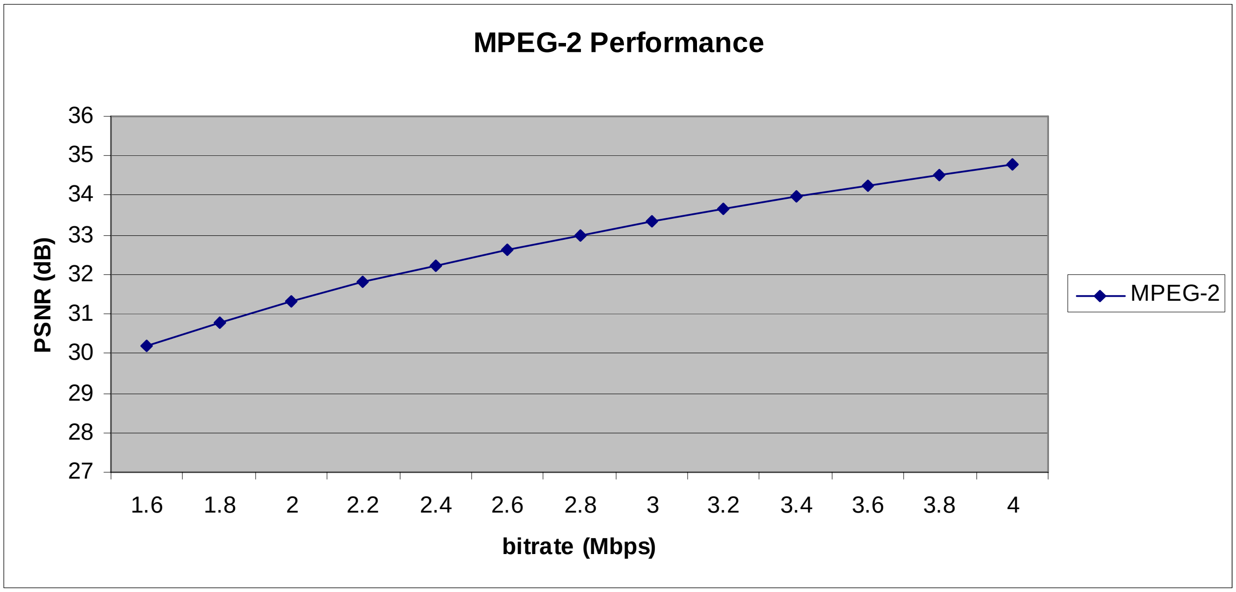


**Figure 1**

Objective measures, however, do not tell the entire story of encoder performance. In addition to the purely objective distortion measures, today's encoders take into account subjective evaluation through the incorporation of human visual system (HVS) models. These models estimate the masking of distortion by content so that encoding bits are preferentially spent to reduce the distortion that is most visible. Encoders also make use of preprocessing to improve their performance on noisy sequences. These adaptive filters remove noise from the original sequence before encoding, both to restore the original content and to avoid allocating excessive bits to encode the typically high frequency noise. Although the use of HVS models and preprocessing are not covered in the encoding standards, they are important to the practical deployment of video encoding.

## MPEG-4 PERFORMANCE

The MPEG-4 and MPEG-2 encoding algorithms are similar in that they both use block based motion compensated prediction, quantized transform coding of residuals, and entropy coding. However, MPEG-4 introduces basic differences in these tools along with additional modes of operation.

The MPEG-4 motion estimation tools have been expanded to include additional block shapes and sizes, and multiple reference frames can be used to predict a macroblock. Field and frame prediction can also be varied on a macroblock basis. A new intraframe spatial prediction mode has also been introduced that has no correspondence to the MPEG-2 coding modes. The DCT transform used in MPEG-2 has been replaced by a smaller integer transform, and VLC entropy coding has been augmented with an optional adaptive binary arithmetic entropy coder (CABAC). The MPEG-4 standard also allows a filter in the encoding loop that helps mitigate encoding artifacts, such as blocking at low encoding rates. Although this does not improve the PSNR performance, it results in more acceptable subjective artifacts. Making full use of these new features enables a 30% to 50% reduction of coding rate for equivalent video quality at the expense of a

5-6 times increase in complexity. Figure 2 shows the performance of MPEG-4 averaged over a broad range of sequences and encoding rates. The input sequences are full D1 resolution and the encoder is operating at main profile@level 3.

It can be seen from the plot that a PSNR of about 34 dB is achieved at an encoding rate of about 2.1 Mbps as compared to 3.5 Mbps for MPEG-2, or a rate reduction of about 40%.

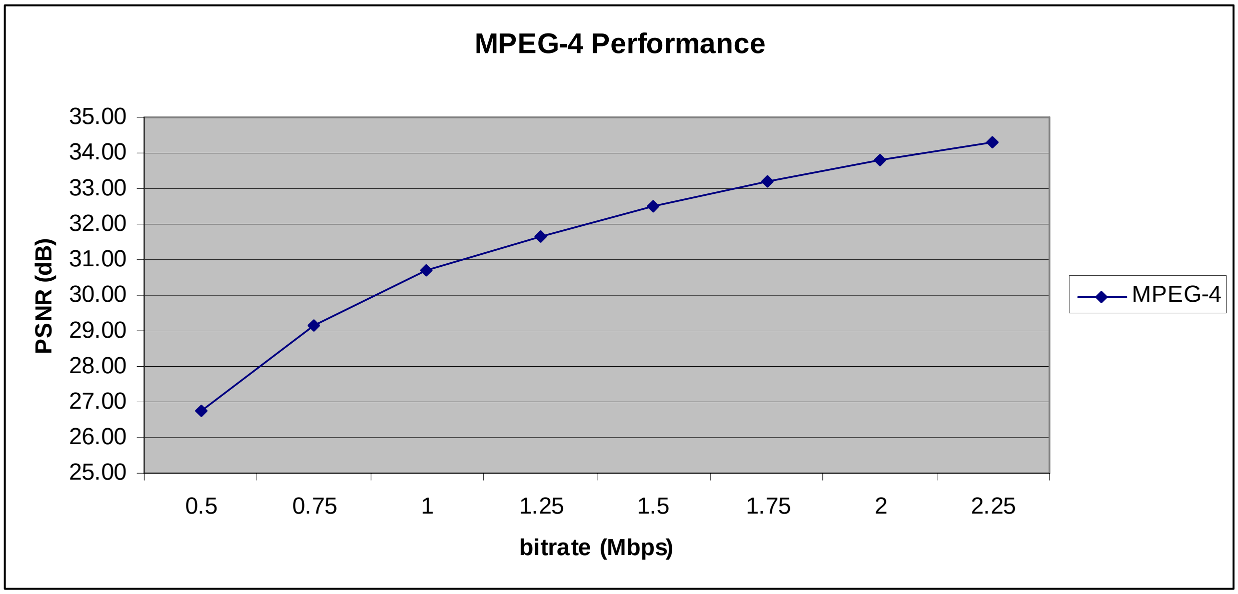


**Figure 2**

As with MPEG-2, MPEG-4 encoding benefits from preprocessing and HVS modeling and identical techniques can be applied to improve the subjective quality.

## SWITCHED NETWORKS AND MULTIPLE CODECS

Switched networks have been deployed in cable networks for video on demand (VOD) applications, and are being deployed to increase the effective capacity for broadcast applications. Switching for VOD enables content to be switched to an individual set top box based on a subscriber request. Switched broadcast enables broadcast content to occupy a portion of the access network only when it is requested by one or more set top box clients. In this case, capacity is increased beyond broadcasting all channels because only a small number of the programs offered are actually requested concurrently.

Switching both for VOD and broadcast can also be used to enable further bandwidth savings by tailoring the requested content based on the capabilities of the requesting set top box. Because of the large number of deployed MPEG-2 set top boxes, both MPEG-2 and MPEG-4 set top boxes will co-exist in cable networks for some time. In order to take advantage of the additional bandwidth savings of MPEG-4 set top boxes, content needs to be available for either on the network when requested. Since switched applications are aware of the requesting client, the delivery can be tailored to the capability of that client, e.g. for VOD only the single requestor need be considered. In the case of switched broadcast all requestors must be considered in order to avoid transmitting the same broadcast content in multiple formats. This can be avoided by transmitting the content in MPEG-2 format only since most MPEG-4 set top boxes can also decode MPEG-2 content.

In the case of VOD, content can be stored in multiple formats on the server, however, broadcast applications require transcoding from the predominant MPEG-2 format to MPEG-4 in real time.

## APPROACHES TO TRANSCODING

Several approaches can be taken to transcode from MPEG-2 to MPEG-4. The lowest complexity approach involves mapping the MPEG-2 encoding parameters into MPEG-4 equivalent representations. This is similar to the encoded domain rate shaping that is used for statistically multiplexing MPEG-2 streams. Encoding gains can be obtained through the intraframe prediction modes and improved entropy encoding available in MPEG-4, however, a large set of tools would be restricted from use limiting the ultimate performance.

An alternative approach is to decode the MPEG-2 content and apply the decoded baseband video directly to an MPEG-4 encoder. This approach does not produce high quality results, as illustrated in Figure 3. This plot compares the average PSNR of original sources that have been encoded with an MPEG-4 encoder, vs. the PSNR of decoded MPEG-2 sequences that have been encoded with the same MPEG-4 parameters. In this case the MPEG-2 sequences were coded at 4 Mbps and 3 Mbps, and their PSNRs were about equal, or greater, than that of the MPEG-4 encoding. This result uses the same set of sequences used in Figures 1 and 2. As the plot shows, the PSNR degrades up to 1 dB in the decode/encode case when the MPEG-2 was coded at 4 Mbps. One of the main reasons for the degradation is due to the fact that the frame type is not maintained between the MPEG-2 and MPEG-4 encodings. The B frames are typically encoded at lower rate,

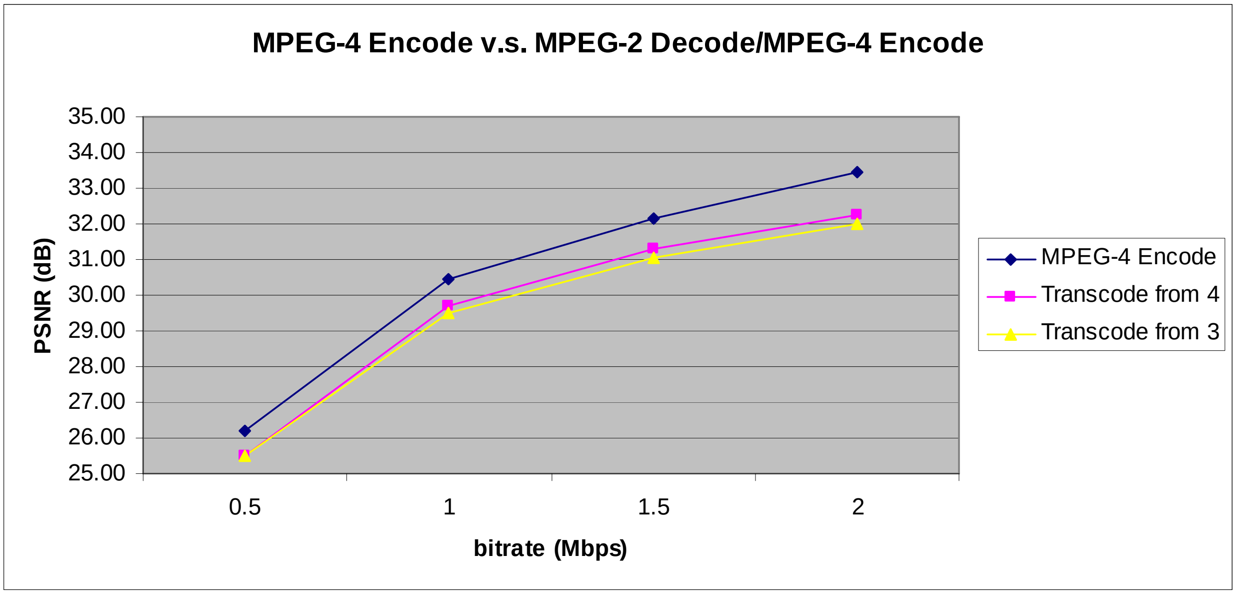


**Figure 3**

and PSNR, than I and P frames in the original MPEG-2 encoding. Without knowledge of the frame type, the MPEG-4 encoder can re-encode the B frames as I or P frames and subsequently use these as reference frames for prediction. This results in lower PSNR in the predicted frames and propagation of this distortion. The results shown at 4 Mbps are for high quality MPEG-2 encoding that maintains good quality for I, B, and P frames. Lower MPEG2 rates, and/or lower quality encoders produce a larger variation in the quality of the different frame types resulting in poorer transcoding results as shown in the 3 Mbps result in Figure 3. As illustrated in the plot, the resulting PSNR of the transcoded sequence degrades further at the lower MPEG-2 rates.

A final method for transcoding also performs decoding to baseband video and re-encoding using an MPEG-4 encoder, however, the MPEG-4 encoder makes use of the MPEG-2 encoding parameters. One example of this is to maintain the frame coding type to avoid the degradation described in the previous method. Referring to the results in Figure 3 again, approximately .2 dB improvement can be gained when transcoding from 4 Mbps MPEG-2 to 2 Mbps MPEG-4. For lower rate and/or lower quality MPEG-2 encoding even larger gains are possible. Passing additional parameters allows for further improvements in transcoding performance, and a reduction in complexity. An example of this is the use of bit allocation in the MPEG-2 encoding as a complexity estimator that can be used for rate allocation in the transcoded sequence. This is similar to two-pass encoding algorithms, however, the necessary information already exists in the MPEG-2 bitstream. This technique improves the transcoding quality without the complexity of a two-pass MPEG-4 implementation.

**Figure 4**

Overall, high quality transcoding and reduced complexity is achieved through full decode/encode with reuse of encoding parameters. This argues for a tightly coupled system that receives MPEG-2 programs, either in SPTS or MPTS, and converts directly into MPEG-4 transport streams.

A final consideration in transcoding is the mitigation of source noise and coding artifacts in the original MPEG-2 encoding. This can be accomplished by filtering the decoded MPEG-2 sequence, either in the transform domain, or in the reconstructed baseband for encode/decode transcoders. High quality MPEG-2 encoders typically apply sophisticated prefilters that remove noise, however, noise filtering improves

**Figure 5**

transcoding when this is not the case. A second source of noise can be introduced by the MPEG-2 encoder. This structured noise can be effectively estimated and adaptively removed to improve the subjective quality of the transcoded sequence. Figures 4 and 5 show the corresponding frame in a transcoded MPEG-2 sequence. Both frames use the same original and transcoded parameters, however, Figure 5 demonstrates the improvements gained through post-filtering of the MPEG-2 sequence.

## CONCLUSION

The need for bandwidth efficiency in cable plants continues to be driven by the increase in service and content offerings. Increasing amounts of high definition content further add to the need for improved efficiency. Switched services will help provide this additional bandwidth and enable a transition to the more efficient MPEG-4 encoding standard, however, this transition will require the coexistence of both MPEG-2 and MPEG-4 services. This coexistence will be enabled by integrated transcoders that provide a high quality, low complexity solution.

## CONTACTS

John Hartung
john@egtinc.com
Santhana Krishnamachari
santhana@egtinc.com

# THE IMPACT OF GOING FIBER SPEEDS OVER COAX FEEDS

Alon Bernstein, Cisco Systems, Inc.
Joe Godas, Cablevision

## *Abstract*

*Data rates over cable networks are constantly increasing due to higher bandwidth demands from consumers as well as competitive pressure. With DOCSIS 3.0 the cable plant can offer speeds that rival those offered by PON (Passive Optical Networks) technologies and meet consumer demands.*

*Building an end-to-end solution that will support these data rates is not only about adding faster network components and updating the user service agreements. This paper will explore the specific impact of supporting data rates in the order of 100 Mbps and above over the cable network.*

## OVERVIEW

With DOCSIS 3.0, the typical traffic profile for a cable subscriber is going to increase from the 10mbps-30mbps range to the 50mbps-100mbps range. These traffic rates compete well against end-to-end optical technologies.

In theory it takes a DOCSIS 3.0 system and a change in the modem configuration file to get to the above rates. In practice, the added throughput has implications on the upper layer application behavior, network design and equipment design. This paper will list what these implications are and how they can be addressed in order to unleash the full potential of DOCSIS 3.0.

Most of the issues described here are not unique to cable. Any service provider network that will serve 100mbps+ speeds will have to deal with them. However, DOCSIS 3.0 is the technology that would allow, for the first time, residential costumers to have access at these rates and so it will be the first time cable service providers manage such a high speed network.

## IMPACT OF HIGHER SERVICE CONTRACTS ON AGGREGATE TRAFFIC

Even before DOCSIS 3.0, in order to compete with telcos, cable Multi-Service Operators (MSOs) have been steadily increasing their traffic contracts from 2.5mbps down/500k up to 30mbps down/5 mbps up. Looking at the lessons learned from past increases will help to infer what will happen with future increases.

While one might expect that the aggregate trend would linearly track the service contract increase, or in other words, that a doubling of a service contract will result in a doubling of the aggregate traffic, it is evident this is not the case.

The following graphs represent aggregate traffic as measured at the internet border in the Cablevision network at different periods over the years.

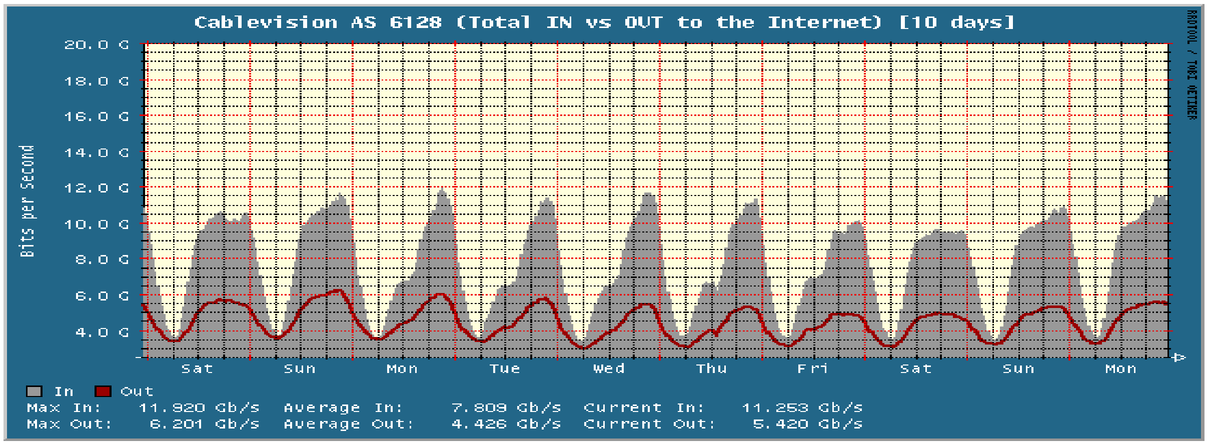


**Figure 1: 1Q2004**

Figure 1 depicts the aggregate traffic during 2004Q1 when an individual contract was 10mbps down, 1mbps up. As can be seen, the traffic peaks are around 12Gbps and the lows are in the 4Gbps range. The gap between the highs and lows is roughly 8Gbps.

18 months later we see traffic demands ramp up quickly as subscribers are added to the system. The modem contracts haven't changed but downstream consumption has almost doubled due to a 45% increase in subscriber acquisition (see Figure 5). Now the peaks are in the 22Gbps range, and the lows at 7Gbps. The spread has increased to 15Gbps, much more then the linear increase in subscriber bandwidth.

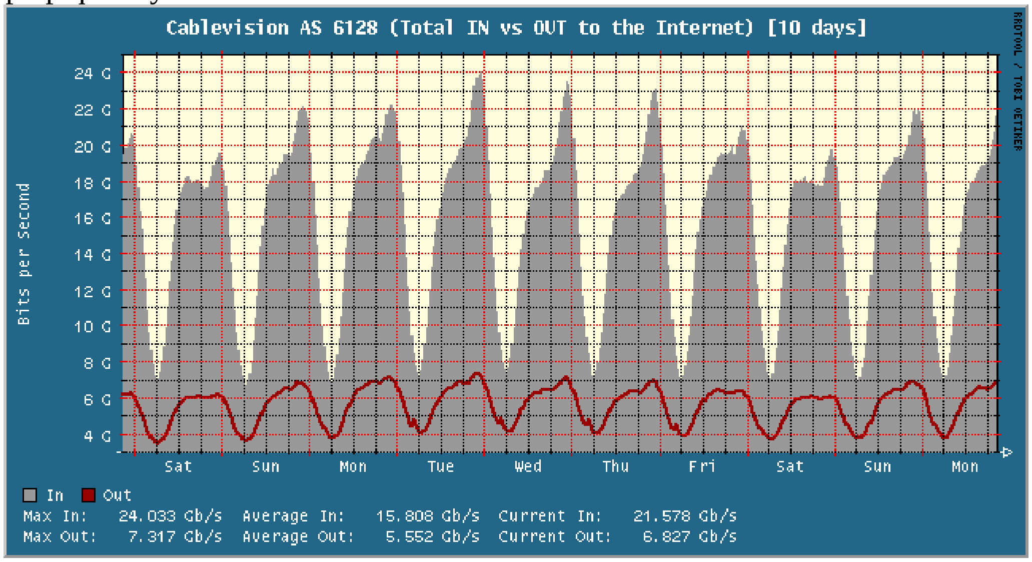


**Figure 2: 3Q2005**

In 2Q2006 (Figure 3) -The traffic contracts for the entire subscriber base were increased to 15mbps down and 2mbps up. The peaks rose to about 40Gbps, and the dips to 12Gbps. The spread is 38Gbps, and keeps growing much faster then expected. In this time period, a 19% increase in subscriber acquisition has produced a whopping 100% increase in traffic consumption. This surge in demand is attributed to a

50% increase in the modem contract burst capability (going from a 10Mbps to 15Mbps downstream contract).

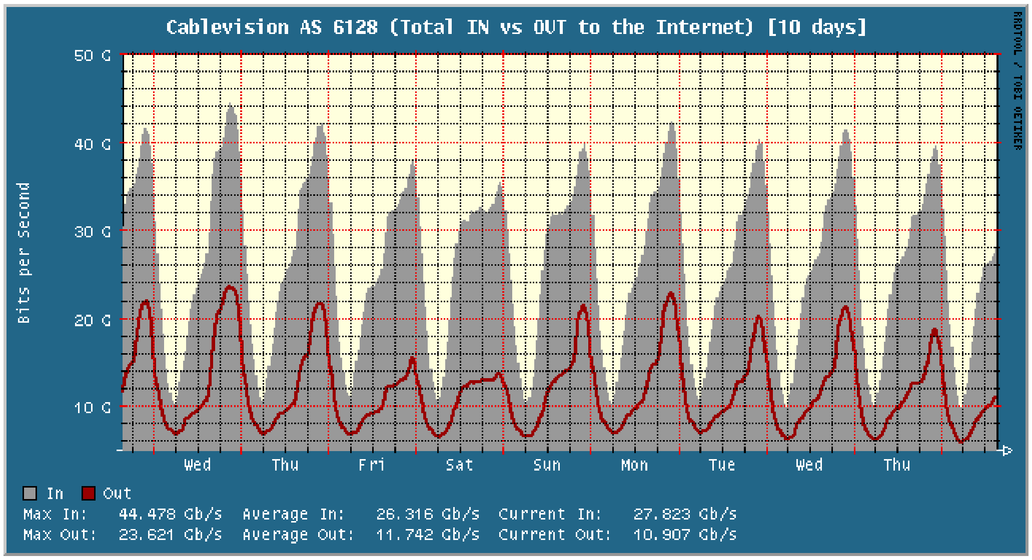


**Figure 3: 2Q2006**

1Q2007 is especially interesting. The first thing to note is that it is barely 6 months since 2Q2006 and corresponds to a subscriber increase of roughly 9%. Curiously enough, while the peaks have increased by 10Gbps, the lows have hardly moved from their 2Q2006 range (about 12Gbps).

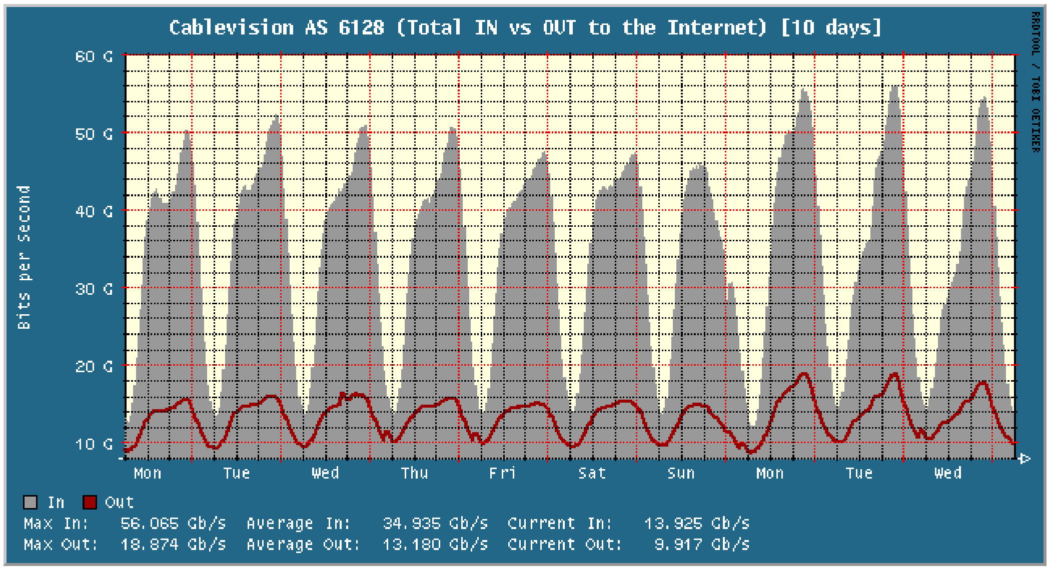


**Figure 4: 1Q2007**

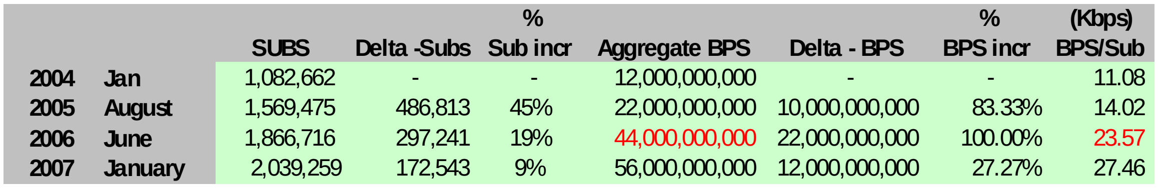
Several observations can be drawn from these set of traffic trends:

1. Traffic rates are increasing rapidly, and non-linearly;

2. Across all graphs, the peaks and the lows occur at about the same time, so they most likely correspond to interactive use of the internet around evening time;

3. If the majority of traffic was peer-to-peer (P2P) then the lows would have been much more pronounced as people tend to leave their P2P clients on through the night. Either people shut down their P2P when not active, or some highly interactive applications (file downloading or real-time streaming) are the dominant behavior during peak time. Higher burstable contracts appear to have the effect of creating higher peak bursts, but not higher sustained use (as noted by stagnating dwell/gap growth).

Another important take away from this graph is that network capacity planning techniques will need to change. The rules previously used for lower bandwidth contracts cannot be extended to higher modem contract rates because of the increased burst capability (thus higher aggregate peak rate) of the traffic flows.



| | | | | % | | | % | (Kbps) |
|---|---|---|---|---|---|---|---|---|
| | | SUBS | Delta -Subs | Sub incr | Aggregate BPS | Delta - BPS | BPS incr | BPS/Sub |
| **2004** | **Jan** | 1,082,662 | - | - | 12,000,000,000 | - | - | 11.08 |
| **2005** | **August** | 1,569,475 | 486,813 | 45% | 22,000,000,000 | 10,000,000,000 | 83.33% | 14.02 |
| **2006** | **June** | 1,866,716 | 297,241 | 19% | 44,000,000,000 | 22,000,000,000 | 100.00% | 23.57 |
| **2007** | **January** | 2,039,259 | 172,543 | 9% | 56,000,000,000 | 12,000,000,000 | 27.27% | 27.46 |

**Figure 5: Usage/Growth stats**

## TCP/IP PERFORMACE IN HIGH SPEED NETWORKS

DOCSIS 2.0 Transmission Control Protocol (TCP) downstream throughput is gated by the number of Acknowledgment (ACK) packets that can be sent on the upstream. This is because TCP tends to self-clock and the rate of the segments sent on the downstream depends on the rate of ACKs sent on the upstream. In other words, if a cable modem can send up to 200 ACKs per second, the downstream rate can not exceed "ACK per second X segment size" Bytes per second, or in a typical case 200*1460*8 = 2.33Mbps. DOCSIS Upstream Concatenation can help, but it does not solve the basic problem of having a limited "ACK rate". See RFC3449 (TCP Performance Implications of Network Path Asymmetry) for more detail on TCP performance limits due to asymmetrical links.

DOCSIS 3.0 allows multiple outstanding requests and so the bottleneck imposed by ACK rates is greatly reduced. The TCP receive window size becomes the new bottleneck. The following text explains the TCP receive window issue at a high level:

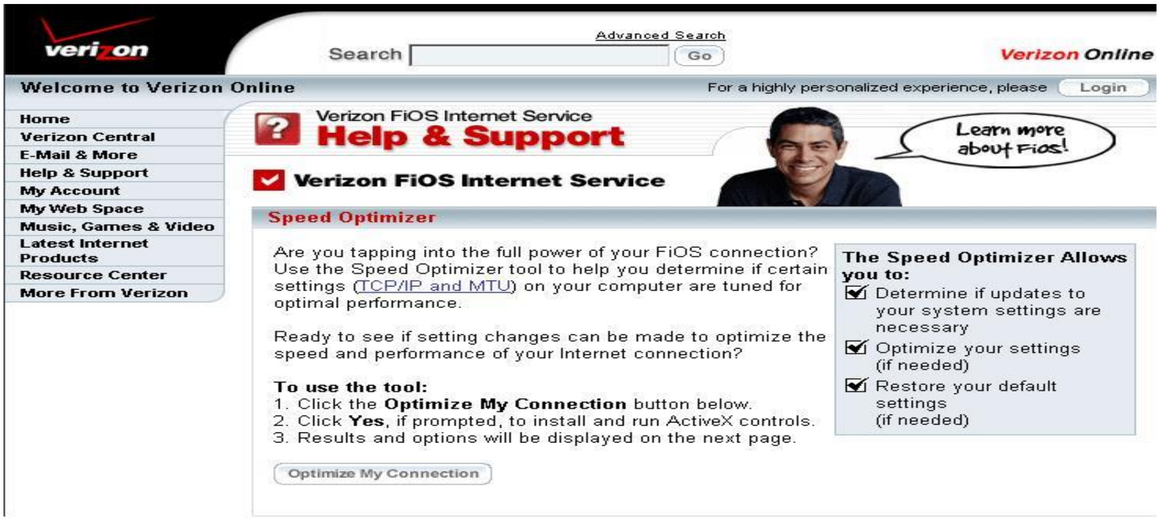
TCP defines a "window" as the amount of bytes that can be sent unacknowledged to a network. They don't need to be acknowledged (yet) because they count as in-transit, meaning that due to the various delays in the network they either did not have a chance to be received by the end point, or the ACK did not have enough time to travel back. In fact, TCP attempts to keep this window at a maximum because that way it "fills the pipe" between the sender and receiver so that there are always bytes in-transmit.

The original TCP definition sets the maximum number of bytes that can be in-transit through the network to 64Kbyte because of the 16bit encoding of the window size in the TCP header. If we assume a very simplistic model where a network introduces a 100ms roundtrip delay (possible on the Internet) and that a single ACK clears a whole 64Kbyte window, then the maximal best case throughput of a TCP session can not exceed 64Kbyte * 100ms * 8 = 52Mbps, which is less then the capability of a DOCSIS 3.0 modem – and is a theoretical upper bound. In a real system it's likely to be even slower.

To make things worse, the default window size in Windows XP is set to 16K, which will limit the maximal throughput in the example about to 52/4 = 13Mbps (even below DOCSIS 2.0 capabilities).

This issue with TCP has been recognized a long time ago, and in 1992 the IETF published RFC1323 that defines TCP/IP extensions to allow larger windows sizes.

To demonstrate that this in not a cable specific problem, even on the web pages of FiOS (Fiber to the home service from Verizon), it is recommended to run an ActiveX component (Figure 6) that sets the proper entries in the windows registry in order to take advantage of the RFC1323 support in Windows.
Note that they are other ways to increase TCP throughput and any "TCP speedup" utility would try setting the same parameters that are set by Verizon.



Are you tapping into the full power of your FiOS connection? Use the Speed Optimizer tool to help you determine if certain settings (TCP/IP and MTU) on your computer are tuned for optimal performance.

Ready to see if setting changes can be made to optimize the speed and performance of your Internet connection?

**To use the tool:**
1. Click the **Optimize My Connection** button below.
2. Click **Yes**, if prompted, to install and run ActiveX controls.
3. Results and options will be displayed on the next page.

**The Speed Optimizer Allows you to:**
- ☑ Determine if updates to your system settings are necessary
- ☑ Optimize your settings (if needed)
- ☑ Restore your default settings (if needed)

**Figure 6: Verizon TCP speed up page**

The following parameters are set:

1. TCP 1323 Extensions - This parameter enables enhancements to the TCP/IP protocol that provide improved performance over high speed connections;

2. TCP Receive Window - This parameter specifies the number of bytes a sender (the source you are downloading from) may transmit without receiving an acknowledgment. Modifying it determines the maximum size offered by the system (appears to be about 400K);

3. MTU (Maximum Transmission Units) - The MTU defines the largest single unit of data that can be transmitted over your connection. The FiOS network requires an MTU of 1492 bytes.

All utilities are available on the web under the general categories of “TCP optimizer”, “TCP speed” etc, change these parameters. In a cable environment, and especially with DOCSIS 3.0, a user should update the operating system TCP defaults to work with TCP 1323 extensions, so that larger window sizes can be define.

## IMPACT ON THE HOME NETWORK

The bottleneck link at the home used to be the internet connection. It was fair to assume that whatever the internet delivered, the home network can consume. With DOCSIS 3.0 this might no longer be the case. For example, if a user has a contract for a burst rate of 100Mbps, but the home wireless router is limited to 54Mbps, it is possible to congest the home network with data.

To allow for proper QoS handling in the event of a home network congestion DOCSIS 3.0 has the option of embedding priority bits in the 1-Byte, 3-Byte or 5-Byte DOCSIS headers. The priority bits are taken into consideration when queuing packets after the packets have been processed on the RF side and are on the way to the home network. Once they are queued to the home network, than in the event of congestion the lower priority packets would be dropped first.

The priority bits are arranged based on the Ethernet 801.D priority bits definition as depicted in Table 1:

| | | **Number of CM output queues** | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| **Traffic Priority** | **0 (Default)** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | **1** | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | **2** | 0 | 0 | 1 | 1 | 1 | 1 | 2 |
| | **3** | 0 | 0 | 1 | 1 | 2 | 2 | 3 |
| | **4** | 1 | 1 | 2 | 2 | 3 | 3 | 4 |
| | **5** | 1 | 1 | 2 | 3 | 4 | 4 | 5 |
| | **6** | 1 | 2 | 3 | 4 | 5 | 5 | 6 |
| | **7** | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

**Table 1 : DOCSIS 3.0 priority encoding**

This table allows vendors to build devices with varying degrees of complexity in terms of how many output queues they have on their home network interfaces. The table defines which priority encoding goes to each queue.

Natually, if there are 8 queues then each gets a dedicated priority queue.

This feature can work in tandem with uPNP and Ethernet type flow control to preserve end-to-end QoS all the way to the home device.

## HIGH SPEED AND POWER CONSUMPTION

Another side impact of higher traffic rates is the increased power consumption by networking equipment. As with other issues raised in this article, this is not a cable specific issue. It's not even a communication industry specific issue. It is a result of the current CMOS technologies, used in most of today's ASICs, reaching their physical limitations.

For example, modern oxides (the insulation layer in CMOS) are at 10-12 angstrom, that's 5 atoms thick! In addition to the fact its hard to go much lower then 5 atoms, there is already a scaling problem, since reducing the density by half would require placing a 2.5 atom thick oxide – clearly not an option…There is also a limit to the minimal voltage that a CMOS circuit can operate in. As a result, there a limit to how much power can be reduced – a limit that impacts CPUs, mobile devices, data centers and telecom.

In the past routers/switches had to look only at the packet headers (Ethernet/VLAN for switches, IP for routers) in order to forward a packet to its proper destination. As data rates went higher, the faster the forwarding rates had to go and the more power was needed to perform it. In addition to this basic forwarding the amount of "heavy lifting" per packet/byte is increasing. A sample of application that require per byte operation are:

1. Deep Packet Inspection: For reasons that relate to filtering certain network traffic type, and to protecting the network against Denial of Service Attacks, the network has to look beyond the L2/L3 portion of a packet in order to figure out what a packet is, not just where it goes to.
2. Video Streams manipulation: High touch video processing, such as transcoding, trans-rating and ad-insertion require extensive byte manipulation.
3. Encryption: While not new to cable networks, encryption is a byte-by-byte operation on a whole packet. The faster the data rates are, the faster the encryption chips have to run and the more power they need.

Hopefully, technology innovations will help reduce the power requirment as much as possible. In addition,new network architectures, such as M-CMTS, allows for distribution of function (L1/L2/L3) which is also a distribution of power. For example, if the M-CMTS packet shelf is located outside the hub, it reduces the power demands on the hub.

## DUAL TOKEN BUCKET

A single token bucket definition (as the one used in DOCSIS 1.1/2.0) places restriction on the peak rate and burst size of a flow. The problem with a single token bucket definition is that for the duration of a burst the flow is not limited to the peak rate, instead, the burst rate can be as high as the total capacity of the link. As long as the burst size is minimal (in the range of a couple of Ethernet frames) this is not a significant issue, but for those customers who define large burst sizes the burst size could become an issue with assuring fairness across a large number of flows. It can further be exacerbated because of the data rates supported in DOCSIS 3.0.

The way to address extended bursts with the DOCSIS 3.0 toolkit is the newly defined "maximum sustained traffic rate" and "maximum downstream traffic rate". These two parameters refer to how fast the traffic can flow during the traffic vs. how fast it can flow once the burst is exhausted. For example, a traffic contract can be:

- Maximum sustained traffic rate is 10mbps

- Maximum downstream traffic rate is 30mbps

- Burst size is 2Mbytes

For this contract, the user can send up to 2Mbytes of data at 30mbps, but if the user sends more then 2Mbytes the CMTS would reduce the rate to 10mbps.

## TRAFFIC PATTERNS IN A HIGH THROUGHPUT NETWORK

With the bandwidth that DOCSIS 3.0 provides it is very likely that a new crop of applications will show up and put to use the added capacity. These new applications will have new traffic patterns, and therefore the network planning can not be derived from the existing traffic patterns. The first section of this paper already demonstrates how unexpected the aggregate traffic can be when individual contracts are increased. This section will discuss some of the applications that might use the bandwidth provided by DOCSIS 3.0 and their impact on the network. In one sentence, the one common theme is that video is the new "killer app"that will drive bandwidth demands.

Today, a majority of internet traffic is related to file sharing. In fact, one may argue that a good percentage of file sharing traffic is video content and therefore is a form of video on demand distribution. In that sense we are already experiencing a bandwidth explosion that is video driven. The next step to file sharing is to stream the video content and play it while it's being downloaded, and this is already being implemented by joost (www.joost.com). The "joost" type of video streaming has a couple of fundamental impacts on network traffic patterns.

1. Since every home is a media source, as well as media consumer, the upstream bandwidth is driven higher.

2. Joost, just like a file sharing application, opens up multiple parallel TCP sessions. That means that it will try to squeeze as much bandwidth from the existing traffic contract without trying to be "fair".

While joost and the multitude of joost-like applications that are sure to follow, represent "over the top" video (meaning video sources that are not originated by the cable MSO), a cable MSO IPTV delivery will have its own impact of bandwidth usage and traffic patterns. Although its not clear at the moment if IPTV will be a significant bandwidth drive in the cable MSO world, its worth exploring the way it is different then over the top delivery.

MSO generated IPTV streams are likely to be carried over RTP (as opposed to TCP). While TCP is a closed loop protocol, RTP is open loop, meaning that the RTP source will not slow down or try to be network friendly if the network is congested. In that case, one might ask why use RTP instead of TCP? The reason is that RTP has a much better story when it comes to delivering real-time content and that is one way to differentiate MSO content from over-the-top content. But because open-loop delivery can not recover as gracefully as TCP, the network has to be over provisioned by much more then with TCP flows. In fact, if IPTV catches on, there could be a major shift from a world where most traffic is TCP to a world where most traffic is RTP: instead of flows that regulate themselves (TCP) a network that has to be over-provisioned to accommodate non-flexible traffic flows (RTP). Also take into account the fact that a video flow is fairly long lived, and the end result is a network that future networks can’t be oversubscribed by the amount used today. Note that this conclusion may be somewhat counter-intuitive since higher bandwidth, in principal, should lead to better statistical multiplexing and higher oversubscription ratios. If the internet traffic remined as it is today, this may have been the case, but it is less likely when considering future applications.

-Alon Bernstein, alonb@cisco.com
-Joe Godas, joe@cv.net

# THE UNIQUE CHALLENGES FACED BY CABLE SYSTEMS SERVING SMALLER MARKETS AS THEY EXPAND TO MEET CONDITIONAL ACCESS AND OCAP REQUIREMENTS

Gary Traver, James Capps
Comcast Media Center

## *Abstract*

*Separable security systems represent a fundamental component of a long-term migration towards an open standards / open architecture world. OCAP, the OpenCable Application Platform, which was developed for the industry by CableLabs*[i]*, is designed to drive the future of interchangeable devices toward this vision for an open-architecture environment.*

*The operational rigor associated with managing devices in this interchangeable world is complex. As both separable security and OCAP evolve, the number of consumer devices able to make use of this open network will grow.*

*Cable systems serving smaller markets must address the expensive regulatory requirement for separable security and the migration toward OCAP at the same time that their customers have more choices for the providers of advanced video services than any other time in history*[ii]*.*

*Because they lack the economies of scale afforded by cable systems serving larger markets, cable operators serving smaller communities need solutions that build upon their existing architecture. In addition, they can meet or exceed the economies of a large-market solution by pooling their resources into a centralized service bureau. However, a one-size-fits-all solution is counter-productive to a competitive marketplace. To be effective, the centralized services bureau will also need to address the unique requirements of each cable system.*

## INTRODUCTION

U.S. cable systems serving exurban areas of the U.S, the first markets to embrace cable television's "community antenna" technology, face the greatest challenge in creating next-generation content distribution networks. As cable nears its 60th anniversary in many of these communities, these cable system operators are tasked with meeting new conditional access requirements that potentially thwart their ability to offer advanced digital video services and threaten their survival. In addition, they must address expensive regulatory requirements at the same time that their customers have more choices for the providers of those services than any other time in history[iii].

While separable security systems represent a fundamental component of a long-term migration towards an open standards / open architecture world, the entire cable industry must respond by complying with a FCC mandate for using separable security to serve all new video customers in the short-term by July 7, 2007. For the purposes of this paper, we are assuming that separable security methodology will evolve over the next several years into a fully software-downloadable security environment.

In addition to managing the conditional access and security processes, another key element of enabling the next-generation content distribution architecture is its ability to support the open standards environment envisioned by OCAP, the OpenCable Application Platform. OCAP, which was developed for the industry by CableLabs[iv], is designed to drive the future of interchangeable devices (i.e., STBs, PCs, TVs, etc.) that can easily connect to a cable operator's network and download that operator's user interface and the key system information that is necessary for the device to navigate within the operator's network. This future downloadable, interchangeable world creates a host of new opportunities. However, the ability to connect any device that is both security-compliant and OCAP-compliant also creates a level of complexity well beyond the scope of anything that is in place today.

Cable operators serving smaller markets must comply with security requirements without the same economies of scale as regional cable systems, which serve an urban/suburban cluster of customers from one headend. For example, the minimum cost of the headend equipment and software required to meet federal requirements for separable security are fixed. Therefore, the investment required to implement those solutions can be several magnitudes higher on a cost per subscriber basis for a small cable operator than they are for a larger cable operation. If the minimum system configuration is sized to meet a 25,000 subscriber base, operators with a 5,000 customer base must pay roughly 5 times the cost per customer, and systems in the 1,000 customer base range must pay 25 times the cost per subscriber in order to offer the same mandated solutions.

Accordingly, the same economic issue that applies to separable security also applies to the larger implementation of OCAP in these exurban markets. The systems necessary to detect, manage and deliver the other critical layers of software to enable the use of consumer-owned devices on the network have very similar requirements and thus very similar cost issues for cable operators serving smaller markets.

Finding an economically viable solution for providing these services requires some key changes to the smaller markets' systems and architecture. To provide the same scale economics as larger systems, the conditional access solutions must be able to pool subscriber bases together from many small, unaffiliated cable systems. It must do so in a manner that maintains the individual configurations and attributes of each small system and maintains the security and integrity of each system's data. The most viable approach to meet these criteria is to create a centralized service bureau type of operation.

The operational rigor associated with managing devices in this interchangeable world is complex. As both separable security and OCAP evolve, the number of consumer devices able to make use of this open network will grow. Attention to maintaining network compatibility must be focused in two directions. The focus must first be **technology forward**, to make sure that the system is capable of supporting new consumer devices as they continue to grow and evolve. It must also be **technology backward**, to ensure the legacy equipment in the cable systems continues to receive the appropriate support. Cable operators serving smaller markets must have solutions that allow them to build upon their existing architecture. The economics of their local businesses do not allow them to engage in a total overhaul of their current content distribution infrastructure.

The country's independent cable operator community must have an affordable solution that complies with federal regulations and additional requirements regarding conditional access security systems by July of this year. While many of these systems have expanded their businesses to offer high speed data and voice services, regulatory requirements pose a real threat to their continued existence. In fact, as the American Cable Association observed: "...many of the small businesses that provide video and broadband services in rural America will cease to exist and the digital divide will actually grow[v]."

As this paper will demonstrate, a solution that is focused on the unique needs of the smaller market in the form of a centralized multi-operator services bureau can allow many cable systems to not only overcome these near-term challenges in conditional access but also provide a long-term strategy for OCAP success. This robust services platform must address many of the requirements for migrating to an all-open standards downloadable environment. By pooling these markets into a national platform, it will also allow them to offer the same access and advantages as other larger cable organizations, enabling smaller system operators to expand the lineup of advanced digital services that they can offer to their customers.

## THE CHALLENGES AND OPPORTUNITIES OF NEXT-GENERATION CONDITIONAL ACCESS

While there are many challenges associated with creating secure and robust systems that meet the federal mandate for separable security, the intent of the mandate is consistent with the cable system operators' need to provide a service that responds to their customer's needs and demands. The migration towards separable security and OCAP-compliant consumer electronic devices increases the number of choices that consumers will have to experience the array of services provided by their cable operators. Further, as this environment begins to proliferate across the entire marketplace, all operators will eventually be expected by their customers to provide support for consumer-supplied devices and the type of portability that they provide. Customers who already have their own downloadable devices will exacerbate this demand even further when they move into another cable system's franchise area.

However, providing the systems and infrastructure required to provision services to a wide array of devices in this future downloadable environment is not an easy or inexpensive task. The complexity of the equipment and the systems required to support an open environment increases significantly over current operational requirements. Without economically and operationally viable alternatives, this migration could result in cable system operators struggling to provide even the most basic services under this next-generation architecture.

From the author's perspective, support for separable security and for OCAP are extensions of the same core mission. They both allow for portability and they both require the consumer device to be connected to the cable operators' network in order to get the necessary software and information required to properly receive the services provided by the operator. Most importantly, the consumer devices must depend on the equipment and systems from both of them to function properly. Therefore, from the standpoint of creating an enabling architecture to provision the devices, it makes sense to combine their functionality into a single service entity.

# ADVANCING TECHNOLOGIES FOR SET-TOP OPERATIONS

## A. CableCARD Implications

Section 629 of the 1996 telecom law resulted in the creation of CableCARD by pressing for separable security[vi]. Responding to this drive, the OpenCable standards were created and specified the CableCARD as a removable security device as follows:

- A Host device (set-top box or integrated television) that provides generic cable tuning and decoding capabilities that are portable across all cable networks.
- A removable CableCARD security module that separates retail delivered set-top box (Host) from proprietary conditional access systems and network messaging.
- An interface between the Host and CableCARD module, that allows for Hosts and CableCARD modules from different vendors to interoperate.

Enabling the CableCARD functionality will require interaction with various downstream systems. Beyond the additional capital outlay required for the CableCARD itself, the smaller operators will now need to ensure that their servers are capable of interacting with the set-top boxes to enable EPG and authorizations.

## B. A Future Alternative to CableCARD – DCAS

A cost-effective, cable company agnostic and conditional access system agnostic alternative to the CableCARD is the Downloadable Conditional Access System (DCAS). DCAS allows the cable operator to download its conditional access system of choice to devices connected to its cable network. It is designed to operate with cable set-top boxes, integrated retail DTVs and other devices that include a secure DCAS microprocessor chip. The DCAS protocol is available to any consumer electronics manufacturer enabling a large market of devices to be available for consumers.

A relevant advantage to DCAS is responsiveness in that it is able to address security breaches more quickly and efficiently. In March of 2007, hackers cracked encryption codes used by Swiss cable operator Cable COM and digital television technology group Kudelski SA, and released them on the Internet[vii]. Web savvy pirates could transfer the published codes to a digital TV decoder called a Dreambox DN 500c, and freely access subscription video on the cable system's lineup, according to press reports. A DCAS model allows such breaches to be addressed immediately and remotely, requiring no new hardware or truck-rolls.

## C. Expanding the Set-top Functions - OCAP

CableLabs originally developed the OpenCable family of standards to provide a common basis for digital cable TV systems in the U.SSupport for devices addressed via a new standard, the OpenCable Application Platform or OCAP. OCAP results in a stack of software residing between applications and the operating system within a consumer electronics device such as a set-top box or OCAP-compliant TV set. OCAP devices can have new information or applications loaded on them, often taking advantage of new two-way capabilities.

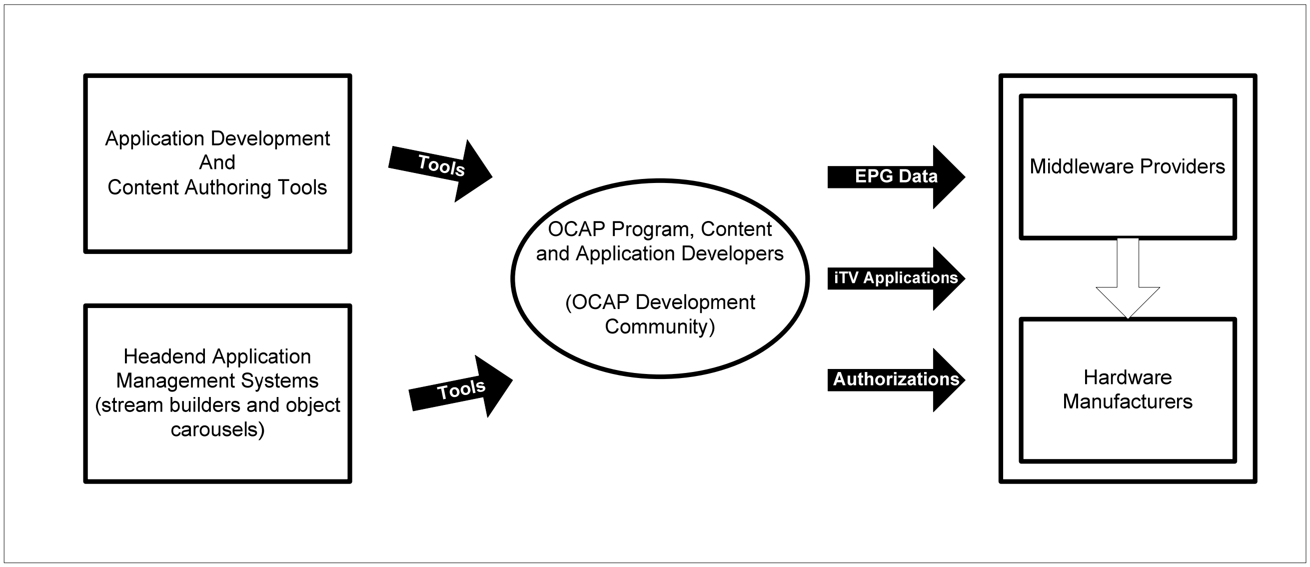


**Figure 1 - High Level Diagram of OCAP Implementation**

Additionally, the OpenCable environment allows multiple vendors to play in the headend space, providing functions and features previously provided by one device or vendor. As with many technological advances, increased choice bears increased complexity. Figure 1 highlights the new approach to set-top box development under an OCAP system. Various providers, vendors and systems now play under a defined specification to enable previously "built-in" functions.

## UNIFYING THE SECURITY AND OCAP IMPLEMENTATION

Separable security and OCAP are currently viewed as independent issues and projects by the industry. However, when examined from the deployment and sustaining operations perspective, they can be viewed as extensions of one another. For an advanced security OCAP compliant device to connect into a cable system's network, it must first receive a series of downloads from both the systems managing the security and the systems managing OCAP functions. It must then continue to receive a series of data streams which provide necessary information and software required to sustain functionality. Therefore, from the standpoint of creating an enabling architecture to provision the devices, it makes sense to combine their functionality into a single service entity.

### A. Separable Security Deployment

The first step an operator will need to undertake is the deployment of an advanced security capability. The implementation of either separable security solution, DCAS or CableCARD, is not a trivial task. Figure 2 shows the messages required to drive the CableCARD. While many of these messages and streams are already a function of existing systems, the OCAP compliant infrastructure may result in each message generator being independent.

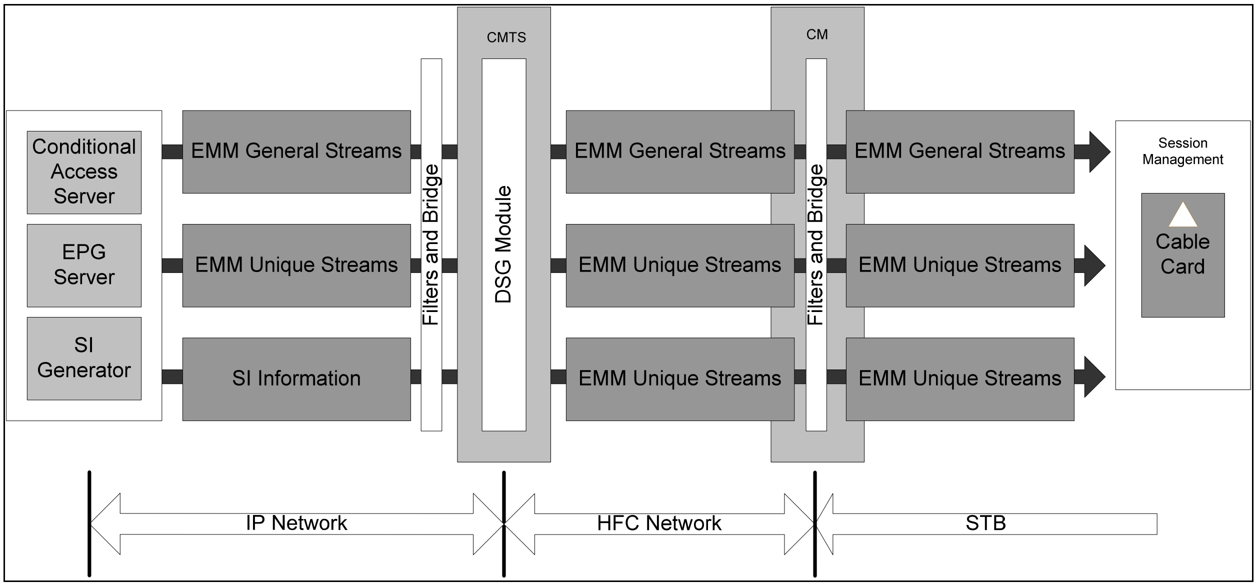


Figure 2 - Device Message Requirements

Independent servers providing EPG (Electronic Program Guide) listings data, and SI (System Information i.e., channel maps) are now necessary to ensure CableCARDs are initialized and operate in the consumer's home.

Figure 3 is the diagram provided by the FCC in its "*Report of the National Cable & Telecommunications Association on Downloadable Security*". This diagram highlights the steps required to initialize the DCAS device. As can be seen, once inserted into the consumer device and connected to the operator's network, the consumer device must communicate with an authentication proxy server located in the headend to initialize use of a system. This DCAS authentication proxy server sends the CAS download key to the set-top. Then the encrypted CAS image is sent to load the set-top box with appropriate encryption software.

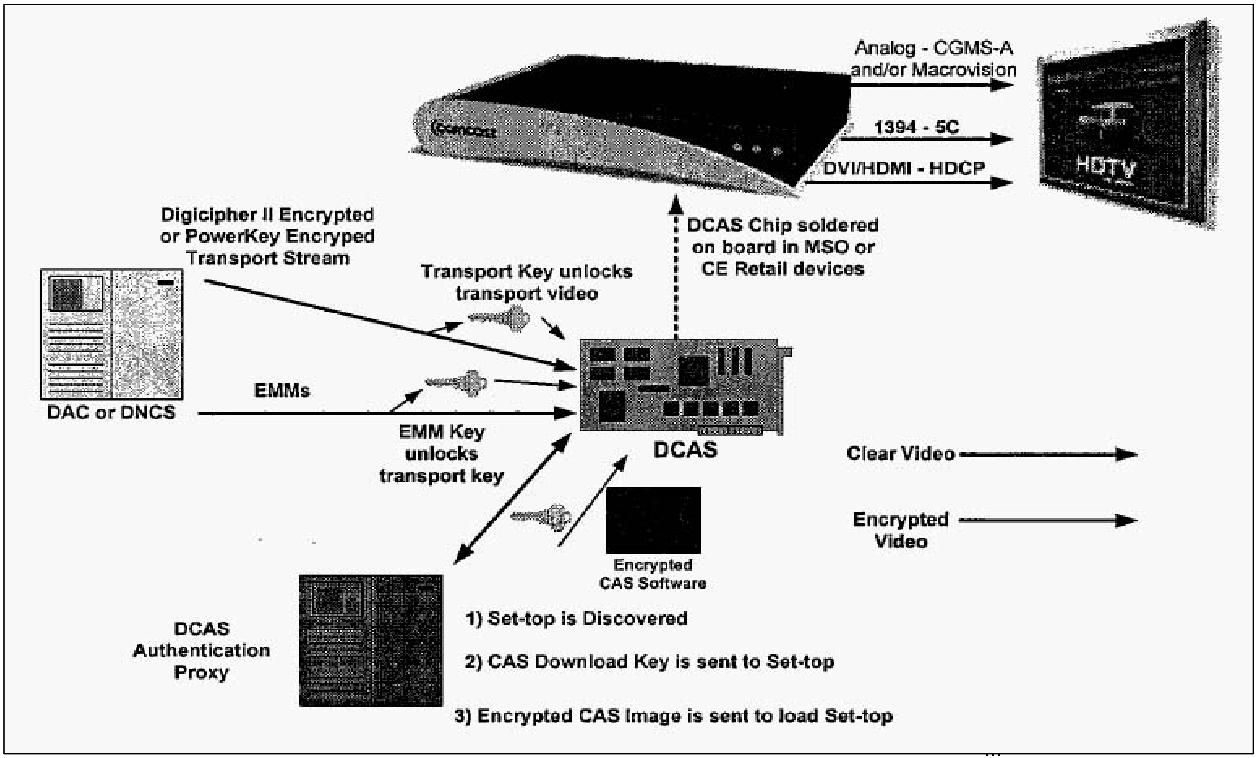


**Figure 3 - DCAS process as envisioned by the FCC**[viii]

Once the security system has linked with, and authenticated, the consumer device, it can then download the series of entitlements and other secure information necessary to work in conjunction with the configuration information included in the services.

Therefore, the first step in providing advanced conditional access requires that several new applications and servers be installed and operated by the operator. To simply initialize and begin communication with the set-top box in a separable security environment, the operator must now be prepared for an increase in equipment and expertise.

B. Basic Set-top Functions under OCAP – Middleware and Application Installation

Following the initialization of the security elements, the consumer devices must be given "personality data" in the form of software. In some cases, the consumer device must acquire the OCAP software stack, a middleware software component developed to enable multiple applications to interact together. Applications, like a programming guide or an on-demand ordering system, sit on top of the middleware. Operating systems (OS) are below it. The job of the middleware is to translate what lies at the root level for what sits above it. This, for example, allows an interactive trigger from a programmer to function without needing to know what version of OS is being used in a particular model of HDTV or set-top box.

Once the stack is available on the device and fully operational, the key software applications that will remain resident in memory in the consumer device will need to be loaded. Key functions such as the main user interface, including the EPG, and other basic functions to the network, like the applications required to receive and decipher channel maps. Figure 13 shows in additional detail the OCAP stack.

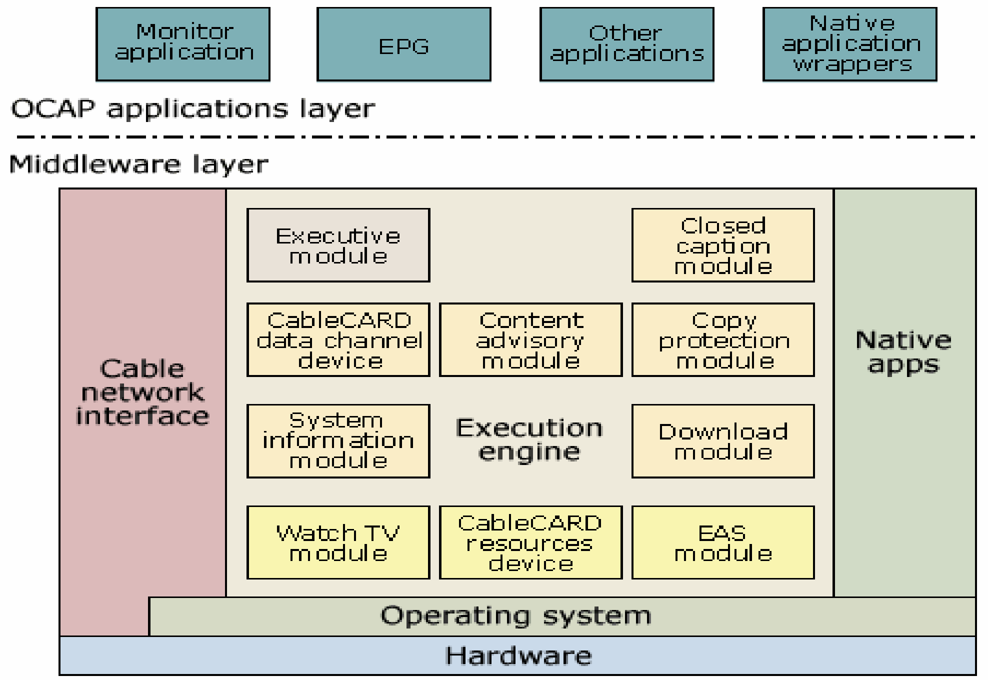


**Figure 4 - OCAP Stack with Applications**

Again, additional hardware will be required in the headend to load the middleware or applications. The software will be deployed using carousel file servers. Two methods are available for delivering these applications into the consumer device via the carousel file server. The first method uses the out of band data channel. The second uses the DOCSIS Set-top Gateway (DSG).

Once the middleware is installed on the set-top box, additional applications that are stored in memory in the consumer device can be installed. Examples of such applications are the main user interface or the EPG. The consumer device also has the capacity to download other applications that will reside in memory during the time that they are in use by the consumer. An example of this type of application is interactive games. The system will require one or more servers to provide these types of applications to the consumer devices.

Comparing this implementation to today's basic cable implementation shows a glimpse of the increased complications that come with the new age of cable television. The exurban operator will be challenged like never before. Even the larger system operators are concerned about the new challenges that lie ahead. In a recent article in *Communications Technology*, Chris Bowick, Cox's CTO, was quoted as saying: "… (OCAP) [is] a very complicated topic, and it's not just new set-top boxes and new software on set-top boxes …From a systems perspective, you have the addressable controller and the DSG that will be used in OCAP deployments for the two-way communications instead of the proprietary S-A and Motorola approaches we have today[ix]."

Implementation of either the downloadable security system or OCAP relies heavily on components not in use today by a cable operator. While every effort will no doubt be made by the numerous vendors involved in developing these systems to ensure smooth interoperability, the complexity of the initial deployment and long term operation will tax even the largest of system operators. Smaller market operators now managing their headends as “lights-out facilities”, requiring little or no daily management or support, will find that the basic activities of simply adding authorizations or new set-top boxes will require complex interactive communications with local and regional security systems. This approach is beyond the technical and financial limits of many rural operators.

An OpenCable implementation, as depicted in Figure 5, is representative of what larger operators will most likely deploy. In addition to providing the fundamental services such as encryption and EPG, these cable systems will expand their competitive service mix by enabling customers with VOD, DVR and OCAP. However, Figure 5 also shows the number of headend servers and services growing substantially.

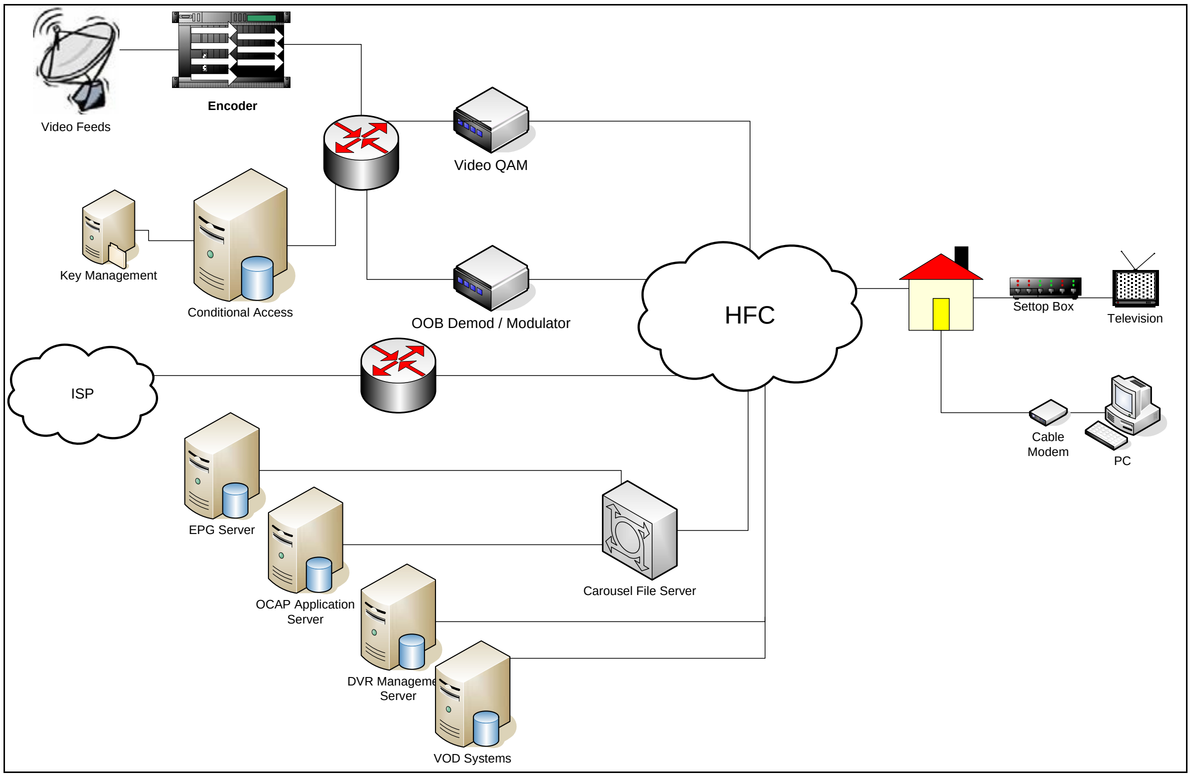


Figure 5 - Complex OCAP Implementation

## DAY TO DAY OPERATIONS - ENTITLEMENT AUTHORIZATION

The small operator's task list has grown under the new FCC and OCAP guidelines.

- First and foremost, the operators must maintain control and function of the existing set-top box population. This value of this investment cannot be understated.
- Second, the operator must respond to FCC-mandated changes surrounding security. Changes in headend gear will be necessary in order to deploy the new security system(s).
- Third, the operator will need to support OCAP product offerings. Advanced services, which will drive revenue, will require the support of OCAP servers in the headend.
- Fourth, the operator must provide continuous updates to support the devices and functions that will be offered as the market continues to evolve.
- Finally, the operator needs to ensure continued operation of this system to authorize and manage these services and functions.

It is this aspect of continued operation that will be most taxing to an operator. New processes and procedures required to add set-tops, replace devices and ensure appropriate billing will overwhelm the lightly staffed smaller cable systems.

### A. Set-top Authorization

As with a non-CableCARD set-top box, the service provider defines the types of rights needed to access the content. Through the CA system in use (DAC, NAS, DNCS) the ECM ciphers with the broadcast key according to the profile and generates a control word (i.e., the content encryption key). The controller sends the ECM and starts streaming the content ciphered with control word.

Following that communication, the CableCard deciphers the ECM (using the broadcast key) and checks the required rights relatively to the list of rights stored in the card. If the rights are correct, the card returns the CW to the decoder that will use it to decipher the received content. When subscriber rights have to be updated, an EMM is sent to the user's CableCard. It is ciphered with the broadcast key and contains the updating information (new rights, loss of rights, increase credit limit, etc.) and the Subscriber_Id.

### B. OCAP Application Authorization and Operations

The previous section addresses the steps to authorize a service. This process is largely handled by the CA system with interfaces to new external devices. The authorization of applications, and the submission of necessary data to those applications, is a much larger and more complicated endeavor. Another look at Figure 5 shows the increased number of servers providing services to the consumer devices. Previous sections discuss the fact that the applications must be loaded on top of the OCAP stack. However, the entitlement of the cable customer to receive this application and the associated data streams must also be managed. Interfaces with billing systems and entitlement servers will be required to enable customers to receive the VOD client and EPG in accordance with appropriate authorizations.

A centralized authorization management system will assist operators greatly. However, because each application is likely to be developed independently, a seamless and unified means by which each carousel file server can be managed and controlled is not likely to exist for several OCAP generations.

## ECONOMIES OF SCALE FOR THE SMALLER SYSTEMS OPERATOR

### A. Centralized Service Bureau

As previously mentioned, the economics for deploying next-generation content distribution systems are very different for cable system operators serving exurban markets within the U.S. In fact, the vast majority of the 7,090 cable systems in the U.S.[x] serve fewer than 25,000 customers and a significant number of them serve fewer than 5,000 customers. The American Cable Association (ACA) represents 930 independent cable systems operators that serve 8 million U.S. cable households via 5,000 systems. This translates to an average of 1,600 customers per headend and ACA indicates that more than 1,000 of these systems serve fewer than 1,000 subscribers[xi]. Many of the ACA's members also participate in the National Cable Television Co-Operative (NCTC). NCTC represents 1,100 independent cable owners of 5,500 individual systems serving more than 10 million subscribers, which also illustrates the number of cable systems serving smaller, exurban communities across the country.

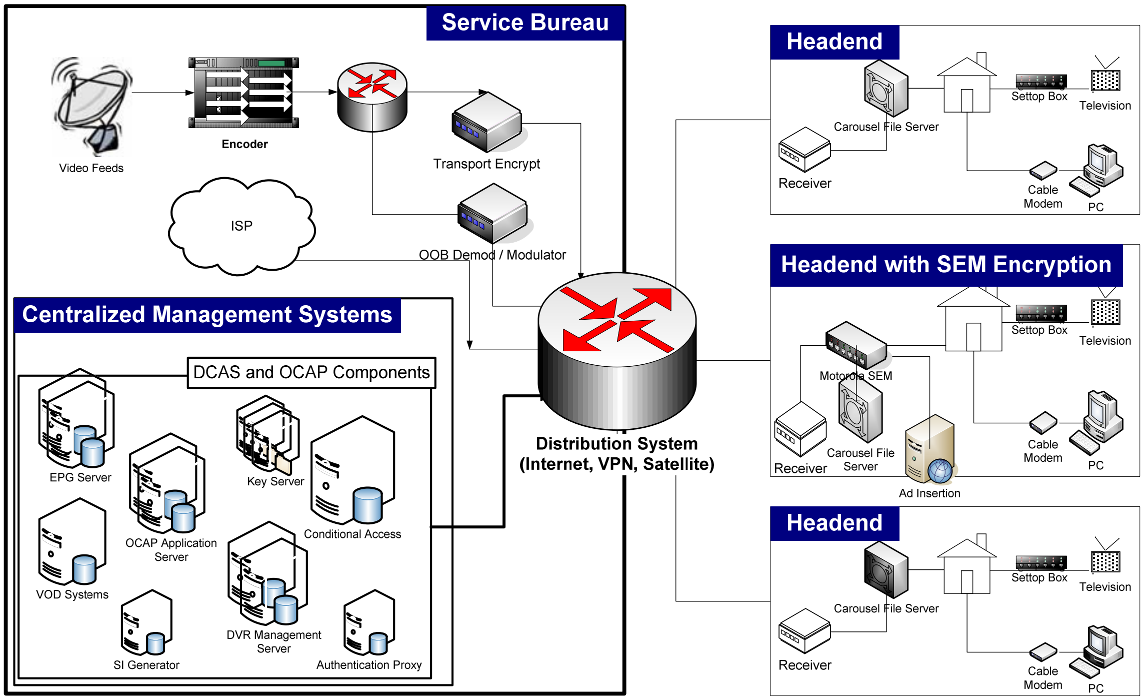


**Figure 6 - Broad Deployment via Centralized Service Bureau**

When we are examining how these markets will comply with the requirement for separable security and OCAP, it is important to note the additional limitations of their current system architecture. In fact, the most recent wave of upgrades in these smaller markets was using 450 MHZ plant and equipment that came onto the market when larger systems upgraded to 750 MHz[xii]. Many of these small-market systems deployed digital video service using the first generation of digital set tops boxes.

The service bureau provides an economically viable solution to cable operators serving smaller markets by pooling many small systems to achieve economies of scale. This creates a single large installation of equipment and applications that mirrors or exceeds the economies of scale found in a large market. Staffing follows the same model. The service bureau can deploy a robust set of operations, technical engineering, and support personnel spread across the entire base of small cable systems. Because the staffing for a centralized approach equates to a small fraction of an FTE on a per system basis, a top-notch staff can be assembled to provide superior services.

In order to address the unique needs of each cable system, it will be important for the service bureau to identify and manage each configuration individually. Additional controls will need to be in place in order to ensure the security and integrity of each cable system's data.

For the service bureau approach to best serve the customer base, it needs to balance the benefits of ubiquity with individuality. The goal of both separable or downloadable security and OCAP is to offer a variety of choices into the marketplace. Offering a one-size-fits-all approach runs counter to the purpose of creating an open marketplace.
Conversely, offering too many choices can increase the complexity of the operation and negatively impact the scale economics.

Regardless, each cable system needs to be treated as an individual entity.
The number of and type of services deployed and the timing of each implementation will be unique for every market. The support structure of the service bureau should be supportive of each cable system’s need to respond to local market demands, as operators move to offer choices that match the competitive alternatives.

Unilateral services such as unified server management, application certification and testing all ensure a quality offering to each headend. It is important that each configuration be validated and tested in a timely fashion. In addition, supplying redundant servers and high throughput systems for the entire customer base will help to ensure high availability and economies of scale.

Services from the centralized service bureau will require downstream connectivity from the service bureau via a satellite or terrestrial fiber-based data feed. The upstream requirements from the head-end to the service bureau can be managed using a secure VPN connection. Cable operators that are currently offering High Speed Internet services and/or VOIP services already have the connectivity to the Internet that is required for this upstream data channel. Now, as a result of these investments in IP-based services, cable system operators serving smaller markets possess many of the prerequisites for using centrally-supported conditional access and OCAP services for expanding their lineup of advanced digital video services.

## Conclusion

Cable system operators, CE manufacturers and retailers, their customers and the regulatory community share a common vision for a fully open-cable content distribution environment that supports a wide array of devices and applications. Achieving this vision for an interchangeable world is a much steeper challenge for cable operators serving rural markets. However, their future financial viability depends upon being able to offer customers the choice, convenience, quality and ease of use that this next generation content distribution platform can provide. A centralized services bureau, which allows operators serving these smaller markets to outsource much of the capital and operating requirements of an OCAP infrastructure, represents an essential element in making the interchangeable world available to everyone.

---

[i] CableLabs' OpenCable initiative. See www.opencable.com
[ii] CED magazine, January 2007
[iii] CED magazine, January 2007
[iv] CableLabs' OpenCable initiative. See www.opencable.com
[v] "Don't Leave Rural America Behind", ACA, May 26, 2005
[vi] First FCC Report and Order: Commercial Availability of Navigation Devices (PDF). FCC (1998-06-24). Retrieved on December 26, 2006
[vii] Sonntagszeitung newspaper, March 18, 2007
[viii] U.S. Federal Communications Commission document: "CS Docket No. 97-80: Report of the National Cable & Telecommunications Association on Downloadable Security"
[ix] Interactive Momentum: Mike Robuck, Communications Technology, April 1, 2006
[x] NCTA - Key Industry Statistics 2006
[xi] ACA report to the FCC
[xii] CED magazine, January 2007

Contact Information:

Gary E. Traver
Senior Vice President and Chief Operating Officer

James A. Capps
Senior Director, Software Engineering

Comcast Media Center
4100 E. Dry Creek Road
Centennial, Colorado 80122
(303) 486-3800
www.comcastmediacenter.com

# WiMAX As A Competitor To Fixed Line Broadband Technology

Tim Burke
Vice President, Strategic Technology, Liberty Global, Inc.
Tony Werner
Executive Vice President and Chief Technical Officer, Comcast Cable

## *Abstract*

*This paper examines the economic and technical capabilities of 802.16e WiMAX technology and contrasts it to fixed line technology.*

*Technical capacity assessment of burst rates, statistical load capabilities and performance under streaming and heavy Peer-to-Peer (P2P) environments are compared. It also identifies frequency re-use plans and antenna sectorization for varying subscriber densities. Performance analysis includes uplink and downlink performance and throughput based upon the available modulation, coding profiles and network configuration settings.*

*Various CPE is compared identifying the trade-off of embedded CPE and network design. The cost and performance trade-offs for mobility are also defined.*

*Economic analysis takes industry averages and builds an end-to-end per subscriber cost analysis based upon various homes passed densities, in building coverage, subscriber speeds and mobility assumptions. The costs include site acquisition, construction, base station, backhaul, CPE and installation costs.*

*The conclusion section contrasts these economic and technical characteristics to fixed line and identifies the areas of opportunity for WiMAX to be competitive.*

## WiMAX History and Design Goals

To better understand the capabilities of WiMAX let's first describe the evolution of the specification and its design goals.

WiMax was created out of an IP industry effort to replicate the success of WiFi technology but to do so in a controlled, wider area macro-environment instead of the distance limited and unlicensed/unregulated Radio Frequency (RF) environment of WiFi. The economic power of WiFi is its ability to incorporate a standardized wireless broadband capability embedded into PC, laptop and portable computing devices. The incentives and economics of this arrangement are compelling for CPE manufacturers, internet access providers and network operators, as the cost of the broadband wireless capability is borne by the user. If WiMAX could utilize the same economic vision of embedded wireless broadband CPE while offering a secure, simple to access method similar to today's wide area cellular networks it would surely come out a winner. In addition, the belief that wireless technologies were overly voice centric, narrow band, circuit switched and solely focused on high speed mobility created a window of opportunity for a technology with roots in the IP and fixed broadband data world.

There were some wrong turns during the development of the WiMAX specification. Early claims aggressively cited high user capacity, over great distances (70 miles) when travelling at very fast speeds (70 mph) were possible by many simultaneous users.

Fortunately, the developers of WiMAX remained steadfast to their vision and in a very persistent manner continued to upgrade the specification and standard through numerous revisions (802.16-2001, 802.16-2004, 802.16-2005). Equally pragmatic was the philosophy of the developers to build upon core technologies and specifications with a proven track record, such as IP, DOCSIS and wideband RF channel technology while at the same time incorporating newly developed technical advancements. Foremost among these enabling wireless technologies built into the specification were OFDMA (Orthogonal Frequency Division Multiple Access) modulation, MIMO (Multiple Input/Multiple Output) antenna technology, advanced coding techniques such as HARQ (Hybrid Automatic Request) and TDD (Time Division Duplex) duplexing methods. In retrospect, their design decisions were foretelling as experts in the cellular world are using the same technologies now in their submissions for the 4G wireless standards.

Right from the start there were some inherent constraints with the WiMAX specification that still exist today. The biggest obstacle being the spectrum ranges in which WiMAX currently operates. The original standard, 802.16-2001, was for the LMDS spectrum (24, 28, 38 GHz), which is only good for line of site microwave type applications. The current 802.16-2004 (Fixed WiMax or 802.16d) and 802.16-2005 (Mobile WiMAX or 802.16e) standards are more appropriate for Non-Line-Of-Site (NLOS) applications as they operate in 2.3, 2.5 & 3.5 GHz ranges. Unfortunately, these frequencies are disadvantaged when trying to match the propagation characteristics of cellular frequencies (.8, .9, 1.8, 1.9, 2.1 GHz).

Despite these challenges the developers of the WiMax standard set out ambitious design goals for a combination fixed broadband and mobile service prior to building the specification. To summarize, the key requirements are[1]:

- High average sector throughput to support > 1 GB /user/month
- High cell edge performance on downlink (> 1,000 kbps) and uplink (> 256 kbps)
- High performance uplink & downlink from NLOS indoor locations
- Support high number of simultaneous users (>150 per sector)
- Low latency to support user experience and real time applications (e.g. VOIP)
- QOS (Quality of Service) to support differentiated services
- Full portability and nomadicity
- High speed mobility (< 120 Km/Hr).

## Description of WiMAX Specification

The WiMAX specification is a double-edged sword. It has many flexible settings and assumptions that allow it to be optimized for a variety of business needs. For instance, if coverage, mobility and symmetrical services are desired then it can be configured to match those needs. Conversely, the network can be optimized for high capacity and asymmetrical capacity. As engineering and product marketing personnel can imagine, the numerous design “knobs” will be a huge benefit for delivering appropriate services to customers. Unfortunately, the flexibility offered by this capability comes at a cost, namely complexity and ultimately interoperability issues.

The biggest economic trade-off for a network operator is the decision to optimize for either capacity or coverage. The flexibility of WiMAX gives a network operator the opportunity to optimize the economics of a network design for either of these key criteria. Capital constrained start-up operators, without the benefit of a customer base, may opt to initially design a coverage based network that minimally meets the capacity requirements.

As the customers come online additional capacity can be added with the appropriate incremental capital. The power of the WiMAX specification is that this incremental capital can be minimized because of the high capacity capability of the standard relative to other wireless standards.

A network operator must first make the traditional business decisions of homes passed (addressable market), anticipated penetration rates, desired product speed, and applications. At that point a critical decision for the operator of a wireless network, like WiMAX, is the choice of terminal devices offered to the customers and the cell edge uplink data speeds. Outdoor fixed antennas, indoor gateways for PC's, PCMCIA cards for laptops, laptops with embedded CPE and handheld/mobile devices offer a variety of market opportunities and different applications. For the network designer each device has a unique link budget that weighs heavily on the service speeds and network costs. The limiting item for a wireless design is typically the uplink data rate at the cell edge. The cell edge is usually indoors through many walls and far away from cell site. The linkage of the choice of CPE and uplink speeds at the cell edge is a critical starting point for determining the network design and how the network should be best configured.

In the WiMAX specification there are a number of key parameters and assumptions that can be adjusted to accommodate either the coverage or capacity needs of a carrier's business plan[8]:

- CPE selection and cell edge uplink budget
- Adaptive modulation assumptions
- Frequency reuse of N=1, N=2, N=3… (where N represents reuse, such as N=1 means the same frequency is used in every adjacent sector/cell)
- Sectorization (Omni, 3 sector, 4 sector… 6 sector)
- Downlink to Uplink Ratios of 1:1, 2:1, 3:1
- Sub-channelization techniques
- Channel bandwidths of 5, 10, and 20 MHz

Antenna Technology[7,14]

  - MIMO (where there are multiple transmit and receive antennas at both the CPE & Base Station)
  - Beam forming Antennas (lock in and track the CPE and null out interference).

Unfortunately, optimizing these parameters can be a "zero sum game" for the operator. Improving one parameter can have a negative effect on others. Particularly disconcerting is the affect of optimizing for coverage and/or mobility which can severely affect capacity and vice versa. As is common in such trade-off's they typically result in economic choices. For instance, spending more money per cell site for sophisticated electronics associated with advanced antenna technology that gives the operator coverage gains versus just adding additional cost and simple base stations/sites.

Since WiMAX is a 4th generation (4G) wireless technology it is useful to understand how WiMAX OFDMA (Orthogonal Frequency Division Multiple Access) technology differs from the previous generations of wireless technology. Figure 1 illustrates the conceptual differences in the frequency, time and power dimensions of multiple access technologies[2]. Multiple access techniques divide channels or voice conversations by either time, frequency or unique identifiers like codes.

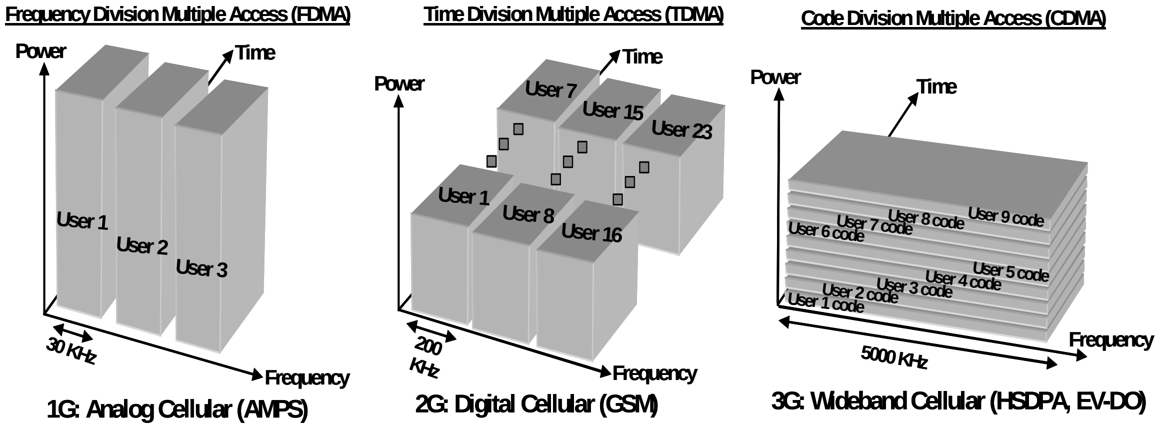


Figure 1: Conceptual Differences of Multiple Access Technologies

In Frequency Division Multiple Access (FDMA) each user receives a unique frequency as shown above in the 1st generation analog cellular example. Advancements in multiple access techniques in the late 1980s led to giving each user a unique time slot within the bandwidth allocation and digitizing the analog voice with vocoders. Time Division Multiple Access (TDMA) 2G digital cellular technologies such as GSM provided for 7 user time slots within a 200 KHz carrier. Finally, in the mid to late 90s Code Division Multiple Access technology reached maturation and was promoted as 3G wideband cellular wireless. CDMA users share time and frequency slots but employ codes that allow users to be separated by the receiver[29]. Contrasting these multiple access techniques to OFDMA is shown in Figure 2.

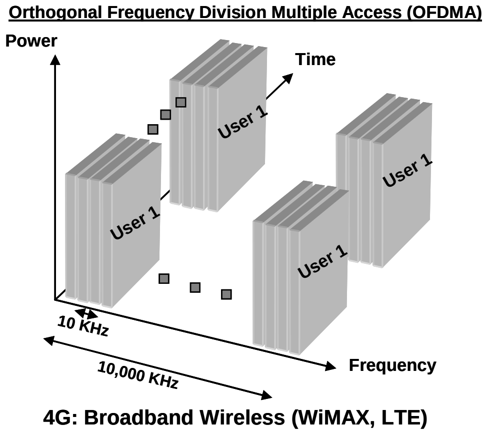


Figure 2: OFDMA

In OFDMA users share tightly spaced subcarriers and time slots. The diagram above shows over 1,000 separate 10 KHz subcarriers in a 10 MHz channel. These subcarriers are orthogonal to each other meaning they are unique and non-interfering. A stream of data from an individual user could be assigned or scheduled a variety of different narrow frequencies and time slots

The next sections will give an overview of the key aspects of the WiMAX specification by focusing in on the critical design parameters and design trade-off's.

## CPE Selection and Uplink Cell Edge Performance

Each CPE device has different output power and antenna gains. Since the uplink is typically the limiting item for the link budget there is a strong relationship between type of device and cell site range. As an illustration, a variety of different cell site ranges based upon the link budget for each of the various CPE devices[9] as shown in Table 1.

| CPE Type | Antenna Gain | Transmit Power | Cell Range |
|---|---|---|---|
| **Indoor** | 7 dB | 27 dB | 2.1 Km |
| **PCMCIA** | -2 dB | 24 dB | 1.0 Km |
| **Laptop** | 2 dB | 27 dB | 1.40 Km |
| **Handset** | -2 dB | 20 dB | .8 Km |

Table 1: CPE Devices & Cell Site Ranges

For simplicity, not all the individual elements and assumptions to the link budget are listed above. For example, in-building margins and fade margins could vary by application and CPE. Additionally, propagation models vary widely for high speed mobility and fixed applications and can have a profound effect on cell radius calculations as shown in Table 1.

As one can see in Figure 3, coverage will vary by CPE. This is mainly a result of the probability of using a higher speed modulation technique (e.g.- 64QAM) at the cell edge with higher gain CPE. Since modulation and coding modes are the key for performance, the next section will describe how modulation is uniquely used in the WiMAX specification.

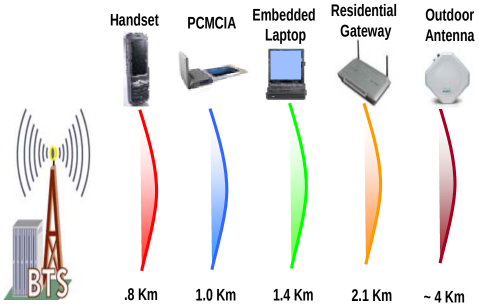


Figure 3: Coverage by CPE

## Adaptive Modulation

Adaptive Modulation is one of the critical characteristics of the WiMAX specification. Adaptive modulation techniques allow an individual subscriber to operate at different modulation rates than an adjacent customer. By having adaptive modulation, the base stations and subscriber stations are able to select at any given time the modulation rate separately for the uplink and downlink which will yield the optimum operation given the current link conditions. Chart 1 is an illustration of the coverage versus capacity tradeoff of using adaptive modulation schemes[17, 18]. At high modulation rates (64QAM) the sector capacity is very high (10 Mb/s) but the range is low (< 1 Km). Likewise at the lower modulation levels a great distance can be achieved but only at the expense of sector capacity.

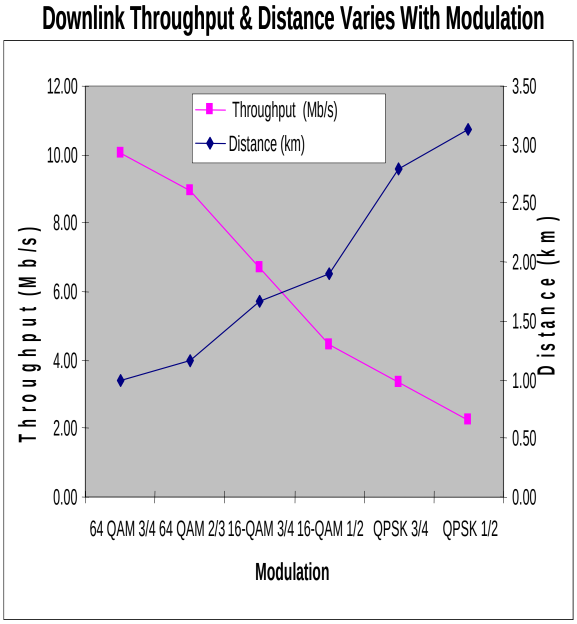


Chart 1: Speed & Distance by Modulation

Switching between modulation rates and coding is based on the Signal to Noise Ratio (SNR). Typically, high modulations are achievable for locations near the base station while cell-edge regions with low SNRs due to distance and potential interference from adjacent cells use low modulations. Fading due to distance causes subscribers at cell-edge to operate at low modulation rates which do not require high SNRs. In the presence of adjacent cell interference, the same subscribers may not be able to meet SNR requirement, and hence, results in lower modulation modes. Generally, NLOS environments suffer from low SNRs due to reflection losses, diffraction losses, multi-path fading and fading due to obstacles' scattering effect. The ability to achieve high modulation at cell edge depends highly on the ability to maintain high SNRs given the demanding link conditions of a NLOS environment. Table 2 summarizes typical downlink WiMAX modulations and SNR levels supported[4,9,10].

| Downlink Modulation and Coding | SNR Required (dB) |
|---|---|
| QPSK ½ | 1.8 |
| QPSK ¾ | 4.0 |
| 16 QAM ½ | 6.8 |
| 16 QAM ¾ | 10.4 |
| 64 QAM ½ | 12.9 |
| 64 QAM 2/3 | 14.2 |
| 64 QAM ¾ | 16.5 |
| 64 QAM 5/6 | 20.6 |

Table 2: SNR By Modulation

Without adaptive modulation, there are two likely situations which can be experienced by subscriber stations:

- Subscriber stations have high modulation rates even when link conditions yield low SNRs, resulting in high BER, and degraded performance
- Subscriber stations have low modulation rates even when link conditions yield high SNRs, hence, resulting to inefficiency.

Under both scenarios, the end result is low performance. By adaptively switching from one modulation rate to another, the system is able to ensure that[9,18,19]:

- Low modulation rates are selected when link conditions yield low SNR, resulting in low BER
- High modulation rates are selected when link conditions yield high SNR resulting in higher performance.

Figure 4 illustrates the concept of how the uplink modulation rate (QPSK ½) at the cell edge is typically the limiting design item and sets the speed of the user experience (256 Kb/s). Adaptive modulation, power and antenna gain of the CPE devices are key drivers of product speeds, cell capacity, maximum subscriber counts and coverage[3.4.16,17].

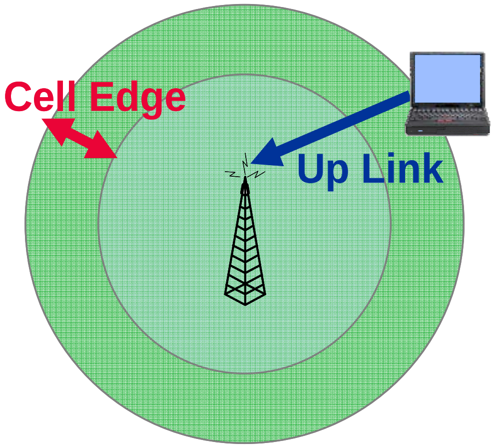


Figure 4: Uplink Cell Edge Representation

At this point it is important to briefly touch on the topic of WiMAX capacity in the adaptive modulation discussion since detailed capacity calculations will be covered in later sections. WiMAX average throughput capacity calculations are a controversial topic as there is little agreement on those numbers because there are so many varying assumptions possible. Suffice it to say, a key aspect of the uplink and downlink WiMAX capacity claims is the assumption on which modulation and coding scheme is used by the customers. For instance, peak or theoretical WiMAX capacities assume there is a single user in a sector and they are operating at the best modulation and coding rates in both the downlink (64QAM $^{5}/_{6}$) and uplink (16QAM ¾) 100% of the time. As can be imagined, high capacity claims can result across some vendors as a result of this assumption. The realism of all users always attaining the highest modulation rate is questionable, which is why we add the words theoretical and single user when using peak capacity numbers. More appropriate capacity claims are under the heading of average throughput or average channel capacity. Here again the adaptive modulation distribution assumption is absolutely critical. For example, the distribution assumption of how many subscribers (or the % of subscribers) operating in each adaptive modulation mode drives the overall average channel throughput number.

<u>Frequency Reuse</u>

The concept of frequency reuse is integral to any wireless and mobility technology. Without this capability the wireless industry would never have matured in a scarce spectrum environment. Basically, in a three sector cell site shown in Figure 5 each sector uses a different frequency. Cellular technologists call this an N=3 reuse pattern.

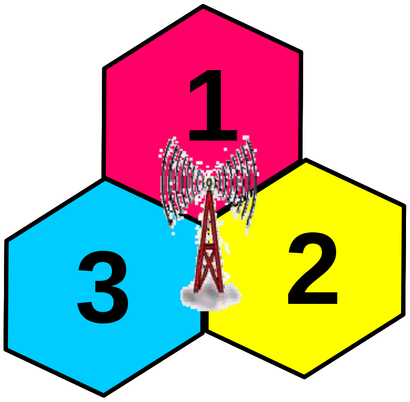
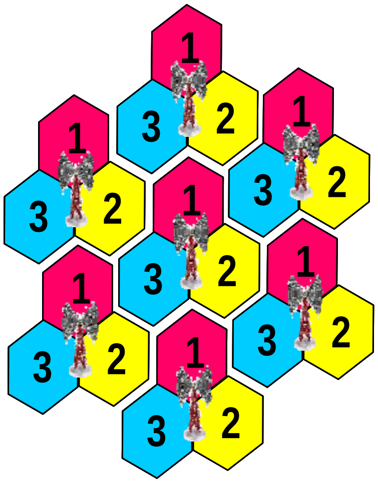


Figure 5: Frequency Reuse (N=3)

This means three different frequencies are used per site. More spectrum is used for N=3 frequency reuse (3X) than an N=1 configuration. Therefore the spectral efficiency (# of bits per bandwidth in Hz) is worse in an N=3 than a N=1 reuse pattern. Figure 6 illustrates the N=3 frequency reuse concept with more than one cell site[29,30].



| Reuse | Interference margin | |
|---|---|---|
| | DL | UL |
| N=1 | 5 dB | 4 dB |
| N=3 | 1 dB | .5 dB |

Figure 6 & Table 3: Frequency Reuse Margins

The general conclusion that can be drawn from a higher reuse pattern is there is less interference from adjacent cells occurring with an N=3 reuse over an N=1 pattern. Lower interference means a higher SNR (Signal to Noise Ratio) which translates into better modulation formats and higher throughput or coverage. Likewise, the link budget is better (lower interference margin) in a higher reuse pattern. Table 3 shows interference margins for various reuse patterns[3,4,7,8]. A 4dB difference between an N=3 and N=1 reuse for WiMAX OFDMA environments is typical.

CDMA technologies can use a N=1 frequency reuse because channels are separated by codes not frequencies. OFDMA systems, such as WiMAX, have that luxury only at a capacity and coverage cost. OFDMA can approach N=1 frequency reuse but still must combat adjacent cell or sector interference.

Fortunately, sub-channelization techniques in WiMAX OFDMA systems can combat N=1 frequency reuse interference. In reality, all WiMAX vendors recommend the use of frequency reuse patterns greater N=1 as this provides the optimum trade-off of spectral efficiency, capacity & coverage. Again, a word of caution is needed here when looking at WiMAX spectral efficiency (bits/Hz) claims as N=1 frequency reuse assumptions are always used. Although N=1 is technically feasible the only way to get that reuse pattern and still attain claimed capacities (average sector throughputs) is to reduce the cell site radius to a very small distance. The frequency reuse, spectral efficiency, coverage tradeoff just described is a perfect illustration of the "zero sum" nature of wireless technology when it comes to capacity versus cell range versus spectrum utilization.

## Sub-Channelization Techniques

One of the most enabling core technologies in the WiMAX specification is the concept of sub-channelization. Dividing the overall channel (e.g.- a 10 MHz channel bandwidth) into sub-channels used only by certain subscribers on the uplink improves overall cell range and uplink capacities tremendously. OFDMA sub-carriers (shown as arrows in Figure 7) are grouped to form Sub-Channels.

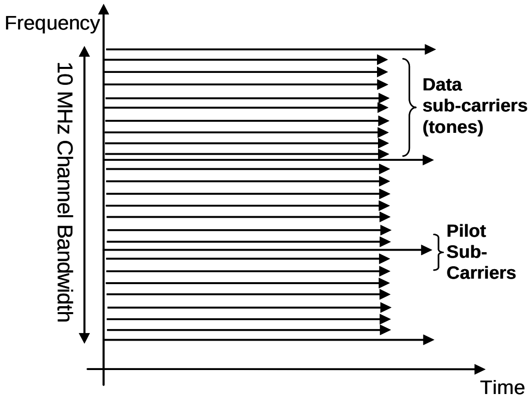


Figure 7: WiMax OFDMA Sub-Channelization

High data rates are attained in OFDMA systems because the information rate is transmitted in parallel over a large number of sub-carriers (1,024 in a 10 MHz channel)[1,8,11]. For instance, the CPE's data stream is divided into several parallel streams of reduced data rates and each sub-stream is modulated and transmitted on separate orthogonal (unique) subcarriers. The high level diagram in Figure 8 illustrates this concept for the transmit portion of a CPE device.

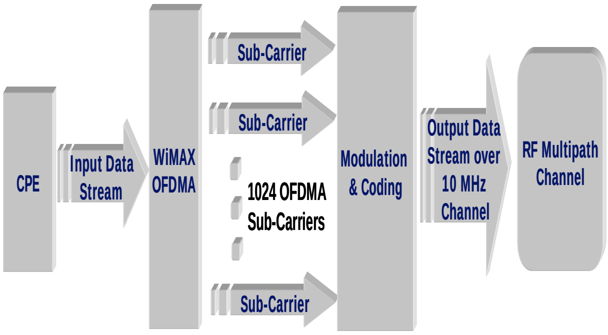
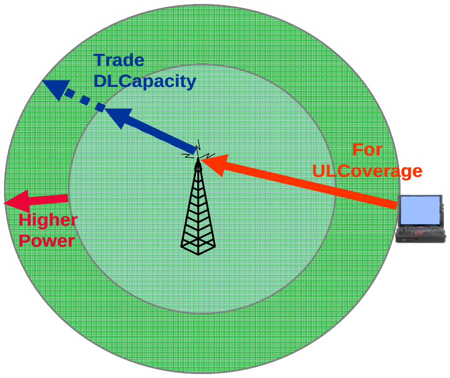


Figure 8: OFDMA Subcarriers

A key enabler in the WIMAX OFDMA system is the MAP (Media Access Protocol) scheduler that allocates bits across various time slots and sub carriers (frequencies). In effect, this function is a very smart control channel that obtains feedback on the channel quality for each user and then tightly schedules and packs user traffic in the optimum time and frequency[2,3].

## Channel Bandwidth and TDD (Time Division Duplex)

TDD (Time Division Duplex) technologies utilize the same slice of spectrum for both uplink and downlink communication. A core advantage of TDD is the ability to allocate more of the available capacity to the downlink than the uplink which is particularly useful for asymmetrical data traffic. An overall high channel capacity can be obtained if a very high DL/UL ratio is assumed. In fact, most WiMAX peak or theoretical capacity claims assume the highest 3:1 ratio as this optimizes for capacity. Unfortunately, there is a coverage penalty for choosing such a asymmetrical ratio. Figure 9 illustrates the trade-off possibilities of TDD spectrum and WiMAX technology. An operator can choose to maximize for coverage and provide less cell sites by adjusting the DL/UL ratio[16,17].



.Figure 9: Dl/UL Trade-off's

Configuring a 1:1 DL/UL ratio gives 1.5 dB of link budget improvement for the uplink[10] over a 2:1 ratio. The higher power from the Base Station makes up the loss in DL link budget. Figure 10 illustrates this phenomenon when a coverage versus capacity calculation is performed[5]. For a new entrant operator without customers it makes sense to initially deploy the network with a coverage optimized TDD DL/UL ratio of 1:1 and incur the capacity "hit" and then change to a more capacity optimized DL/UL ratio later on. Table 4 shows the DL and UL capacities for each of the possible TDD ratios [9,10].

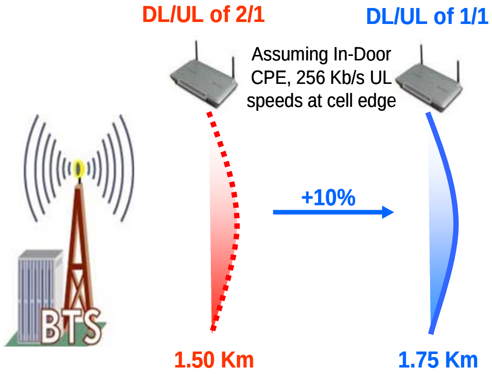


Figure 10: DL/UL Representation

**Average Sector Throughput (Mb/s)**

| TDD Ratio (DL/UL) | Downlink Capacity | Uplink Capacity |
|---|---|---|
| **1/1** | 5.57 | 4.76 |
| **2/1** | 7.28 | 3.30 |
| **3/1** | 8.25 | 1.99 |

5 MHz Channel Bandwidth, with 2x2 MIMO

Table 4: DL/UL Ratio's

The technical details on why this effect occurs goes back to the concept of sub-channelization and the subcarriers in OFDMA technology[6,8,11]. The UL coverage gain in going from a 2:1 ratio to a 1:1 TDD Ratio (DL/UL) occurs because with a 1:1 ratio, there are more uplink symbols (time slots) available for a subscriber data rate (256 Kb/s). Specifically, 21 UL symbols versus 15 UL symbols. The fewer tones (data sub-carriers) needed to be allocated in the uplink results in a reduced uplink signal bandwidth (670 KHz versus 930 KHz). The same power across a smaller bandwidth results in an improved sensitivity. This means the uplink coverage improves while maintaining the subscribers uplink data rate of 256 Kb/s. The frequency and time domain representation in Figure 11 illustrates this concept[9,12,13].

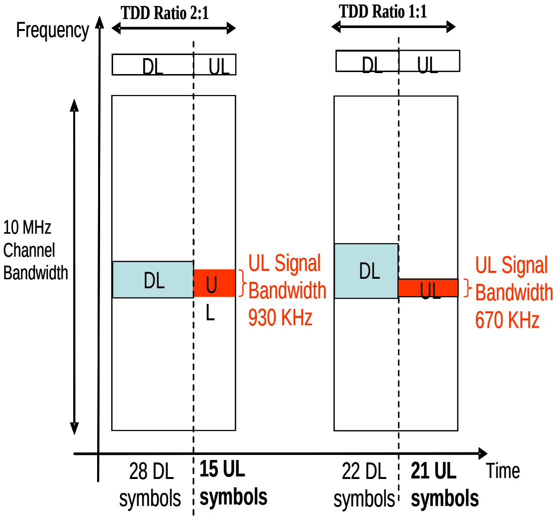


Figure 11: Details of DL/UL Trade-off's

## Channel Bandwidth Selection

WiMAX is unique from previous wireless standards in that it allows for flexible or scalable channel bandwidth choices by the operator while maintaining the same interoperable specification. The ability to have the same CPE operate in either a 5 MHz or 10 MHz channel is extremely valuable as the network can grow in capacity without adding equipment or changing out CPE. Although the RF front-end hardware of current WiMAX CPE and Base Stations are not capable of handling 20 MHz channels the specification is capable of growing to a 20 MHz channel size. Interestingly the 4G cellular specifications being created (called LTE for Long Term Evolution) envision a 20 MHz OFDMA channel [1,3.]

A non-intuitive concept of OFDMA and WiMAX that is foreign to the world of CDMA is the concept of being able to increase capacity by increasing the channel bandwidth without losing coverage. Increasing from a 5 to 10 MHz channel adds considerable network capacity to a sector (cell). Either greater speeds can be provided to the same number of subscribers or more subscribers can be served at the same product speeds. Increasing from 5 to 10 MHz bandwidth causes a 3 dB downlink loss in the link budget[3] as shown in Figure 12. If the full channel is used then less BTS power is applied across the larger 10 MHz channel. There is no loss in the link budget for the up link because of sub-channelization. The bandwidth of each sub-channel is the same regardless of the total channel bandwidth therefore the same amount of power from the CPE is applied across the uplink bandwidth. Since the uplink is usually the most limiting item for coverage there is no reduction in cell range when WiMax capacity is increased (as shown in Table 5) by going from 5 to 10 MHz. In the rare situation that the design was downlink limited the network would lose coverage when going from a 5 MHz to 10 MHz channel bandwidth[11,12,30,31].

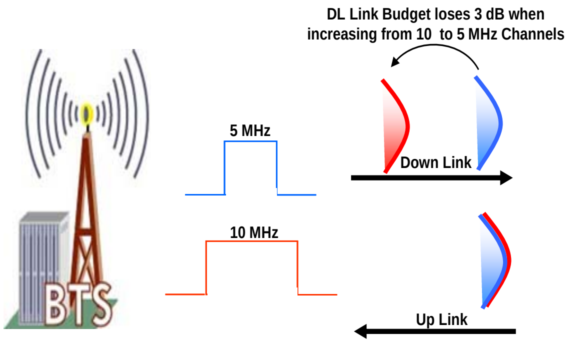


Figure 12: Channel Size Trade-off's

**Throughput improves when increasing channel size from 5 to 10**

**Average Sector Throughput (Mb/s)**

| Channel Bandwidth | Downlink Capacity | Uplink Capacity |
|---|---|---|
| **5 MHz** | 7.28 | 3.30 |
| **10 MHz** | 13.81 | 6.71 |

2/1 DL/UL TDD ratio with 2x2 MIMO

Table 5: Channel Size Capacity

## Advanced Antenna Technology

Because WiMAX utilizes TDD and OFDMA technologies it has some inherent advantages when it comes to smart antenna technology. In TDD operation the same RF channel is used for both transmit and receive so the RF channel link conditions at any point in time are known. This allows for fast, closed-loop type adjustments by the base station and CPE that can increase performance. Additionally, OFDMA is not as susceptible to frequency selective fading as other technologies, which plays to the strengths of advanced antenna techniques[6]. In the specification WiMAX has many smart antenna technologies incorporated into the standard as options. There are so many options and unspecified vendor implementations that it adds a level of complexity to the standard that will surely cause many interoperability issues with CPE devices. If the interworking aspects can be resolved over time some extremely beneficial capacity and coverage gains will be reaped with these technologies. There are various technologies that focus on either improving cell site range such as Adaptive Antenna Systems (AAS or Beamforming) or increasing capacity such as MIMO (Multiple Input Multiple Output) Spatial Multiplexing (SM) technology.

Increasing capacity is the main benefit of MIMO. MIMO can improve capacity by transmitting parallel data streams using multiple antennas at both transmitter & receiver. For example, if there are two transmit antennas then each will carry ½ of the total data in the same spectrum at the same time. Likewise the receive antennas demodulate and combine in the same way. Therefore twice the capacity is possible. MIMO uses multipath to its advantage and works best in urban environment where the signals transmitted by the antennas bounce off buildings and take many paths before they reach the multiple receive antennas. If the received signals are uncorrelated they can be combined in many ways (e.g.- Maximum Ratio Combining or MRC) to increase performance. If two transmit antennas are at the BTS and two receive antennas at the CPE it is called 2x2 MIMO in the downlink[7,14].

MIMO capable CPE will have two receive antennas and one transmit antenna while base stations will initially have two transmit and two receive antennas evolving to 4 transmit and receive antennas on the base station side. If the CPE and Base Station have only one transmit and one receive antenna then it is called a SIMO or SISO (Single Input Multiple Output or Single Input Single Output) configuration. It will take many years to get CPE prices low enough to accommodate two transmit antennas on the uplink from the CPE.

WiMAX will be using mainly 1x2 and 1x4 MIMO capabilities on the uplink. MIMO transmit diversity puts the extra antenna at the base station instead of the CPE in order to keep the subscriber unit costs low. Figure 13 illustrates the typical WiMAX 2x2 MIMO configuration downlink and 1x2 MIMO uplink.

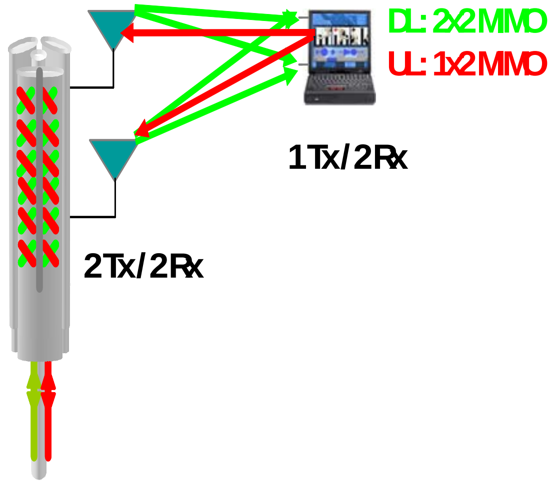


Figure 13: 2x2 MIMO

Adaptive Antenna Systems (AAS) or Beamforming improve link budget & reduce interference. Several antennas (at the base station only) form large directional arrays with a narrow beam width. The higher antenna gain improves range and the narrow beam reduces interference to and from other cells. In the same way signals are added to improve gains, unwanted signals or interference can be subtracted to create a null directed towards interferers and thereby reduce interference. Finally, Adaptive MIMO Switching or AMS is the technology that adjusts the downlink to choose the best advanced antenna option for the current RF conditions. For instance AMS can dynamically switch from a MIMO capability to an AAS mode dynamically[6]. Because of the complexity of AAS, MIMO and AMS these capabilities will be proven out in during later phases of the WiMAX industry certification and interoperability testing.

WiMAX Capacity Claims

The WiMAX Forum and other vocal proponents of WiMAX technology typically state peak data rate capacities. Although interesting numbers these claims are very theoretical as they represent the peak rate a single user operating in perfect conditions can obtain at the highest modulation and coding rate. In addition, MAC (Media Access Control) layer overheads for functions such as resource allocation and access controls are not taken out of the number. Overall, stating peak rates and comparing them with other wireless and fixed technologies is not very useful. Table 6[1,5,23,24,26] lists the aggressive WiMAX Forum capacities on the far right column for a 10 MHz channel, 3:1 DL/UL ratio and N=1 frequency reuse configuration. Peak, average sector throughput and spectral efficiency numbers are listed for downlink and uplink in SIMO and MIMO environments.

The middle column illustrates much more realistic assumptions that reflect the real world mobile wireless environment of today. As an example the 14.1 Mb/s MIMO downlink claim by the WiMAX forum versus the 9.0 Mb/s industry average is a 35% reduction in capacity due to changing assumptions.

| **WIMAX Capacity**<br>N=1 reuse, 10 MHz BW, 3:1 DL/UL ratio | | **WiMAX Industry**<br>More Conservative Assumptions | **WiMAX Forum**<br>Aggressive Assumptions |
|---|---|---|---|
| Downlink Peak Data Rate (Mbps)[1] | SIMO | 23 | 23 |
| | MIMO | 46 | 46 |
| Uplink Peak Data Rate (Mbps) | SIMO | 3.5 | 4.03 |
| | MIMO | 3.5 | 4.03 |
| Downlink Avg. Sector Throughput (Mbps) | SIMO | 7.5 | 8.8 |
| | MIMO | 9.0 | 14.1 |
| Uplink Avg. Sector Throughput (Mbps) | SIMO | 1.1 | 1.38 |
| | MIMO | 1.5 | 2.19 |
| Downlink Spectral Efficiency (b/Hz) | SIMO | 1.0 | 1.21 |
| | MIMO | 1.2 | 1.93 |
| Uplink Spectral Efficiency (b/Hz) | SIMO | .40 | .55 |
| | MIMO | .60 | .88 |

Table 6: Capacity Comparison

A large contributor to higher capacity numbers is the type of traffic model used in the simulation. For example, a simulation that sends 100% large packets (called full buffer) will have a much smaller MAC layer overhead count therefore making the average throughput capacities larger. If more realistic call & traffic models are used that represent a mix of large and small packet type traffic (such as web browsing) the overhead counts will be higher and result in lower net throughput capacities[27].

Additionally, a lower gain, less complex type CPE receiver is used instead of the more advanced non linear receiver assumed by the WiMAX Forum[1,2,3,8]. Certainly these advancements may be possible in future years but it is yet to be proven in commercial deployments.

Lower power base stations that provide 2 to 4 watts at the antenna were assumed because the cost and size of a high power amplifier may be difficult to implement. The WiMAX Forum is assuming a 20 watts at the ground assumption[1,8] which is a less practical RF implementation for tower configurations. The effect of higher power base station causes a larger distribution of subscribers obtaining the higher modulation rates which increases channel throughput capacity. It is very important to realize that average throughput capacity claims assume a mix of modulation modes.

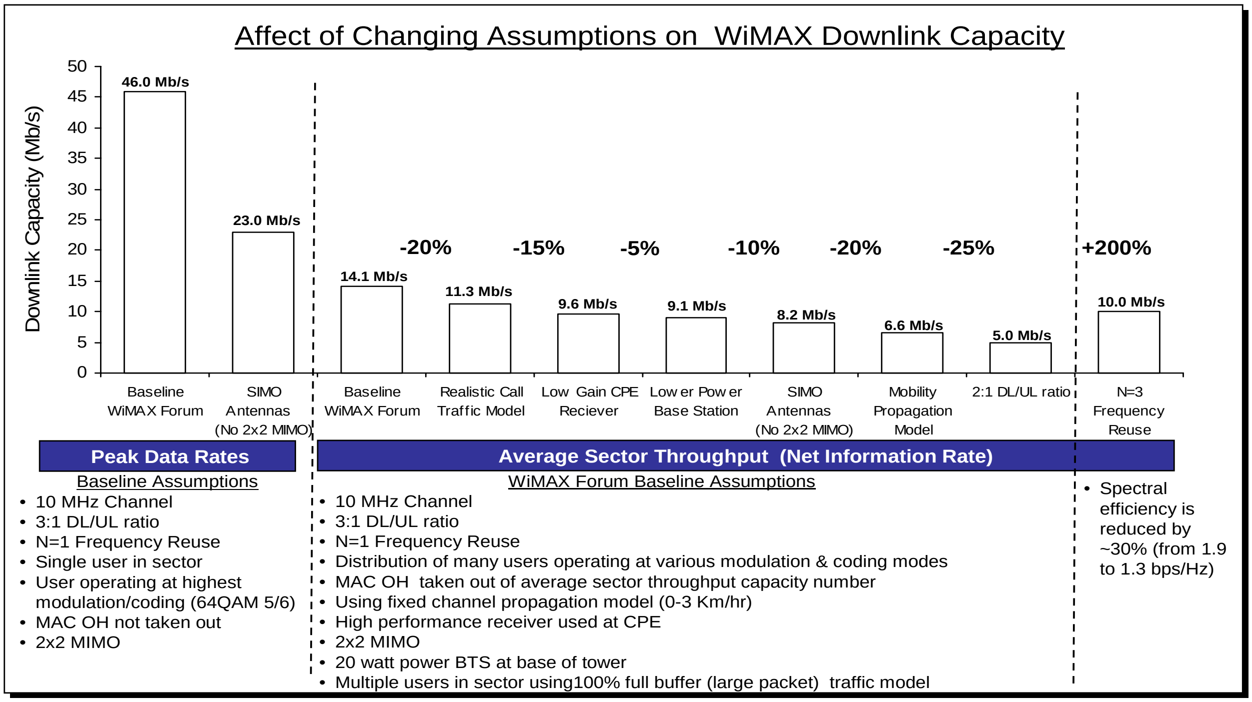


Chart 2: Impact of Assumption Changes on WiMAX Downlink Capacity

Therefore, if one simulation assumed a majority of users get 64QAM versus a second simulation which assumed a lower percentage of the users are able to obtain the highest modulation rate then a lower capacity claim will result in the second simulation.

Chart 2[3,5] gives a more detailed breakdown of how on a broad-guage percentage basis varying the key assumptions mentioned above will affect the WiMAX downlink capacity numbers. Additional assumption changes that may reduce average sector throughput from the baseline WiMax Forum numbers are; no MIMO antenna capability, using a full mobility channel propagation model and reducing the downlink to uplink ratio. WiMAX deployments with a simpler SIMO antenna scheme or perhaps rural applications where little multi-path fading exists will severely hamper the MIMO gains anticipated in computer simulations. Additionally, MIMO gains have not been fully proven when a user is operating at low modulation rates such as QPSK ½, which will be the case at cell edge operation[6]. Regarding the affect of mobility, when a 30 Km/hr channel propagation model [27,28,29] is used instead of a mix of channel models skewed towards a fixed environment, (e.g.- 0 to 3 Km/Hr used by the WiMAX Forum), the anticipated gains and ultimately sector capacities are reduced[16,17,22].

Finally, if either a more symmetrical user service or a network design optimized for coverage is required then a 2:1 DL/UL (or even 1:1) ratio will be required and impact capacity. In fact, a subscriber average uplink service level of 256 kb/s combined with a moderate VoIP mix of services and the need to optimize for range will quickly drive the operator to design to a 1:1 DL/UL ratio because of uplink requirements[25,26,27]. In these situations a large overall average sector throughput reduction for the downlink will result as shown in the chart above. This is an illustration of a classic design trade-off where uplink demands result in a downlink capacity “hit”.

On the positive side, extremely large capacity gains are possible by increasing the frequency reuse to an N=3 scheme and reduce interference as shown above. There are tradeoffs to this design decision particularly, the negative effect on spectral efficiency. Although the average sector throughput capacity increase is greater than 100% with an N=3 frequency reuse, it also means three times the spectrum is used. The spectral efficiency bits per hertz ratio is upwards of 30% lower for N=3 over N=1 because of the much larger total spectrum used[9,16,30].

One could infer from Chart 2 that a wireless operator with a large amount of contiguous spectrum and a willingness to use an N=3 frequency reuse and build a very dense network (< 1.5 Km cell radius) could offer some reasonable downlink capacities. For instance, the 9.1 Mb/s sector capacity (before MIMO was removed) could feasibly double to a 15 to 17 Mb/s average sector throughput. The operator would have to be willing to forego offering mobility (change from a mobile to a fixed propagation model) and take the uplink capacity restriction associated with a 3:1 DL/UL ratio. Although a WiMAX 20 MHz channel bandwidth is quite a bit down the road in developmental time frames, downlink sector capacities under similar assumptions might some day be able to approach 25 to 30 Mb/s downlink.

To summarize, channel bandwidth, frequency reuse and downlink to uplink ratios are very big drivers of capacity. As noted in Chart 2, MIMO antennas, base station power, CPE receive sensitivity, traffic models and channel propagation model assumptions can also skew results in a variety of directions.

In conclusion, one sees many differing capacity claims regarding WiMAX. It is feasible that all these claims are the result of very accurate and sophisticated simulations, but are not necessarily comparable, (nor realistic) as the key underlying assumptions will most likely vary.

Good historical references from other new wireless technologies, such as CDMA, exist. When CDMA was first deployed the claims on capacity and coverage were based on simulations. As commercial deployments were rolled out it became apparent that many claims were either based on faulty assumptions, never achievable, or would take many years of fine tuning and enhancements to reach. Over time (10 to 12 years), CDMA technology met its original claims and well exceeded the prior wireless technologies in capacity, coverage and performance. It is the opinion of the authors that in the same way OFDMA will surpass its predecessor wireless technology.

## WiMAX Subscriber Speeds and Services

Translating WiMAX channel capacity numbers we have just gone through into a useful understanding of the number of subscribers that can be served in a WiMAX cell site and network requires standard traffic engineering dimensioning. Paramount to this analysis is an understanding of the likely product mix, and market statistics as shown illustratively in Table 7 & 8.

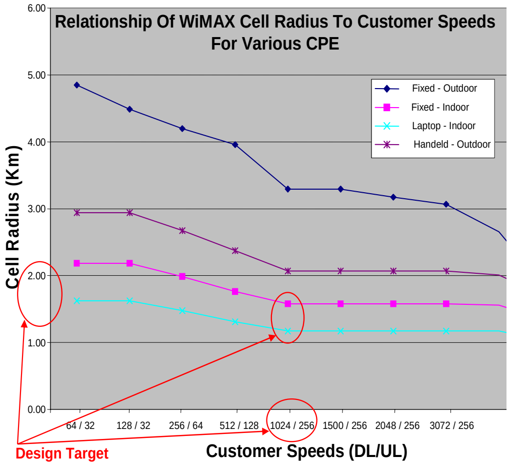
These design assumptions are needed to determine the cell site radius, number of sites and sectors and subscribers served per sector. Desired product speeds, market density and subscriber penetration levels are particularly important starting points as is the overbooking ratio (or concurrency) assumption. Since the small packet size and large header requirements of VoIP services are a burden on capacity, the VoIP penetration level is likewise necessary.

| Component | Assumption |
|---|---|
| Total Area | 443 km² |
| CPE | Indoor Gateway |
| Market Penetration | 10% |
| Total Homes in Area | 186,000 |

| Product Speeds (DL/UL) | Take Rate | Overbooking DL/UL | Avg. Kbps |
|---|---|---|---|
| 1024 / 256 | 10 | 40 / 15 | 25.0 / 6.4 |
| 512 / 128 | 40 | 25 / 15 | 20.5 / 5.1 |
| 256 / 64 | 50 | 20 / 15 | 12.8 / 3.2 |
| Wtd. Avg | 100 | | 17.1 / 4.3 |

Table 7&8: Product Mix and Market Statistics

As customer speed requirements increase the network goes from coverage limited to capacity constrained and the cell site radius is reduced to meet the demand. Chart 3 shows the effect of desired customer speeds and CPE devices on cell site radius[13,16]. In particular, speeds in the uplink offered to customers can severely impact the cell radius calculations. The key drivers of this graphical representation are the typical SNR's and associated adaptive modulation rates required at the cell edge for both DL and UL link budgets[1,5,17,18]. It is easy to see that the link budget is uplink limited for all but the outdoor CPE case as cell radius is flat until the uplink customer speeds are increased. The design target for WiMAX networks is optimum at 1 Mb/s DL & 256 Kb/s UL for indoor CPE.



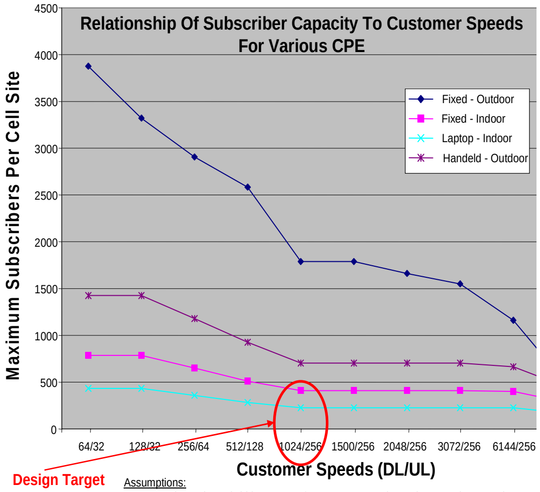
Assumptions:

- 5 MHz Channel Bandwidth, 2.3 GHz frequency, 2:1 DL/UL ratio, SISO
- 50% Data Penetration, 10% Voice (VoIP) Penetration, 729E VoIP coder
- Market Density: 420 HP's per Sq. Km.

Chart 3: Cell Radius to Customer Speeds

In the same way a graphical representation of total subscriber capacity (Chart 4) in a 3 sector cell under the same varying customer speeds and CPE device scenarios[10,11,12,16,17]. Once again, SNR's and adaptive modulation rates required at the cell edge for both downlink and uplink are the critical calculators. Given the desired customer speeds at the cell edge, customer product speeds and overbooking assumptions a target number of subscribers per cell and sector can be calculated.

The conclusion of charts 3 and 4 is that a WiMAX three sector cell site has the capacity to serve approximately 125 customers per sector (382 subs / 3) at 1 Mb/s downlink and 256 Kb/s uplink when the cell radius is 1.5 Km and the CPE is a fixed indoor residential gateway. Even the subscribers at the cell edge and operating at the lowest modulation rate will be able to receive a 1 Mb/s downlink service. As wireless system design is very probabilistic and uplink limited the modeling used here focuses on the probability of serving customers at the cell edge with 256 Kb/s uplink service.



Assumptions:

- 5 MHz Channel Bandwidth, 2.3 GHz frequency, 2:1 DL/UL ratio, SISO (no MIMO)
- 50% Data Penetration, 10% Voice (VoIP) Penetration, 729E VoIP coder
- Market Density: 420 HP's per Sq. Km.

Chart 4: Cell Capacity to Customer Speeds

It is certainly possible to serve either more customers at lower speeds or fewer customers at higher downlink speeds. Using the same assumptions a 10 MHz channel would serve between 200 and 300 customers per sector with the similar service levels.

Another interesting effect that can be concluded from these charts is if a dense network is initially built then downlink speeds from 1 to 5 Mb/s can be provided to customers. The uplink on the other hand is extremely sensitive to increasing customer service levels and uplink sensitive VoIP type services. Optimizing for uplink is possible but only with a coverage or downlink capacity negative impact.

As the penetration levels increase capacity will need to be added by either growing to a 10 MHz channel bandwidth, adding more sectors and using more frequency, upgrading to MIMO CPE (mainly in urban areas), going to a higher DL/UL ratio (at the expense of adding new sites to cover the shrinking cell radius) or cell splitting (adding new cell sites). If the above scenario wasn't already at an N=3 frequency reuse increasing the frequency reuse would also help with both range and capacity increases.

Although the average service level offered to an individual customer is sized to 1 Mb/s downlink speeds certainly an individual user could burst to the entire sector average throughput number. It is unrealistic that a single CPE will have access to a fully unoccupied channel. Most of the current version CPE hardware is rate limited to 3 to 5 Mb/s downlink speeds and it is reasonable to assume most network operators will tightly control via software the downlink and uplink speeds of wireless users.

It is appropriate to reiterate in this section that average sector throughput numbers, such as the 7 Mb/s number, is truly an average. If an individual user is in perfect channel conditions and operating in a 64QAM adaptive modulation mode then that user could in theory attain a 10 Mb/s speed even though the average throughput is 7 Mb/s. Conversely, cell edge customers may not be able to burst above their stated 1 Mb/s threshold service level.

## Coverage Considerations

The old cellular adage "coverage is king" certainly applies to WiMAX as well. It is a particularly appropriate mantra in the initial launch stages when the network has not added any customers and the operator is uncertain where and when they will appear.

There are many variables that affect coverage. Certainly the frequency and power limits set by the spectrum being used are foundational variables. Network and WiMAX specification settings we have mentioned are critical. Base station antenna height, TDD DL/UL ratios, frequency reuse, CPE type and smart antenna have the most impactful engineering design decisions. Customer propositions agreed to in the business case can vary cell site range quite extensively as well. For instance, uplink speeds available to subscribers at the cell edge, mobility versus fixed services, probability of attaining coverage at the cell edge and in-building versus outdoor coverage are key factors. All ten of the variables mentioned above are baked into the link budget along with the chosen vendors receive sensitivities and transmit power gains of their equipment to arrive at a maximum allowable path loss for both the downlink and uplink separately. Typically the uplink will be the most limiting so that path loss will be used to calculate the cell radius for a typical terrain (suburban, urban). Chart 5 provides an indication of the degree each of the technical variables affect cell radius[9,10,11,16,21,22].

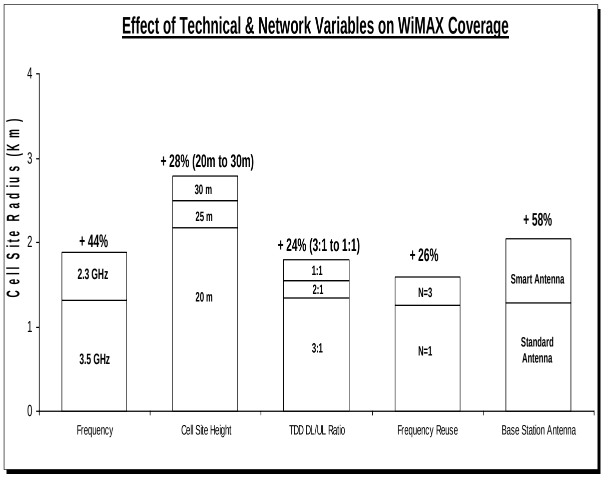


Chart 5: WiMAX Coverage Impact - Technical

Going from a 3.5 GHz to a 2.3 GHz spectrum operation can increase cell radius approximately 44%, while a 10 meter antenna height increase can easily provide a 28% improvement. As mentioned in previous sections reducing the DL/UL ratio from 3:1 to 1:1 can grow cell site range from 1.4 Km to 1.74 Km. Frequency reuse and the associated gains in link budget, by reducing interference, will increase range 26% for a 5 MHz channel at 2.5 GHz. Finally, the cell radius gains from various advanced antenna technology varies tremendously. A typical industry average can provide a 58% improvement in coverage.

Most new entrants opted for range when making decisions on the variables in Chart 5. A choice of using either 3.5 GHz or 2.3 GHz spectrum is an easy decision, as the lower frequency has better propagation. Additionally, a higher 25 meter cell site height, 1:1 UL/DL ratio and N=3 frequency reuse choices will optimize for coverage over capacity and spectral efficiency.

Chart 6 provides a similar representation of the impact on coverage for various customer impacting variables[9,10,11,16,22].

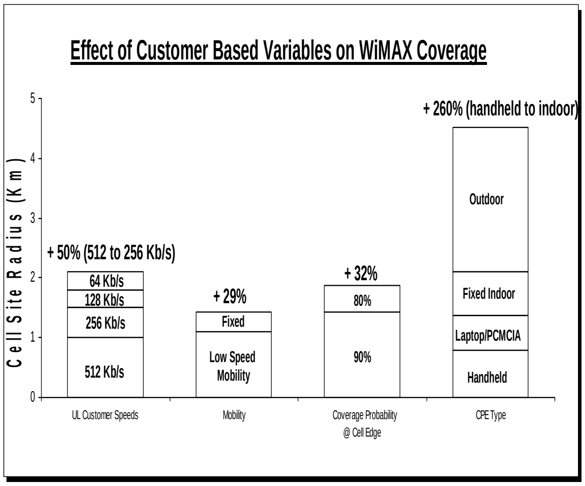


Chart 6: WiMAX Coverage Impact - Customer

Uplink speeds offered to the customer at the cell edge can cause widely fluctuating cell radius. Offering a service tier to customers at 256 Kb/s instead of 512 Kb/s can increase cell size by 50%. Cellular networks are typically designed for 64 Kb/s indoor coverage, which creates cell sizes twice the size of a 512 Kb/s designed WiMAX network.

Choosing a mobility based path loss channel propagation model introduces fading margins in the link budget that can easily add almost 30% cell reductions. Building in for overlapping cell coverage for handover also adds additional cells.

Another subtle tool used by RF engineers to increase range is to set the threshold of reliability at the cell edge to a level lower than the industry standard 90%. 80% coverage probability means 20% of all users at the cell edge will not get the minimum level of uplink service required by their service tier (e.g.-256 kb/s).

As mentioned in the very beginning of the paper the choice of CPE devices drive a wide variance in cell radius. In a 2.5 GHz design, the cell radius can grow from 1.76 Km in a suburban area for indoor CPE to 4.65 Km if the CPE is mounted outdoors. Rural terrain has an even larger effect, as a 2.32 Km indoor configuration grows to 6.55 Km with an outdoor installation.

Wireless broadband cell coverage engineering certainly contains a large amount of trial and error and has many elements of art as well as science. The end result is a somewhat variable end product offered to customers. Some subscribers will have an over- engineered service while others in difficult terrain or well inside a building made of difficult materials are unknowingly given a poor level of service.

Performance Considerations

WiMax will have many performance issues as the technology becomes developed and deployed. The most challenging aspects will be uplink related performance.

Maintaining uplink speeds and capacity at the cell edge with small packet size applications such as VoIP will be challenging. OFDMA technology performs much better than CDMA or TDMA schemes because of sub-channelization. By splitting the entire 10 MHz channel bandwidth across many subscribers each customer uses only a small subset of subcarriers with a far lower power than if it had to transmit over the entire bandwidth[3,4,5]. Even with that advantage the uplink capacity is an extremely limited resource and inefficient small packet applications such as VoIP and web browsing will take its toll on capacity. Both industry simulations and testing confirmed some major uplink capacity reductions associated with small size packet call traffic over full buffer type traffic[1,2,3,27].

Packet loss, bit error rate (BER) and packet error rates (PER) as the system is loaded and stressed will be important performance issues to watch as the technology develops and commercial networks are deployed. To the extent the complicated MAP scheduler function operates and achieves the optimal performance of the WiMAX specification will determine the BER stability.

Latency could become an issue with WiMAX as the inherent nature of TDD systems causes a longer round trip delay associated with the transmit and receive paths sharing the same swath of spectrum. Industry testing to see if actual performance matches the anticipated 50 ms specification[8,25,27] will be a closely monitored performance data point.

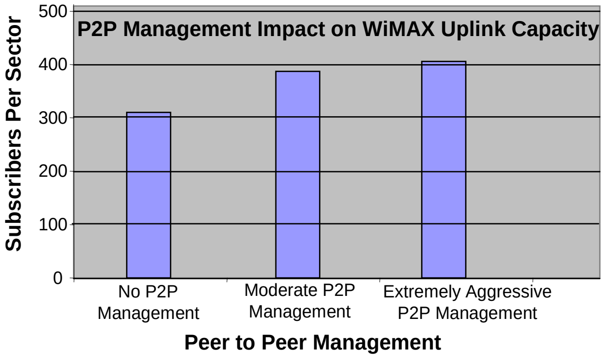
As mentioned in the previous section, the variability of product speeds (both downlink and uplink) seen by customers in the real world will be a marketing and technical challenge for operators. Predictive tools will help to proactively solve these issues but there will be no substitute for having ancillary CPE such as window mounted antennas and higher power devices to assist the operator in a variance of customer service levels.

Mobility will have its own share of performance issues and will be left for the later stages of WiMAX development. The specification focused on the MAC and PHY layer so handover and other key mobility functions will require greater definition in the core network to reduce interoperability issues.

MIMO technology has yet to be proven in the field. In particular, much is riding on the capacity gains associated with this feature. Very visible industry tests will be made to measure the gains associated with suburban and rural terrains, where minimal multipath exists, and to understand the gains with operation in lower modulation and coding modes.

Interoperability will be the key performance issue. The WiMAX specification has many options, settings and specification vagueness subject to vendor interpretation and implementation. Overall the complexity and choices comprised in the standard will invariably lead to substantial interoperability challenges. Particularly challenging is CPE to base station inter-working. Initial WiMAX chipset vendors are small companies and their coordination with many large base station vendors could unfortunately lead to design errors in the world of a specification with many choices.

Finally, it will be important to manage peer to peer traffic, especially in the uplink, as is common in the fixed broadband networks. Chart 7 depicts the capacity gains possible by closely managing a WiMAX network uplink resource. The chart indicates moderate to aggressive peer to peer management can raise overall subscriber capacity by 30%.



Assumptions:
10 MHz, 2:1 DL/UL, 3072 Kb/s DS, 256 Kb/s US, 3 Sectors,
20% data & voice penetration, 440 HP/Km2 density, 60% DL & 85% UL P2P Traffic

Chart 7: Peer to Peer Management

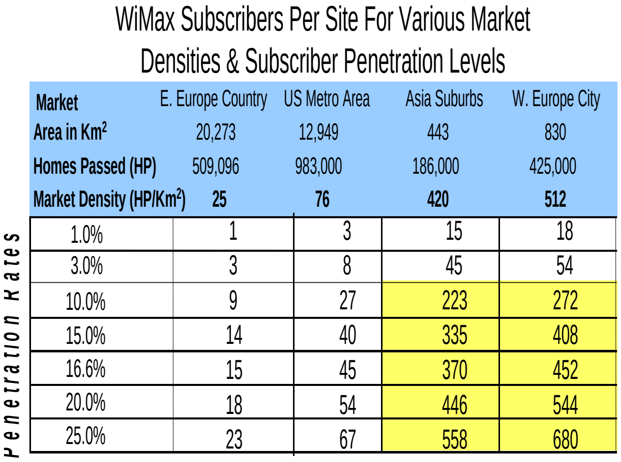
In conclusion, maintaining a high level of performance in a wireless broadband network will be challenging. Mobility and broadband have been historically two mutually exclusive design parameters. Wireless engineers are good at either optimizing a network for either mobility or separately optimizing for broadband capacity. It will be a new and challenging task to optimize for both objectives. From a cost perspective it may become an uneconomical design as there will be a requirement to design for the lowest common denominator. Mobility propagation models, cell overlap (for handover) and high speed fixed uplink services inside the home will all trigger a very dense network design.

Secondly, the fact that a wireless network has a very tight linkage between the uplink and downlink means you can not treat the engineering and design of these two segments separately. Optimizing for the downlink will surely have an effect on the uplink and vice versa. Typically, broadband networks that can separate these two dimensions are easier and more economical to manage performance levels.

Building on this second point, performance in wireless networks is all about tradeoffs. The classic analogy for wireless is that performance management is a fixed triangle, where coverage is at the apex, capacity is at one base and quality (or service levels) is at the third corner. Optimizing for one of the three corners always puts pressure and reduces the capability on one or both of the other corners. Non-negotiable trade-offs result in such an environment where there is always a losing metric with regards to performance.

## WiMAX Economics

The relationship of capacity, coverage and performance to economics is also a very important trade-off. One of the main drivers of WiMAX economics is the penetration rates and market density of the customers served. As Table 9 indicates, wireless networks with small cell radii, such as WiMAX, the linkage of market density and penetration levels to the subscribers served is critical[2,10].



WiMax Subscribers Per Site For Various Market Densities & Subscriber Penetration Levels

| | Market | E. Europe Country | US Metro Area | Asia Suburbs | W. Europe City |
|---|---|---|---|---|---|
| | Area in Km$^2$ | 20,273 | 12,949 | 443 | 830 |
| | Homes Passed (HP) | 509,096 | 983,000 | 186,000 | 425,000 |
| | Market Density (HP/Km$^2$) | 25 | 76 | 420 | 512 |
| Penetration Rates | 1.0% | 1 | 3 | 15 | 18 |
| | 3.0% | 3 | 8 | 45 | 54 |
| | 10.0% | 9 | 27 | 223 | 272 |
| | 15.0% | 14 | 40 | 335 | 408 |
| | 16.6% | 15 | 45 | 370 | 452 |
| | 20.0% | 18 | 54 | 446 | 544 |
| | 25.0% | 23 | 67 | 558 | 680 |

Table 9: Subs Per Site for Various Market Densities

Typically cities with market density greater than 400 HP/Km$^2$ such as the Western European and Asian suburbs shown in Table 9 have the best per site subscriber levels. Because WiMAX has the potential to offer very reasonable channel and subscriber capacities (compared to other wireless technologies) at short ranges, high density/high penetration conditions are very conducive to economic success.

This point is best proved by first looking at the WiMAX network cost structure. As Chart 8 illustrates the main costs associated with the upfront network build are construction related[2,19,21,23,32,33]. Typically, site acquisition, legal, civil works, tower costs, power & electrical costs are almost 50% of the total costs. This is true of most wireless networks but is amplified in a WiMAX network because the actual base station costs are reasonable. The second largest expenditure is for all the costs associated with the base station electronics. When the microwave, installation, rigging (cabling on a tower or rooftop) test & turn-up, battery back-up, environmental (A.C. or heat exchanger), cabinets and ancillary equipment are added together they account for 30% of the total costs.

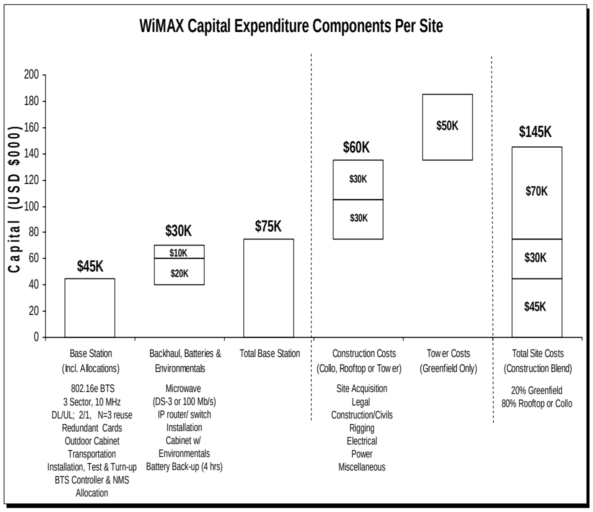


Chart 8: WiMAX Capital Cost Management

It is important to realize that the costs shown in Chart 8 can vary widely based upon the vendor's equipment and design. Certainly the type of antenna, the allocation for the base station controller function (called ASN or Access Services Network in the standard), redundant cards, spares and software feature/functionality will drive a wide range of costs.

Similarly, the construction assumptions can cause a large cost variance. Foremost in the construction economics is the degree an operator can forego the cost of building towers by obtaining collocation agreements with existing tower owners and landlords of rooftops. For the purposes of Chart 8, a 80% / 20% split of collocations to "Greenfield" new builds of towers is assumed.

One important learning we obtained in the understanding of the cost components and how to minimize expenditures was the need to minimize construction costs by obtaining a low-power, small footprint base station and its ancillary equipment. If large base station and microwave backhaul electronics are required then large shelters, air conditioning, additional power and larger battery back-up gear is needed. All these things drive higher construction costs as now cranes are needed, additional electrical work from the power company and larger concrete footings will be triggered. In the harsh environmental conditions of some rural networks the payoff of having "hardened" microwave and base station electronics is substantial in order to avoid these add-on costs.

Applying the up front Capital costs to the market density and penetration numbers discussed previously will lead to a cost per subscriber metric shown in Chart 9 [2,19,21,23,32,33]. Key to these calculations is the capacity, traffic engineering, coverage and customer performance metrics illustrated in prior sections. The resultant cost per sub is $550 where over 40% of the costs is for CPE. Constructions costs contribute 28% and the base station, microwave and ancillary equipment split the remaining 30%. The cost structure assumes a dense 1.3 Km cell radius, moderate market density (440 HP/Km$^2$) and penetration level (20%). Finally, the network was designed for 90% coverage at the cell edge with 1 Mb/s downlink and 256 Kb/s uplink service levels.

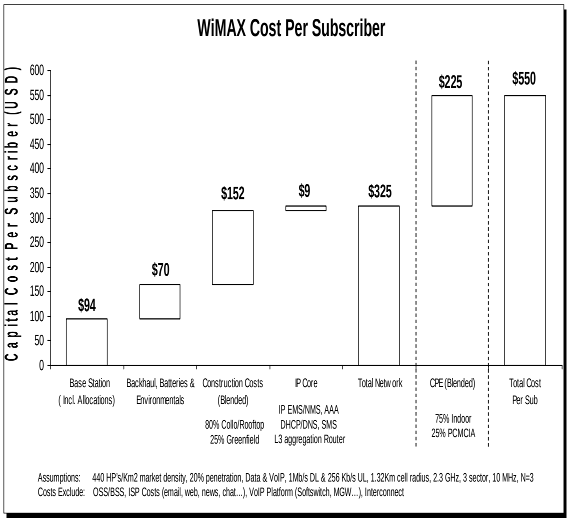
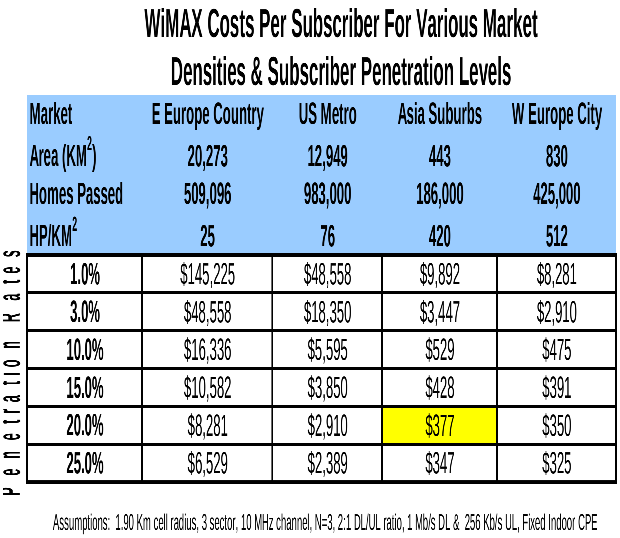


Chart 9: WiMAX Cost Per Subscriber

An interesting sensitivity to the cost per subscriber number mentioned is the effect market densities and penetration rates have on this number. Table 10 provides a variety of cost per sub numbers for the illustrative markets already discussed[2,10].



WiMAX Costs Per Subscriber For Various Market Densities & Subscriber Penetration Levels

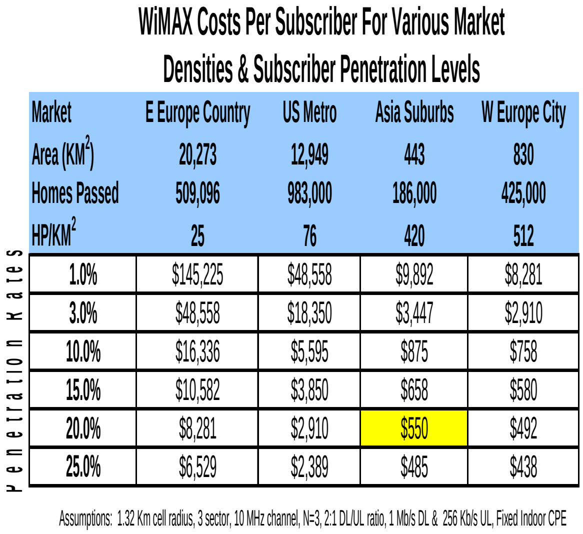
| Market | E Europe Country | US Metro | Asia Suburbs | W Europe City |
|---|---|---|---|---|
| Area (KM$^2$) | 20,273 | 12,949 | 443 | 830 |
| Homes Passed | 509,096 | 983,000 | 186,000 | 425,000 |
| HP/KM$^2$ | 25 | 76 | 420 | 512 |
| Penetration Rates | | | | |
| 1.0% | $145,225 | $48,558 | $9,892 | $8,281 |
| 3.0% | $48,558 | $18,350 | $3,447 | $2,910 |
| 10.0% | $16,336 | $5,595 | $529 | $475 |
| 15.0% | $10,582 | $3,850 | $428 | $391 |
| 20.0% | $8,281 | $2,910 | $377 | $350 |
| 25.0% | $6,529 | $2,389 | $347 | $325 |

Assumptions: 1.90 Km cell radius, 3 sector, 10 MHz channel, N=3, 2:1 DL/UL ratio, 1 Mb/s DL & 256 Kb/s UL, Fixed Indoor CPE

Table 10: Cost per Sub for Various Market Densities

Not surprisingly the cost per subscriber numbers begin to look quite beneficial for highly dense markets where operators believe they can secure many customers. For example, a Western Europe city with over 500 homes passed per square kilometer can reduce the cost per subscriber metric by 20%. .

Another spin on this sensitivity is to look at the same metric while just varying the cell radius assumption[10,12,13,16] as shown in Table 11. For instance, 128 Kb/s UL speeds with 80% cell edge probability could possibly result in a 1.9 Km cell radius. Assuming all other design assumptions are the same the cost per subscriber metric improves tremendously.



WiMAX Costs Per Subscriber For Various Market Densities & Subscriber Penetration Levels

| | Market | E Europe Country | US Metro | Asia Suburbs | W Europe City |
|---|---|---|---|---|---|
| | Area (KM$^2$) | 20,273 | 12,949 | 443 | 830 |
| | Homes Passed | 509,096 | 983,000 | 186,000 | 425,000 |
| | HP/KM$^2$ | 25 | 76 | 420 | 512 |
| Penetration Rates | 1.0% | $145,225 | $48,558 | $9,892 | $8,281 |
| | 3.0% | $48,558 | $18,350 | $3,447 | $2,910 |
| | 10.0% | $16,336 | $5,595 | $875 | $758 |
| | 15.0% | $10,582 | $3,850 | $658 | $580 |
| | 20.0% | $8,281 | $2,910 | $550 | $492 |
| | 25.0% | $6,529 | $2,389 | $485 | $438 |

Assumptions: 1.32 Km cell radius, 3 sector, 10 MHz channel, N=3, 2:1 DL/UL ratio, 1 Mb/s DL & 256 Kb/s UL, Fixed Indoor CPE

Table 11: Cost per Sub for Various Market Densities

The business case will now start to look beneficial at low penetration rates (say 10%) in the more densely populated areas. As in any sensitivity analysis there are a lot of variables to change and WiMAX certainly does not have its shortage of moving parts.

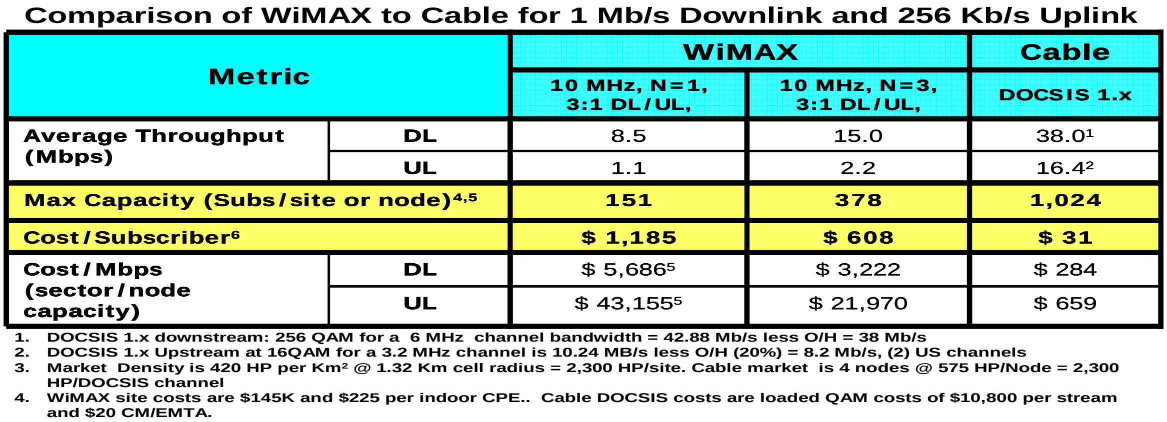
All the previous economic analysis assumed a network designed to serve an indoor high power/high gain type CPE. The costs to serve this CPE design result in $325 per subscriber of network costs and $225/sub of CPE costs. The vision of WiMAX is for customers to access a broadband network with a laptop containing an embedded WiMAX chipset and software. The economics of a laptop CPE network that obtains 1 Mb/s downlink and 256 Kb/s uplink will cause CPE costs to go down to approximately $100/sub. Unfortunately the associated network costs will rise to the $500/sub level in order to support a smaller cell radius. Therefore, in total, the costs will go up and no savings are obtained.

In conclusion, the economics of WiMAX can be compelling if the network operator has the wherewithal to spend a large sum of upfront capital to build a very dense network. This expenditure is only justifiable if there is a fairly compact addressable market and it is underserved. The definition of underserved must mean that downlink speeds of 1-3 Mb/s are not readily available from competitors and that uplink speeds of 256 Kb/s are sufficient. Additionally, underserved probably means that there is a sufficient pent up demand for mobility and portability by customers when using broadband services. And finally, these needs outweigh the inconsistent service levels a subscriber will receive because of the inherent variability of wireless technology.

## Comparison of WiMAX to Fixed Broadband

Contrasting WiMAX capacities to the well known Cable High Speed Data (HSD) capabilities of DOCSIS is a very useful exercise. Table 12 compares the capacities and economics of a DOCSIS 1.x network with two different WiMAX configurations. It should be noted that although DOCSIS 1.x was chosen most operators have deployed DOCSIS 2.0 and DOCSIS 3.0 installations are expected to start in 2008. Both advancements will provide even higher capacities to the Cable side of this comparison.

A 10 MHz WiMAX channel bandwidth operating at a 3:1 DL/UL ratio with either an N=1 and N=3 frequency reuse are used. A common customer product speed offering of 1 Mb/s downstream and 256 Kb/s upstream is assumed. Although these are low speeds for a cable network, because it appears to be the “sweet spot” for WiMAX networks, we will focus first on this scenario.



**Comparison of WiMAX to Cable for 1 Mb/s Downlink and 256 Kb/s Uplink**

| Metric | | WiMAX | | Cable |
|---|---|---|---|---|
| | | 10 MHz, N=1, 3:1 DL/UL, | 10 MHz, N=3, 3:1 DL/UL, | DOCSIS 1.x |
| Average Throughput (Mbps) | DL | 8.5 | 15.0 | 38.0[1] |
| | UL | 1.1 | 2.2 | 16.4[2] |
| **Max Capacity (Subs/site or node)[4,5]** | | **151** | **378** | **1,024** |
| **Cost/Subscriber[6]** | | **$ 1,185** | **$ 608** | **$ 31** |
| Cost/Mbps (sector/node capacity) | DL | $ 5,686[5] | $ 3,222 | $ 284 |
| | UL | $ 43,155[5] | $ 21,970 | $ 659 |

1. DOCSIS 1.x downstream: 256 QAM for a 6 MHz channel bandwidth = 42.88 Mb/s less O/H = 38 Mb/s
2. DOCSIS 1.x Upstream at 16QAM for a 3.2 MHz channel is 10.24 MB/s less O/H (20%) = 8.2 Mb/s, (2) US channels
3. Market Density is 420 HP per Km² @ 1.32 Km cell radius = 2,300 HP/site. Cable market is 4 nodes @ 575 HP/Node = 2,300 HP/DOCSIS channel
4. WiMAX site costs are $145K and $225 per indoor CPE.. Cable DOCSIS costs are loaded QAM costs of $10,800 per stream and $20 CM/EMTA.

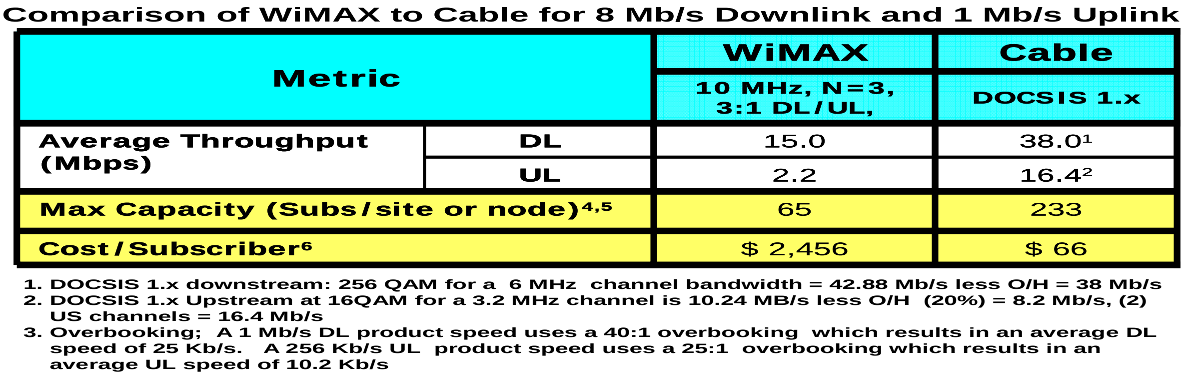
Table 12: Comparison of WiMAX to Cable

As the numbers indicate in Table 12 the DOCSIS network has 6.5 and 2.5 times the subscriber capacities of the respective WiMAX networks. These results used a common HSD capacity traffic model that assumed a DS overbooking of 40:1 resulting in an average speed of 25 Kb/s and an US overbooking of 25:1 leading to a 10.2 Kb/s average speed. The WiMAX capacity numbers appear slightly different from previous discussions in this paper because a more realistic 30% P2P traffic assumption and equivalent circuit theory tier deduction was accounted for in the calculations.

The comparative economics illustrate a similar DOCSIS advantage of $31/sub versus $608/sub for WiMAX. In addition to the larger capacity advantage for DOCSIS, the lower cost for a QAM leads to the much lower cable cost per subscriber number. Assuming WiMAX is a new entrant, a fully loaded three sector WiMAX site is compared with an embedded Cable plant. A more detailed comparison will want to delve into the impact of adding cable upgrade costs or even a new build into the analysis.

Looking at the higher capacity customer proposition of an 8 Mb/s downstream and 1 Mb/s upstream product in Table 13 clearly indicates the capacity and cost advantage of DOCSIS over wireless broadband.

DOCSIS 1.x with two 3.2 MHz upstream channels allocated results in a 3.5 times improvement in capacity over WiMAX. In the WiMAX case the upstream is clearly the limiting item on the number of subscribers serviceable. Furthermore, the cost per subscriber advantage of DOCSIS is amplified as a $66/sub cost is associated with the Cable network while $2,456/sub is needed to build a WiMAX network. The DOCSIS capacity advantage for serving higher product speeds is very apparent in this example. If a cost per Mb/s metric is calculated and compared as shown in the first chart the WiMAX network disadvantage for both the downstream and upstream is clearly identified. The upstream economic limitations of WiMAX becomes even more pronounced when the $22,000 WiMAX metric is contrasted to the $659 cost per Mb/s DOCSIS number.



**Comparison of WiMAX to Cable for 8 Mb/s Downlink and 1 Mb/s Uplink**

| Metric | | WiMAX | Cable |
|---|---|---|---|
| | | 10 MHz, N=3, 3:1 DL/UL, | DOCSIS 1.x |
| Average Throughput (Mbps) | DL | 15.0 | 38.0[1] |
| | UL | 2.2 | 16.4[2] |
| Max Capacity (Subs/site or node)[4,5] | | 65 | 233 |
| Cost/Subscriber[6] | | $ 2,456 | $ 66 |

1. DOCSIS 1.x downstream: 256 QAM for a 6 MHz channel bandwidth = 42.88 Mb/s less O/H = 38 Mb/s
2. DOCSIS 1.x Upstream at 16QAM for a 3.2 MHz channel is 10.24 MB/s less O/H (20%) = 8.2 Mb/s, (2) US channels = 16.4 Mb/s
3. Overbooking; A 1 Mb/s DL product speed uses a 40:1 overbooking which results in an average DL speed of 25 Kb/s. A 256 Kb/s UL product speed uses a 25:1 overbooking which results in an average UL speed of 10.2 Kb/s

Table 13: Comparison of WiMAX to Cable

Overall, it is clear from this analysis that a fixed broadband solution such as DOCSIS has the ability to more easily scale US and DS capacity to meet the growing capacity needs of customers. In fairness to WiMAX capabilities, the costs and capacities shown here do offer a nomadic and portable capability not possible with fixed networks.

## Conclusions

The WiMAX standard extends the state of the art for wireless technology. It lays the foundation for a true 4th generation wireless technology that offers both a broadband and mobility customer proposition. In any new technology the advantages are relative to what it is being compared to. As a comparison to existing wireless technology, WiMAX will exceed the capacity, throughput and economic capabilities of its CDMA predecessors while still offering full portability and mobility.

It is the opinion of the authors that the technical and economic viability of WiMAX versus fixed line alternatives like DOCSIS is very targeted. WiMAX seems to make sense in environments where there is a lack of viable competitor offerings of customer downlink speeds greater than 3 Mb/s and uplink speeds better than 256 Kb/s. Additionally, minimum market density environments and realistic target penetration rates must be attainable for the economics to work.

Overall, as a competitor to fixed line services, like Cable's HSD offerings, WiMAX does not have the speeds or economics to compete against an in place fixed line broadband technology. The uplink challenges, variability of service levels to customers and lack of downlink speeds for video will constrain its competitiveness against fixed line alternatives,

Once the entire WiMAX ecosystem is built out and technology, such as MIMO, advances it may offer a wireless displacement option for certain customers. If, however, speeds and consumption continue to grow, as they have historically, wireless may have a difficult time meeting the home requirements of consumers.

Finally, for fixed line operators, the use of WiMAX could serve as an extremely complementary lower speed tier that augments the high speed fixed capabilities currently available. The ability of an MSO to provide customers with a portable broadband service that also provides mobility can be a very compelling proposition.

## REFERENCES

1. WiMAX Forum web site, *"Technical Overview and Performance Evaluation"*, March 2006 http://wimaxforum.org

2. WiMAX Forum web site, “*The Business Case for Fixed Access in Emerging Markets”*, March 2005 http://wimaxforum.org
3. WiMAX Forum web site, “*WiMAX End to End Network System Architecture”*, April 2005 http://wimaxforum.org
4. WiMAX Forum web site, “Considerations for Fixed Wireless in 2.5, 3.5 GHz”, June, 2005 http://wimaxforum.org
5. WiMAX Forum web site, “*Mobile WiMAX a Comparative Analysis”*, March 2006 http://wimaxforum.org
6. Jeffrey Andrews, "Orthogonal Frequency Division Multiple Access" *IEEE Journal*, July 29, 2006
7. Diversity, Interference& Spatial Multiplexing in MIMO Mobile WiMAX”, E. Biglieri, 2007
8. IEEE Standard 802.16e – 2005 and Amendment & Corrigendum to 802.16-2004, IEEEE LAN/MAN Standards Committee, February 28, 2006
9. Alcatel Technology White Paper, “Universal WiMAX”, September14, 2006
10. Alcatel Technology White Paper, “Extending the DSL Penetration to Mobile WiMAX”, 2Q, 2005
11. Alcatel Technology White Paper, “4G”, September 14, 2006
12. Alcatel Technology White Paper, “WiMAX, Making Ubquitous High Speed Data a Reality”, June 28, 2004
13. Alcatel Technology White Paper, “Universal Broadband Access Going Wireless and Mobile”, 2Q05
14. Nortel White Paper, “MIMO, Key Technology Choice in Deploying WiMAX”, October 10, 2006
15. Nortel WiMAX Solution Brief, March, 2007
16. Nortel White Paper, “Considerations for Deploying WiMAX at various Frequencies”, March, 2006
17. Nortel Technical Journal, “Wireless Access”, July, 2006
18. Intel Technical Journal, “SOFDMA in 802.16e”, Vol 8, Issue 3, 2004
19. Intel, “The Business Case for WiMAX”, February 2007
20. Intel Technical Journal, “Multiple Antenna Technology in WiMAX Systems”, Issue 3, 2004
21. HSBC Global WiMAX, “WiMAX Anti-Climax”, Richard Dineen, January 3, 2006
22. Rethink Research Associates, “Wireless Watch” “February 19, 2007
23. Rethink Research Associates, “WiMAX Business Models; How to Make Money in WiMAX” “October 17, 2005
24. Marvedis, “WiMAX and Broadband Worldwide Market Trends and Analysis” “September, 2006
25. Navini White Paper, “VoIP over WiMAX” “November, 2006
26. SRTelecom White Paper, “WiMAX Capacity” “November, 2006
27. Ericsson White Paper, “Basic Concepts of HSDPA” February, 2007
28. Qualcomm White Paper, “The Economics of Mobile Wireless Data” August, 2006
29. Qualcomm White Paper, “A Wireless Primer”, Jeff Belk, September, 2005
30. Alvarion White Paper, “ Understanding the Radio Technology of Mobile WiMAX” 2006
31. Alvarion White Paper, “ Comparing Mobile WiMAX 3G and Beyond” 2006
32. CableLabs, “Cable Optimized Mobile Wireless Broadband”, November 4, 2005
33. CableLabs, “Cost Comparison of Mobile Wireless Technologies”, June, 2005

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