MOVING TO AN ALL-IP, ALL-AVC HFC ACCESS: OPPORTUNITIES AND CHALLENGES:

Carol Ansley and Mark Bugajski, ARRIS Group

Paper Outline:

1. A typical MSO HFC downstream usage – "staying the course, or not"?
   a. "Staying the course" path option
   b. Flash-transition to all-AVC in HFC
   c. Bandwidth usage comparison

2. VBR H.264 AVC over DOCSIS3.0 option benefits
   a. VBR SPTS UDP streaming
   b. HTTP FMP4 streaming

3. Demarcation Gateway (DGW)
   a. Architecture
   b. Support for legacy home CPE
   c. Building blocks

4. Conclusions

Stay the Course or Jump Ahead?

Multiple System Operators (MSOs) have evolved over the last ten years from delivering only video services, originally in analog, to providing a triple play of video, voice and data. Along the way their capital investments have been directed at each service in turn, usually success-based. But this paradigm is starting to show its age, and may become unsustainable if recent trends in HSD growth and increasing competition from Telco and satellite sources continues.

A typical HFC plant today provides some analog services, a large digital tier of broadcast linear programming, and an expanding tier of narrowcast video and data services. Voice services take a small portion of upstream HSD bandwidth, but are generally negligible in the downstream direction. The remainder of the upstream bandwidth is used for HSD and STB return signals.

With the ever increasing pressure to provide more HD program content, particularly from the satellite providers, MSOs must find ways to increase the number and quality of video programming options for their subscribers. There is also substantial pressure on the HSD capability over the HFC plant to match FTTx deployments with their high downstream and upstream capabilities.

To increase video capacity many operators have deployed, or are considering deploying, switched digital video deployments, as well as decreasing the amount of spectrum allocated to inefficient analog services. Both of these approaches present public relations obstacles and require capital outlays. For example, to deploy switched digital services one must install a control plane to manage the switching of the channels and deploy narrowcast QAMs for each service group, as well as commit operational resources to carefully monitor and balance the SDV program loads over the channels committed to the narrowcast SDV services. To reduce analog services would seem to be a simpler option technically, but it has business implications due to must-carry agreements, franchise agreements, and other business arrangements that may prevent a channel from being moved to a digital tier or removed outright. In the aftermath of the digital transition, MSOs have acquired some subscribers who specifically wanted to be able to continue to use their analog televisions. To accommodate these implications, MSOs have sometimes chosen to also deploy DTAs as a part of the analog removal, which requires capital investment.

Another trend that is having a profound effect on MSO capital and operational investments is the spread of personal video recording technology. Recently it was reported that PVR technology has spread to greater than 40% of North American households. The consumer has become accustomed to being able to time shift their consumption of video content. To accommodate these expectations, the MSOs have had to deploy advanced settop boxes that can provide these services, with the requisite increase in their subscriber capital costs. Another option that championed by at least one operator is moving the PVR functionality into the headend servers. This option decreases the capital outlay necessary for widespread PVR service offerings, but does have a concomitant increase in bandwidth usage as subscribers move from the broadcast program stream over to the unicast streams.

One last area that is expected to provide increasing pressure on the MSO's plant is the upstream speeds required to remain competitive in HSD service. DSL and FTTx services are typically symmetric, while the present cable plant is profoundly asymmetric. To address the asymmetry will require either extensions to the high end of the cable plant to provide increased upstream service, or changes to the downstream plant range, converting some portion of it to upstream use. Either of these changes has implications for the customer premise equipment, and the basic hardware of the HFC plant. The MSOs have not identified a preferred solution, but market pressures are expected to eventually force the MSOs to choose one of these alternatives.

Recent FCC proposals may also lead to fundamental changes in the home network architecture for the delivery of video and data services. The recently released Broadband Plan calls for the industry to rapidly adopt a Gateway architecture with the goal of providing a more open marketplace for CPE equipment. This proposal has a long way to go before it is adopted, but it would be best for MSOs to determine how they might cost effectively comply with that new requirement if it survives the legislation and oversight process.

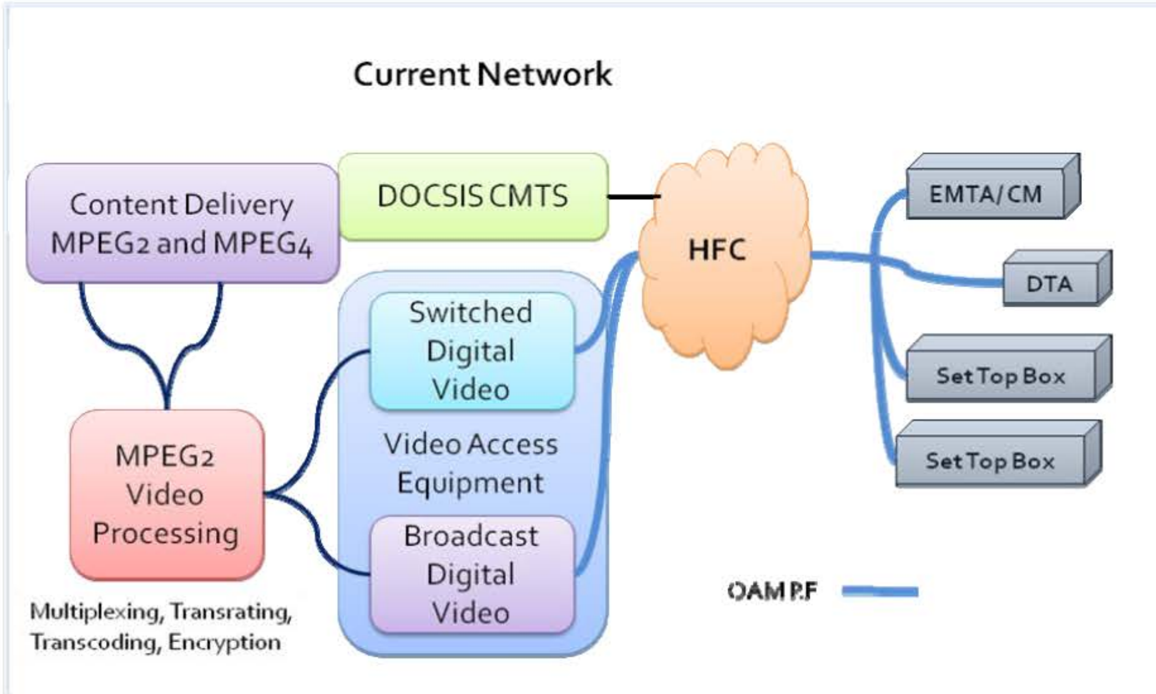
This paper compares the costs of continuing these activities with a seemingly radical alternative and rapid movement to an IP to MPEG gateway called the Demarcation Gateway to allow the legacy equipment to continue without replacement and which separates the network evolution from that within the subscriber's residence. It will compare the head-end and access changes necessary to provide the market demands of advanced video services and increased HSD bandwidth.

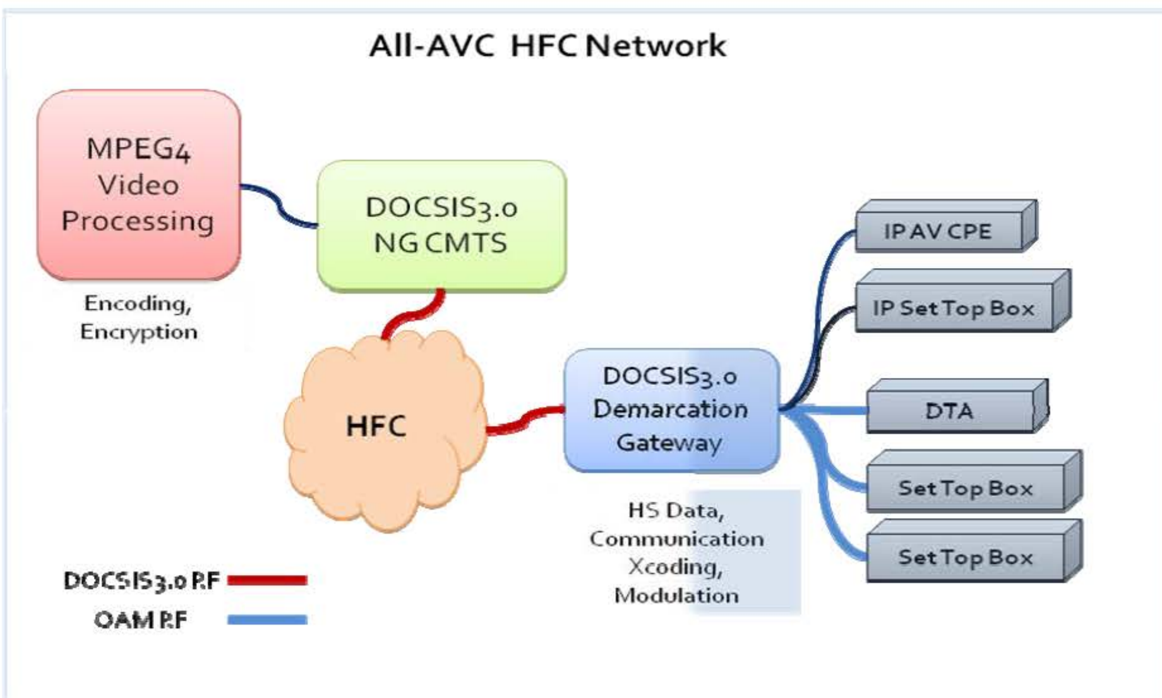
There have been proposals made periodically to move to a deployment model that makes use of a gateway to interface between the outside plant network and the subscriber's premises. The Demarcation Gateway proposed in this paper can disconnect the outside plant from the inside network, offering several advantages. The MSO network can evolve separately from the Home network. RF ingress from sources within the home can be blocked from outside network, improving network performance and reliability. Recent regulatory pronouncements also indicate that, within the US at least, regulators may impose a Gateway-based architecture on the industry. Assuming for the moment that this initiative has a fair chance of success, this paper analyzes some advantages to be had in accepting this direction, and using it to reach an advantageous long-term position.

The overall video industry has been developing and utilizing new compression technologies beyond the traditional MPEG2 encoding that is most common in MSO plants today. The H.264 AVC codec offers substantial bandwidth reduction over the MPEG2 codec in use today, but the millions of legacy set top boxes deployed already have hindered the widespread deployment of AVC within the CATV industry. AVC has been used to great effect within the satellite industry so decoders, encoders and robust CPE silicon solutions are readily available at reasonable cost points, but it has been the upfront cost of swapping out the CPE devices that made the business case for that swap out unacceptable within today's financial markets, even though there are many technical advantages.

The Demarcation Gateway architecture eliminates the legacy swap-out concern while allowing the MSO network to take advantage of the latest industry developments.

The network block diagram below shows the new architecture. The transport mechanism over the HFC transitions from MPEG2 over QAMs to IP Video over DOCSIS. Also there are advantages to transitioning from the constant bit rate (CBR) encoding commonly used for SDV deployments now to a variable bit rate encoding scheme that the wider data transmission pipes over DOCSIS 3.0 have enabled.





The new architecture presents an opportunity to take advantage of recent technological improvements (technology update) in several key areas without wholesale stranding of the current Capex investments. It will be possible to:

- Gradually and seamlessly transition to all-IP transmission in the access for full service and equipment convergence

- Use VBR formatted H.264 AVC encoded streams for significant improvements in content transmission efficiency
- Deploy highly integrated SoICs that enable development of a Demarcation Gateway
- Instantiate the viewing device specific high quality, while cost effective transcoding inside the DGW

We will next expand on some of these technical advantages ….

VARIABLE BIT RATE H.264 AVC CONTENT DELIVERY OVER IP ACCESS.

As stated earlier, the H.264 AVC encoding brings about significant expansion of the compression toolkit and allows the HD 1080i content to be encoded for streaming at average bit rates ranging from 5 to 8 mbps. The average bit rate is dependent on how the encoders (transcoders) are provisioned and what type of content is being encoded. As with MPEG2, content with static screens and slow-action (talking-heads) can be encoded at the lowest bit rate and sports programming requiring the highest bit rate for the same format.

ARRIS has performed very extensive studies on methods of unicast (VoD) streaming using Variable Bit Rate (VBR) SPTS formats that represent a departure from the traditional Constant Bit Rate (CBR) method. The VBR transport streams are compliant with the MPEG specs but have not to-date been used in the HFC networks because of the complexities and challenges associated with implementation of streaming servers and the downstream spectrum management. The 6MHz channelization of the HFC RF spectrum creates 40mbps delivery "pipes" that are too narrow to enable "self-averaging", the natural (not forced) statistical multiplexing of peaks and valleys of VBR streams.

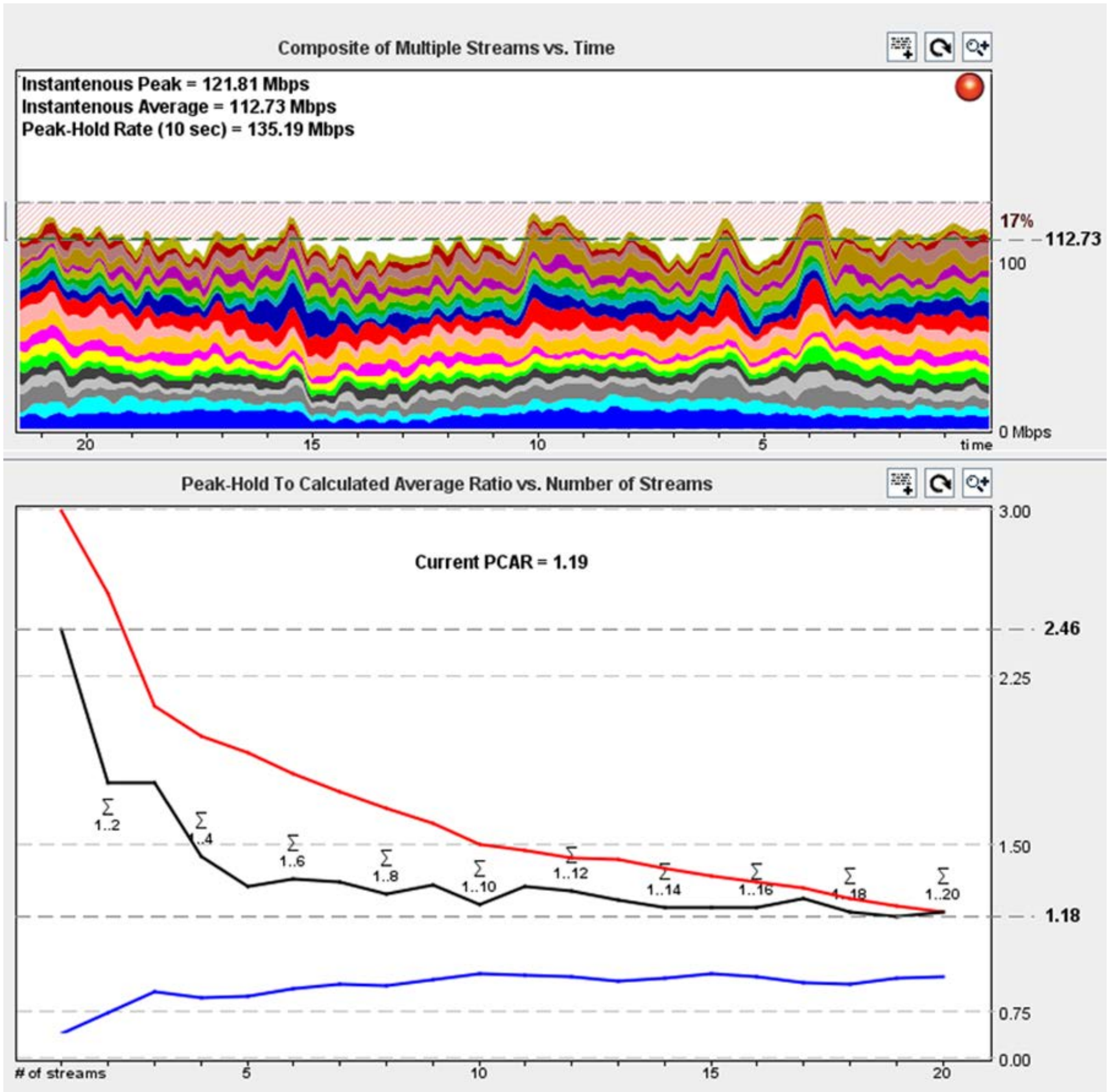
The channel bonding feature/functionality of DOCSIS3.0 breaks the 40mbps limit and allows for establishing delivery pipes that are 4, 8, 16 or more times wider that a single 6MHz channel. We have developed sophisticated research and simulation tools that allow us to simultaneously stream hundreds of combinations of continuous and fragmented VBR streams. We studied combinations of the most variable content (movie trailers and commercials) as well as several categories of time-shifted cable programming. During our research, the streaming sessions were run over extended periods of time to investigate random occurrence of cumulative peaks of the combined streams. The results of these studies have been published and presented at previous cable conferences.

In summary, VBR streams when allowed to self-average inside a D3.0 bonded-downstream pipe create a combined flow that is fairly predictable (as measured by the peak to average (PAR) indicator) and the bandwidth allocation for this flow is manageable with simple provisioning rules. Actual cable content was transcoded from MPEG2 to H.264 AVC format and used for the purpose of these studies.

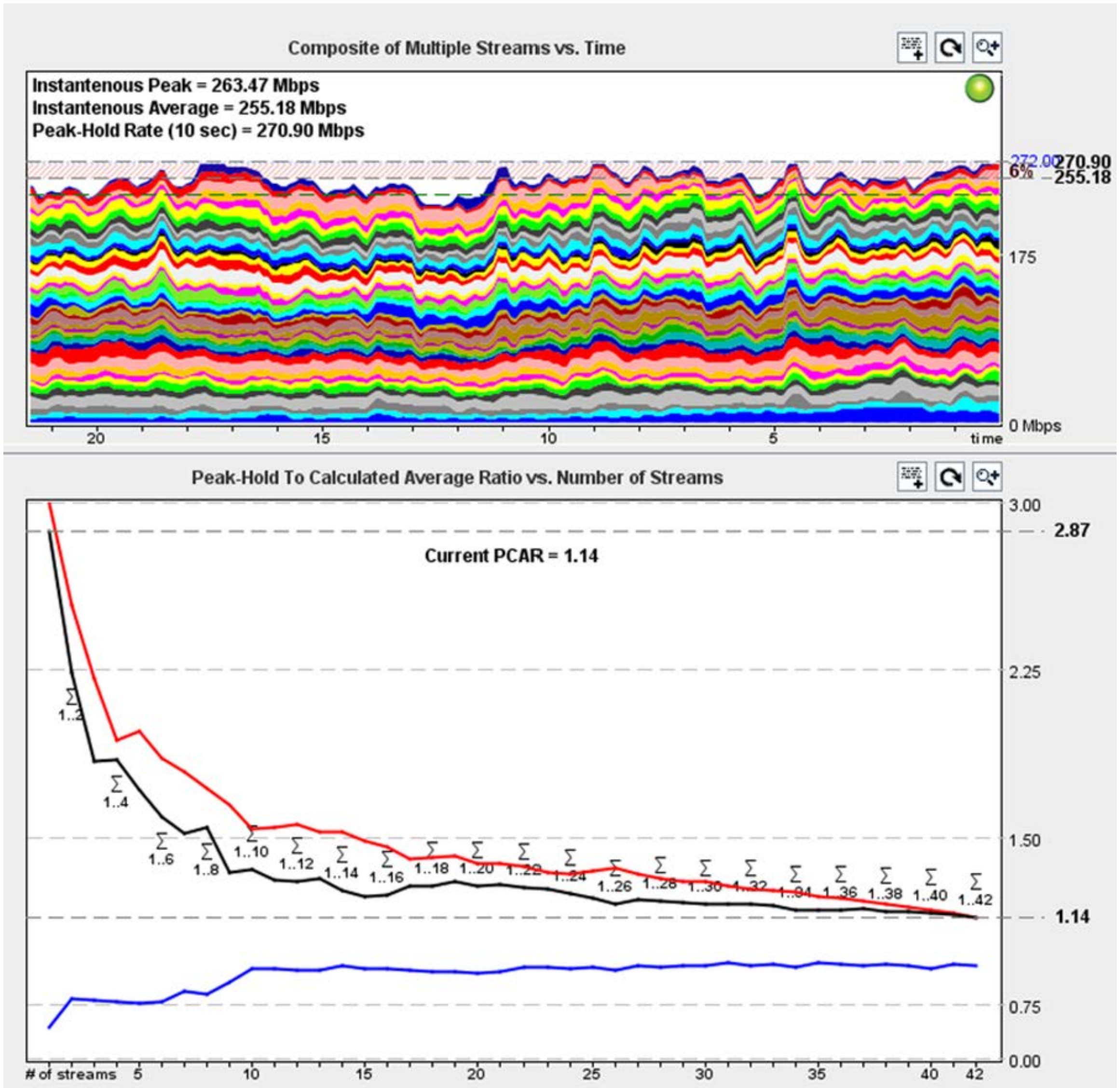
The following plots illustrate the results of one of our studies. A Peak to Calculated Average (PCAR) parameter has been defined as a ratio of the peak bit rate (calculated as a peak-hold inside a 10ms time window) of the VBR stream to the average bit rate that is calculated from the media file size divided by the duration of its playback. Both parameters are easily obtainable for on-demand assets and for the multicast (or time-shifted TV) and can be predicted from the encoder settings and type of content streamed.

The plots below developed during the study illustrate that:

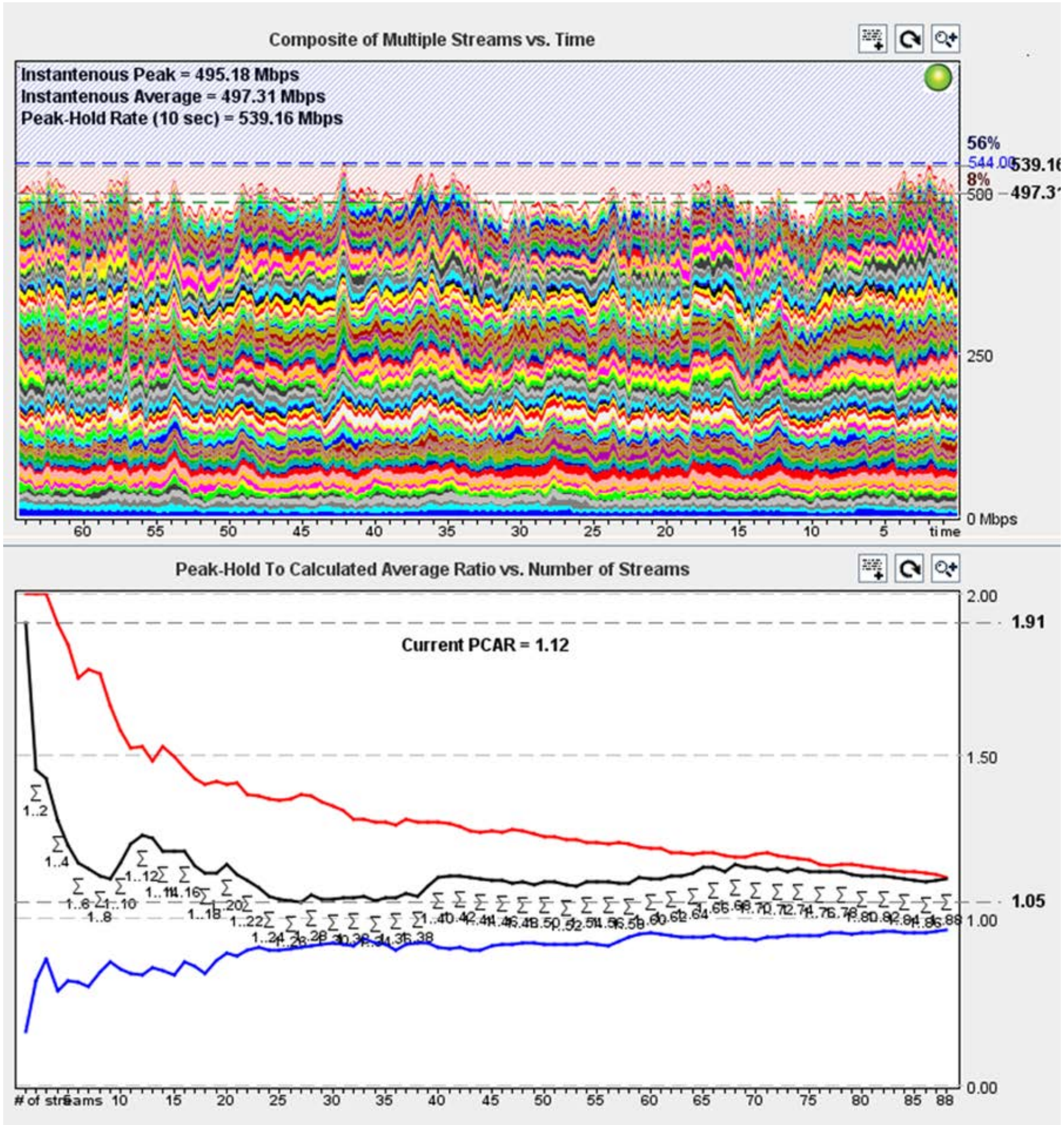
- More streams can be safely combined in a larger bonded pipe than in a smaller bonded group (as shown on X-axis of the plots)
- The PCAR variability is asymptotic to a value of 1 as more VBR streams are being combined thus assuring that peaks will not exceed the maximum bandwidth available



Plot 1. Example of 20 VBR streams inside one quad-bonded D3.0 pipe. The PCAR value (max hold value in red, min hold value in blue) is plotted vs. the number of streams combined.



Plot 2. Example of 42 VBR streams inside one octal-bonded D3.0 pipe. The PCAR value (max hold value in red, min hold value in blue) is plotted vs. the number of streams combined.



Plot 3. Example of 88 VBR streams inside one sixteen-bonded D3.0 pipe. The PCAR (max hold value in red, min hold value in blue) is plotted vs. number of streams combined.

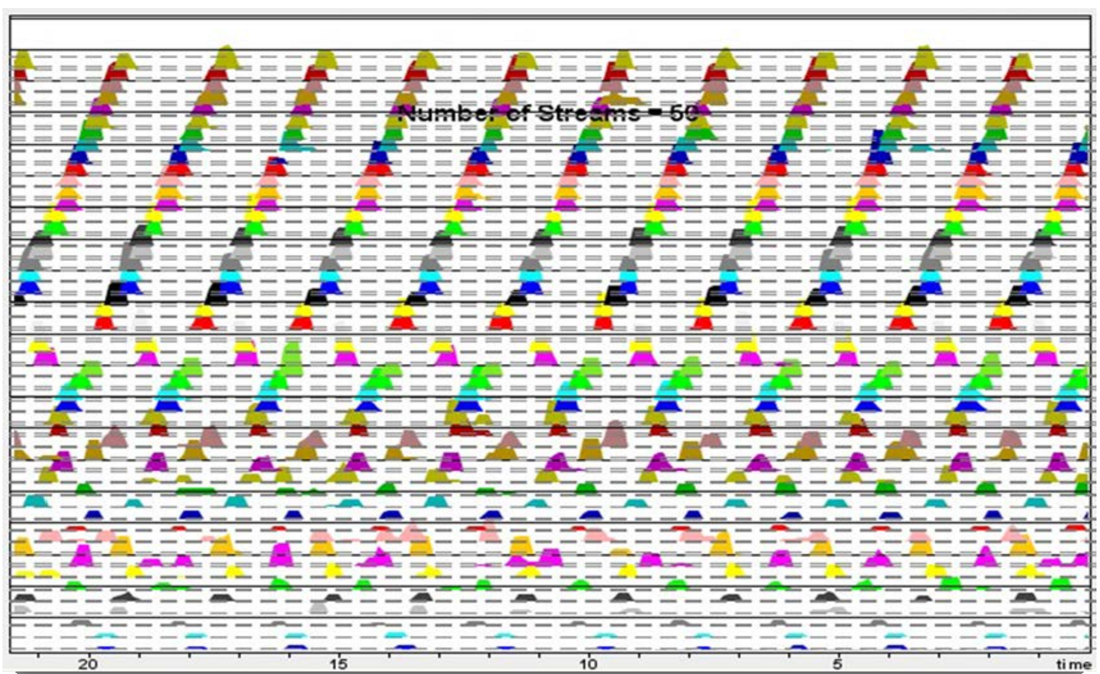
In our study, multiple clients receiving the VBR streams were simulated in Java environment. The streaming method was linear and used the UDP IP protocol. The streams' time positions were regularly randomized to simulate user interaction with the content.

Peak to calculated average (PCAR) of individual streams was typically between values of 2 and 3 as the encoders that generated the content under study were set to 12mbps peak rate.

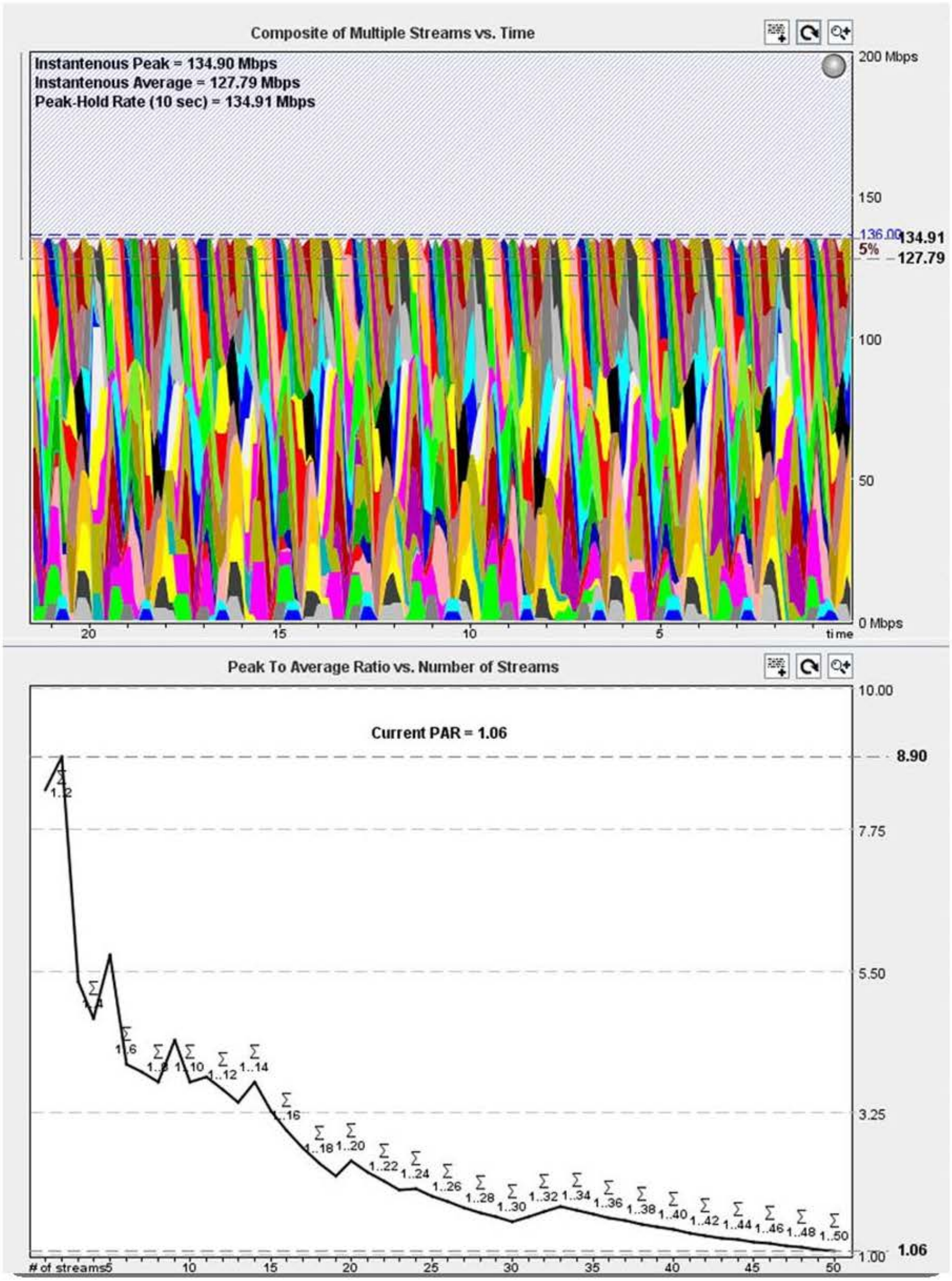
The emerging, standards-based Fragmented MP4 method of streaming of VBR formatted "chunks" of content lends itself even more to delivery over D3.0 CMTS than the continuous VBR (MPEG2TS UDP). The sizeable video buffers inside FMP4 compliant clients allow for increased flexibility of the interarrival time of the fragments. The player will wait for the fragments arriving from the server for much longer durations of time, while continuing to play content from its buffer. As a result, the streaming process is very tolerant to queuing in the CMTS and the overall Content Distribution Network (CDN). The queuing function, a built-in DOCSIS feature, acts as a traffic shaper that uses all transmit opportunities to send the variable fragments (as they arrive from the FMP4 servers) towards the client devices. As shown on the two plots below, the CMTS can act as an elastic buffer that spreads and smoothes the peaks of the combined downstream flows.

The FMP4 studies showed that the Peak to Calculated Average Ratio (PCAR) values of many sequential bursts (responses to clients' "get" requests) that are combined inside the quad or octal bonded D3.0 pipe stays very close to the ideal value of 1 for long durations of time. Plots 4 and 5 below illustrate how peaks of transfers of individual content fragments are smoothed out by bandwidth shaping applied to the entire flow.

Plot 4 shows the bursts for individual client devices and Plot 5 shows all of the bursts combined inside a quad-bonded downstream pipe. The PCAR value of the combined FMP4 flow traverses faster towards the value of 1 as more bursts are combined than in the case of VBR UDP streaming described earlier in this section.



Plot 4. A view on 50 FMP4 individual http transfer flows towards the client devices combined inside the wide-band D3.0 downstream pipe. The peaks are smoothed-out by a queuing mechanism.



Plot 5 50 FMP4 flows (consisting of regular bursts) to individual clients combined inside a quad-bonded downstream

In managing the UDP or FMP4 VBR streaming it is very critical that the average value of the combined flow (calculated or estimated as a sum of the averages of the individual streams over an extended time (continuous or fragmented)) does not exceed the bandwidth available in the DOCSIS pipe. Simply speaking, the laws of FIFO "physics" have to be respected and any bandwidth smoothing buffer will overflow and some packets dropped if more packets enter the FIFO than can be sent to the HFC.

Ours VBR streaming studies from analyzing many streams of various CATV content can be summarized as follows:

- The average bit rate (ABR) of H.264 AVC 1080i stream encoded with a real-time commercial coder ranges between 5 and 9 mbps. This range can be shifted lower by 1 mbps average when off-line, file based encoding is performed
- The instantaneous value of Peak to ABR of any stream varies between 0.6 and 3. Again, for off-line encoding the PAR value can reach a much higher value.
- Combining of multiple streams reduces the variability of P/ABR of the composite flow (VBR self-averaging effect) to the range of 0.9 – 1.1 for 500 mbps flows
- Efficiency of VBR self-averaging strongly depends on several factors:
  - Number of VBR streams combined, up to a point of diminishing returns
  - Encoder (transcoder) settings, maximum bit rate and the strength of compression
  - Mix of content in the combined flow (ratio of talking heads to movies and to sports)
- A "light touch" of delay applied to some or all streams further smoothes out the peaks and maximizes the effects of self averaging while preventing loss of video packets. Attention must be paid to keep the delay at a value lower than that of any receiving CPE
- The fragmentation of VBR content (files or live streams) and delivery of such "chunks" in response to regular HTTP "get" requests by clients with very deep (up-to 20 seconds) input buffers further maximizes efficiency of content delivery over a D3.0 system.
- DOCSIS overhead has to be included in the actual HFC bandwidth provisioning but it does not have a significant impact on the improvements in the overall transmission efficiency.

VBR bandwidth can be managed in the backbone and on the HFC by using the simplified provisioning parameters

- A conservative average value of 8mbps can be used for provisioning VBR VoD/SDV (H.264 AVC @ 1080i content) over octal-bonded D3.0 downstream facility. This number is significantly lower than today's rate of 15mbps of HD VoD MPEG2 encoded CBR streams.
- The same rate of 8 mbps rate can be safely used for provisioning of Broadcast VBR. Today's 3 into 1 HD stat-muxing results in an average bit rate of 13.3mbps MPEG2 CBR (one third of ~40mbps MPTS)
- Use of jumbo Ethernet frames (14 and higher MPEG2TS packets in a frame) further improves the transmission efficiency (reducing the 3% (ETH+IP+UDP) overhead)

Significant additional savings in CAPEX and OPEX, or expenditure avoidance can be realized with the transition to the H.264 AVC because no video processing related to improving efficiencies of content delivery using 6MHz channelized HFC will be needed:

- No stat-muxing will be required for Broadcast streams to fit a maximum number of them into 40mbps bandwidth of a single 6MHz channel
- No grooming will be needed to CBR for VoD/SDV services

The content will traverse through the delivery system in its "native" format as generated by the encoders. Its quality will be maintained as generated at the source.

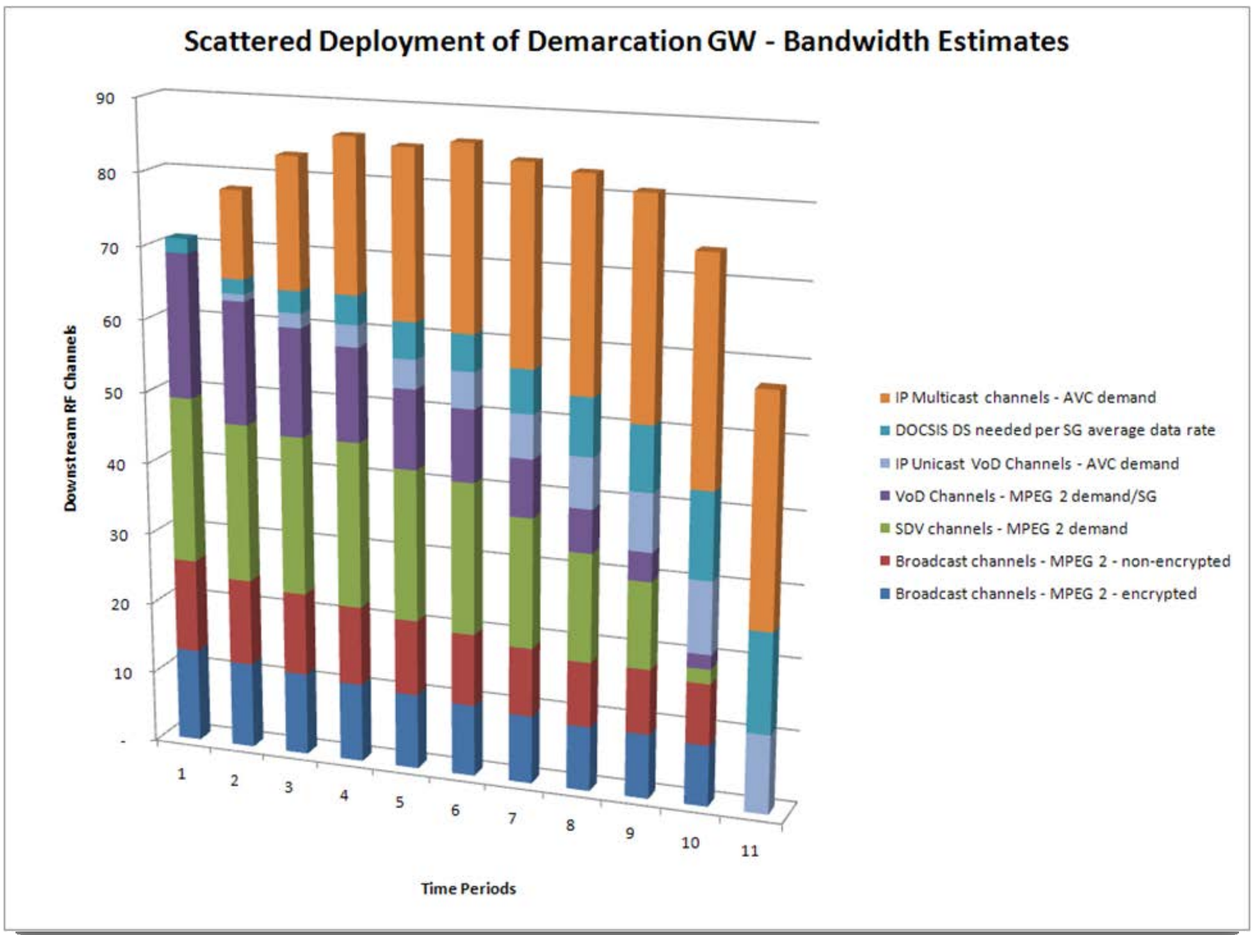
RF Bandwidth Savings/Recovery Strategy Opportunities

According to our on-going analysis of bandwidth usage trends by large MSOs, continuing with the traditional deployment paradigm may yield a possible bandwidth map as shown in the following diagram. The current MPEG2 channel distribution is contrasted with a possible MPEG4-based channel distribution that would provide all current subscribers with the same services, except that the video is delivered over using H.264 AVC over IP.

It is clear and intuitive that once the operator has transitioned from MPEG2 to H.264 AVC over DOCSIS, the bandwidth usage gains and recovery can be very significant thus allowing the operator to expand current offerings and aggressively grow new services.

The transition to an all-IP AVC environment is the most difficult part of the strategy.

As illustrated below, a gradual transition that is based on slowly "retiring" the MPEG2-only STBs, while introducing the new AVC CPE unfortunately results in a substantial simulcasting penalty. The serving areas with a mix of customers who have been moved to IP Video, and those who have not, still must carry a duplicate suite of video services. Within each node this could lead to a transition with increased RF usage over a long period of time.

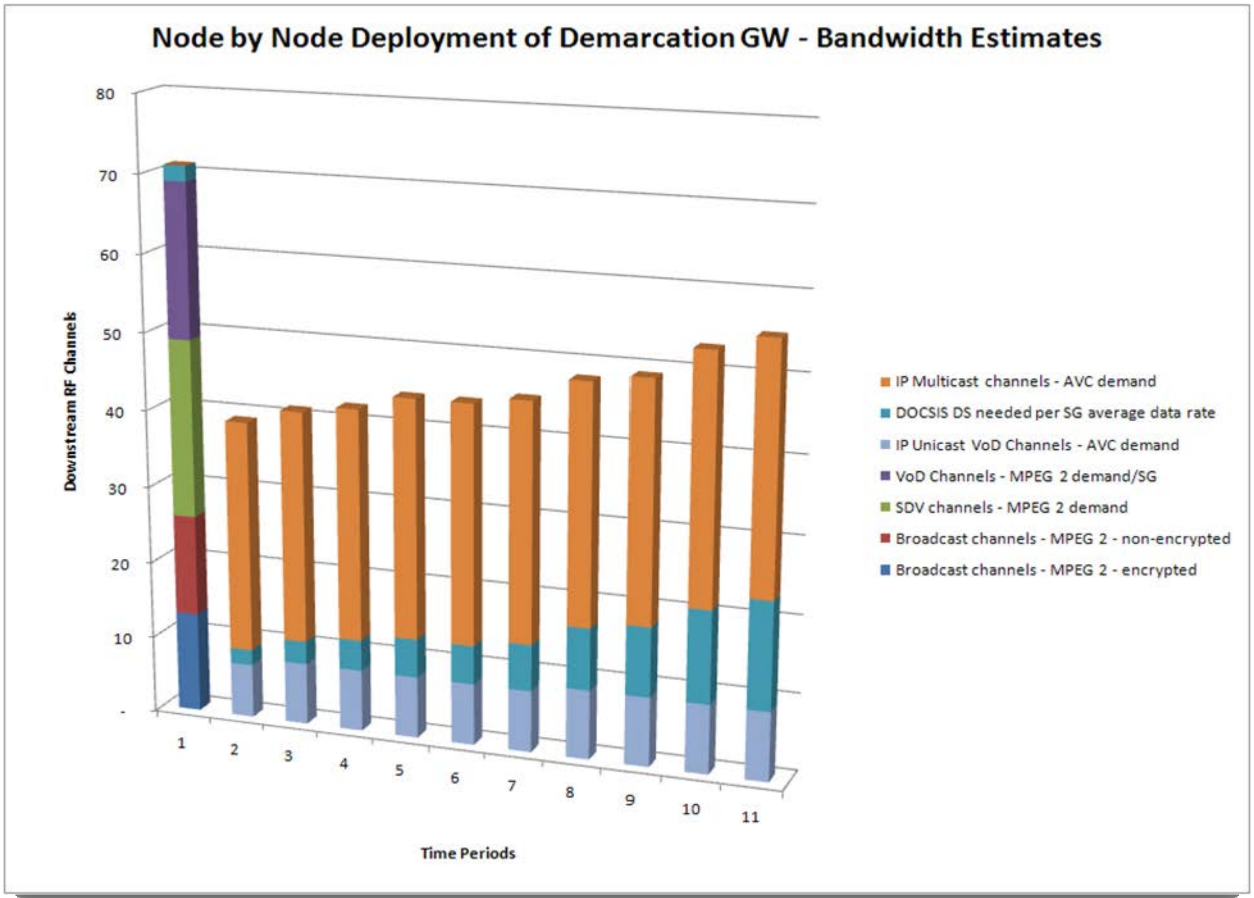


Such a gradual transition does not make efficient use of capital or operational investments.

In contrast, a node-by-node, wholesale change to VBR AVC in the access strategy can be enabled by pre-installation of the Demarcation Gateway in the entire serving area followed by a flash cut-over to all-AVC HFC. Such a strategy will allow the operator to switch quickly from one channel map to another and start expanding the offerings without a need for simulcasting.

The Demarcation Gateways can be deployed within a targeted node, without disturbing the existing services with the help of an RF pass-through path. The operator may deploy the DGWs outside of customer premises at the HFC/home demarcation point, thus eliminating the need to work around subscribers' schedules and improving overall operational efficiencies later on.

When the head-end AVC infrastructure is ready, and the DGW deployment is completed, the DGWs can be commanded to disconnect the pass-thru path and begin providing modulated RF signals to the home STBs to replace the MPEG2 services that are now decommissioned in the HFC.



The biggest advantage of this approach is when the costs of the transition are considered. The ability to avoid replacing the existing in-home CPE provides a significant cost benefit. Lower cost IP STBs can still be deployed once the Demarcation Gateways are in place, but that investment is not required immediately. The STB replacement investment can be transformed into a success-based selective investment that happens at a slower pace.

THE DEMARCATION GATEWAY

In the early 1990's ARRIS pioneered a two-way HFC specific telephony Network Interface Device as a part of the Cornerstone product portfolio. More than ten million Voice Ports were installed on the side of houses served by MSOs around the world. Design of such an active demarcation point of deployment device offers engineering challenges that were proven to be surmountable. Combining the know-how from the best of telco and cable RF worlds and some hard-core engineering work resulted in a device that was five 9's robust while consuming less than a Watt of power most of the time.

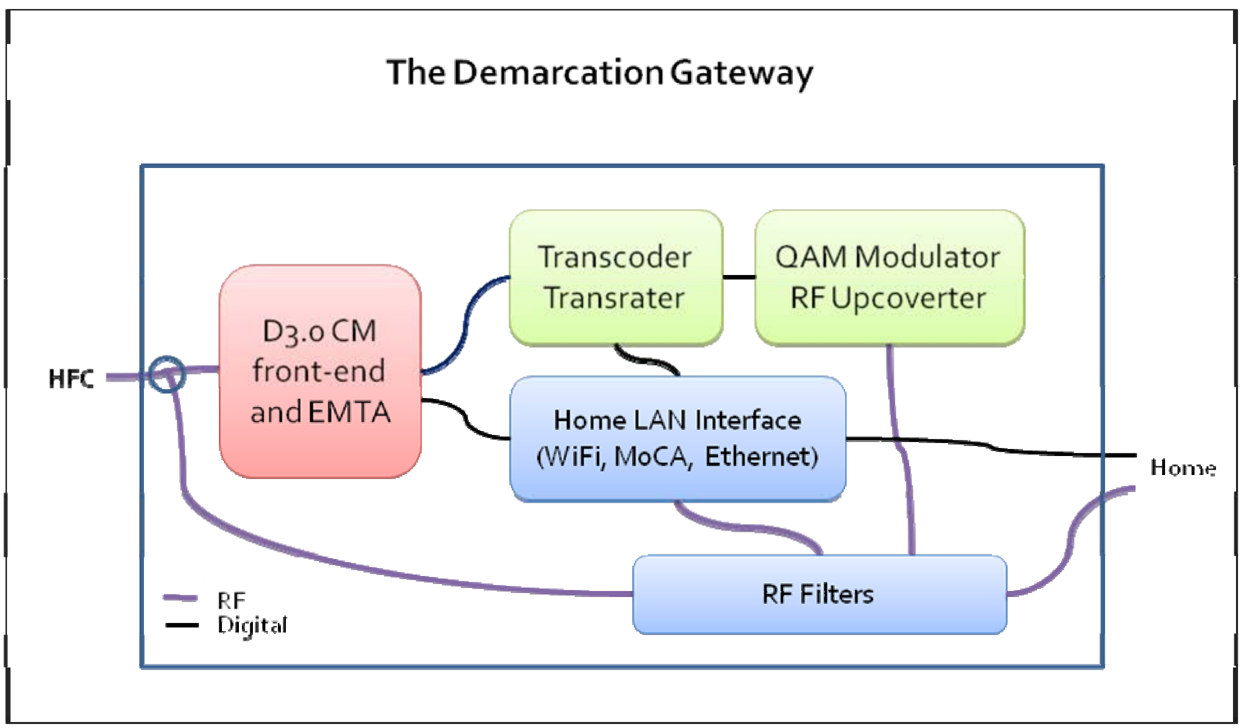
The subsequent proposals and even some trials have attempted to address a whole home Gateway device, and have even explored on-site transcoding possibilities for MPEG2 support, but cost and performance have been barriers that prevented serious consideration of these ideas.

The relentless advance of silicon integration and processing capabilities has overcome these barriers. The same advances that have empowered the latest generation of DOCSIS3.0 capabilities can also be used to enable the Demarcation Gateway. We all can relate to the shining example of integration and true multi-functionality of Apple portables -- today's functionality would require racks of equipment to implement not so long ago…

The bonded channels of DOCSIS3.0 have opened a wide pipe to deliver all content over "bursty" IP to the home. The resulting statistical multiplexing gains due to the change from standard MPEG digital channels to bonded DOCSIS channels are significant. The technology that enables these RF advances was developed in response to D3.0, taking advantage of technical developments in the cellular phone and wireless technology markets.

Silicon integration has also been continuing on the video processing front. Only a few years ago an MPEG2 to MPEG4 HD transcoder was a top of the line 2RU encoder, then it was shrunk into a stand-alone co-processor chip next to a set top box video processor. The upcoming generation of set top box silicon will commonly include this feature within the main set top chip. The processing requirements of a Demarcation Gateway will be met at a reasonable price point, far less than would be required to replace the existing MPEG2 Set top boxes within the home. It is worth stating that the bandwidth at home is significantly cheaper than on a crowded HFC network and the transcoder can generate much higher b/w MPEG2 stream for the same resolution as today's rack-mount, HE based xcoders

The Demarcation GW can be architected as illustrated below:



Key components of the DGW are in place already although implemented as separate subsystems and located in different HFC network elements and home devices.

The two new functional blocks, additions to what we can find in currently designed cable transport GWs, are the micro-QAM and RF up-converter and the transcoder/transrater.

The QAM+RF design can be easily "borrowed" from next generation super-dense EQAM devices currently under development by several vendors. Our estimates show that the two 6MHz channels that can be generated in this subsystem will provide sufficient transmission capacity for the very popular, multi-tuner HD DVR. The digital modulator can synthesize QAM carriers directly onto the desired RF channels, thus avoiding the need for a complex and power-hungry RF upconverter.

Integration of AVC to MPEG2 transcoding is critical to enabling the continual use of tens of millions of MPEG2-only STBs. The functionality proposed for integration within the DGW is already widely deployed in home AV appliances from CE vendors like Sharp, Hitachi, Toshiba, NEC, etc. Such coding conversion features are being widely deployed in the home for a variety of reasons.

(a) Storage expansion on PVRs to increase storage capacity on a fixed-size HDD

(b) Enabling smart phones to tap to HD broadcast (scaling)

(c) Allowing PVR based content to be viewed on smart phones

The prices of stand-alone silicon chips that can perform these functions are moving below $20. Integration with other building blocks on a same die of a SoIC will translate into a significantly lower cost of the transcoding function. Silicon fab house such as TSMC and UMC have consistently been shrinking silicon geometries from 180nm through 130nm, 90nm, 65nm to 40nm now. As these geometries decrease, we are seeing that the silicon area to integrate the many hardware blocks has shrunk approximately 20 times with a transition from 180nm to 40nm. This brings an astounding opportunity to integrate; power and cost reduce the needed functionality onto a single SoIC.

The cable industry has the footprint, a large deployed MPEG2 STB base and if united, can create production volume requirements for its vendors which can make the investment of creating a DMG on a hardened SoIC a viable business case.

Conclusions

ARRIS believes that today, more than ever before (and enough) "stars have aligned" for a successful development of a Demarcation GW. Let us re-iterate the convincingly influencing factors for making the DGW (as outlined above) case:

- The deployment of DGW may free-up 50% of the HFC RF spectrum, the strongest "weapon" that MSOs possess. Delivery of VBR formatted, AVC encoded content in the DOCSIS3.0 downstream bonded "pipe" brings about this saving.
- Use of the NG DOCSIS3.0 technology in the HE (to drive the DGW's front-end), pioneered by among others, the Comcast's CMAP initiative, will bring the access equipment cost down as a significant step, an order of magnitude similar to the transition from D1/2.0 to D3.0.
- The DGW may enable a flash transition to the long-awaited, all-IP access on a serving area by serving area or node by node basis.
- The DGW installed on the side of the house will allow an operator to non-intrusively and transparently meet all communications and entertainment needs of the customer in a powerfully "sticky" way.
- The technology for fast-prototyping and fine-tuning the DGW is here now.
- The silicon geometry to integrate the DGW functional blocks is already in place at the foundries that serve the cable industry.