Separation of Video Processing and Multiplexing: The Key to Network Scaling of HD-VOD and other Bandwidth Intensive PersonalizedTV® Services

Ron Gutman, CTO
Lorenzo Bombelli, VP – Product Strategy and Management

Imagine Communications, Inc.

Abstract

Cable operators need to rapidly expand bandwidth capacity to accommodate the emerging personalized and time-shifted viewing paradigm and the accelerating availability and demand for HD content. Linear programming, traditionally broadcast to subscribers using efficient high-end closed-loop Variable Bit Rate (VBR) technology, will increasingly be delivered over on-demand networks which, using current Constant Bit Rate (CBR) technology, consume approximately 50% more bandwidth without matching the video quality of high-end broadcast VBR delivery.

Although there have been a number of proponents of on-demand VBR, architectures and solutions based on broadcast-oriented rate-shaping technology inherently rely on performing both video processing and statistical multiplexing on each on-demand stream. Such approaches result in lower video quality than high-end broadcast VBR and a cost structure that is not viable for wide scale deployment in on-demand applications.

In this paper we propose a new architecture that enables on-demand VBR by separating the video processing of each video asset from the downstream multiplexing of each stream to achieve the same video quality as broadcast VBR and eliminate the unnecessary and significant cost of re-processing the same asset multiple times. This "never cover the same ground twice" approach makes VBR viable for on-demand services, and delivers the following value to system operators:

- *Clear and sustainable competitive advantage in video quality*
- *Up to 50% gain in HFC bandwidth*
- *Reduced overall CAPEX and OPEX*
- *Uniform and controlled video quality across the entire network*
- *Improved system availability*
- *Increased service velocity*

ENABLING ON-DEMAND VBR

Limitations of Traditional VBR Technology

While VBR technology can provide significant bandwidth gains and video quality improvements, a number of important factors have impeded its adoption in on-demand applications.

During initial VOD rollouts, the challenge and complexity of developing the underlying infrastructure, combined with the relatively low initial demand for on-demand services, made optimizing video quality and bandwidth lower priority.

After initial rollouts, adoption of traditional VBR solutions was further deferred because existing VBR solutions based on broadcast-oriented rate-shaping technology could not generate enough incremental value, both in terms of video quality and bit rate reduction, to justify the incremental investment.

A key architectural limitation of these broadcast-oriented solutions is that they apply both video processing and multiplexing on each on-demand stream, which leads to two key shortcomings.

The first is that broadcast-oriented rate-shaping solutions attempt to compensate for the inherent cost associated with having to apply video processing to each stream by using bit rate as a less complex yet less accurate proxy for video quality. As a result, these solutions have no mechanisms to establish and set the optimal bit rate to achieve a desired video quality benchmark, and tend to sacrifice video quality in order to meet bit rate targets.

The second and related shortcoming is economic. Despite attempts to reduce cost per processed stream in broadcast-oriented rate-shaping solutions, it is impossible to overcome the inherent architectural limitation of applying video processing on a per-stream basis by reducing component cost. As a result, broadcast-oriented VBR solutions have been unable to meet the system cost requirements for scalable deployment on all streams.

One proposed approach to mitigate this limitation is to add an intelligent session management layer which would dynamically assign video processing resources only to over-provisioned QAM channels, relying on predictable and static user behavior patterns to allocate resources. While this approach does partially alleviate the cost and density burden inherent in a broadcast-oriented rate-shaping architecture, it does not provide the scalability required in a largely user-demand driven delivery network and is therefore a short term workaround to a sub-optimal architecture, rather than a transformative long term solution.

As a result of these shortcomings, although VBR has the potential to deliver considerable value in the rapid expansion of on-demand services, a new approach is required to make VBR viable in on-demand networks.

A New Approach for On-Demand VBR

We propose the following approach to on-demand VBR to make it viable and accelerate its adoption in cable:

1. Decouple video processing from statistical multiplexing to enable optimal technical and economic placement of each function in the network.
2. Process each asset one time upon ingest using high-end VBR technology rather than once for each stream using rate shaping technology.
3. Utilize the highest quality video processing, giving priority to video quality rather than bit rate to ensure that quality is highly controlled and not compromised.
4. Optionally, create a simple linkage between the video processor and the SRM that enables the SRM to incorporate quality in bandwidth allocation decisions.

The Case for Asset-based Video Processing

Decoupling video processing from statistical multiplexing and applying video processing on a per-asset instead of a per-stream basis enables significant improvements in both the video quality performance and the economics of on-demand VBR. Take the following example. A cable system with a half million subscribers and a deployed base of one million digital set-top boxes has a VOD system designed for 10% peak utilization, meaning that during peak usage, no more than 100,000 VOD or unicast standard definition streams can be active simultaneously on the network. The operator has taken into account user patterns across the system and has deployed an intelligent session management layer to enable over-provisioning of rate-shapers, and sufficient broadcast-oriented rate-shapers to process 50% of deployed streams. In effect, the operator has deployed 50,000 unique and distinct video processors into the network that can be assigned to user sessions as required.

Now consider a "super" video processor, with 100 times the processing power of a rate-shaper, applied on a per-asset basis at the ingest point of the on-demand library. Also suppose that at the ingest point the system requires a content refresh rate of 10,000 hours per month. Assuming real-time processing, this would require 10,000/30/24 = 14 processors.

Even assuming 20 processors to allow overhead for peaks in refresh rates, per-asset pre-processing with "super" video processors would only require 2,000 video processing units, saving an equivalent of 48,000 video processing units. These gains are further amplified if asset-based processing is applied in regional clusters of 2-3 million subscribers or national clusters of 5-20 million or more subscribers.

Video Quality Benefits

The allocation of video processing on a per-asset, rather than per-stream, basis represents a fundamental shift in the approach to on-demand VBR. With 100 times more processing power available for each asset, asset-based processing can take advantage of advanced video processing techniques that are not economically viable when applied on a per-stream basis. Look-ahead configurations can be maximized, two-pass and even multi-pass encoding can be applied, the search for more precise motion vectors can be enhanced, scene changes can be more accurately identified, and the diligence applied to noise filtering can be both infinitely granular and wholly comprehensive.

Networking and System Benefits

In addition to the quality benefits described above, shifting to asset-based video processing has other important system benefits.

The ratio of processing power required for video processing vs. multiplexing is approximately 98:2. Therefore, separating video processing from multiplexing not only unlocks the potential for higher video quality on the video processing side, but also enables significant reductions in cost, density, power consumption and system delay on the statistical multiplexing side.

Moreover, since assets are pre-processed, they can be encrypted upstream from the multiplexer (after the video pre-processing), and then distributed, multiplexed and transmitted to the subscriber without compromising the content protection or digital rights management.

Also, by consolidating video processing operations and distributing the multiplexing function, system operators can enjoy a new level of quality and operational control, ensuring that all subscribers have same high quality viewing experience regardless of the particular and immediate usage patterns in their neighborhood.

IMPLEMENTATION

Statistical Multiplexing Efficiency

The benefits of statistical multiplexing of VBR signals are well known and understood. The bandwidth efficiency benefits amounts to up to 30% gain in channel capacity. When paired with the unlocked potential of a separated video processing engine, the efficiency of the VBR statistical multiplexer can reach as high as 50%, but not without some exceptions.

Our research suggests that when unconstrained video processing is combined with constant quality VBR multiplexing, the aggregate rates of the individual programs will peak above an allocated channel rate approximately 15% of the time. The two examples given below plot instantaneous bit rates for two different 30 minute standard definition programs and provide a histogram for the number of seconds at each bit rate. In both examples, the content is coded at a constant "ICE-Q®" quality score of 96, indicating a quality level of NMD (No Material Degradation), providing a constant quality VBR signal that matches the quality of the source.

The complexity difference between the two programs is reflected in their average and maximum bit rates as well as their Effective Bit Rate (EBR). The EBR is a new metric that indicates how much bandwidth is required to maintain constant quality for any given asset. In Example A, the input CBR rate is 3.75 Mbps. The content is an HBO Sex in the City segment with an average VBR bit rate of 2.3 Mbps, standard deviation of 0.45 Mbps, a maximum bit rate of 3.32 Mbps and an EBR of 2.78 Mbps.

Example A

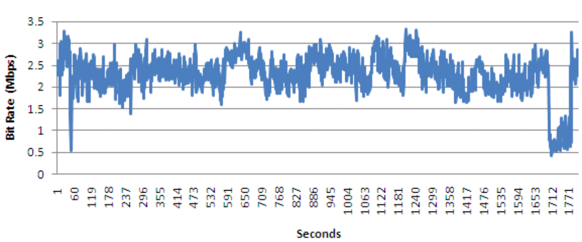

Figure 1: HBO: bit rate over time

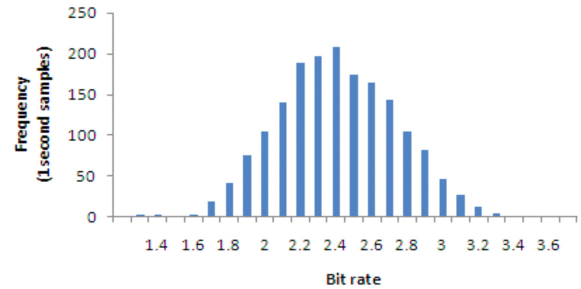

Figure 2: HBO: bit rate histogram

In Example B, the input CBR rate is also 3.75 Mbps. The content is a nature documentary. The average bit rate is 2 Mbps, the standard deviation is 0.4 Mbps, the maximum bit rate is 3.67 Mbps and the EBR is 2.46 Mbps.

Example B

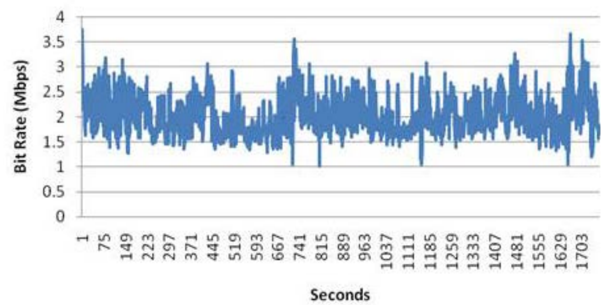

Figure 3: Africa: bit rate over time

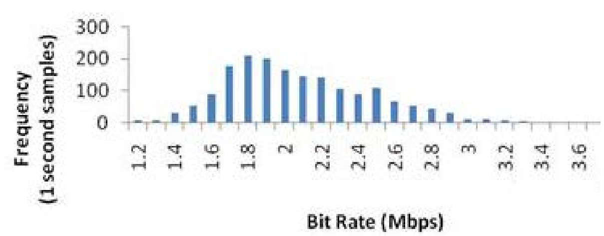

Figure 4: Africa: bit rate frequency histogram

In Example A, the percentage of seconds that are above the EBR is 12%. In Example B, the percentage is 15%. Similar analysis has been performed on a larger sample of 40 distinct HD and SD clips of varying content type and complexity. VBR bit rate distribution in this sample shows a clear statistical match to normal distribution. Overall, the SD clips have an average VBR bit rate of 2.13 Mbps, standard deviation of

0.389 and an average EBR of 2.41 Mbps. If the available QAM bit rate is 38 Mbps (allowing some reserve) the average number of streams per QAM is 15.8 according to the EBR average. Assuming optimal time-shifting and buffer management at the edge multiplexer, the probability of the aggregated average bit rate exceeding the QAM channel bit rate at 16:1 density is 0.7%. Similarly, the probabilities of streams exceeding the QAM bit rate for 3 HD plus 3 SD streams, 2 HD plus 8 SD streams, and 1 HD plus 12 SD streams are 2.4%, 2.5% and 1.3% respectively. Clearly, more streams per QAM provide larger statistical gain.

The sensitivity of video and audio communications is such that a ~2% error results in a completely unusable product. Therefore, some additional compression is required on 2% of the content.

Second-Pass Asset Processing

An additional benefit of the massive amount of video processing enabled in the asset processing stage is that all of that processing can be brought to bear in a second pass of processing whereby the variable bit rate signal can be more aggressively compressed a second time to create an alternate VBR version of the signal to the predominant constant quality "V1s." This alternate ("V2") can then be transmitted to the edge multiplexer and inserted during the 2% of the time in which it is required.

In this architecture, the statistical multiplexer performs the critical process of rate management without needing to perform rate-shaping or re-compression of the signal. And because the V2 version of the signal has been compressed using the full capability of the high-end video processor, it is of higher quality than can be provided by a broadcast-oriented rate-shaping platform, which is severely constrained by cost and density requirements.

Improved Quality and Interoperability

It is important to note that the asset-based approach for content processing can be applied to improve video quality as well as bandwidth efficiency. This is accomplished by starting with a higher input bit rate and higher quality encoded source as the input to the asset-based processing stage, for example 7 Mbps CBR for SD. The resulting EBR for the processed asset will be higher to support the higher quality level, resulting in a channel capacity that is equivalent to today's 10 streams per QAM, with each CBR at 3.75 Mbps.

These processed streams will appear to the network as typical 3.75 Mbps CBR. The video bit rate inside the 3.75 CBR transport rate, however, will be variable. In the valleys of the VBR rate, additional information can be stored that allows statistical multiplexers to recreate the higher quality program derived from the 7 Mbps source file. The file containing the extra layers of information must conform to standard VOD specifications such that servers, pumps, QAMs and asset management systems will process the asset without requiring any additional processing or format conversion. The same formula can be used for high definition content, substituting 25 Mbps input sources for the 15 Mbps CBR content prevalent today. Using this method, standard definition VOD assets can be "DVD quality," while HD-VOD assets can approach "Blu-ray quality."

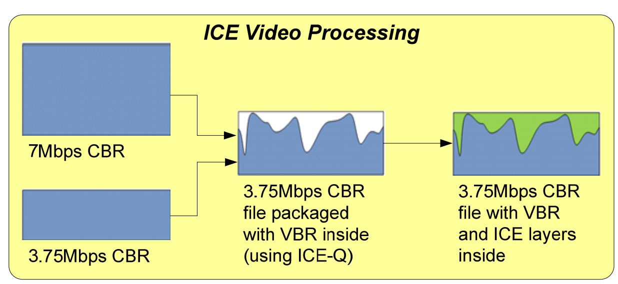

Figure 5: Improving VOD quality with ICE layers

Video Quality Aware SRM

Optionally, with minor modifications, the Session Resource Manager (SRM) can play a vital part in further optimizing bandwidth while maintaining the highest video quality in an on-demand VBR system.

From a video coding perspective, content varies in complexity. Sports content typically includes high temporal complexity (motion) while other high-complexity content is spatially detailed or includes spatial subtlety (ramp-type edges) that requires advanced processing. Other content such as static images, dark backgrounds or non-moving head-and-shoulders video is less complex and can be processed much more easily.

The fundamental problem with providing the same bit rate for all streams without taking into account content complexity is that it results in sub-optimal picture quality and bandwidth efficiency. Just as 3.75 Mbps may be inappropriate (overkill) to adequately encode a head-and-shoulders news clip, the 3.75 Mbps rate can be equally inappropriate (underkill) for encoding a complex scene. Even with virtually unlimited encoding, it can be impossible to maintain quality during dramatic scene changes, fast motion and high spatial complexity if strictly constrained to 3.75 Mbps.

For example, the processing-rich encoding in Example A resulted in an Effective Bit Rate (EBR) of 2.78 Mbps. If one tries to multiplex 16 or even 15 such streams, severe degradation in video quality is virtually guaranteed, especially if done by a processing-limited rate-shaping device. The problem is amplified in HD because the aggregated standard deviation is much higher due to the lower number of streams per QAM.

A simple solution to this problem is to provide the EBR to the SRM thereby enabling the SRM to make QAM allocation decisions based on quality with a guarantee that total channel capacity will not be exceeded.

The EBR is equal to the combination of the average VBR bit rate combined with the maximum V2 bit rate, thus guaranteeing the quality of the multiplex at any given time.

EBR = Max(Average(V1),Max(V2))

The EBR value is calculated on a per-asset basis and could be delivered to the SRM using the ADI interface making the standard SRM quality aware and enabling it to load-balance according the quality rather than bit rate. Note that this approach does not require adding more elements to the session management flow.

Furthermore, while the EBR approach described above is very simple and effective, it is possible to further optimize the solution. For example, the asset video processor could provide the SRM with the averages and standard deviations of V1 and V2 per asset:

μ – Asset average bit rate
σ – Asset standard deviation

In this case the SRM would open a session if both of the following conditions are true:

V1 – Normal_Dist(38, M_1, S_1, 1) $\geq$ 99%

V2 - Normal_Dist (38, M_2, S_2, 1) $\geq$ 99.999%

Where:
$M = \sum_m \mu(1)$
$S = \sqrt{\sum_m \sigma^2(1)}$

Therefore, enabled by the EBR (Effective Bit Rate), and using very simple methods that can be enhanced and optimized by independent SRM vendors, the Session Resource Manager (SRM) can actively load-balance the QAM sessions according to video quality rather than stream counts, thereby guaranteeing constant video quality to all sessions while further optimizing bandwidth utilization.

In other words, the SRM is transformed into a video-quality aware device that can ensure and control video quality while delivering more streams at any given system capacity.

CONCLUSION

The growing demand for bandwidth-intensive PersonalizedTV services such as HD-VOD, switched unicast, and Internet video (whether streaming or downloads) will drive substantial HFC bandwidth pressure.

While on-demand VBR can deliver significant video quality improvements and bandwidth gains, current solutions based on broadcast-oriented rate-shaping technology have significant shortcomings that challenge their viability for scale deployment.

We propose that while edge multiplexers must perform rate management to accommodate the small percentage of the time when aggregated bit rate peaks exceed the available channel rate, it does not follow that those same edge devices must also be burdened with the computational intensity of video processing platforms. By relying on edge video processing, we unnecessarily impair service quality by sub-optimally forcing video quality decisions to lower common denominator devices.

In contrast, by separating the video processing and performing it at the video server ingest point on a per-asset basis, we ensure not only a consistently higher video quality during times of non-contention, but also a dramatically improved picture during the most critical periods of aggregate peak rate channel utilization.

In addition to enabling superior video quality and bandwidth efficiency, a separated asset-based processing architecture is uniquely equipped to scale with advancements in video and audio coding, while maintaining centralized control over delivered video quality, optimizing the cost and density of the edge multiplexing solutions, and minimizing delay in the end user experience.

In summary, an on-demand VBR approach that separates advanced video pre-processing from statistical multiplexing offers a compelling solution for scale deployment that delivers the full value of on-demand VBR and provides system operators a new and powerful tool in the ongoing battle for bandwidth and the consumer.