

# Monetising video traffic

Duncan Potter, chief marketing officer at Edgware, shares his thoughts on ensuring Quality of Experience and assuring wholesale monetisation with a Distributed Video Delivery Network (D-VDN)

Since two tin cans were strung together to make a simple communication device, the technical community has been faced with what appear to be conflicting needs and requirements within the same systems. With the recent and ongoing dramatic increase in both video traffic and the variety of devices that can be used to watch video and the additional pressure on operators to build their own wholesale or CDN networks, the operator community has been faced with the same problem.

## The 'Session-less' Issue

The basic problem can be stated fairly simply: at the time when measurement and assurance of Quality of Experience is becoming more important across the IP network, the protocols in use have moved from a consistent bit rate transmission to a variable bit rate that is *under the control of the client*. This is especially significant when considering operator CDN or wholesale networks where the operator must be able to show specific metrics in order to bill a content provider for access to the operator subscriber base.

To understand the conundrum requires a quick tutorial in HTTP based adaptive streaming.

## HTTP-Based Adaptive Streaming

Adaptive streaming is a hybrid delivery method that acts like streaming, but is based on HTTP progressive download. It's an advanced concept that uses HTTP rather than a new protocol. Both IIS Smooth Streaming, Apple Live Streaming and Adobe Flash Dynamic are examples of adaptive streaming. Even though these technologies use different formats, and encryption schemes, they all rely on HTTP as the transport protocol and perform the media download as a long series of very small progressive downloads, rather than one big progressive download.

In a typical adaptive streaming implementation, the video/audio source is cut into many short segments ('chunks') and encoded to the desired delivery format.

Chunks are typically two-to-10-seconds long. At the video codec level, this typically means that each chunk is cut along video GOP (Group of Pictures) boundaries (each chunk starts with a key frame) and has no dependencies on past or future chunks/GOPs. This allows each chunk to later be decoded independently of other chunks. In effect this means that there is no direct 'session' associated with the protocol unlike RTSP that is traditionally used in managed network environments such as IPTV or Cable.

The encoded chunks are hosted on a HTTP Web server. A client requests the chunks from the Web server in a linear fashion and downloads them using plain HTTP progressive download. As the chunks are downloaded to the client, the client plays back the sequence of chunks in linear order. Because the chunks are carefully encoded without any gaps or overlaps between them, the chunks play back as a seamless video.

The 'adaptive' part of the solution comes into play when the video/audio source is encoded at multiple bit rates, generating multiple chunks of various sizes for each two-to-10-seconds of video. The client can now choose between chunks of different sizes. Because Web servers usually deliver data as fast as network bandwidth allows them to, the client can estimate available bandwidth and decide to download larger or smaller chunks ahead of time. The size of the playback/download buffer is fully customisable dependent upon the actual player or client software in use.

Once you have even a basic understanding of this approach, it becomes fairly clear that there is a problem: If there is no real session associated with the communication, and the operator has deployed distributed network caches to optimise the effect of video on their network, *how does the operator report the QoE of a specific session or even the quality of video that has been delivered to a specific*

*client over a certain time period?*

The following diagrams show the issue and possible solutions.

In the diagram below, the top row is video, with the different colours denoting different bit rates. This bit rate schema is generally done at the time of encoding although there are some new technology advancements to be able to perform this on a more dynamic and distributed basis. These schemes do not alleviate this issue of monitoring and reporting and in fact could cause additional complexity due to the fact that they still need to be cached and then transmitted as close to the viewer as possible but now the number of components to an asset is fluid rather than fixed.

## Possible Solutions

There are three potential answers to this issue:

- Implement a log-based monitoring system that pulls all the log files from all possible locations and especially from the video servers and attempt to build an effective correlation between individual files being transmitted and the client that is viewing those files.

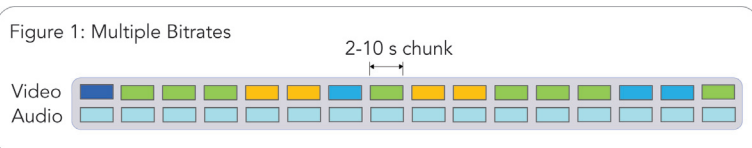
This method is attractive because it is fairly simple but means an enormous amount of horsepower in the servers running the analysis and correlation software. In the most simplistic of examples, a 90 minute soccer match encoded at five different bit rates would mean 13,500 files.

- For the distributed servers that are actually transmitting the video information to the client to play an active role in aggregating the information. Clearly this has the advantage that it can be done fairly simply as the video server already knows the video quality that is being requested every two seconds from the client and can report in real time what bit rate is being transmitted and therefore what QoE a particular client may be experiencing.
- The most likely scenario is that the system will necessarily become a hybrid, which means that it is essential to select video servers that can be distributed, have the ability to manage the asset propagation and can do the statistics aggregation.

## A Differentiated Solution

The main differentiator here is that the video servers must be capable of using their 'video awareness' and proximity to the client to create a virtual 'session' that enables the operator to report the actual Quality of Experience or at least the bit rate being consumed by the client in a way that is meaningful to the content provider.

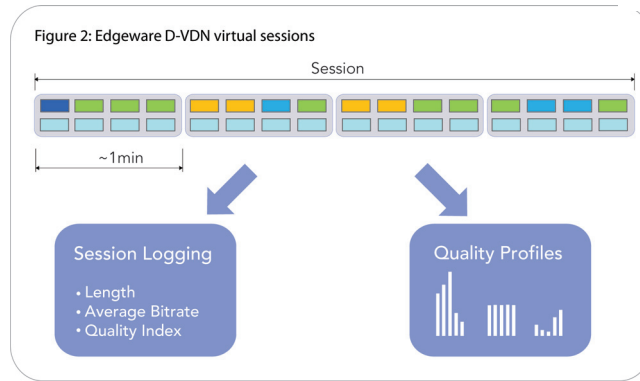
Once a session concept has been introduced, see Figure 2, it is possible to report usage data tied to sessions and content, not fragments.



Parameters of interest could be session length, average bitrate and engagement. Engagement is typically measured as session length vs. asset length, ie it shows how much of the content that was actually watched. This type of measurement again requires ‘video awareness’, the system must understand what is an asset and that it has a certain length. Just knowing what fragments that have been delivered is not enough.

To get the full QoE picture when using adaptive streaming technologies, one piece is still missing. Session averages don’t give the full picture since the actual user experience is controlled by the client by requesting different quality throughout a viewing session. Two users with the same average bitrate could have very different user experience, for example: the first user could use the lowest available bitrate for the first half of the movie and the highest bitrate for the last part, while another user could use the average bitrate throughout. To capture these kinds of user behavior the actual use of the different bitrates must be captured, without having to log every single fragment delivery. One way to do this is to build histograms of the use of all the available bitrates and thereby create ‘quality profiles’.

The used bitrates can be sampled at a reasonable rate, eg every minute, to reduce the amount of data while still providing enough detail. Such quality profiles could then be used to differentiate the two use cases described above, but even more importantly, they could be used to describe the actual delivered quality over some subset of sessions. These subsets



could for example be ‘all sessions to a certain device type’, ‘all sessions over a certain network segment’ or ‘all sessions for some specific content’. Using these quality profiles makes it possible for an operator to show its content providers what quality their users are experiencing.

**Conclusions**

In the future, it is highly likely that we will see the development of standards for aggregated statistics of this type allowing operators to simplify both their frameworks, implementations and procurement cycles. However, this will always remain a complex field, as it is possible to extrapolate almost endlessly into different areas, manipulating these types of statistics for almost any

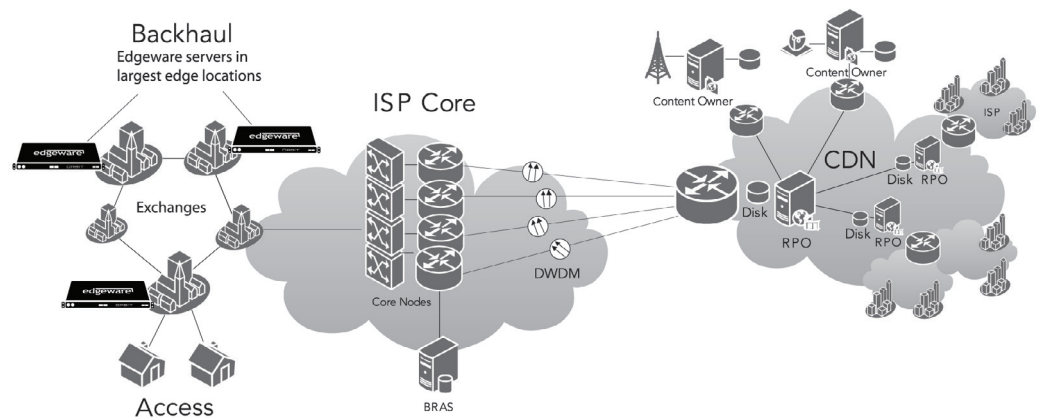
operator application.

Nevertheless, the basic conundrum still remains: how to create a session where there is no actual session establishment at one or more level of the protocol stack, and how to ensure that whatever is flowing over that session, arrives successfully, is viewed effectively without causing the viewer frustration caused by the network and finally, is reported in such a way that the operator can both report the actual

QoE that the viewer experienced and possibly even use that information for dynamically adjusting the Quality of Service (QoS) parameters of the underlying network infrastructure to enable an even greater consistency in viewer QoE.

At this stage, the optimum solution that can report in real time is to build a truly video-aware distributed delivery network that understands the nature of the file type and its contents being requested by the client. This is likely to remain the best and most effective solution for this particular problem and raises some interesting questions in the discussions around transparent caching vs. video delivery networks, as operators continue to look to monetise, rather than just contain the traffic running across their network.

Figure 3: Video aware network infrastructure



**The Advantages of an Edgware Distributed Video Delivery Network (D-VDN)**

- Overlay video delivery network – no ‘rip and replace’ of existing infrastructure
- ‘Pay as you grow’ technology reduces initial CAPEX
- Sophisticated asset propagation and integrated management system simplifies implementation of QoE, retail and wholesale video delivery network services
- Ultra low power consumption and no additional cooling requirement means deployable anywhere across the network
- Integrated wholesale CDN capabilities provide additional monetisation opportunities
- Reduced network and video delivery OPEX + additional wholesale revenue opportunities means dramatically faster ROI.

