Measuring Quality of Experience for Over-the-Top Video Services

August 2011
Avvasi Technical Press
AI-WP-427032-002

Avvasi Confidential
Introduction ......................................................................................................................................................... 4
OTT Video: Media Streaming Fundamentals ........................................................................................................... 5
OTT Video Content Lifecycle ................................................................................................................................. 5
Streaming OTT Video .................................................................................................................................................. 6
Progressive Download ................................................................................................................................................ 6
Real-Time Streaming Protocol ............................................................................................................................... 7
Real-Time Messaging Protocol ............................................................................................................................. 7
Adaptive or Dynamic Streaming ............................................................................................................................ 7
HTML5 Video ........................................................................................................................................................... 7
Peer-2-Peer TV .......................................................................................................................................................... 8
Video Conferencing and Video Chat .......................................................................................................................... 8

Video Quality Issues ................................................................................................................................................... 8
Capture ......................................................................................................................................................................... 8
Authoring/Encoding .................................................................................................................................................. 8
Transmission .............................................................................................................................................................. 8
Playback ....................................................................................................................................................................... 9
Common Video Quality Issues ............................................................................................................................... 9
Stalling ......................................................................................................................................................................... 9
Blurring ....................................................................................................................................................................... 9
Blocking ...................................................................................................................................................................... 9
Jerkiness ...................................................................................................................................................................... 10
Damaged Blocks ....................................................................................................................................................... 10
Loss of Service .......................................................................................................................................................... 10
Loss of Audio-Video Synchronization ....................................................................................................................... 10

Video Quality Measurement .................................................................................................................................... 10
Subjective Quality Assessment .................................................................................................................................. 11
Objective Quality Metrics ........................................................................................................................................ 11
Full-Reference ........................................................................................................................................................... 11
No-Reference ........................................................................................................................................................... 12
Reduced-Reference ................................................................................................................................................... 13
Broadcast vs. OTT Video Quality Measurement ......................................................................................................... 14
QoS vs. QoE Measurement ...................................................................................................................................... 14
Network Impact on Video Quality ............................................................................................................................ 15
QoS Architectures ...................................................................................................................................................... 15

Avvasi Video Quality Scores ...................................................................................................................................... 16
OTT Protocol Support .................................................................................................................................................. 16
Independent Delivery and Presentation QoE Reporting ............................................................................................... 16
Q-VUE Delivery Quality Score .................................................................................................................................. 17
Q-VUE Presentation Quality Score ........................................................................................................................... 18
Effects of Adaptive Streaming on Quality Scores ..................................................................................................... 18
Q-VUE Combined Quality Score ................................................................................................................................ 19

Advantages of Q-VUE for Video QoE Measurement ................................................................................................... 20
Glossary ......................................................................................................................................................................... 21
INTRODUCTION

Over-the-Top (OTT) video traffic volumes continue to double every year and are predicted to represent a full two-thirds of all mobile traffic by 2014. This staggering growth rate is fueled by the unstoppable popularity of OTT video content from sites such as YouTube and Netflix, and the proliferation of smart devices with larger screens and more powerful processors which are well suited to play short- or long-form video content.

The lines between traditional television and mobile broadband are blurring as smart televisions support built-in Wi-Fi connectivity and smartphones and tablets support 1080p video and include HDMI output ports. This convergence not only drives a hybridization of media devices, it skews viewer expectations. Mobile viewers are quickly coming to expect the highly reliable, high-quality delivery they demand from traditional broadcast television or IPTV services.

Mobile and fixed broadband service providers face the challenge of satisfying subscriber expectations while struggling to manage the growing volume of OTT video traffic on their networks. While video services offer the promise for new revenue streams, poor video Quality of Experience (QoE) can quickly become a leading cause of subscriber churn.

As video traffic has grown on mobile and fixed broadband networks, the tools to measure the quality of video have not developed sufficiently to provide an accurate and scalable (i.e. network-wide) means of measuring how subscribers experience video and their resulting satisfaction levels. Traditional broadcast methods of measuring video quality are highly accurate but not scalable. Network QoS KPIs on the other hand are scalable but do not accurately reflect QoE. Conventional IPTV video quality measurement standards cannot cope with the multitude of technologies and standards in the OTT environment.

Furthermore, as video traffic growth continues to outpace increases in network capacity, video QoE is degrading. Figure 1 illustrates that during congested periods of the day, sessions exhibiting poor QoE can represent over 50% of overall video traffic on the network. With video traffic representing roughly 50% of all traffic, this means that—at times when bandwidth is most expensive—over 25% of it is arguably being wasted, as it is being used to deliver a poor experience.

This paper explores the need for a scalable technology that can provide perceptual QoE scores that accurately represent what every subscriber on the network is experiencing. Furthermore, the paper provides an explanation of how QoE can be measured for OTT video in an accurate and scalable manner, and includes background concepts relating to the OTT video lifecycle, streaming technologies and common video quality issues.

REFERENCES

OTT VIDEO: MEDIA STREAMING FUNDAMENTALS

With the popularity of OTT video sites, such as YouTube and Netflix, video has become pervasive and easy to access for all subscribers. But few people ever stop to consider how it gets created and eventually viewed on a mobile device.

OTT VIDEO CONTENT LIFECYCLE

The lifecycle of any video content begins with content creation. Once a video has been shot or captured, the content may be edited and one or more video clips may be authored from that edited content. Authoring typically involves compressing the video (and audio) to reduce the file size, followed by wrapping the compressed video (and audio) in a specific file type (or container). The file is then hosted on a (media) server or Content Delivery Network (CDN) where other users can access it. From the server or CDN, the file can then be streamed to a client, where it is decoded and played on a device for viewing.

The source, and therefore the quality of the content, can range from user-generated videos shot using a smartphone to major studio productions shot by a professional camera crew using commercial-grade equipment. With social media sites, video tends to be user-generated which means the content is typically shot using lower-grade equipment, and the resulting source quality tends to be lower.

Content is authored, either by a user who does their own editing, or automatically. The processes that automatically author the content are often hidden to the user as they are done directly on the capture device or transparently as part of uploading to the content provider. YouTube uses the latter approach and encodes the video using a codec, such as H.264. Encoding reduces the bandwidth required to stream the video content over a network in real-time. All of the information pertaining to this video (metadata, video and audio) is then put into a media container, such as FLV, MP4, MPEG2-TS or WebM. Some sites may encode and store the content in multiple formats, with different resolutions, bitrates, codecs and containers.

The video file is stored on a server or CDN where users can access it. The video file is either downloaded (as is the case with iTunes) or streamed (as is the case with Netflix). In order to watch downloaded video, the entire file must be received before playback can begin, which can take a long time. The file is stored ‘more permanently’ and is available for future consumption. For streaming video, playback begins almost immediately after the user requests the content. The file is typically stored in a ‘more temporary’ location, and is generally not available for future consumption.

Figure 2: The OTT video lifecycle
It is important to note that video files tend to be streamed and not downloaded so that playback can begin before the entire file has been received. When a subscriber requests the video content it is delivered from the content source by streaming across a packet network. The client buffers sufficient incoming data to enable continuous real-time decoding and playback. This process may sound simple enough, but it is complex and full of opportunities for issues to arise; issues that affect the quality of the delivery and/or the quality of the presentation. Even a small percentage of packet loss can cause quality degradation. Some streaming technologies include mechanisms for the device to switch to a lower bitrate profile (such as adaptive streaming) in order to use less bandwidth on the network and increase the probability of successful delivery.

**STREAMING OTT VIDEO**

In order to watch video on an internet-connected device, several components are required. A media server on the network (or CDN) provides a source for the video content, which is streamed over the network to an internet-connected device with a client, which can receive and display the video.

The content should be streamed in real time or faster. That is to say, the video data must be sent at the same rate or faster than the rate required to sustain real-time playback. When the content is streamed faster than the playback rate, video data accumulates in the client’s buffer. This buffering helps prevent playback interruptions such as stalling and can compensate for changes in network throughput. More recently, adaptive streaming technologies have been introduced which enable clients to respond to changes in network throughput and switch to lower bitrate streams when the network is congested.

With sufficient network throughput, a client receives the video data at a faster rate than playback. Therefore brief outages or reductions in throughput can be tolerated without impacting QoE, as long as the buffer stays full. However, during times of congestion or poor connectivity, the video buffer may become empty which will result in stalling (hourglass) and therefore poor QoE. If an adaptive streaming protocol is in use, the client can try switching to a lower bandwidth stream, which may reduce stalling, but will degrade visual quality through a reduction in resolution and bitrate.

There are numerous OTT video streaming technologies, below is a description of the most popular protocols.

**PROGRESSIVE DOWNLOAD**

Progressive download uses standard HTTP/TCP to stream (not download) content to the client as quickly as possible, maximizing buffering potential for smooth playback. This protocol is relatively simple and widely adopted across the Internet, most notably by YouTube. Progressive download is suited to unicast applications only and does not support multi-casting (for live events). The most popular container formats delivered via progressive download are FLV and MP4.
REAL-TIME STREAMING PROTOCOL
Real-time Streaming Protocol (RTSP), along with Real-time Transport Protocol (RTP) and Real-time Transport Control Protocol (RTCP), is commonly used to deliver live and on-deck content as well as Video-On-Demand (VOD) services. RTSP is used to establish and control the media session, to issue commands during the session, and is delivered over TCP. Most RTSP servers use RTP to deliver the media streams, typically over unreliable UDP. The media is therefore prone to significant quality degradation due to packet loss. Because of this, another approach called RTSP interleaved (which interleaves the RTP and RTCP data with the RTSP data) can be used. Instead of having one flow for RTSP and separate flows for audio and video tracks, a single RTSP/TCP flow is used. The RTSP data is sent as is, while RTP and RTCP are multiplexed through virtual channels.

REAL-TIME MESSAGING PROTOCOL
Real-time Messaging Protocol (RTMP) is a protocol developed by Macromedia (now Adobe) for streaming audio, video and data to a Flash player. Common variants include RTMPE (encrypted), and RTMPS, which works over an SSL connection. RTMP is supported by Flash Media Server and Flash clients.

ADAPTIVE OR DYNAMIC STREAMING
With adaptive streaming, the client detects network bandwidth availability and dynamically switches across multiple streams of differing bitrates to seamlessly deliver the content. These protocols are founded on the premise that smooth delivery is the biggest contributor to overall high video QoE. How the client decides which stream to select is specific to the client. Some clients are more aggressive and will select the best quality stream first; whereas, others are more conservative and will select lower-quality streams and monitor performance before switching to improve quality.

There are many examples of this technology including HTTP Live Streaming (HLS), HTTP Dynamic Streaming, Microsoft Silverlight Smooth Streaming, and Netflix Streaming Service.

The impact of dynamic streaming protocols is to offer a real-time tradeoff between the visual fidelity of the video and the throughput. However, due to the fact that clients only become aware of network congestion after the fact, dynamic streaming tends to be reactive and causes a high degree of visual quality variation, which in itself can lead to an overall worse QoE for the subscriber.

HTML5 VIDEO
HTML5 augments and expands the HTML standard to include a method to natively embed video on a website. This approach eliminates the dependence upon third-party browser plug-ins. HTML5 is supported by newer browsers such as Internet Explorer 9, FireFox 3.5, Safari 3.0, Chrome and Opera. While the standard is open, there are competing interests, the standard is in flux, and browser
vendors are free to support any video format they feel appropriate. YouTube uses HTML5 to deliver content to Apple iOS devices such as the iPhone and iPad. HTML5 video can use various container formats including WebM, MP4 and HLS.

**PEER-2-PEER TV**

Peer-2-Peer TV (P2PTV) delivers media over multiple peer connections. In P2PTV, each client downloads a video stream while concurrently uploading that stream to other P2P users. This approach is akin to a real-time BitTorrent. Streams are typically time-delayed by several minutes compared to the original source content. Video quality is a factor of the number of subscribers in the peer network, with quality improving as the number of subscribers increases. There are many P2PTV networks including PPLive, SopCast, StreamTorrent, Veetle, and SwarmPlayer.

**VIDEO CONFERENCING AND VIDEO CHAT**

Video chat applications such as Skype or Apple FaceTime are introducing a whole new set of OTT video use cases. The key difference between video chat and streaming is that video chat needs to be delivered at a very low latency in order to satisfy real-time two-way communication and it must be streamed bi-directionally. Popular video chat services include Skype (proprietary, RC4 encrypted signaling protocols) and Apple's FaceTime (SIP and RTP based).

**VIDEO QUALITY ISSUES**

Every step of the video content lifecycle can contribute to video quality issues, affecting the subscriber’s QoE.

**CAPTURE**

Poor video capture can be the result of a poor capture environment, e.g. lighting, a low-quality lens, poor focus, low resolution, camera motion, etc. With the exception of sophisticated pre- and post-processing of the captured video, it is very difficult for any future step in the lifecycle to improve the quality of poorly captured content.

**AUTHORING/ENCODING**

The authoring step can introduce additional quality issues due to the use of lossy compression algorithms, which are necessary in order to bring the required bandwidth down to usable levels for real-time streaming of the video content. For audio, these quality issues can appear as a result of reducing the sampling rate or number of channels relative to the original (captured) content as well as the codec itself. Some of these audio artifacts include ringing, echo, drop-outs and hissing. For video, these quality issues appear as a result of reduced bitrate, resolution or frame rate relative to the original (captured) content. Some of these artifacts include blocking, blurring, jerkiness, trailing artifacts, ‘mosquito’ noise, ringing, contouring, beating and breathing.

**TRANSMISSION**

There are two major network factors that affect video quality: congestion and connectivity. The volume of data required to deliver media (with acceptable QoE) is significantly more than for voice or other data forms such as email or static web content. The maximum amount of traffic that can be simultaneously delivered to subscribers represents the total capacity of the network. Based on the number of concurrent subscribers in a given cell sector, backhaul link or otherwise limited aggregation point, and the amount of network traffic that they generate, this can lead to congestion (where demand exceeds capacity). Congestion can lead to dropped packets, delayed delivery of data or even service interruption. For TCP-based (i.e. reliable) non-adaptive streaming sessions this can result in long delays in initial playback as well as stalling. For TCP-based adaptive streaming sessions this can also result in long delays in initial playback but the stalling is mitigated (though not necessarily eliminated) by switching to clips authored with lower fidelity and therefore lower bandwidth requirements.

On wireless networks, signal issues due to coverage, handoff, interference or resource contention can
lead to degraded throughput and therefore produce similar video quality issues as those that appear under congestion. In this case though, the quality degrades only for the subscribers experiencing the signal issues and not necessarily for all subscribers on that network node. Under congestion, all subscribers on a particular network node are impacted.

**PLAYBACK**

Due to the diversity of device types and display resolutions, the playback device itself has a significant impact on the subscriber’s perception of video quality issues. On smaller screens, artifacts become imperceptible at common viewing distances based on visual acuity limitations. Artifacts noticeable on an HD display can be imperceptible on smartphone-type devices. Increasing display resolutions on mobile devices for example, due to the emergence of tablet devices, increases the minimum fidelity (and therefore bandwidth) requirements necessary to satisfy subscriber video quality expectations.

**COMMON VIDEO QUALITY ISSUES**

As discussed above, there are many issues that contribute to poor video quality and affect the QoE. The most significant issues are described in more detail below.

**STALLING**

Stalling (sometimes referred to as re-buffering) typically occurs during reliably-delivered media sessions. When the network fails to deliver the media content in time for playback, due to insufficient throughput, playback will stall while the client waits for additional content. Generally the client waits to receive a certain amount of content in its internal buffer before resuming playback.

**BLURRING**

Blurring refers to a lack of sharpness and is the result of insufficient detail for the display size and resolution. It is often due to content encoded at low resolution that is displayed in high resolution on the playback device, e.g. in full-screen mode. As device screen sizes increase, content will need to be captured and encoded at higher resolutions (and therefore require more bandwidth) to mitigate this issue. Blurring can also be caused by pre-processing prior to encoding, where low-pass filtering may be applied to smooth the content details, permitting more aggressive compression. Finally, blurring can be caused by encoding stages, including quantization as well as deblocking and denoising filters.

**BLOCKING**

The main cause of blocking artifacts is the application of overly aggressive quantization by the encoding algorithm to blocks of pixels. Typically this occurs when the encoding algorithm is trying to compress to an aggressively low bitrate. The result is that finer details within blocks of pixels and subtle differences in the values of neighboring pixels are lost due to the removal of high frequency components in the transform domain. This makes all the pixels within the block (in the spatial domain) appear to have the same or similar value.
JERKINESS
Jerkiness refers to a class of temporal artifacts that result in a perceived lack of smoothness or continuity of motion. This can be due to a reduced frame rate during capture or encoding. This artifact will typically persist throughout the media session and the subscriber may grow accustomed to it and even accept it, particularly at lower resolutions and/or on smaller device screen sizes.

Other causes of jerkiness include loss of video content during unreliably-delivered (generally UDP-based) media sessions, frequent stalling during reliably delivered (generally TCP-based) media sessions, and other client- or device-related performance issues during playback. These result in more sporadic jerkiness, which is very difficult for a subscriber to grow accustomed to.

DAMAGED BLOCKS
When video content is lost or corrupted during transmission and cannot be retransmitted or reliably recovered (generally UDP-based), clients may take different approaches when it comes to displaying the pictures containing the damaged blocks. This is generally referred to as error concealment. On some clients, the damaged block may be omitted entirely, while on others it may be replaced by spatially or temporally neighboring blocks or combinations thereof. Worse, this initial artifact or discontinuity continues to propagate temporally and spatially until the entire picture is refreshed (e.g. via an I-frame).

LOSS OF SERVICE
Loss of service refers to the failure to deliver video content and is akin to dropping a voice call. This is caused by the congestion and signal issues previously described. Generally the media session is terminated and cannot be recovered.

LOSS OF AUDIO-VIDEO SYNCHRONIZATION
Audio-video synchronization issues are easily noticed in scenes with dialog where lip movements do not match the timing of audio delivery. This can be introduced in the capture or authoring stage, although this is rare. More often, this is due to packet loss in unreliably delivered (i.e. UDP-based) media sessions or client- or device-related issues during playback.

VIDEO QUALITY MEASUREMENT
Mobile and fixed broadband service providers face the challenge of satisfying subscriber expectations while struggling to manage the growing volume of video traffic on their networks. Video services offer the promise for new revenue streams. Inability to measure (and assure) video QoE makes it difficult to offer revenue-generating video services. Moreover, adding visibility into video QoE is key to managing increasing amounts of subscriber churn due to consistently poor QoE.

As OTT video traffic has grown on mobile and fixed broadband networks, the technologies and solutions used to measure video quality have not developed sufficiently to provide an accurate and scalable means of determining the level of subscriber satisfaction when consuming this OTT video. This section of the paper looks at how video quality has traditionally been measured and scored and the applicability of such techniques to the OTT video domain.
SUBJECTIVE QUALITY ASSESSMENT

The ‘gold standard’ for assessing media quality is subjective experiments. These represent the most accurate method for obtaining quality scores and ratings. In subjective video experiments, a number of viewers—typically 15-30—are asked to watch a set of clips and rate their quality. There are a wide variety of subjective testing methods and procedures, which are beyond the scope of this paper. The most common/concise way to reflect the result of the experiment is through the average rating over all viewers. Note that, in some cases, additional data processing, including normalization and outlier removal, may be required. This average rating is referred to as a Mean Opinion Score (MOS), shown in Table 1. One well-known application of MOS score principles is in the evaluation of voice call quality, based on various speech codecs and transmission parameters.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Quality</th>
<th>Impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Excellent</td>
<td>Imperceptible</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
<td>Perceptible, but not annoying</td>
</tr>
<tr>
<td>3</td>
<td>Fair</td>
<td>Slightly annoying</td>
</tr>
<tr>
<td>2</td>
<td>Poor</td>
<td>Annoying</td>
</tr>
<tr>
<td>1</td>
<td>Bad</td>
<td>Very annoying, unwatchable</td>
</tr>
</tbody>
</table>

Table 1: MOS scores

It is always challenging to quantify a qualitative characteristic because perception is individualistic and generally conveyed only as an opinion based on shared comparisons. Subjectivity and variability of viewer ratings cannot be completely eliminated. Subjective experiments try to minimize these factors through precise instructions, training and controlled environments. It is still important to remember that a quality score is a noisy measurement that is defined by a statistical distribution rather than an exact measurement.

OBJECTIVE QUALITY METRICS

Objective quality metrics are algorithms designed to characterize the quality of video and predict subjective quality or viewer MOS. There are a wide variety of objective quality metrics, from both academia and industry standardization activities. These metrics can be categorized as being full-reference, reduced-reference, or no-reference, based on the amount of information required about the reference video.

FULL-REFERENCE

Full-reference (FR) quality measurement techniques, illustrated in Figure 9, compare a transformed version of the video to a reference version of the video. The transformed version is typically the video as output from some system, which could be an encoder, transcoder, lossy channel or other video processing system, while the reference version is the input to the system. They operate in the spatial (i.e. pixel) domain as opposed to the compressed domain.

Figure 9: Full-Reference measure
These measures are generally very accurate at reflecting how closely the transformed video resembles the reference video, and some of the more complex methods also try to find common artifacts such as blocking, blurring and related artifacts. Typically, none of the other lifecycle stage impairments are accounted for.

When measuring video quality in a lab, this approach makes sense for several reasons:

- Scalability is not required and computational complexity can be very high as the measurement is being performed on a few streams.
- The reference video is usually accessible.
- Precise, often manual, spatial and temporal alignment of the reference and transformed video can be performed.
- The delivery network is reliable and uncongested. Therefore accounting for transmission impairments is not a requirement.

However, when measuring OTT video in a network, the above conditions do not apply.

- There are potentially many concurrent video sessions, so scalability to an entire subscriber base is required, thus computational complexity must be constrained.
- Access to the reference video is difficult if not impossible.
- Automatic spatial-temporal alignment is error-prone and computationally expensive.
- The delivery networks are much less reliable and often congested, thus it is important to incorporate transmission impairments into the quality measurement.

Several popular full-reference measures are described below. They all operate in the spatial domain and require access to the reference video. As such, they have all the deficiencies identified above related to full-reference measures when it comes to measuring the quality of OTT video.

**PEAK SIGNAL-TO-NOISE RATIO**

Peak Signal-to-Noise (PSNR) ratio is a measure that quantifies how much a signal is degraded or corrupted by distortion or ‘noise’; the higher the ratio, the better the quality of the signal. In the case of compressed video, the distortion is the loss of information introduced by a lossy encoding process. In the case of transmission channels, the distortion is the loss of information introduced by a lossy channel.

**STRUCTURAL SIMILARITY INDEX**

Structural Similarity Index (SSIM) is based on the principle that a human visual system is highly trained to identify shapes, thus the metric focuses on the amount of structural similarity between video frames. This is counter to most other metrics (like PSNR) which are based on differences between images. This score is typically in the range of -1 to 1, with 1 being the best score.

**VIDEO QUALITY METRIC**

Video Quality Metric (VQM) is a video quality measure based on a human visual model of perceived effects of blurring, jerkiness (local and global), noise, and video distortions. It is computationally intensive procedure composed of four steps: calibration, feature extraction, parameter calculation, and final score calculation. A unique VQM model is used for different scenarios, for example two different models are needed for television and video conferencing.

**PERCEPTUAL EVALUATION OF VIDEO QUALITY**

Perceptual Evaluation of Video Quality (PEVQ) is another video quality measurement based on a human visual model of the perceived effect of spatial and temporal distortions. It is another very computationally intensive procedure, composed of similar steps as VQM: calibration/alignment, perceptual difference calculation, classification of differences, and final score calculation. It has been standardized as part of ITU-T J.247. It employs five indicators and uses region-of-interest (ROI) to limit complexity.

**NO-REFERENCE**

No-reference (NR), also referred to as zero-reference, quality measurement techniques, illustrated in Figure 10, do not compare transformed to reference content. Rather, no-reference techniques estimate quality by analyzing only the post-encoded content, using algorithms and heuristics that are based on indicative encoding parameters and/or inferred
Measuring Quality of Experience for Over-the-Top Video Services

encoding artifacts. There are two sub-categories of no-reference approaches:

- Bitstream-based methods, which typically parse various headers and payloads to varying depths.
- Pixel-based methods which fully decode the compressed video to baseband, are superior at detecting and quantifying encoding artifacts.

No-reference measures are not as accurate as full-reference, however they are generally less computationally complex and are therefore more scalable in terms of deployment in a service provider network. One attractive attribute is that computational complexity can be traded off against accuracy by controlling the depth of parsing. Access to reference content is not a requirement. Similar to full-reference measures, conventional no-reference measures do not account for transmission impairments. However, given their relatively low computational complexity, they can be extended to incorporate network impairments and still provide acceptable scalability and performance.

There are many no-reference approaches currently under study and/or development, both within standardization bodies as well as academia. Within the standards community, ITU-T SG-12 P.NAMS and P.NBAMS are developing non-intrusive parametric models for the assessment of performance of multimedia streaming. The former uses only header information while the latter uses the codec bitstream. None have been approved or widely adopted to this point.

**REDUCED-REFERENCE**

Reduced-reference (RR), also referred to as partial-reference, quality measurement techniques, illustrated in Figure 11, are a compromise between

![Figure 10: No-Reference measure](image1)

![Figure 11: Reduced-Reference measure](image2)
the full-reference and no-reference approaches, in which only partial information about the reference video is available for quality estimation. Reduced reference can be quite suitable in situations where the overhead of storage or transmission of the reference video is prohibitive but the accuracy of a no-reference approach is too low. The disadvantage of this technique is that additional storage or transmission of side information is necessary. This side information typically includes parameters summarizing the quality of the reference video. The main advantage of this approach is that the quality parameters for the reference video are computed only once, making it much more scalable than full-reference approaches. The main problem with reduced-reference quality measurement is that most OTT video does not contain the side information required by this approach.

**BROADCAST VS. OTT VIDEO QUALITY MEASUREMENT**

Some of the key differences between broadcast video and OTT video are outlined in Table 2 below. These differences illustrate why expensive and complex solutions that make sense in the broadcast video world are not viable for OTT video, outside of isolated use cases (e.g. lab trials or ‘shoot-outs’).

Effective OTT video quality measurement requires normalization of quality scores for a wide variety of content, streaming technologies, and display devices. Timely support of emerging formats and devices is essential. Consideration of the impact of network impairment on QoE is crucial. These requirements suggest a scalable, low-complexity approach, i.e. no-reference.

**QOS VS. QOE MEASUREMENT**

Quality of Service (QoS) metrics or Key Performance Indicators (KPIs) are often used interchangeably with or to infer QoE. For example, if enough data is delivered in time (high network throughput) then a high QoE may be inferred. While there is a relationship between QoS and QoE, it is not a direct relationship. Good QoS KPIs often occur yet QoE is unsatisfactory. QoS measurement tends to be focused on the quality

<table>
<thead>
<tr>
<th>Element</th>
<th>Broadcast Video</th>
<th>OTT Video</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source content and formats</td>
<td>• Limited</td>
<td>• Vast</td>
</tr>
<tr>
<td></td>
<td>• Small set of known standards</td>
<td>• Large number of changing ‘standards’</td>
</tr>
<tr>
<td>Delivery channel</td>
<td>• Broadcast</td>
<td>• Unicast</td>
</tr>
<tr>
<td></td>
<td>• Reliable</td>
<td>• Unreliable</td>
</tr>
<tr>
<td></td>
<td>• Uncongested</td>
<td>• Congested</td>
</tr>
<tr>
<td>Ecosystem</td>
<td>• Unified</td>
<td>• Fragmented</td>
</tr>
<tr>
<td></td>
<td>• Single entity has end-to-end control</td>
<td>• Multiple competitive entities serving user</td>
</tr>
<tr>
<td>Client viewing device</td>
<td>• Limited number of vendors</td>
<td>• Large number of vendors</td>
</tr>
<tr>
<td></td>
<td>• Static</td>
<td>• Rapidly changing capabilities</td>
</tr>
<tr>
<td></td>
<td>• Similar</td>
<td>• Widely varying devices</td>
</tr>
<tr>
<td></td>
<td>• Small set of known standards</td>
<td>• No unified standards</td>
</tr>
<tr>
<td>Measurement Solution scalability requirements</td>
<td>• Can be deployed “per-channel”</td>
<td>• Should be deployed “per concurrent subscriber”</td>
</tr>
<tr>
<td></td>
<td>• Low scalability needs</td>
<td>• High scalability needs</td>
</tr>
<tr>
<td>Measurement solution computational requirements</td>
<td>• Tightly constrained feature set</td>
<td>• Wide and evolving feature set</td>
</tr>
<tr>
<td></td>
<td>• High complexity remains cost-effective</td>
<td>• High complexity no longer cost-effective</td>
</tr>
</tbody>
</table>

| Table 2: Broadcast vs. OTT and video quality measurement considerations | 14 |
of a network, whereas QoE tends to be focused on user intent and performance of an application (video playback). User intent and expectation vary significantly depending on the characteristics of the client device, the type of content they are viewing and many other variables.

**NETWORK IMPACT ON VIDEO QUALITY**

Traditional IP networks (without an end-to-end QoS architecture) provide best-effort service over a common, shared infrastructure. Any link or node in the network can experience congestion. The primary mitigation strategy is to drop packets proactively. Reliable networking protocols account for this and have built-in congestion avoidance algorithms and retransmission mechanisms.

Wireless networks pose additional problems due to radio transmission issues. This leads to increased latency and packet loss. In TCP-based delivery, this manifests itself as sender-side timeouts and retransmissions. In UDP-based delivery it appears as severe client-side jitter and packet loss are experienced. Packet delay, loss and retransmissions can significantly impact video delivery as the throughput within a session can vary greatly depending on conditions within the network and on the radio link.

In TCP-based delivery, poor network conditions during media streaming cause low network throughput, which, in turn, may cause the client’s media input buffer to empty during playback. Essentially, some segment of media data does not arrive at the client in time to be played back at the appropriate time. In this case, playback stalls, waiting until enough content has been received and the buffer replenished. For TCP-based protocols that do not allow the quality level to adapt to the available throughput, delays in playback, either at the beginning of the session, or in the form of stalls during the session, are the only symptoms of an impaired network that the subscriber can observe. Therefore, the details of these delays and stalls are the only factor that matters in determining how the network has impacted the QoE.

For adaptive streaming protocols delivered over TCP, network congestion will cause the client to request a lower bitrate stream, which will result in a reduced amount of stalling, and a reduced visual fidelity of the video.

For unreliable UDP-based delivery, in the event of packet loss, there are no retransmissions, resulting in many different quality issues. Dropped, corrupted, and late packets are effectively also lost. These issues manifest themselves as damaged blocks, loss of audio-video synchronization, stalling, jerkiness, or combinations thereof.

**QOS ARCHITECTURES**

Network QoS refers to the ability of a network to provide traffic prioritization and resource management mechanisms. These traffic management mechanisms are required by applications that intend to deliver real-time or near-real-time services, such as voice over IP (VoIP) and latency-sensitive online gaming and video communications services. This is true in general, but especially true in times of network congestion when the probability of packets being delayed or dropped is much higher. For QoS to work, an end-to-end approach is required where all nodes that the data traverses must have appropriately configured mechanisms to achieve the desired characteristics for the application data flow.
Measuring Quality of Experience for Over-the-Top Video Services

It is possible to define, configure and manage a QoS architecture within a single network domain, for example, a broadband access network where all nodes are under the control of a single administration. In the case of OTT video, the video server or CDN node is usually located outside the service provider network and the client is connected to the Internet via an Internet service provider. The video data will transit via several autonomous networks on the way from the server to the client. QoS mechanisms are not generally deployed in the Internet core and do not transit across administrative domains either. Access networks do not generally provide methods for applications or servers to request network QoS from them. Therefore, achieving an end-to-end QoS control is not currently possible in OTT video delivery. For this reason most networks (with the exception of enterprise networks) are only using best-effort delivery and rely on over-provisioning of resources to ensure that packets do not suffer unnecessary delays or get dropped. However, there are no guarantees, and often, in times of peak usage, network links can become congested. Transmission issues can be experienced anywhere on the path between the client and the server. As such, these issues are difficult to locate and troubleshoot and can often be outside the control of the service provider.

AVVASI VIDEO QUALITY SCORES

Traditional methods of measuring video QoE are not suitable for OTT video since they are too computationally intensive for real-time monitoring of unicast streaming services, and do not address the variety of streaming protocols and devices in the OTT ecosystem. Avvasi’s patent pending Q-VUE system supports all major OTT protocols and enables real-time monitoring of video sessions and associated video quality over an entire network, including the largest tier-1 broadband and mobile networks, and the technology.

Separate delivery and presentation quality scores, and a combined quality score (along with extensive demographic information), are reported for every media session. Avvasi’s video quality scores are objective measures of subjective quality that correlate to traditional MOS scores for video as previously described. These quality scores can be integrated offline into the service provider support systems or in real-time into service provider charging and policy systems, to enable powerful use cases for operations, marketing and services teams.

OTT PROTOCOL SUPPORT

Avvasi’s OTT video QoE measurement technology supports all of the most popular OTT video protocols, including progressive download, partial progressive download, HTTP chunked download, RTMP, HLS, Silverlight Smooth Streaming, Netflix Streaming, RTP/RTSP, and others. Avvasi’s combined quality score provides a normalized QoE score that has the same meaning no matter which protocol is used. This enables a rich set of analytics use cases that allows service providers to view their network traffic broken down by QoE across all protocols.

INDEPENDENT DELIVERY AND PRESENTATION QOE REPORTING

In order to deploy monetized video services, the ideal service assurance technology provides an overall end-to-end video QoE score that encompasses the video quality as streamed from the source (CDN or streaming server), the degradation of the quality by the delivery network, and the impact of the device screen size. The end-to-end QoE score provides a measure of the user’s satisfaction with the service, considering all measurable factors that contribute to the experience.

In today’s fragmented OTT video ecosystem, there is a lack of collaboration between service
providers, content providers and aggregators, and device manufacturers. Therefore, it is important to understand QoE from several different perspectives. Avvasi has developed multiple perceptual QoE scores, including a Delivery Quality Score, a Presentation Quality Score, and a Combined Quality Score, described further below.

**Q-VUE DELIVERY QUALITY SCORE**

Avvasi’s patent pending Delivery Quality Score, or DQS, reflects the impact of network delivery on QoE and purposely ignores the source quality. Using deep packet inspection (DPI) techniques, media flows are monitored and inspected for all subscribers and sessions in real time. For each media session, a patent pending model of the client buffer fullness and player state is maintained. It is therefore possible to determine time, duration, and frequency of stall events which, as described previously, are the primary symptoms of network impairment that the subscriber can observe in the case of reliable (i.e. TCP-based) delivery.

Figure 13 illustrates buffer fullness over time for two sample media sessions. One media session containing stall events is labeled Poor DQS. Another media session containing no stall events is labeled Good DQS. The figure is also annotated with the player state for both media sessions, including Buffer Initialization, Playback, and Stall.

Under all network conditions a player begins in the Buffer Initialization state, filling with enough video data (frames) to provide some minimum amount of uninterrupted playback. Once enough data is accumulated, playback commences. During the Playback state, the client buffer is simultaneously being filled (data arriving via the network) and drained (data consumed via playback). Based on the difference between the fill and drain rates, the client buffer fullness increases or decreases over time.

In this figure, the client buffer for the Good DQS media session always contains sufficient data to support continuous playback, thus the DQS for this media session will be high. On the other hand, on two occasions, the client buffer for the Poor DQS media session empties completely, causing playback to stall. As discussed previously, stalls are the primary symptoms of network impairment that subscribers observe, thus the DQS for this media session will be low. Note that, in the stall state, a player typically requires some meaningful amount of video data to accumulate in its buffer (similar to during the Buffer Initialization state) prior to resuming playback, so that some further minimum amount of uninterrupted playback can be provided.

These events are incorporated into a patent pending subscriber model of satisfaction with the media session, which is used to compute the DQS. Avvasi’s approach applies best practices from professional video broadcast standards, including the use of subjective testing methods using test subjects to evaluate viewer satisfaction with the video session,
and is based on a MOS score of 1-5, as described previously. These results are then correlated to the subscriber model, which allows for automatic and objective measurement of the subjective quality. Some observations that must be incorporated into the model include:

- A stall of the same duration at the start or at the end of a video session does not have the same impact on QoE.
- Large numbers of stalls at the beginning of a video session have a different impact on QoE than regularly occurring stalls throughout the session.

There are many other examples and observations that must be incorporated. Simply reporting on the overall number of stalls or stall frequency per playback minute is insufficient to provide a reliable representation of QoE. The model must be tested with and correlated to numerous artifact scenarios, using a representative sample of viewers, to arrive at an accurate Delivery Quality Score.

**Q-VUE PRESENTATION QUALITY SCORE**
Avvasi’s patent pending Presentation Quality Score, or PQS, is a measure of the quality of the media session with respect to the display device, and purposely ignores the network impact. Using DPI techniques media flows are monitored and analyzed for all subscribers and sessions in real time. For each media session certain relevant KPIs are extracted, such as:

- Video codec resolution, bitrate, frame rate, motion vectors and quantization parameters.
- Audio codec, bitrate, sampling rate and number of channels.
- Playback device and related screen size.

These parameters are incorporated into a no-reference bit-stream model of satisfaction with the audio-visual quality of the media session. Avvasi’s approach applies best practices from professional video broadcast standards, including the use of subjective testing methods to evaluate viewer satisfaction with the video session, and is based on a MOS score of 1-5. Figure 14 illustrates an individual frame from a sample clip at two different quality levels, and the associated PQS scores. These results are then correlated to the subscriber model, which allows for automatic and objective measurement of the subjective quality. Some observations that must be incorporated into the model include:

- Typically, audio fidelity is reasonably high. However, poorly sampled audio or a low number of channels significantly degrades QoE even with very high-quality (i.e. HD-like) video.
- Often, the exact same media session produces a significantly different (lower) QoE when displayed on a larger, tablet-style device versus a smaller, smartphone-style device.

There are many other examples and observations that must be incorporated. The model must be tested with and correlated to numerous types of content and display devices, using a representative sample of viewers, to arrive at an accurate Presentation Quality Score.

**EFFECTS OF ADAPTIVE STREAMING ON QUALITY SCORES**
When measuring video QoE for adaptive streaming protocols, such as HLS, there is an additional challenge in that delivery and presentation quality are being altered in real time in response to network conditions, whereby the player is trying to achieve the optimal balance.
Adaptive streaming technologies rely on the content being provided at several quality levels. Generally, adaptive streaming players begin playback at a low quality level. This allows for a low startup delay, and also lets the player learn the current network throughput. If the player estimates that the network throughput can sustain continuous playback at a higher bitrate (i.e. quality level), then the player will switch to a higher quality level. The estimation algorithm that determines switching is player dependent.

Figure 15 illustrates two media sessions for a content source available at three bitrates and corresponding quality levels: 0.6 Mbps (low), 1.4Mbps (medium), and 2.0 Mbps (high). The media session labeled “Good QoE” has ample throughput. It quickly switches to and maintains the highest available quality level, thus the QoE for this media session will be high. The media session labeled “Poor QoE” experiences reduced throughput. It can only switch to higher quality levels for short periods of time. Due to the overall lower quality and multiple noticeable quality switches, the QoE for this media session will be low.

**Q-VUE COMBINED QUALITY SCORE**

The combined quality score represents the complex interactions of playback continuity and source quality relative to the display device and, therefore, is a function of both the Delivery and Presentation Quality Scores. It conveys the user’s overall satisfaction with the media session, considering all measurable factors that contribute to the experience.

Figure 16 illustrates in a simplified manner how media session bitrate affects the Presentation and Delivery Quality Scores and, therefore, the combined quality score. A higher encoding bitrate generally produces a higher Presentation Quality Score. However, this imposes greater requirements on network throughput. Often, the higher encoding bitrate tends to increase network impairments, and therefore produces a lower Delivery Quality Score for the media session. On the other hand, streaming a media session with a lower encoding bitrate and, therefore, a lower Presentation Quality Score, tends to improve the Delivery Quality Score.
ADVANTAGES OF Q-VUE FOR VIDEO QOE MEASUREMENT

The list below summarizes the advantages of Avvasi’s approach to QoE measurement.

- **Clientless**: Q-VUE computes multiple quality scores without having to install software on the client device. This means Avvasi analytics are scalable and can measure the QoE across multiple clients in real time, and on a large network.

- **No-Reference**: Q-VUE quality scores are calculated using a no-reference, bitstream-based approach, which means they do not require original source video, and are therefore low complexity and highly scalable for large network deployments.

- **Accurate Measurement**: Q-VUE quality scores measure the original source content quality, the effects of any network degradation, and account for the impact of device screen size and capabilities. These scores are objective quality metrics, and correlate well to the subjective quality or MOS scores obtained from subjective testing.

- **Separate and Combined Scores**: Q-VUE quality scores are reported as separate scores and a combined score, to enable service providers and content providers to understand in detail what aspects of the ecosystem are impacting the end-to-end QoE. This supplies actionable information to address QoE issues in the field.

- **Scalable Network- and Traffic-Wide**: Q-VUE scales easily to an entire network, including the largest tier-1 broadband networks. It is not based on a sampling approach. The technology supports all major OTT protocols.

- **Rich Analytics Integration**: Q-VUE provides QoE scores integrated into a rich analytics suite such that QoE metrics can be viewed by device type, media site, network location, subscriber, and many other parameters. This enables powerful business intelligence, network and service planning, troubleshooting, and service assurance use cases.

- **Troubleshooting Integration**: Avvasi enables QoE score-reporting in near real time through a Usage Detail Record (UDR) technique that can be integrated with troubleshooting tools to help service provider operations teams to address QoE issues proactively, quickly and efficiently.

- **Policy Control Integration**: Avvasi provides QoE score reporting that can be integrated in real time with policy control solutions to enable advanced service creation based on QoE SLAs.

- **CRM Integration**: Avvasi provides QoE score reporting that can be integrated into CRM systems to enable correlation of QoE with subscriber churn.
GLOSSARY

Adaptive streaming  Streaming technologies that dynamically adapt the delivered video quality by switching between streams of varying bitrates to maintain continuous playback while maximizing quality under changing network conditions. Popular implementations include HTTP Live Streaming, Adobe Dynamic Flash Streaming, 3GPP DASH, Microsoft Smooth Streaming. Also referred to as Dynamic Streaming.

Adobe Flash  Formerly Macromedia Flash (sometimes referred to just as Flash), is a dominant form of web content, typically installed via a browser plug-in. Flash is a multimedia platform that is used to add animation, video, and interactivity to web pages.

Audio-video synchronization  A media playback quality issue where playback timing of the audio track does not align to the action in the video track.

Blocking artifact  A common quality issue in compressed video where a block of neighboring pixels that are coded as a single entity are displayed having the same or similar values.

Blurring  A common video quality issue characterized by a lack of sharpness that is generally due to insufficient video resolution for the size of the display area. It is often the result of a subscriber choice to expand low resolution content to full-screen mode for playback, but may also be caused by pre-processing prior to encoding or encoding stages, including quantization as well as deblurring and denoising filters.

Client Buffer/ Buffering  A segment of memory used by a media playback application to temporarily store received media data prior to decoding and playback, as a means of steadying the rate of output for a fluctuating input rate.

Codec  A physical implementation of the encoder-decoder capable of reading (decoding) and/or writing (encoding) streams compliant with the syntax of a given standard. Popular video codecs include H.264 (MPEG-4 AVC), MPEG-2, MPEG-4, and On2 VP6.

Container  A meta-file format that specifies how to identify and interleave the various multimedia components. Content includes the encoded video and audio streams, and metadata such as thumbnails and author information. Popular containers include Audio-Video Interleaved (AVI), Flash Video (FLV), Windows Media (WMV), QuickTime (MOV) or MP4.

Decoding  The process by which a codec implementation decompresses media content for playback.

Delivery QoE  A measure that reflects how the network delivery impacts subscriber QoE and purposely ignores the source quality. For reliable (i.e. TCP-based) delivery, it is based on a model of the client buffer state and an estimation of playback delays and stall events that are experienced by the subscriber.

Download  A broad category of video content transmission by which the entire file must be transferred to the client and stored locally before the file can be played. These files are typically stored ‘more permanently’ and are available for future consumption.
<p>| <strong>DPI</strong> | Deep Packet Inspection. A form of packet processing that examines the header and the application payload of the packets as they pass through an inspection point. Conventional DPI focuses primarily on packet headers and signatures, searching for protocol non-compliance, viruses, spam, intrusions or other predefined criteria. |
| <strong>Encoding</strong> | The process by which a codec implementation compresses media content for compact storage or transmission. |
| <strong>FaceTime</strong> | A video calling software application and related protocol developed by Apple for supported mobile devices running Apple iOS (using SIP and RTP). |
| <strong>Flash Video (FLV)</strong> | A dominant container designed for streaming audio/video content. |
| <strong>Frame rate</strong> | The rate at which a series of images (or frames) of a video are displayed. Television uses several frame rates including 25, 30, 50, 60 frames per second (fps), whereas film content commonly uses 24 fps. Internet content may use any of these frame rates and sometimes uses lower frame rates, such as 12 or 15, especially for smaller resolutions and low bitrate encoding. |
| <strong>Full-reference</strong> | A broad category of techniques used to measure video quality by comparing a transformed version of the content to a reference version of the content. |
| <strong>H.263</strong> | ITU H.263, which forms the most basic profile of MPEG-4 Part 2. A legacy video codec still used for some content on YouTube. |
| <strong>H.264</strong> | ITU H.264 / MPEG-4 Part 10. Currently the most popular video codec used in most media streaming applications. It provides significant improvement in coding efficiency over earlier codecs, such as H.263, MPEG-2 and MPEG-4 Part 2. |
| <strong>HLS</strong> | HTTP Live Streaming. A media streaming communications protocol implemented by Apple as part of their QuickTime X and iOS software systems. HLS segments the stream into a sequence of small HTTP-based file downloads. As the stream is played, the client may select from a number of different alternate streams containing the same material encoded at a variety of bitrates, allowing the streaming session to adapt to the available throughput. HTTP Live Streaming is capable of traversing any firewall or proxy server that lets through standard HTTP traffic, unlike UDP-based protocols such as RTP. |
| <strong>HTML5</strong> | The fifth revision of the HTML standard that defines a language for structuring and presenting Web content. In particular, HTML5 adds a video tag to directly embed video into Web content without the need for third-party browser plug-ins. HTML5 has been hampered by a disagreement about which media formats should be supported. |
| <strong>IPTV</strong> | Internet Protocol TV. A system through which television services are delivered over a packet-switched network instead of traditional broadcasting or satellite networks. |
| <strong>Jerkiness</strong> | A common video quality issue where the video appears to rapidly start and stop, caused by a low frame rate or insufficient network throughput. |
| <strong>Jitter</strong> | The variation in latency causing some packets to take longer than others to reach their destination or arrive out of order. |
| <strong>Loss of service</strong> | A catastrophic failure to deliver video content caused by severe network bandwidth limitations, intra-service hand-offs or dead spots. |</p>
<table>
<thead>
<tr>
<th><strong>Macroblock</strong></th>
<th>A square of pixels within a frame of video, grouped together for compression purposes. Macroblocks are commonly 16x16 luma samples in size.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Macroblocking</strong></td>
<td>A common video quality issue where colorful or dark artifacts appear as a result of a partial loss of the data required to reconstruct the image.</td>
</tr>
<tr>
<td><strong>MOS</strong></td>
<td>Mean Opinion Score. The numerical identification of the perceived quality of presented audio or video content after compression and/or transmission. This number is generally the average of the results of subjective listening and/or viewing experiments, involving a significant number of subjects, typically 15 to 30. It is important to remember that a quality score is a noisy measurement that is defined by a statistical distribution rather than an exact measurement. Objective quality metrics are algorithms designed to characterize the quality of audio and/or video and predict subjective quality or MOS.</td>
</tr>
<tr>
<td><strong>MP4</strong></td>
<td>A media container file format defined in MPEG-4 Part 14. Due to per-frame metadata that must be transmitted as a single syntax element for the entire file, the format is best suited for storage and download and not ideal for streaming.</td>
</tr>
<tr>
<td><strong>Network Latency</strong></td>
<td>The delay in time that a packet takes to traverse the network from one host to the other.</td>
</tr>
<tr>
<td><strong>No-reference</strong></td>
<td>Also referred to as zero-reference or reference-free, this is a measure of video quality which does not require reference content for comparison.</td>
</tr>
<tr>
<td><strong>OTT Video</strong></td>
<td>Over-the-Top video. Refers to data which originates from a non-operator source, but travels over the operator’s network. Popular source include YouTube and Netflix.</td>
</tr>
<tr>
<td><strong>P2P TV</strong></td>
<td>Peer-2-Peer TV. A broad category of services that use multiple peer connections to deliver video content to members of the peer community.</td>
</tr>
<tr>
<td><strong>Pixel</strong></td>
<td>A pixel is the smallest representable unit for graphic displays and in digital images. The density of pixels arranged on a screen determines the display resolution.</td>
</tr>
<tr>
<td><strong>Presentation QoE</strong></td>
<td>A measure of the perceived audio/video quality of content presented in a media session with respect to the playback device that purposely ignores the network impact.</td>
</tr>
<tr>
<td><strong>Progressive download</strong></td>
<td>The most commonly used streaming protocol, which uses standard HTTP/TCP protocols to deliver video content while allowing playback to begin prior to the completion of the download.</td>
</tr>
<tr>
<td><strong>PSNR</strong></td>
<td>Peak Signal-to-Noise Ratio. A full-reference measure of signal quality that quantifies how much a signal is degraded or corrupted by distortion or ‘noise’; the higher the ratio, the better the quality of the signal. In the case of compressed video, the distortion is the loss of information introduced by the lossy encoding process.</td>
</tr>
<tr>
<td><strong>QoE</strong></td>
<td>Quality of Experience. Refers to a subjective opinion of a customer's experience with a particular service. This paper is primarily focused on QoE as related to video, specifically OTT video.</td>
</tr>
</tbody>
</table>
QoS  Quality of Service. Refers to the ability to prioritize and manage traffic resources on a network for data flows associated with a higher level application or service. Various QoS mechanisms exist and are typically implemented at the OSI layer 2 or layer 3 levels. For IP networks Diffserv is the prevalent architecture versus the legacy Intserv model.

Resolution  Refers to the spatial dimensions of a video frame, image or display device. Resolution is typically measured as the width (in pixels) by the height (in pixels).

RTCP  Real-time Transport Control Protocol. A network control protocol that is used in conjunction with RTP to provide out-of-band control information and statistics about the delivery of the media data in the associated RTP flows. For example, RTCP may be used to provide the server statistics on the amount of packets lost from the RTP flows.

RTP  Real-time Transport Protocol. RTP is a standardized packet format used for delivering multimedia data over IP networks in real time. It is most commonly delivered over UDP, and supports both multicast and unicast. RTP is often used in conjunction with RTSP and RTCP to provide streaming services commonly used to deliver on-demand content of live events.

RTMP  Real-Time Messaging Protocol. A protocol developed by Macromedia (now Adobe) for streaming audio, video and data to a Flash player. Common variants include RTMPE, which is encrypted, and RTMPS, which works over an SSL connection.

RTSP  Real-time Streaming Protocol. An application-layer protocol used to establish and control media sessions by allowing clients to issue commands, such as pause and play, to a media server. The actual media data is transmitted in RTP flows that are advertised to the client and initiated via the RTSP conversation.

Stalling  A common video quality issue where the video freezes during playback, typically caused by insufficient network throughput.

Streaming  A broad category of video content transmission by which playback can occur during the download process. The content is typically stored in a temporary location, and is generally not available for future consumption.

VP6/VP7  A pair of proprietary video codecs developed by On2 Technologies. VP6 was the primary codec used for Flash Video starting in Version 8. Later versions of Flash added support for H.264. Use of VP6 is now fading as Google no longer supports the codec since the On2 acquisition. VP7 was an extension of VP6 that has been used for Skype video.

VP8  Originally a proprietary video codec developed by On2 to replace VP7, the project was open-sourced by Google after they acquired On2 in 2010. The codec was put forward by the WebM project to be used for HTML5, along with Ogg Vorbis audio and the WebM container. Its coding efficiency is generally considered to be slightly worse than H.264.

WebM  An open-source media container format based on the Matroska format and developed by the WebM project for use in HTML5. WebM files contain VP8 video and Ogg Vorbis audio.
ABOUT AVVASI
Avvasi is the award-winning company helping service providers assure, improve and monetize video. Service providers are deploying Avvasi solutions on the Xperium platform to understand the network impact of new devices and subscriber usage patterns, plan and dimension network expansion, while reducing subscriber churn and maximizing revenue. Based in Waterloo, Ontario, Avvasi is a venture backed, privately held company.