WHAT DOES "GOOD" LOOK LIKE?

Toward an industry standard for measuring OTT video quality

Akamai

AKAMAI.COM

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This is the second phase of a broader research project from Akamai, the leading content delivery network, investigating issues around the quality of IP-delivered video streams.

The first phase was an extensive study undertaken on behalf of Akamai by Sensum into how we as viewers react to video quality; it focused on how quality of experience impacts commercial success. The resulting white paper demonstrated the impact of buffering, for example, on viewer engagement.

This new paper extends the broader project by asking what "good" actually looks like in terms of streamed video. Akamai worked with Eurofins Digital Testing to examine how the perceived quality of streamed video is impacted by a series of factors, including the video device or player, the content genre viewed, the encode profile and finally, the network conditions typically experienced by users.

The initial tests established the optimal profile per device and per content genre. The next tests introduced challenging network conditions — including bandwidth variance, packet latency, and outages — to determine their impact on these profiles. The final element was to analyse the key quality indicators, noting the significance of consistency of throughput — with less buffering on all devices at the highest bitrates.

In summary, the results provided a clear indication among optimal video fidelity, content genre, and viewing device, with screen size a notable factor. Noting the bitrates required and benchmarking against Akamai's data on the state of the Internet across multiple territories, the results also showed that the current Internet infrastructure is capable of delivering high-definition complex video content to most viewer devices in most Western European markets with little or no buffering.

As the IP video market matures, future phases of the research will explore other challenges around delivering high-quality video, including the need to scale efficiently. Akamai welcomes further discussion and input into this research from clients and industry stakeholders.

INTRODUCTION

Video quality matters.

IP-delivered video has, in the last decade, become an increasingly significant part of the media landscape. Despite linear broadcast still accounting for the majority of long-form episodic TV viewing today, there is an inexorable shift by audiences in many countries towards IP delivering much of their entertainment mix.

This trend poses a number of critical questions for the TV and video industry. Is the Internet capable of delivering high-quality video streams consistently and at TV-audience scale? From a viewer's perspective, can video delivery via the Internet exceed the viewing experience enjoyed through linear TV? And from a distributor's perspective, can good video be delivered in a cost-efficient way while improving the way an audience engages with entertainment programming?

It might help if the industry could agree on what "good" looks like when it comes to video delivered over the internet. We know that quality of Experience (QoE) is important — to viewers, to advertisers, and to OTT distributors. Akamai's recent research with Sensum demonstrated — using a variety of sophisticated experiential research techniques — the link between high-quality video streaming and viewer engagement in story lines. In summary, viewers watching higher fidelity video streams without being interrupted by buffering are more engaged emotionally in a story and spend more time watching that content. They are also less likely to abandon either a piece of video or, indeed, a subscription service. Moreover, QoE with streaming has positive impact on brand perception and the propensity for viewers to recommend a particular service.

VIDEO QUALITY IMPACTS THE BOTTOM LINE

Of course, such findings have a clear commercial implication for providers of OTT video services.

Poor quality video experiences — such as buffering, stalling, or low-resolution video — can negatively impact a distributor's ability to deliver value against its charter or monetise its content. With costs of subscriber acquisition high and the threat of churn a constant concern, it is increasingly important for companies delivering video content via IP to ensure that quality is as high as it can be.

For OTT distributors, however, delivery costs are also a concern. OTT services typically operate on small margins due to rights costs. While they recognise the value of high-quality video streams to viewer engagement and retention, they must also deliver these as cost-efficiently as possible. Creating and storing multiple encode profiles for each piece of content, for example — to allow for multiple variations in network throughput — can impact operating margins. As such, the renditions being stored must be optimal and offer value to the viewer experience.

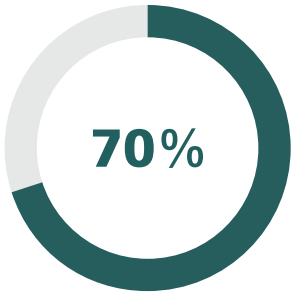
The quandary for product and technical teams within OTT distributors, then, is that they need to keep workflow costs down while at the same time improving the quality of the video experienced by their viewers across a wide range of devices.

High-quality video, however we define it, is no longer a "nice to have". In a rapidly evolving competitive landscape where large global OTT distributors continue to raise the bar through their investment in video quality, it is now a "must have" for commercial or public service OTT services that want to attract and retain viewers, advertisers, and subscribers.

VIEWERS INCREASINGLY EXPECT HIGH-QUALITY VIDEO STREAMS

Viewers in the mature TV markets demand more from their OTT services. They are used to a minimum of HD TV delivery via broadcast and expect similar (or better) quality video streams from OTT services. "Good enough" — for example, a video stream that equates, at best, to an SD broadcast feed — may no longer be good enough.

As audiences spend more of their viewing time watching episodic long-form content delivered via IP, they expect that premium video content, regardless of the screen or the route to that screen, should be of linear quality. This is challenging when we consider that — despite the popularity of video viewing on mobile devices — viewers in mature markets like to watch long-form episodic content and films on their TV screen. Netflix claims, for example, that 70% of its streams end up on Smart TVs rather than on phones, tablets or PCs. In such a scenario, audiences expect broadcast TV and OTT video to deliver the same quality of experience across all devices, including large-screen TVs.



of Netflix's streams end up on Smart TVs rather than on phones, tablets or PCs.

THERE IS NO INDUSTRY STANDARD FOR WHAT "GOOD" VIDEO LOOKS LIKE

Is that step up in quality achievable and affordable? How do we measure it? What does "good" look like? These seem like simple questions, but while the broadcast industry has established codes of practice around image fidelity, for those delivering video via the Internet, no such shared understanding exists.

Most video service providers focus on a set of core key performance indicators (KPIs) when considering video Quality of Experience (QoE). The favoured metrics have evolved as online video has matured, but three are typically cited: rebuffering ratio, the ratio between the rebuffering duration and the actual duration of video that played; video start time (VST), the time between a user requesting a video asset and it appearing on the screen; and bitrate, the rate at which bits are transferred from one location to another, measured in kilobits per second (Kbps) or megabits per second (Mbps).

These are all important in understanding the performance of a video service, but while these are the favoured metrics, they are not universal: Different organisations have different KPIs. As the industry matures, and as audience expectations rise, we need to consider a more holistic and authentic way to measure the quality of Internet-delivered video that also captures the fidelity of the viewers' video experience.

While the absence of buffering is very important, so is the visual quality of the video the viewer experiences. Subjective methods for measuring that quality are not necessarily helpful. But the quality of the streamed video can be measured, using objective and rigorous methodology.

Akamai sought to test the quality of online video streams and to understand the impact of that quality of different variables in terms of network conditions, content genre, and the device used by the viewer. The findings — of how we define good quality video and how it can be achieved across a range of devices, genres, and network conditions — will be of value to any business seeking to understand the optimal way to deliver high-quality video streams to viewers, at scale, and in the most cost-efficient way.

METHODOLOGY

Measuring perceptual quality.

Akamai has been helping publishers and content providers deliver high-quality video via the Internet for the last two decades. Like its customers, Akamai wants to understand what "good" Internet video looks like. Currently, much of the focus in measuring QoE is on the absence of negative factors such as rebuffering or video start failure, but as the industry matures there is a need for a more positive measure of video quality that shows how close it can get to the best possible image. However, it needs to be objective and repeatable to enable comparisons to be made between different streams and over time.

MULTIPLE FACTORS IMPACT THE QUALITY OF VIDEO VIEWED

Akamai worked with Eurofins Digital Testing, a leading testing provider, to construct a rigorous series of experiments. The aim was to not only measure video quality in an objective way, considering the quality of the video perceived by the viewer, but also to understand the impact on that quality of different factors, including the different devices used to access it.

When measuring and defining video quality, multiple factors can impact the quality of a video stream. Between the viewer requesting it and that content appearing on the screen in front of them. Issues can arise throughout a complex video supply chain with multiple components, but Akamai identified four factors in particular that can affect the ability to deliver the optimal video experience.

CONTENT GENRES

Different genres have different spatial and temporal demands during an encode to ensure an optimal viewing experience and, therefore, exert different requirements on the video delivery infrastructure. As such, the testing reflected three different video genres: talking heads, drama, and high action to understand the relative importance of what "good" looks like.

PLAYER/DEVICE

Do different connected devices for streaming video deliver similar quality experiences to viewers, all other elements being equal? Eurofins and Akamai agreed to test using a selection of popular devices — including smartphones, Smart TVs, tablets, set-top boxes, and PCs — to see the extent to which the device affected the quality of the video stream, even when other factors were constant.

NETWORK CONDITIONS

It was important to understand how each video stream is impacted further by challenging network conditions, so every combination of player, content, and bitrate was tested under both normalised and challenging network conditions.

BITRATE LADDER

Content providers have to create and store multiple versions of each video file, using different bitrates to allow the player to adjust to differing network conditions. Lower bandwidth availability will require lower bitrate profiles in order to deliver the best possible video experience at that time, and players typically switch between different profiles as required. Identifying the optimal encode profile to deliver high-quality video stream at the lowest bitrate is, as one TV executive responsible explained, "gut-based, but we want it to be evidence-based."

PERCEPTUAL QUALITY MEASURES THE BEST POSSIBLE IMAGE

For each scenario, though, it was necessary to answer an apparently simple question: What does an optimal video stream looks like? Akamai used the concept of Perceptual Quality (PQ) — essentially measuring the video fidelity seen by the viewer. They also used the tried-and-tested perceptual quality software SSIMPLUS® to determine the extent to which various pieces of video met this optimal standard, depending on the bitrate, the device used, the frame rate, the resolution, the dynamic range, and the content genre, as well as adverse network conditions.

SSIMPLUS® is a well-known Quality of Experience (QoE) measure that is thought to most closely represent the human perception of video quality, referenced against the intended PQ. Other algorithms are used in this space — notably VMAF — but given the parameters of this test, SSIMPLUS® was identified as the most practical and best tool to use. It ranks video output on a score of 0-100 dependent on how close the content viewed is to the absolute best quality the eye could see. An SSIMPLUS® score translates as follows: Excellent (81-100), Good (61-80), Fair (41-60), Poor (21-40), and Bad (1-20). A score of 80+ is generally considered the equivalent of an HD TV broadcast.

A ROBUST, REPEATABLE TESTING PLATFORM CREATES MEANINGFUL DATA

Eurofins and Akamai developed a robust platform to repeatedly execute the same test on different devices and to control both the bitrate and the environment the device experienced. After input from broadcaster executives responsible for video delivery on a daily basis, Eurofins and Akamai also investigated the dynamic between temporal and spatial metrics by testing examples of content representing different genres: a basic, low-movement clip of talking heads; a clip typical of mid-range drama with some action but mainly muted colours; and a hard-to-encode piece of underwater footage, including lots of colour and shimmering light.

In order to emulate conditions typically found in the Internet, the test also sought to replicate different network conditions for different broadband environments.

The test platform was then able to extract the resulting video from the different combinations of device, genre, and bitrate, and provide a quality score, using SSIMPLUS®. It captured the output versus a known reference to extract a score for each video (identified by using QR codes embedded in each frame) to assess quality in an objective way.

The aim was to measure the video quality experienced by the viewer in an objective way. In addition, the platform had to provide the ability to run the same test again and again to build up a volume of data that is statistically meaningful.

RESEARCH FINDINGS

Best QoE vs best video quality

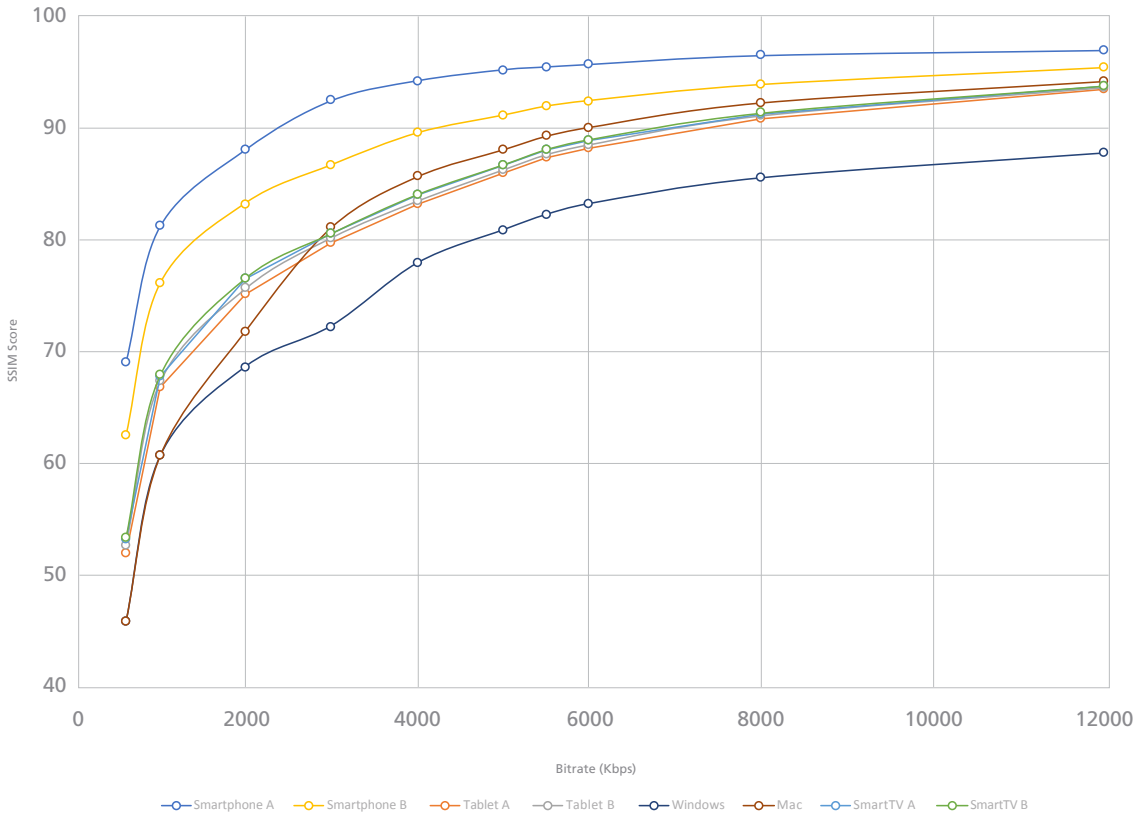
In the first phase of this research, testers sought to understand the impact on objective video quality — Perceptual Quality — of both content genre (the temporal and spatial qualities of the video asset itself) and the device or player through which the video is delivered.

The same piece of video content — in the case shown below, complex encode footage — was played through a range of devices at different bitrates. By embedding QR codes in each frame, testers were able to measure the perceptual quality of each combination of device and bitrate to understand how each device responded.

Surprisingly, the results showed significant variation in the bitrate required to deliver a high-definition image (with a SSIMPLUS® score of 80 or more) to different devices, even when the content and the network conditions were the same. For example, the two smartphones tested achieved optimal perceptual quality scores with bitrates of less than 3Mbps but had significant perceptual quality variation between them. In contrast, the Smart TVs on test required encode profiles of beyond 6Mbps to achieve a similar score. Desktop PCs on test required bitrates of around 5Mbps, but again showed significant variation in their perceptual quality.

The Perceptual Quality of hard-to-encode video content, by device and by bitrate

Figure 1a

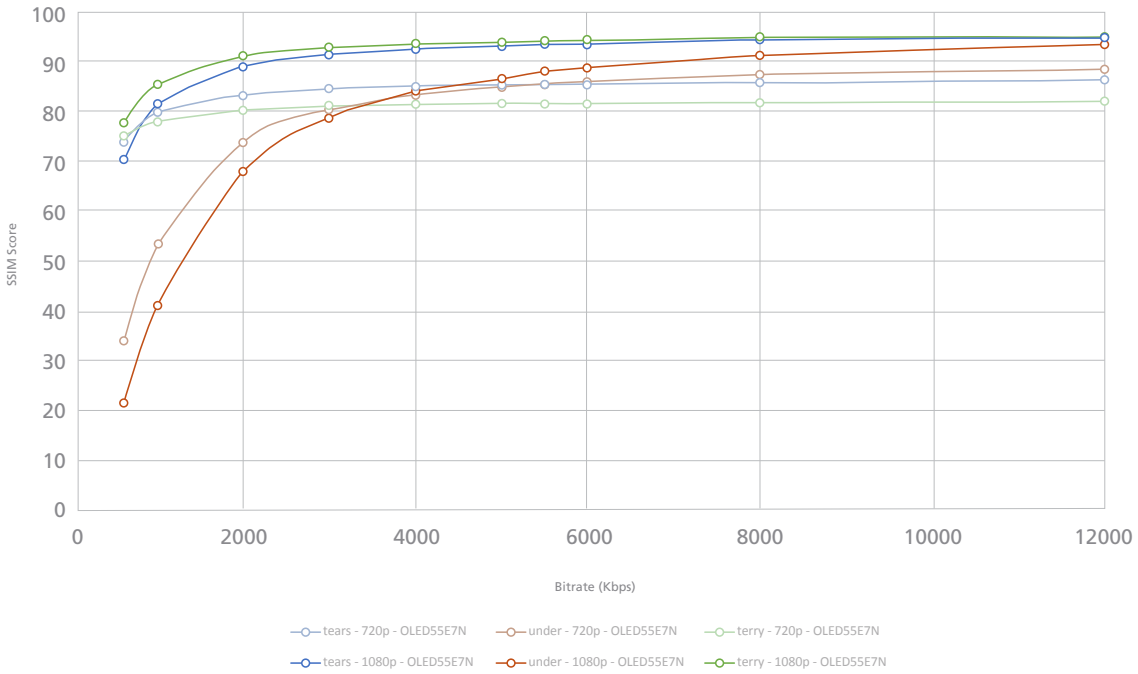


Such findings have implications for OTT distributors seeking to optimise video content for a range of different devices and players, where the cost of creating and storing multiple encode profiles for each content asset is significant.

Moreover, the variance of content genres on a given device was also highly evident. In the results below, we can see that on a large-screen TV the lower complexity content immediately benefited from being encoded in 1080p with excellent HD quality being achieved at 2Mbps. Mid-range quality content at 1080p also derived an excellent quality rating at 3Mbps. Neither genre benefited at the lower rates from encoding at 720p. As previously stated, higher complexity content required a bit rate of a minimum of 6mbps but again did not benefit at higher bit rates when encoded at 720p.

Perceptual quality of all content genres on a large screen TV

Figure 1b

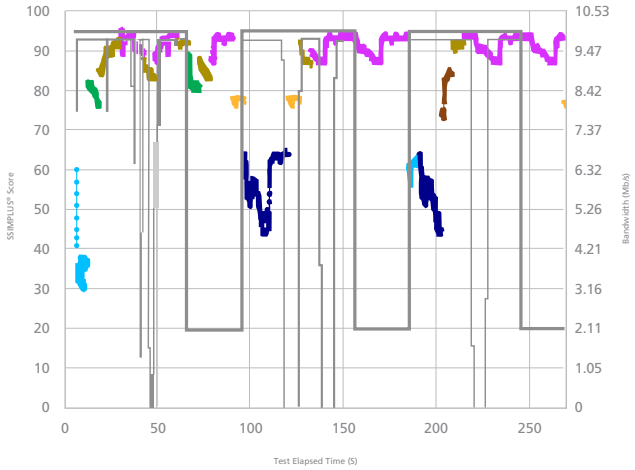


However, for the first phase of tests, network conditions were constant, which is not always a given in real-world conditions. In order to further understand how multiple variables can impact the video quality experienced by the end user, Eurofins and Akamai then applied a range of challenging network variants —including outages — to further test the various combinations of player and content and related these to the common KPIs seen in the market.

In the charts below, the coloured lines show the variation in SSIMPLUS® score over time (on the y-axis) during video playback; the different colours represent the different network variants selected. The accompanying table summarises the performance of the particular combination of device and content. For example, as shown in Figure 2, the Smart TV client showing the hard-to-encode footage averaged an SSIMPLUS® score of 83, which is nominally good, but following a second sustained drop in throughput consistency the player buffered,

A Smart TV client playing back hard-to-encode content

Figure 2

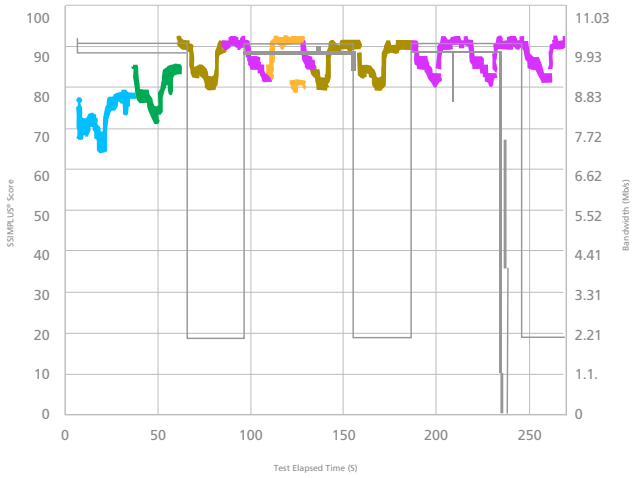


Metric	Score
Time to start playback (from receiving play command)	1.14
Highest quality variant played	8
Time to highest quality variant played (receiving play command)	25.12
Time to optimal bitrate (from receiving play command)	25.12
Number of content switches	20
Number of buffering events	1
Duration of buffering events	3.94
Average SSIMPLUS® Score	83
Worst SSIMPLUS® Score	30

The set-top-box device tested using the same content and network conditions (see Figure 3) delivered a similar average SSIMPLUS® score — 85 — using similar bandwidth (between 8 and 10 Mbps). Crucially, however, it coped better with the challenges, with the score never dropping below 65, unlike the Smart TV which dropped to 30 ("poor"). Our previous research with Sensum highlighted the importance of consistency to viewers and the negative impact of buffering on their engagement with the content. Given the growth in revenue forecast for advertising delivered via Smart TVs, inconsistency of throughput is a key factor to be addressed. An example is moving to protocols such as UDP for video delivery due to its inherent consistency benefits in order to deliver an optimal experience for viewers and advertisers.

A set-top box client playing back hard-to-encode content

Figure 3



Metric	Score
Time to start playback (from receiving play command)	1.91
Highest quality variant played	6
Time to highest quality variant played (receiving play command)	103.79
Time to optimal bitrate (from receiving play command)	N/A
Number of content switches	7
Number of buffering events	0
Duration of buffering events	0
Average SSIMPLUS® Score	85
Worst SSIMPLUS® Score	65

These variations demonstrate a need to understand the relationship among the device, the consistency of video throughput, and the content genre to ensure viewers are receiving the best possible experience.

DIFFERENT DEVICES DELIVER DIFFERENT VIDEO QUALITY

The testing devised by Eurofins for Akamai demonstrated that different players and devices deliver very different video quality, even when the content and network conditions are identical. To deliver the best possible video experiences, service providers need to understand the characteristics of the end-user devices they are targeting, as well as the characteristics of the content. Moreover, due to the variable nature of devices, the consistency of video throughput (i.e., the ability to deliver a stream at a constant bitrate through a CDN) is crucial to ensuring the highest quality video experience for viewers. By understanding how real-world devices consume the content and the strategic relationship among content, CDN, and device, it is possible to either improve the QoE or reduce buffering significantly.

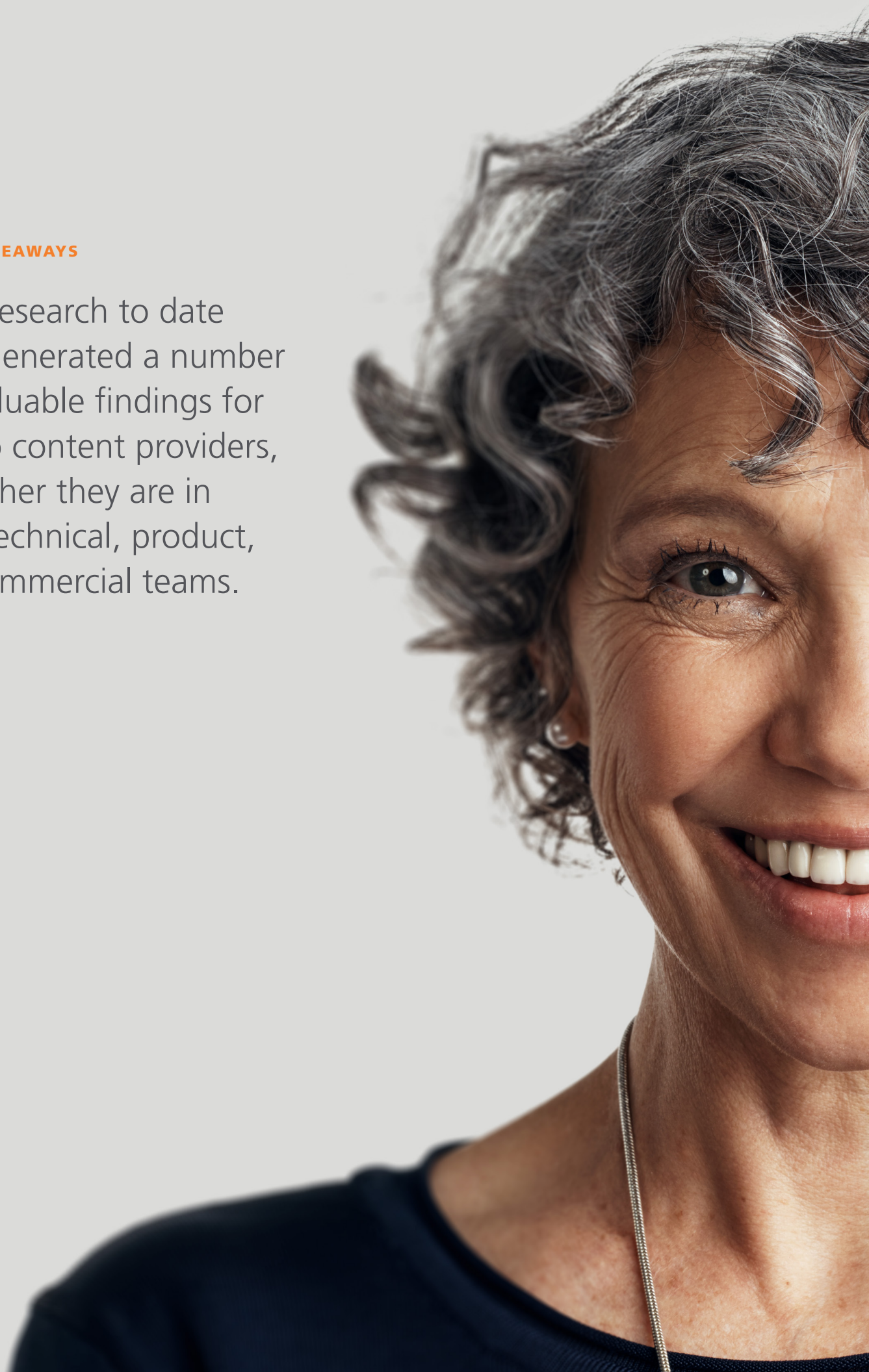
QOE METRICS MUST EVOLVE TO INCLUDE VISUAL QUALITY

The research also demonstrated that QoE assessments that over-index on start-up time and buffering time deliver misleading results. Using these metrics alone, the "best" client would be the one that typically selects only the lowest bit-rate variant in the bit-rate ladder. As Akamai demonstrated previously in its research with Sensum, the objective perceptual quality of the video matters and must be taken into account when measuring QoE. Great stories will always drive an audience to watch. The role of technology is to enhance that experience and increase viewer engagement.

Measuring objective video quality accurately in a way that correlates to the viewers' video experience is challenging to do. But by conducting such analysis in carefully controlled laboratory conditions — using actual content, actual encoding profiles, simulated network conditions and real-world devices and applications — it is possible to achieve a better understanding of the optimal system setup.

KEY TAKEAWAYS

The research to date has generated a number of valuable findings for video content providers, whether they are in the technical, product, or commercial teams.



VIDEO QUALITY IMPACTS THE BOTTOM LINE

It seems almost unnecessary to say that quality of experience is important to viewers of OTT video content. Yet given the tough economics of the sector, many OTT service providers have sought to deliver content at lower bitrates to avoid buffering issues or the cost associated with higher bitrate files. Thus, many OTT viewers receive streams that equate, at best, to a standard definition TV broadcast. But this carries risks, too. Akamai's research into the impact of delivering lower-fidelity streams demonstrates the link between video quality and engagement. As viewers who are used to HD TV come to expect similar levels of quality from their OTT providers, providers who fail to do this may find increased churn from unhappy viewers and advertisers. Optimising video quality equates to better engagement, which impacts derived value, subscriber churn, and advertising revenue.

ALL VIDEO PLAYERS ARE NOT EQUAL

A striking takeaway from this research is the degree to which different devices deliver different levels of video quality, even when tested under identical conditions. It is worth noting, for example, that smartphones delivered a consistently high-quality video stream — even when network conditions were especially challenging — and did this noticeably better than most of the other players tested. This factor also helps us to understand the importance of stream consistency. The industry has typically examined metrics such as throughput in minute detail, and while this is important, the tests show that a smoother flow through the Internet is much more important for QoE KPIs. A more consistent stream, irrespective of bitrate, resulted in less buffering. The use of transport protocols such as UDP rather than TCP can dramatically improve stream consistency and should be considered when delivering professional video content.

UPDATE YOUR KPIS — THERE IS MORE TO VIDEO QUALITY THAN BUFFERING

The KPIs typically used by OTT providers — including rebuffering ratio, video start time, and video start failures — are valuable, but do not provide a full picture of video quality. For example, combinations of player, genre, and bitrate that performed strongly in terms of those KPIs did not always deliver the highest levels of Perceptual Quality. As the industry matures, so must the metrics used to measure performance. Any meaningful and holistic measure of video quality for OTT services should now include a PQ score as an objective measure of picture quality in addition to the current KPIs.

IF THERE IS A QUALITY ISSUE, IT MAY NOT BE THE CDN!

OTT content providers spend significant resources addressing quality issues with their service, and it can be challenging to identify where — within an increasingly complex video supply chain — a problem lies. This research supports the idea that there are many variables in the chain that can create issues for the viewing experience. The more we know about potentially troublesome combinations of player, content, and encode profile, for example, the easier it may become to anticipate and resolve quality issues in future.

EVIDENCE-BASED PROVISIONING CAN CREATE SIGNIFICANT SAVINGS

There are more immediate ways in which this research is of value to OTT service providers. Creating a bitrate ladder, comprising all the encode profiles deemed necessary for each piece of video content, involves a large degree of guesswork, and, often, a large number of profiles that need to be created and stored. Through an evidence-based understanding of the likely impact of genre, device, and bitrate on the viewer's experience of a particular video file, OTT providers can potentially reduce the number of bitrate profiles they need to create and store for each video asset, generating significant savings.

THE INTERNET CAN DELIVER HIGH-QUALITY VIDEO, AT SCALE, IN MOST MARKETS!

Amidst the growth of OTT consumption in recent years, there is still a question mark over whether the Internet is mature and robust enough to support significant further migration from broadcast to an IP delivery model. Another output from this research was measuring Akamai's overall traffic data and local last-mile connection speeds to identify the ability of each market to deliver the video equivalent of HD content at scale at the optimal bitrates. Based on Akamai's own traffic data in assessing the percentage of users able to access at least 8Mbps, for example, and the known devices and network conditions, the answer in most mature markets of Western Europe, North America and Asia Pacific is yes!

Percentage of connections by speed, selected countries, Q1 2017

Figure 4

Country	Europe			Americas		Asia Pacific	
	Germany	UK	Sweden	USA	Canada	Japan	S. Korea
Above 4Mbps	90%	92%	94%	90%	90%	93%	98%
Above 10Mbps	53%	60%	56%	67%	61%	73%	85%

ABOUT EUROFINS DIGITAL TESTING

Eurofins Digital Testing is the world's leading end-to-end Quality Assurance (QA) service provider for Digital TV Operators and device-testing specialists, operating globally with facilities in the UK, Belgium, USA, Poland, Sweden, and Hong Kong. We provide specialised on-site test resources, testing tools, and services to validate digital media delivery systems and device conformance for multiple standards and operators across the world. For more information on Eurofins Digital Testing's methods for validating video streaming QoE, please visit www.eurofins-digitaltesting.com

ABOUT SSIMPLUS®

The SSIMPLUS® metric was developed by Emmy Award-winning researcher Professor Zhou Wang and his Ph.D. students, Drs. Abdul Rehman and Kai Zeng, at the University of Waterloo. With two decades of research behind it, the structural similarity algorithm has reinvented the way that video quality is measured. SSIMPLUS®' genius is that it can predict what the video consumer will see, and can do so at every step in the video delivery chain. It literally puts the human visual system into software. SSIMPLUS® is built for the digital age and is a proprietary algorithm of SSIMWAVE Inc.

Akamai secures and delivers digital experiences for the world's largest companies. Akamai's intelligent edge platform surrounds everything, from the enterprise to the cloud, so customers and their businesses can be fast, smart, and secure. Top brands globally rely on Akamai to help them realize competitive advantage through agile solutions that extend the power of their multi-cloud architectures. Akamai keeps decisions, apps and experiences closer to users than anyone — and attacks and threats far away. Akamai's portfolio of edge security, web and mobile performance, enterprise access and video delivery solutions is supported by unmatched customer service, analytics and 24/7/365 monitoring. To learn why the world's top brands trust Akamai, visit **www.akamai.com**, **blogs.akamai.com**, or **@Akamai** on Twitter. You can find our global contact information at **www.akamai.com/locations**. Published 09/18.