Quality of Experience for HTTP Adaptive Streaming Services

Ozgur Oyman and Sarabjot Singh, Intel Corporation

ABSTRACT

The growing consumer demand for mobile video services is one of the key drivers of the evolution of new wireless multimedia solutions requiring exploration of new ways to optimize future wireless networks for video services towards delivering enhanced quality of experience (QoE). One of these key video enhancing solutions is HTTP adaptive streaming (HAS), which has recently been spreading as a form of Internet video delivery and is expected to be deployed more broadly over the next few years. As a relatively new technology in comparison with traditional push-based adaptive streaming techniques, deployment of HAS presents new challenges and opportunities for content developers, service providers, network operators and device manufacturers. One of these important challenges is developing evaluation methodologies and performance metrics to accurately assess user QoE for HAS services, and effectively utilizing these metrics for service provisioning and optimizing network adaptation. In that vein, this article provides an overview of HAS concepts, and reviews the recently standardized QoE metrics and reporting framework in 3GPP. Furthermore, we present an end-to-end QoE evaluation study on HAS conducted over 3GPP LTE networks and conclude with a discussion of future challenges and opportunities in QoE optimization for HAS services.

INTRODUCTION

With the introduction of smartphones like the iPhone™ and Android™ based platforms, the emergence of new tablets like the iPad™, and the continued growth of netbooks, ultrabooks, and laptops, there is an explosion of powerful mobile devices in the market that are capable of displaying high-quality video content. In addition, these devices are capable of supporting various video streaming applications and interactive video applications like videoconferencing, and they can capture video for video sharing, video blogging, video Twitter™, and video broadcasting applications. Cisco predicts that mobile traffic will grow by a factor of 26 until 2015 (almost double every year), and that mobile traffic will be dominated by video; for example, by 2015, various forms of video will exceed 90 percent of global consumer traffic, and almost 66 percent of the world's mobile traffic will be video.[1] As a result, future wireless networks will need to be optimized for the delivery of a range of video content and video-based applications.

However, video communication over mobile broadband networks today is challenging due to limitations in bandwidth and difficulties in maintaining high reliability, quality, and latency demands imposed by rich multimedia applications. Even with the migration from 3G to 4G networks — or RAN and backhaul upgrades to 3G networks — the demand on capacity for multimedia traffic will continue to increase. As subscribers take advantage of new multimedia content, applications, and devices, they will consume all available bandwidth and still expect the same quality of service that came with their original service plans — if not better. Such consumer demand requires exploration of new ways to optimize future wireless networks for video services toward delivering higher user capacity to serve more users and also deliver enhanced quality of experience (QoE) for a rich set of video applications.

One of the key video-enhancing solutions is adaptive streaming, which is an increasingly promising method to deliver video to end users allowing enhancements in QoE and network bandwidth efficiency. Adaptive streaming aims to optimize and adapt the video configurations over time in order to deliver the best possible quality video to the user at any given time, considering changing link or network conditions, device capabilities, and content characteristics. Adaptive streaming is especially effective in better tackling the bandwidth limitations of wireless networks, but also it allows for more intelligent video streaming that is device-aware and content-aware. While adaptive streaming technologies support only 17 percent of the Internet video traffic today, the adaptive streaming portion of Internet video is anticipated to grow at an average of 77 percent a year toward supporting 51 percent of Internet video by 2015, according to a recent study by TDG Research.[2]

---

[1] *See the following whitepapers from Cisco Visual Networking Index: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html, http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html*

[2] *Source: TDG Research, http://www.tdgresearch.com*

0163-6804/12/$25.00 © 2012 IEEE

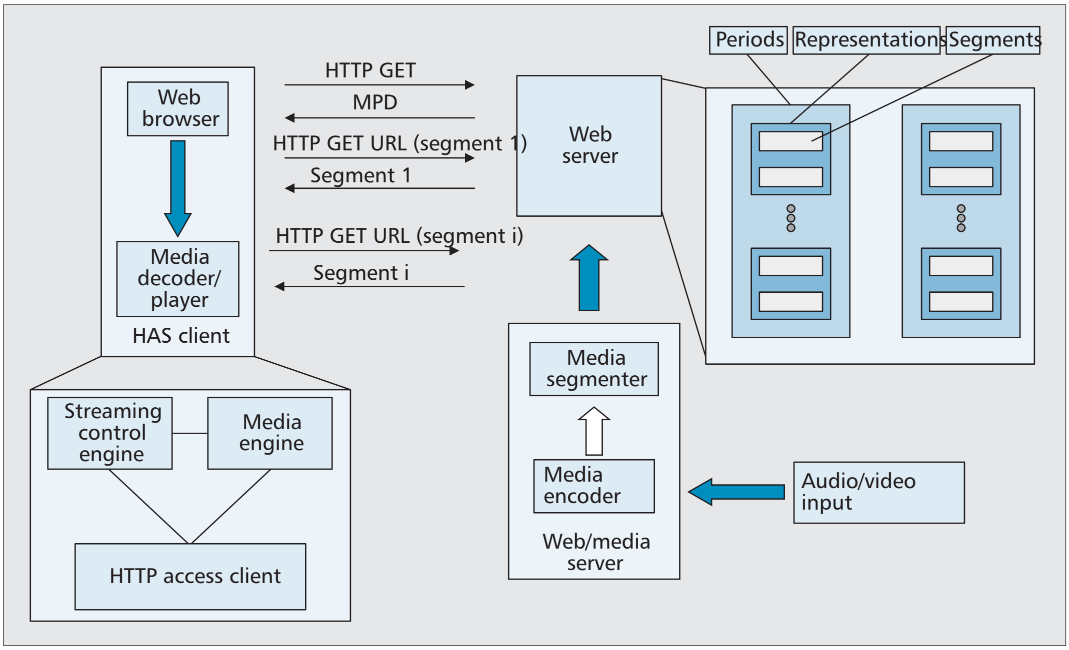


Figure 1. *HAS framework between the client and web/media server.*

> *Most of the expected broad adoption of adaptive streaming will be driven by new deployments over the existing web infrastructure based on the hypertext transfer protocol (HTTP), and this kind of streaming is referred here as HTTP adaptive streaming (HAS).*

Most of the expected broad adoption of adaptive streaming will be driven by new deployments over the existing web infrastructure based on HTTP [1], and this kind of streaming is referred here as HTTP adaptive streaming (HAS). HAS follows the pull-based streaming paradigm rather than the traditional push-based streaming based on stateful protocols such as the Real-Time Streaming Protocol (RTSP) [2], where the server keeps track of client state and drives the streaming. In contrast, in pull-based streaming such as HAS, the client plays the central role by carrying the intelligence that drives the video adaptation, since HTTP is a stateless protocol. Several important factors have influenced this paradigm shift from traditional push-based streaming to HTTP streaming, including:

- Broad market adoption of HTTP and TCP/IP protocols; they support the majority of the Internet services offered today.
- HTTP-based delivery avoids NAT and firewall traversal issues.
- A broad deployment of HTTP-based (non-adaptive) progressive download solutions already exists today, which can conveniently be upgraded to support HAS.
- The ability to use standard/existing HTTP servers and caches instead of specialized streaming servers allows reuse of the existing infrastructure, thereby provides better scalability and cost effectiveness.

Accordingly, the broad deployment of HAS technologies will serve as a major enhancement to (non-adaptive) progressive download methods, allowing for enhanced QoE enabled by intelligent adaptation to different link conditions, device capabilities, and content characteristics.

HAS has already been spreading as a form of Internet video delivery with the recent deployments of proprietary solutions such as Apple HTTP Live Streaming, Microsoft Smooth Streaming, and Adobe HTTP Dynamic Streaming.[3] In the meantime, standardization of HAS has also made great progress with the recent completion of technical specifications by various standards bodies including the Third Generation Partnership Project (3GPP), Motion Picture Experts Group (MPEG), and Open IPTV Forum (OIPF) [3–7]. Going forward, future deployments of HAS are expected to converge through broad adoption of these standardized solutions referred to as dynamic adaptive streaming over HTTP (DASH).

As a relatively new technology in comparison with traditional push-based adaptive streaming techniques, deployment of DASH and associated HAS techniques presents new challenges and opportunities for content developers, service providers, network operators and device manufacturers. One of these important challenges is developing evaluation methodologies and performance metrics to accurately assess user QoE for HAS services, and effectively utilizing these metrics for service provisioning and optimizing network adaptation. In that vein, this article provides an overview of HAS concepts and recent DASH standardization, and reviews the recently adopted QoE metrics and reporting framework in 3GPP standards. Furthermore, we present an end-to-end QoE evaluation study on HAS conducted over 3GPP LTE networks and conclude with a discussion of future directions and challenges in QoE optimization for HAS services.

HAS Concepts and Standardization Overview

The HAS framework between a client and a web/media server is depicted in Fig. 1 for the typical use case of Internet video streaming over the browser. The media preparation pro-

---

[3] *Related whitepapers can be found at the following links: Microsoft Smooth Streaming: http://www.microsoft.com/download/en/details.aspx?id=17678, Adobe HTTP Dynamic Streaming: http://www.adobe.com/products/httpdynamicstreaming/, Apple HTTP Live Streaming: http://developer.apple.com/library/ios/documentation/networkinginternet/conceptual/streamingmediaguide/StreamingMediaGuide.pdf*

HAS provides the ability to the client to fully control the streaming session, i.e., it can intelligently manage the on-time request and smooth playout of the sequence of segments, potentially adjusting bitrates or other attributes.

cess generates segments that contain different encoded versions of one or several of the media components of the media content. The segments are then hosted on one or several media origin servers typically, along with the media presentation description (MPD) that characterizes the structure and features of the media presentation, and provides sufficient information to a client for adaptive streaming of the content by downloading the media segments from the server over HTTP. The MPD describes the various representations of the media components (bit rates, resolutions, codecs, etc.) and HTTP URLs of the corresponding media segments, timing relationships across the segments and how they are mapped into media presentation. Based on the MPD metadata information, clients request the segments corresponding to their selected representations using HTTP GET or partial GET methods with byte ranges (this essentially imitates streaming via short downloads).

HAS provides the ability to the client to fully control the streaming session; that is, it can intelligently manage the on-time request and smooth playout of the sequence of segments, potentially adjusting bit rates or other attributes. The client can automatically choose initial content rate to match initial available bandwidth without requiring negotiation with the streaming server and dynamically switch between different bit rate representations of the media content as the available bandwidth changes. Hence, HAS allows faster adaptation to changing network and wireless link conditions, user preferences, and device states (e.g., display resolution, CPU, memory resources). Such dynamic adaptation provides better user QoE, with higher video quality, shorter startup delays, fewer rebuffering events, and so on.

Standardization of HAS techniques has been driven by various standards bodies including 3GPP, MPEG, and OIPF. Here we focus on HAS standardization in 3GPP, where HAS was standardized by the 3GPP SA4 Working Group, with the activity beginning in April 2009 and Release 9 work with updates to TS 26.234 and TS 26.244 completed in March 2010. During Release 10 development, a new specification, TS 26.247, on 3GPP DASH was finalized in June 2011. The scope of 3GPP DASH specification includes a normative definition of a media presentation (for DASH access client), a normative definition of the segment formats (for media engine), a normative definition of the delivery protocol used for the delivery of segments (HTTP/1.1), and an informative description on how a DASH client may use the provided information to establish a streaming service.

QUALITY OF EXPERIENCE IN 3GPP DASH

The development of QoE evaluation methodologies, performance metrics and reporting protocols play a key role for optimizing the delivery of HAS services. In particular, QoE monitoring and feedback are beneficial for detecting and debugging failures, managing streaming performance, enabling intelligent client adaptation (useful for device manufacturers), and allowing for QoE-aware network adaptation and service provisioning (useful for the network operator and content/service provider). Having recognized these benefits, both 3GPP and MPEG bodies have adopted QoE metrics for HAS services as part of their DASH specifications. Moreover, the 3GPP DASH specification also provides mechanisms for triggering QoE measurements at the client device as well as protocols and formats for delivery of QoE reports to the network servers. Here, we shall describe in detail the QoE metrics and reporting framework for 3GPP DASH, while it should be understood that MPEG has also standardized similar QoE metrics in MPEG DASH.

In the 3GPP DASH specification TS 26.247, QoE measurement and reporting capability is defined as an optional feature for client devices. However, if a client supports the QoE reporting feature, the DASH standard also mandates the reporting of all of the requested metrics at any given time; that is, the client should be capable of measuring and reporting all of the QoE metrics specified in the standard. It should also be noted here that 3GPP TS 26.247 also specifies QoE measurement and reporting for HTTP-based progressive download services, where the set of QoE metrics in this case is a subset of those provided for DASH.

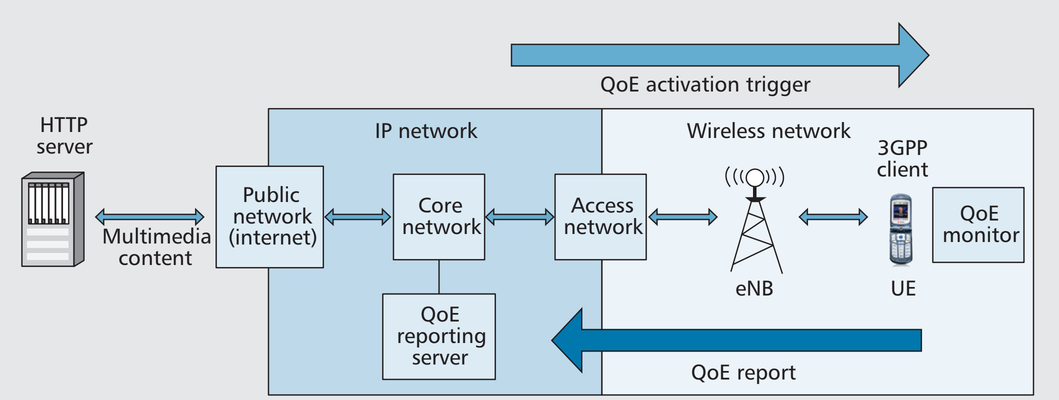
Figure 2 depicts the QoE monitoring and reporting framework specified in 3GPP TS 26.247, summarizes the list of QoE metrics standardized by 3GPP in TS 26.247, and indicates the list of metrics applicable for DASH/HAS (adaptive streaming) and HTTP-based progressive download (non-adaptive). At a high level, the QoE monitoring and reporting framework is composed of the following phases:

- A server activates/triggers QoE reporting, requests a set of QoE metrics to be reported, and configures the QoE reporting framework.
- A client monitors or measures the requested QoE metrics according to the QoE configuration.
- The client reports the measured parameters to a network server.

We now discuss each of these phases in the following subsections.

ACTIVATION AND CONFIGURATION OF QOE REPORTING

3GPP TS 26.247 specifies two options for the activation or triggering of QoE reporting. The first option is via the `QualityMetrics` element in the MPD, and the second option is via the OMA Device Management (DM) QoE Management Object. In both cases, the trigger message from the server would include reporting configuration information such as the set of QoE metrics to be reported, the URIs for the server(s) to which the QoE reports should be sent, the format of the QoE reports (e.g., uncompressed or gzip), information on QoE reporting frequency and measurement interval, percentage of sessions for which QoE metrics will be reported, and access point name (APN)



3GPP QoE metric	Progressive Download	DASH / HAS
HTTP request/response transactions	Yes	Yes
Representation switch events	No	Yes
Average throughput	Yes	Yes
Initial playout delay	Yes	Yes
Buffer level	Yes	Yes
Play list	Yes	Yes
MPD information	No	Yes

Figure 2. *QoE metrics and reporting framework for 3GPP DASH and progressive download.*

QoE monitoring and feedback are beneficial for detecting and debugging failures, managing streaming performance, enabling intelligent client adaptation (useful for device manufacturer) and allowing for QoE-aware network adaptation and service provisioning (useful for the network operator and content/service provider).

to be used for establishing the packet data protocol (PDP) context to be used for sending the QoE reports.

QOE METRICS FOR DASH

The following QoE metrics have been defined in 3GPP DASH specification TS 26.247, to be measured and reported by the client upon activation by the server. It should be noted that these metrics are specific to HAS and content streaming over the HTTP/TCP/IP stack, and therefore differ considerably from QoE metrics for traditional push-based streaming protocols.

HTTP request/response transactions: This metric essentially logs the outcome of each HTTP request and corresponding HTTP response. For every HTTP request/response transaction, the client measures and reports:

- Type of request (e.g., MPD, initialization segment, media segment)
- Times for when the HTTP request was made and corresponding HTTP response was received (in wall clock time)
- HTTP response code
- Contents in the byte-range-spec part of the HTTP range header
- TCP connection identifier
- Throughput trace values for successful requests

From HTTP request/response transactions, it is also possible to derive more specific performance metrics such as the fetch durations of the MPD, initialization segment, and media segments.

Representation switch events: This metric is used to report a list of representation switch events that took place during the measurement interval. A representation switch event signals the client's decision to perform a representation switch from the currently presented representation to a new representation that is later presented. As part of each representation switch event, the client reports the identifier for the new representation, the time of the switch event (in wall clock time) when the client sends the first HTTP request for the new representation, and the media time of the earliest media sample played out from the new representation.

Average throughput: This metric indicates the average throughput that is observed by the client during the measurement interval. As part of the average throughput metric, the client measures and reports:

- Total number of content bytes (i.e., the total number of bytes in the body of the HTTP responses) received during the measurement interval
- Activity time during the measurement interval, defined as the time during which at least one GET request is still not completed
- Wall clock time and duration of the measurement interval
- Access bearer for the TCP connection for which the average throughput is reported
- Type of inactivity (e.g., pause of presentation)

Initial playout delay: This metric signals the

Looking into the future, enabling optimized delivery of HAS services over wireless networks will require developing specialized end-to-end network protocols, architectures and algorithms for enhancing service capacity and user QoE.

initial playout delay at the start of the streaming of the presentation. It is measured as the time from when the client requests the fetch of the first media segment (or sub-segment) to the time at which media is retrieved from the client buffer.

Buffer level: This metric provides a list of buffer occupancy level measurements carried out during playout. As part of the buffer level metric, the client measures and reports the buffer level that indicates the playout duration for which media data is available starting from the current playout time along with the time of the measurement of the buffer level.

Play list: This metric is used to log a list of playback periods in the measurement interval, where each playback period is the time interval between a user action and whichever occurs soonest of the next user action, the end of playback, or a failure that stops playback. The type of user actions that trigger playout may include a new playout request, resume playout from pause, or user-requested quality change. For each playback period, the client measures and reports the identifiers of the representations that were rendered and their rendering times (in media time) and durations, playback speed relative to normal playback speed (e.g., to track trick modes such as fast forward or rewind), and reasons why continuous playback of this representation was interrupted (e.g., due to representation switch events, rebuffering, user request, or end of period, media content, and a metrics collection period).

MPD information: This metric allows for reporting information on the media presentations from the MPD so that servers without direct access to the MPD can learn the media characteristics. Media representation attributes on bit rate, resolution, quality ranking, and codec-related media information including profile and level can be reported by the client via this metric.

QOE REPORTING PROTOCOL

In 3GPP DASH, QoE reports are formatted as an Extensible Markup Language (XML)[4] document complying with the XML schema provided in TS 26.247. The client uses HTTP POST request signaling (Internet Engineering Task Force [IETF] RFC 2616) carrying XML-formatted metadata in its body to send the QoE report to the server.

QOE EVALUATION OF HAS OVER 3GPP LTE NETWORKS

Toward demonstrating performance enhancements from HAS techniques over HTTP-based progressive download solutions, we now evaluate end-to-end QoE over a dynamic system-level simulator for the LTE air-interface based on a MATLAB-based software platform with detailed abstractions of application, transport, medium access control (MAC), and physical layers. The QoE metric of interest for the analysis described here is the rebuffering percent, which is defined as the percentage of the total presentation time in which the user experiences rebuffering due to buffer starvation. Our motivation for focusing on this metric is driven by a recent study conducted by Conviva, where rebuffering has been identified as the single most dominating QoE impairment.[5] In a 3GPP DASH-based implementation of QoE metrics in the client device, this metric can be computed via monitoring the buffer status and play list metrics introduced in the previous section.

In our end-to-end QoE evaluation, we consider five variable bit rate (VBR)-encoded video clips with different bit rate requirements and rate-distortion characteristics hosted at the HTTP server with multiple versions of each video clip available at different quality levels in the peak signal-to-noise ratio (PSNR) range of 26–39 dB. Each user in the LTE network randomly requests one of the video clips. For the HTTP-based progressive download (non-adaptive) use case, the requested video is served at a fixed quality level (e.g., a target PSNR level), while for adaptive streaming (DASH/HAS), users may consume varying qualities of video based on the working of our adaptation algorithm, which selects the optimal quality/bit rate representation among the available video clips based on monitoring of user experience via 3GPP-based QoE metrics described in the previous section.

The client player starts playback with initial startup delay of 1 s. In both cases of HTTP-based progressive download and HAS, the client player requests the video at a higher fetch rate during the buffering mode (playback buffer under a specified threshold), while the fetch rate is lower during the streaming mode (playback buffer above the specified threshold). The rate adaptation algorithm tries to match the source rate to the link throughput, which is estimated using the throughput value averaged over the last P IP packets downloaded by the client, or

$${}_{hrpt} = \frac{1}{P} \sum_{i=L_P-P+1}^{L_P} \frac{S_{packet}(i)}{T_{download}(i) - T_{fetch}(i)}, \qquad (1)$$

where L_P is the index of the last packet downloaded by the client, $T_{download}(i)$ is the time at which packet i enters into client queue, $T_{fetch}(i)$ is the time when it enters the server queue, and $S_{packet}(i)$ is the packet size for packet i. We focus on the QoE evaluation over the LTE air interface, and ignore delays and losses over the wired connection between the server and base station. The effect of TCP retransmissions would be reflected in R_{thrpt} as the download time would increase drastically for the corresponding packets. In the streaming mode, the maximum bit rate video supported by the client's throughput estimate is fetched with representation level of

$$Q_{rep}^{spp} = \arg \max_i \; b_i; \; b_i \leq R_{thrpt} \;\; \text{over } i = 1,2,\ldots N \qquad (2)$$

where b_i denotes the bit rate of encoded video of representation level i and N denotes the highest quality or representation level. While in the buffering mode, in our implementation, the client keeps reducing the fetched bit rate by one

---

[4] *For further information on the XML language, see http://www.w3.org/XML/*

[5] *See the following article for further details: http://www.videonuze.com/blogs/?2011-11-15 percent2009:06:19/Buffering-Video-Engagement-s-1-Enemy/&id=3285*

quality level stepwise or otherwise to the bit rate governed by supported quality, subject to the minimum bit rate available for the video, resulting in

$$Q_{rep}^{fetch} = \max\left(1, \min\left(Q_{rep}^{prev} - 1, Q_{rep}^{spp}\right)\right) \qquad (3)$$

where Q_{rep}^{prev} the is the representation level of the latest received chunk. Thus, in the buffering mode quality level is never scaled up. Encountering playback buffer starvation, the client enters rebuffering mode while stalling the playback. The playback resumes after 0.5 s if the playback buffer is non-zero; otherwise, the rebuffering mode continues for another 0.5 s, and so on. The operation of TCP is abstracted in the form of transport layer retransmissions and rate throttling upon packet losses.

Multiuser resource allocation over the orthogonal frequency-division multiple access (OFDMA)-based downlink Long Term Evolution (LTE) air interface is performed based on the well-known proportional fair scheduling principles. Only half of the available bandwidth of the 10 MHz LTE system is assumed to be reserved for the video streaming service, while the remaining half is assumed to be dedicated for other services (e.g., voice and data services). Table 1 summarizes the LTE system simulation parameters.

Figure 3 shows the distribution of rebuffering percent across users with fixed rate and adaptive streaming. The QoE enhancement from adaptive streaming is apparent from the plot in which, with fixed rate streaming over LTE at a target PSNR of 37 dB and only 20 users in the system, the 95th percentile value of rebuffering percent is 5 percent, whereas the corresponding value for an LTE system with HAS-based adaptive streaming and twice the load (i.e., with 40 users) is less than 1 percent. This empirical data demonstrates that HAS-based adaptive streaming reduces the occurrences of rebuffering events significantly in comparison with HTTP-based progressive download techniques, even under heavier system loading conditions, and finds the QoE-optimal capacity-quality trade-off by adjusting the quality/bit rate levels of the videos in a QoE-aware fashion (with QoE metric being the rebuffering percent). This is an intuitively expected outcome, given the significantly varying link quality among the users in the LTE network, leading to frequent occurrences of rebuffering with HTTP-based progressive download in the absence of any video quality/bit rate adaptation, especially when the network is unable to support the fixed bit rate during moments of low throughput caused by unfavorable link conditions. In contrast, with DASH/HAS-based adaptive streaming, each client device can dynamically select the video representations that ensure continuous playback while also optimizing quality that could be achieved for the given link throughput, and such adaptation capability ensures finding the best possible compromise between high video quality and minimal occurrences of rebuffering events, and delivering enhanced QoE to a larger number of LTE clients.

Parameters	Assumption
Channel model	3GPP Case 1 with 3D antenna pattern, SCM-UMa (15 degrees angular spread)
Downlink transmit power	46 dBm
MIMO Mode	4 × 2 SU-MIMO for the downlink
Cellular Layout	Hexagonal grid, 19 cell sites, 3 sectors per site
Distance-dependent path	loss L=I + 37.6$\log_{10}$(.R), R in kilometers, I=128.1
Lognormal Shadowing	Similar to UMTS 30.03, B 1.141
Shadowing standard deviation	8 dB
Number of antennas at UE	2
Number of antennas at eNB	4
Antenna configuration at UE	Co-polarized antennas
Antenna configuration at eNB	Co-polarized (0.5λ spacing)
Outer-loop for target FER control	10 percent FER for 1st HARQ transmission
Link adaptation	MCSs based on LTE transport formats according to 3GPP TR 36.213
HARQ scheme	Chase combining
DL overhead	3 for PDCCH
UE speed	3km/h
Scheduling granularity	5 RB subband
Receiver type	MMSE-IRC
Feedback mode	Wideband PMI based on LTE 4-bit CB, subband CQI
Intersite Distance	500 m
User distribution	Users dropped uniformly in the entire cell

Table 1. *LTE simulation parameters.*

CONCLUDING REMARKS ON FUTURE CHALLENGES AND OPPORTUNITIES IN QOE OPTIMIZATION FOR HAS SERVICES

Looking into the future, enabling optimized delivery of HAS services over wireless networks will require developing specialized end-to-end network protocols, architectures, and algorithms for enhancing service capacity and user QoE. Such development must also ensure access-specific optimizations for HAS services, which clearly requires different approaches and methods compared to those for traditional push-based (e.g., RTSP-based) streaming services, observing the client-driven nature of HAS services and that QoE for HAS services is measured via different performance metrics (as discussed earlier). Below we further elaborate on two specific research vectors in this space.

HAS-SPECIFIC CROSS-LAYER ADAPTATION ALGORITHMS

This optimization relies on tight integration of the HAS/HTTP-specific media delivery with network-level and radio-level adaptation and QoS mechanisms in order to jointly determine the best video, transport, network and radio configurations in a link-aware, device-aware and content-aware fashion toward realizing the highest possible end user QoE. An example cross-layer adaptation architecture and the associated open systems interconnection (OSI) communication layer stack for HAS-specific optimizations is depicted in Fig. 4. The end-to-end video delivery system configurations that could be jointly optimized via such cross-layer cooperation and QoE monitoring/feedback (including sharing of the MPD with appropriate network entities) include the following parameters:

- Video level: Bit rate, frame rate, resolution, codecs, and so on
- Transport level: Sequence and timing of HTTP requests, number of parallel TCP connections, HAS segment durations, frequency of MPD updates (e.g., for live services)
- Radio and network level: Bandwidth allocation and multiuser scheduling, target QoS parameters for the core network and radio access network, modulation and coding scheme (MCS), OFDMA time/frequency resource/burst allocations

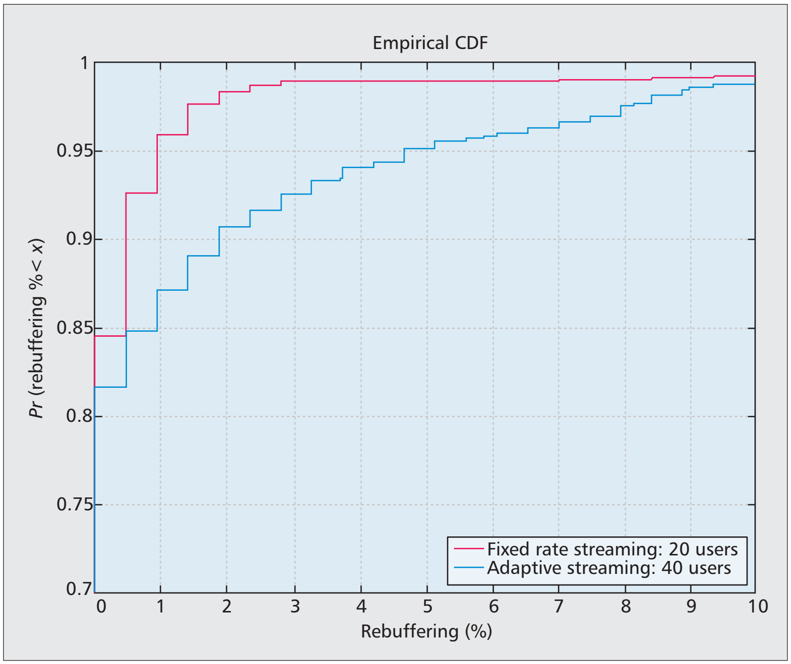


Figure 3. *Distribution of rebuffering percent across the users with HTTP-based progressive download (fixed-rate streaming) and DASH/HAS-based adaptive streaming.*

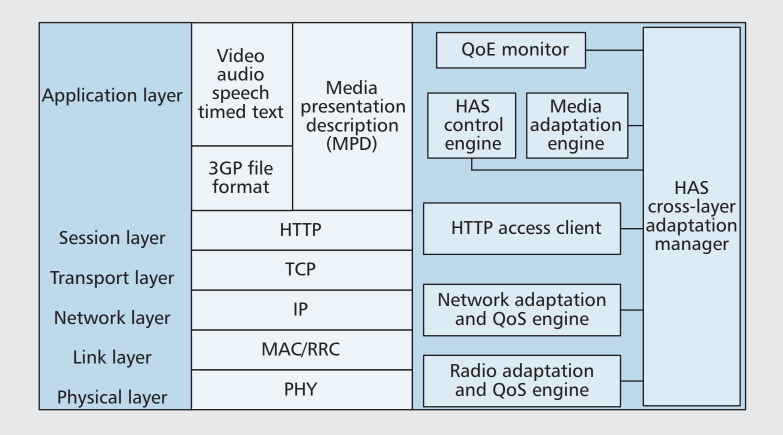


Figure 4. *Example cross-layer adaptation architecture for HAS-specific QoE optimizations.*

QoE-aware HAS service optimization problems emerging from this kind of a cross-layer cooperation framework include investigation topics such as QoE-aware radio resource management and scheduling, QoE-aware service differentiation and QoS prioritization, and QoE-aware heterogeneous networking, including QoE-aware handoff and WiFi offloading.

HAS-SPECIFIC QOS DELIVERY AND SERVICE ADAPTATION

In the context of 3GPP LTE-Advanced systems, an important target for HAS-specific optimizations is quality of service (QoS) delivery and service adaptation, which are beneficial in order to optimally manage limited network resources toward enhancing network capacity utilization and providing better QoE to the end user. In particular, the current policy and charging control (PCC) architecture [8–10] for 3GPP networks only has mechanisms to handle QoS delivery and service adaptation for traditional RTSP-based adaptive streaming services. New QoS delivery and service adaptation methods should be devised and adopted specifically targeting HAS-based multimedia services over 3GPP radio access network (RAN) and core IP network architectures.

Figure 5 depicts an example PCC architecture delivering end-to-end QoS support for HAS services with the added new capability to parse or interpret the MPD in order to gain information on the application-layer parameters for HAS. In the current PCC architecture, the application function (AF) interacts with the applications requiring dynamic policy and charging control. Hence, in order to provide QoS for HAS services, the AF should have the ability to extract session information from the MPD, map it into the appropriate audio-video parameters (AVPs), and provide the AVPs to the policy and charging rules function (PCRF) over the Rx reference point. The PCRF combines the HAS-related session information from the AVPs received over the Rx reference point and the input received from the Gx and Gxa/Gxc reference points with user-specific policies data from the subscription profile repository (SPR) to form session-level policy decisions, and provides those to the policy and charging enforcement function (PCEF) and bearer-binding and event-reporting function (BBERF).

According to the current HSD specification 3GPP TS 26.247, the key MPD attributes and elements to be used for deriving the AVP/QoS mapping rules may be retrieved by the operator network (e.g., server hosting the AF) by triggering of QoE reporting mechanisms from the 3GPP clients supporting this feature. Other methods for gathering MPD attributes and elements in the AF server may also be utilized. In this context, it is important to investigate mapping optimizations between HAS-specific application-layer information (e.g., that contained in the MPD) and the AVPs and associated set of target QoS parameters for the RAN and core network, such as QoS class identifiers (QCIs) and DiffServ/DSCP parameters, in order to deliver the best QoE for HAS services. While an initial set of MPD-aware QoS mapping rules has recently been adopted in TS 26.247 Release 10, this is currently an active investigation topic for 3GPP Release 11 in the SA4 working group and may also have implications for 3GPP specifications maintained by other working groups, including SA2 and CT3.

ACKNOWLEDGMENTS

Beneficial discussions and feedback from the following colleagues at Intel are gratefully acknowledged: Jeffrey Foerster, Apostolos Papathanassiou, Debdeep Chatterjee, Srinivasa Somayazulu, Chris Cavigioli, and Suman Sharma.

REFERENCES

[1] IETF RFC 2616, "Hypertext Transfer Protocol-HTTP/1.1," June 1999.

[2] IETF RFC 2326, "Real-Time Streaming Protocol (RTSP)," Apr. 1998.

[3] 3GPP TS 26.234 v10.2.0, "Transparent End-to-End Packet Switched Streaming Service (PSS); Protocols and Codecs," Release 10, Sep. 2011.

[4] 3GPP TS 26.244 v10.1.0, "Transparent End-to-End Packet Switched Streaming Service (PSS); 3GPP File Format (3GP)," Release 10, June 2011.

[5] 3GPP TS 26.247 v10.1.0, "Transparent End-to-End Packet Switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming Over HTTP (3GP-DASH)," Release 10, June 2011.

[6] ISO/IEC DIS 23009-1, "Information Technology — Dynamic Adaptive Streaming Over HTTP (DASH) — Part 1: Media Presentation Description and Segment Formats," Aug. 2011.

[7] OIPF Specification Vol. 2a, "HTTP Adaptive Streaming," v. 2.1, June 2011.

[8] 3GPP TS 23.203 v11.3.0, "Policy and Charging Control Architecture," Sept. 2011.

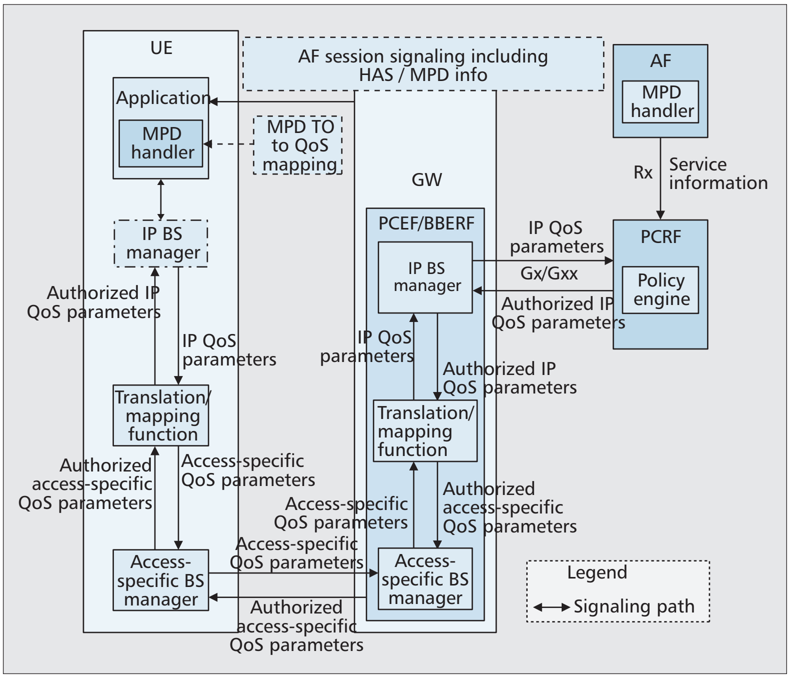


Figure 5. *Example PCC architecture to deliver end-to-end QoS for HAS services.*

[9] 3GPP TS 29.213 v11.0.0, "Policy and Charging Control Signalling Flows and Quality of Service (QoS) Parameter mapping," Sept. 2011.

[10] 3GPP TS 29.214 v11.2.0: "Policy and Charging Control over Rx Reference Point," Sept. 2011.

BIOGRAPHIES

OZGUR OYMAN (ozgur.oyman@intel.com) is a senior research scientist in the Wireless Communications Laboratory of Intel Labs. He is currently in charge of video over LTE research and standardization. He is author or co-author of over 70 technical publications, resulting in Best Paper Awards at IEEE GLOBECOM '07, ISSSTA '08, and CROWNCOM '08. His service includes TPC chair roles for technical symposia at IEEE WCNC '09, ICC '11, WCNC '12, and ICCC '12. He also serves as an editor for *IEEE Transactions on Communications*. He holds Ph.D. and M.S. degrees from Stanford University, and a B.S. degree from Cornell University.

SARABJOT SINGH (sarabjot@utexas.edu) received his B.Tech. degree in electronics and communication engineering from the Indian Institute of Technology Guwahati in 2010 and was awarded the President of India Gold Medal. He is currently working toward his Ph.D. at the University of Texas at Austin where his research is focused on the QoE-aware design and analysis of wireless systems. He has held internships at Intel Corporation in Santa Clara, California, Qualcomm Inc. in San Diego, California, and the Indian Institute of Science Bangalore.