Dynamic Adaptive Streaming over HTTP – Design Principles and Standards

Thomas Stockhammer
Qualcomm Incorporated

ABSTRACT

In this paper, we provide some insight and background into the Dynamic Adaptive Streaming over HTTP (DASH) specifications as available from 3GPP and in draft version also from MPEG. Specifically, the 3GPP version provides a normative description of a Media Presentation, the formats of a Segment, and the delivery protocol. In addition, it adds an informative description on how a DASH Client may use the provided information to establish a streaming service for the user. The solution supports different service types (e.g., On-Demand, Live, Time-Shift Viewing), different features (e.g., adaptive bitrate switching, multiple language support, ad insertion, trick modes, DRM) and different deployment options. Design principles and some forward-looking considerations are provided.

1. INTRODUCTION

Internet access is becoming a commodity on mobile devices. With the recent popularity of smart phones, smartbooks, connected netbooks and laptops the Mobile Internet use is dramatically expanding. According to recent studies [6], expectations are that between 2009 and 2014 the mobile data traffic will grow by a factor of 40, i.e., it will more than double every year. Figure 1 shows that the video traffic will by then account for 66% of the total amount of the mobile data. At the same time mobile users expect high-quality video experience in terms of video quality, start-up time, reactivity to user interaction, trick mode support, etc., and the whole ecosystem including content providers, network operators, service providers, device manufacturers and technology providers need to ensure that these demands can be met. Affordable and mature technologies are required to fulfil the users' quality expectations. One step into this direction is a common, efficient and flexible distribution platform that scales to the rising demands. Standardized components are expected to support the creation of such common distribution platforms.

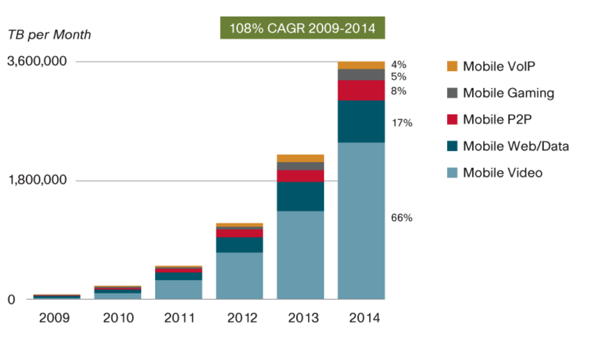

Figure 1 Video Will Account for 66 Percent of Global Mobile Data Traffic by 2014 (Source [7], Figure 2)

Traditional streaming generally uses a stateful protocol, e.g., the Real-Time Streaming Protocol (RTSP): Once a client connects to the streaming server the server keeps track of the client's state until the client disconnects again. Typically, frequent communication between the client and the server happens. Once a session between the client and the server has been established, the server sends the media as a continuous stream of packets over either UDP or TCP transport. In contrast, HTTP is stateless. If an HTTP client requests some data, the server responds by sending the data and the transaction is terminated. Each HTTP request is handled as a completely standalone one-time transaction.

Alternatively to streaming, progressive download may be used for media delivery from standard HTTP Web servers. Clients that support HTTP can seek to positions in the media file by performing byte range requests to the Web server (assuming that it also supports HTTP/1.1 [3]). Disadvantages of progressive download are mostly that (i) bandwidth may be wasted if the user decides to stop watching the content after progressive download has started (e.g., switching to another content), (ii) it is not really bitrate adaptive and (iii) it does not support live media services. Dynamic Adaptive Streaming over HTTP (DASH) addresses the weaknesses of RTP/RTSP-based streaming and progressive download.

2. DESIGN PRINCIPLES

HTTP-based progressive download does have significant market adoption. Therefore, HTTP-based streaming should be as closely aligned to HTTP-based progressive download as possible, but take into account the above-mentioned deficiencies.

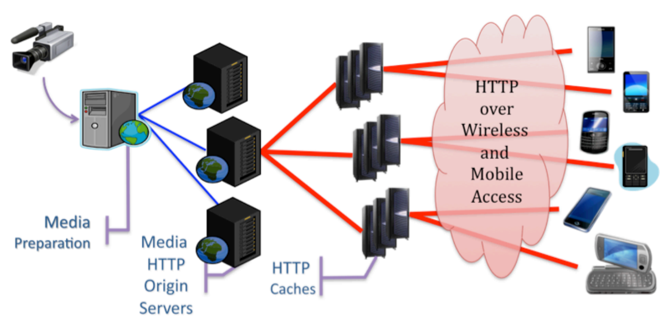

Figure 2 Example Media Distribution Architecture

Figure 2 shows a possible media distribution architecture for HTTP-based streaming. The media preparation process typically 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). The media origin server is preferably an HTTP server such that any communication with the server is HTTP-based (indicated by a bold line in the picture). Based on this MPD metadata information that describes the relation of the segments and how they form a media presentation, clients request the segments using HTTP GET or partial GET methods. The client fully controls the streaming session, i.e., it manages the on-time request and smooth playout of the sequence of segments, potentially adjusting bitrates or other attributes, for example to react to changes of the device state or the user preferences.

Massively scalable media distribution requires the availability of server farms to handle the connections to all individual clients. HTTP-based Content Distribution Networks (CDNs) have successfully been used to serve Web pages, offloading origin servers and reducing download latency. Such systems generally consist of a distributed set of caching Web proxies and a set of request redirectors. Given the scale, coverage, and reliability of HTTP-based CDN systems, it is appealing to use them as base to launch streaming services that build on this existing infrastructure. This can reduce capital and operational expenses, and reduces or eliminates decisions about resource provisioning on the nodes. This principle is indicated in Figure 2 by the intermediate HTTP servers/caches/proxies. Scalability, reliability, and proximity to the user's location and high-availability are provided by general-purpose servers. The reasons that lead to the choice of HTTP as the delivery protocol for streaming services are summarized below:

1. HTTP streaming is spreading widely as a form of delivery of Internet video.
2. There is a clear trend towards using HTTP as the main protocol for multimedia delivery over the Open Internet.
3. HTTP-based delivery enables easy and effortless streaming services by avoiding NAT and firewall traversal issues.
4. HTTP-based delivery provides reliability and deployment simplicity due as HTTP and the underlying TCP/IP protocol are widely implemented and deployed.
5. HTTP-based delivery provides the ability to use standard HTTP servers and standard HTTP caches (or cheap servers in general) to deliver the content, so that it can be delivered from a CDN or any other standard server farm.
6. HTTP-based delivery provides the ability to move control of "streaming session" entirely to the client. The client basically only opens one or several or many TCP connections to one or several standard HTTP servers or caches.
7. HTTP-based delivery provides the ability to the client to automatically choose initial content rate to match initial available bandwidth without requiring the negotiation with the streaming server.
8. HTTP-based delivery provides a simple means to seamlessly change content rate on-the-fly in reaction to changes in available bandwidth, within a given content or service, without requiring negotiation with the streaming server.
9. HTTP-based streaming has the potential to accelerate fixed-mobile convergence of video streaming services as HTTP-based CDN can be used as a common delivery platform.

Based on these considerations, 3GPP had identified the needs to provide a specification for a scalable and flexible video distribution solution that addresses mobile networks, but is not restricted to 3GPP radio access networks (RANs). 3GPP has taken the initiative to specify an Adaptive HTTP Streaming solution in addition to the already existing RTP/RTSP-based streaming solutions and the HTTP-based progressive download solution.

Specifically the solution is designed

- to support delivery of media components encapsulated in ISO base media file format box structure,
- to address delivery whereas presentation, annotation and user interaction is largely out-of-scope,
- to permit integration in different presentation frameworks.

The 3GPP sub-group SA4 working on codecs and protocols for media delivery started the HTTP streaming activity in April 2009 and completed the Release-9 specification work early March 2010. The 3GPP Adaptive HTTP Streaming (AHS) has been integrated into 3GPP Transparent end-to-end Packet-switched Streaming Service (PSS). Specifically, 3GPP TS 26.234 [1] (PSS Codecs and Protocols) clause 12 specifies the 3GPP Adaptive HTTP Streaming solution, and 3GPP TS 26.244 [2] (3GP File Format) clauses 5.4.9, 5.4.10, and 13 specify the encapsulation formats for segments. The Release-9 work is now under maintenance mode and some minor bug fixes and clarifications were agreed during the year 2010 and have been integrated into the latest versions of 3GPP TS 26.234 and 3GPP TS 26.244.

The solution supports features such as

- fast initial startup and seeking,
- bandwidth-efficiency,
- adaptive bitrate switching,
- adaptation to CDN properties,
- re-use of HTTP-server and caches,
- re-use of existing media playout engines,
- support for on-demand, live and time-shift delivery services,
- simplicity for broad adoption.

3GPP has also sought alignment with other organizations and industry fora that work in the area of video distribution. For example, as the Open IPTV Forum (OIPF) based their HTTP Adaptive Streaming (HAS) solution [8] on 3GPP. 3GPP recently also addressed certain OIPF requirements and integrated appropriate features in the Release-9 3GPP Adaptive HTTP Streaming specification. Also MPEG's draft DASH solution is heavily based on 3GPP's AHS. Finally, 3GPP has ongoing work in Release-10, now also referred to as DASH. This work will extend the Release-9 3GPP AHS specification in a backward-compatible way. Close coordination with the ongoing MPEG DASH activities is organized.

3. 3GPP Adaptive HTTP Streaming

3GPP Adaptive HTTP Streaming, since Release-10 referred to as as 3GP-DASH, is the result of a standardization activity in 3GPP SA4 Figure 3 shows the principle of the 3GP-DASH specification. The specification provides

- a normative definition of a Media Presentation, with Media Presentation defined as a structured collection of data that is accessible to the DASH Client through Media Presentation Description,

- a normative definition of the formats of a Segment, with a Segment defined as an integral data unit of a media presentation that can be uniquely referenced by a HTTP-URL (possibly restricted by a byte range),
- a normative definition of the delivery protocol used for the delivery of Segments, namely HTTP/1.1,
- an informative description on how a DASH client may use the provided information to establish a streaming service for the user.

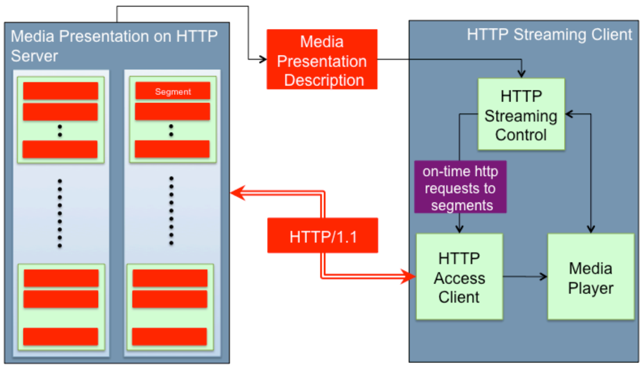

Figure 3 Solution overview – 3GP-DASH

DASH in 3GPP is defined in two levels:

1. Clause 12.2 in TS 26.234 [1] provides a generic framework for Dynamic Adaptive Streaming independent of the data encapsulation format for media segments.
2. Clause 12.4 in TS 26.234 [1] provides a specific instantiation of this framework with the 3GP/ISO base media file format by specifying the segment formats, partly referring to the formats in TS 26.244 [2].

This approach makes the framework defined in 3GPP extensible, for example to any other segment formats, codecs and DRM solutions.

3G-DASH supports multiple services, among others:

- On-demand streaming,
- Linear TV including live media broadcast,
- Time-shift viewing with network Personal Video Recording (PVR) functionalities.

Specific care was taken in the design that the network side can be deployed on standard HTTP servers and distribution can be provided through regular Web infrastructures such as HTTP-based CDNs. The specification also leaves room for different server/network-side deployment options as well as for optimized client implementations.

The specification also defines provisions to support features such as

- Initial selection of client- and/or user-specific representations of the content,
- Dynamic adaptation of the played content to react to environmental changes such as access bandwidth or processing power,
- Trick modes such as seeking, fast forward or rewind,
- Simple insertion of pre-encoded advertisement or other content in on-demand and live streaming services,
- Efficient delivery of multiple languages and audio tracks,
- Content protection and transport security, etc.

4. Ongoing Work in DASH

3GP-DASH defines the first standard on Adaptive Streaming over HTTP. Specific design principles have been taken into account that enables flexible deployments when using the formats defined in 3GP-DASH. Major players in the market, including those that offer proprietary solutions today, participated in the development of the specification. 3GP-DASH also serves as baseline for several other organizations, in particular the Open IPTV Forum and MPEG. Especially MPEG [7] is considering backward-compatible extensions to the 3GP-DASH specification to integrate additional media such as multiview or scalable video coding. Furthermore, initial efforts in interoperability testing have started. Currently there is great hope that the foundations laid in 3GP-DASH build the core package of an industry-standard for Dynamic Adaptive Streaming over HTTP (DASH).

One of the important aspects in MPEG and other fora will be a suitable definition of Adaptive Streaming Profile including media codecs, formats as well as selected options from the rich DASH specification. Envisaged profiles may focus on live services and On-Demand services, services targeting existing set-top boxes that for example are based on MPEG-2 TS processing, etc. In addition, prototyping and interoperability efforts are about to start.

5. ACKNOWLEDGMENTS

Many thanks to all the colleagues in 3GPP SA4 and MPEG DASH for the collaboration on the matter and their contributions to a hopefully successful and widely deployed standard.

6. REFERENCES

[1] 3GPP TS 26.234: "Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs".

[2] 3GPP TS 26.244: "Transparent end-to-end packet switched streaming service (PSS); 3GPP file format (3GP)".

[3] IETF RFC 2616: "Hypertext Transfer Protocol – HTTP/1.1", Fielding R. et al., June 1999.

[4] ISO/IEC 14496-12:2005 | 15444-12:2005: "Information technology – Coding of audio-visual objects – Part 12: ISO base media file format" | "Information technology – JPEG 2000 image coding system – Part 12: ISO base media file format".

[5] IETF RFC 3986: "Uniform Resource Identifiers (URI): Generic Syntax", Berners-Lee T., Fielding R. and Masinter L., January 2005

[6] Cisco White Paper: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2009-2014, http://bit.ly/bwGY7L

[7] ISO/IEC JTC1/SC29/WG11 N11338, Call for Proposals on HTTP Streaming of MPEG Media, April 2010, Dresden, Germany.

[8] OIPF Specification Volume 2a - HTTP Adaptive Streaming V2.0, 2010/09/07.