MPEG's Dynamic Adaptive Streaming over HTTP (DASH) - An Enabling Standard for Internet TV

Thomas Stockhammer
Qualcomm Incorporated
User Frustration in Web-based Video

- Video not accessible
  - Behind a firewall
  - Plugin not available
  - Bandwidth not sufficient
  - Wrong/non-trusted device
  - Wrong format
- Fragmentation
  - Devices
  - Content Formats
  - DRMs
- Low quality of experience
  - Long start-up delay
  - Frequent rebuffering
  - Low playback quality
  - No lip-sync
  - No DVD quality (language, subtitle)
- Expensive
  - Eats my bandwidth
  - Need a dedicated device
  - etc.

let’s do something...!?
DASH 2011

- Apple HLS 2008
- MS SS 2008
- OIPF 2009
- MPEG 2010
- 3GPP 2009
- W3C 2011?
- others

Delivery Format

Confidence
DASH in a Nutshell

- **What:** Video streaming solution where small pieces of video streams/files are requested with HTTP and spliced together by the client. Client entirely controls delivery.
- **Why:** reuse widely deployed standard HTTP servers/caches for scalable delivery, e.g. existing Internet CDNs; traverse NAT/Firewalls; simple rate adaptation; fixed-mobile convergence; convergence of services, etc.
- **Use case:** Accessing OTT video streaming services over any access network to any device
MPEG DASH ISO/IEC 23009-1

• MPEG DASH ISO/IEC 23009-1 technically frozen in August 2011
• Timeline and Activities
  – Draft International Standard (DIS) 23009-1 publicly available
  – 2 months balloting period until October 2011
  – Parallel approval process for extensions to
    • ISO base media FF to support DASH 14496-12/AMD 3
    • Common Encryption 23001-7
  – Continuous coordination with 3GPP and other SDOs (DECE, OIPF, etc.)
  – Conformance and Reference Software activities kicked off (see WD 23009-2)
  – Licensing and promotional efforts ongoing – see last slide
• Good news: Converging standard for adaptive streaming on the way

Convergence = Confidence
(Some) DASH Design Principles

- **DASH is not:**
  - system, protocol, presentation, codec, middleware, client specification
- **DASH is an enabler**
  - provides **formats** to enable efficient and high-quality delivery of streaming services over the Internet considered as one component in an e2e service
  - System definition left to other organizations (SDOs, Fora, Companies, etc.)
- **It attempts to be very good in what is to be addressed by the standard**
  - Enables reuse of existing technologies (containers, codecs, DRM etc.)
  - Enables deployment on top of HTTP-CDNs (Web Infrastructures, caching)
  - Enables very high user-experience (low start-up, no rebuffering, trick modes)
  - Enables selection based on network and device capability, user preferences
  - Enables seamless switching
  - Enables live and DVD-kind of experiences
  - addresses global and regulatory deployment issues
  - Moves intelligence from network to client, enables client differentiation
  - Enables deployment flexibility (e.g., live, on-demand, time-shift viewing)
  - Provide simple interoperability points (profiles)
  - provides convergence with existing proprietary technologies in this space
What is specified – and what is not?

Media Presentation on HTTP Server

Segment

Resources located by HTTP-URLs

HTTP/1.1

Media Presentation Description

DASH Client

DASH Access Engine

on-time http requests to segments

HTTP Access Client

Media Engines
• MPD and Index Information for DASH Access client
  – Core specification aspects of DASH
• Initialisation and Media Segments for Media engine
  – Reuse of existing container formats and easy conversion
  – Small adaptations may be necessary for usage in DASH
Media Presentation Data Model

- Media Presentation Description (MPD) describes accessible Segments and corresponding timing

**Media Presentation**
- Period, start=0s
  - ...
- Period, start=100s
  - ...
- Period, start=295s
  - ...
- ...

**Segment Info**
- Initialization Segment
  - http://www.e.com/dash-5

**Adaptation Sets**
- **Adaptation Set 1**
  - video
- **Adaptation Set 2**
  - audio
- ...

**Representation 1**
- **Segment Info**
  - duration=10s
  - Template: ./dash-5-$Number$
- **Representation**
  - bandwidth=500kbit/s
  - width 640, height 480
- ...

**Representation 2**
- **Segment Info**
  - duration=10s
  - Template: ./dash-5-$Number$
- **Representation**
  - bandwidth=250kbit/s
  - width 640, height 480
- ...

**Media Segments**
- Media Segment 1
  - start=0s
  - http://www.e.com/dash-5-1
- Media Segment 2
  - start=10s
  - http://www.e.com/dash-5-2
- Media Segment 3
  - start=20s
  - http://www.e.com/dash-5-3
- Media Segment 20
  - start=190s
  - http://www.e.com/dash-5-20

- Splicing of arbitrary content
- Selection of Components/Tracks
- Select/Switch of Bandwidth
Key feature – Common Timeline

• Representations in one Period share common presentation timeline
  – presentation time of access unit within the media streams is mapped to the global common presentation timeline
  – enables synchronization of different media components and seamless switching of different coded versions of the same media components

• Other timelines
  – segment availability times (mapped to UTC clock)
  – internal media decode time (not exposed on DASH level)
**Profile Identifier**

Type: ‘Live’ or ‘On-Demand’

**Adaptation Set:** Set of switchable Representations

**Descriptors**

**Representation:** Encoded version of a media component

**Period:** Time sequence of Media Presentation
Video/Audio Parameters

Codecs, Container

Bandwidth

Sub-Representations

URL Construction

Common Base

Playlist-based

Template-based
MPD Information

- Redundant information of Media Streams for the purpose to initially select or reject Adaptation Sets/Representations
  - Examples: Role, Codec, DRM, language, resolution, bandwidth
- Access and Timing Information
  - the HTTP-URL(s) and byte range for each accessible Segment
  - the earliest next update of the MPD on the server
  - the segment availability start and end time in wall-clock time
  - the approximated presentation start time and duration of a Media Segment in the media presentation timeline
  - for live service, playout start instructions such that segments will be available in time for fluent playout in the future
- Switching and splicing relationships across Representations
- not much more ...
Accessing Segments

- Multiple Base URLs
  - same information can be accessed at multiple locations
  - Redundancy, client-side load balancing, parallel download

- Byte range access with regular GETs
  - mapping to byte ranges needs to be done in CDNs
  - includes environments for which direct access to HTTP stack is not possible (browser-plugins)
Descriptors

- Content Protection (2 schemes defined)
- Role (1 scheme defined)
  - caption, subtitle, main, alternate, supplementary, commentary, dub
- Accessibility (Role scheme may be used)
- Rating
- Viewpoint
- Frame Packing (2 schemes defined)
- Audio Channel Configuration (1 scheme defined)
<MPD>
  <Period>
    <AdaptationSet mimeType="video/mp4" group="1">
      <Role schemeIdUri="urn:mpeg:DASH:role:2011" value="main"/>
      <Viewpoint schemeIdUri="urn:mpeg:DASH:viewpoint:2011" value="vp1"/>
      <Representation id="11" bandwidth="1024000">...</Representation>
      <Representation id="12" bandwidth="512000">...</Representation>
    </AdaptationSet>
  </Period>
  <AdaptationSet mimeType="video/mp4" group="1">
    <Role schemeIdUri="urn:mpeg:DASH:role:2011" value="alternate"/>
    <Viewpoint schemeIdUri="urn:mpeg:DASH:viewpoint:2011" value="vp2"/>
    <Representation id="21" bandwidth="1024000">...</Representation>
    <Representation id="22" bandwidth="512000">...</Representation>
  </AdaptationSet>
  <AdaptationSet mimeType="audio/mp4" group="2">
    <Role schemeIdUri="urn:mpeg:DASH:role:2011" value="main"/>
    <Role schemeIdUri="urn:mpeg:DASH:role:2011" value="supplementary"/>
    <Viewpoint schemeIdUri="urn:mpeg:DASH:viewpoint:2011" value="vp1"/>
    <Representation id="31" bandwidth="1280000">...</Representation>
    <Representation id="32" bandwidth="640000">...</Representation>
  </AdaptationSet>
  <AdaptationSet mimeType="audio/mp4" group="2">
    <Role schemeIdUri="urn:mpeg:DASH:role:2011" value="alternate"/>
    <Role schemeIdUri="urn:mpeg:DASH:role:2011" value="supplementary"/>
    <Viewpoint schemeIdUri="urn:mpeg:DASH:viewpoint:2011" value="vp2"/>
    <Representation id="41" bandwidth="1280000">...</Representation>
    <Representation id="42" bandwidth="640000">...</Representation>
  </AdaptationSet>
</MPD>
Segment Indexing

• Provides binary information in ISO box structure on
  – Accessible units of data in a media segment
  – Each unit is described by
    • Byte range in the segments (easy access through HTTP partial GET)
    • Accurate presentation duration (seamless switching)
    • Presence of representation access positions, e.g. IDR frames

• Provides a compact bitrate-over-time profile to client
  – Can be used for intelligent request scheduling

• Generic Data Structure usable for any media segment format, e.g. ISO BMFF, MPEG-2 TS, etc.

• Hierarchical structuring for efficient access

• May be combined with media segment or may be separate
Segment and Subsegment Index for MPEG-2 TS
Media Segments

- Contains the actual segmented media streams
- additional information to map segment into media presentation timeline for switching and synchronous presentation with other Representations
- For ISO BMFF, contains one or more movie fragments
- Can be short (≈1-10 sec) and long (≈10sec – 2h)

<table>
<thead>
<tr>
<th>Segment duration</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>• Suitable for live</td>
<td>• Large number of files</td>
</tr>
<tr>
<td></td>
<td>• commonality with live</td>
<td>• Large number of URLs</td>
</tr>
<tr>
<td></td>
<td>• High switching granularity on segment level</td>
<td>• Fixed request size</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• switching granularity on segment level</td>
</tr>
<tr>
<td>Long</td>
<td>• Small number of files</td>
<td>• Need for Segment Index</td>
</tr>
<tr>
<td></td>
<td>• Small number of URLs</td>
<td>• Difference from Live</td>
</tr>
<tr>
<td></td>
<td>• High switching granularity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Flexible request sizes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Improved cache performance</td>
<td></td>
</tr>
</tbody>
</table>
Live Presentation

• Live Services enabled
  – Generation of Segments on-the-fly
  – Access of only a subset of the Segments within a time window
  – Server/Network may offer Segments only for a certain time window
  – Update of MPD to describe new Segments and/or new Periods, such that the updated
    MPD is compatible with the previous MPD to ensures that
    • clients may immediately begin using the new MPD without synchronisation with
      the old MPD, since it is compatible with the old MPD before the update time; and
    • the update time needs not be synchronised with the time at which the actual
      change to the MPD takes place: i.e. changes to the MPD may be advertised in
      advance
  – Media Presentation is described by the initial MPD and all updates.
  – With URL templates, updating of MPD generally not necessary
  – Client and server are expected to be synchronized to UTC time.
• Time-shift viewing and network PVR functionality seamlessly enabled
  – Segments may be accessible on the network over a long time.
Profiles

- Set of restrictions on the offered Media Presentation (MPD & Segments)
- can also be understood as permission for DASH clients that only implement the features required by the profile to process the Media Presentation
- Profiles defined in ISO/IEC 23009 (as below). More restrictions may be added
ISO Base media file format On Demand

MPEG-2 TS simple
Summary: DASH Selected Feature List

• Live, On-Demand and Time-shift services
• Independency of request size and segment size (byte range requests)
• Segment formats
  – ISO base media FF and MPEG-2 TS
  – guidelines for integrating any other format
  – Are codec independent
• Support for server and client-side component synchronization (e.g., separate and multiplexed audio and video)
• **Support for efficient trick mode**
• Simple splicing and (targeted) ad insertion
• Multiple base URLs for the same content
• Clock drift control for live sessions
• DASH metrics for reporting the session experience
• Profile: restriction of DASH and system features (claim & permission)
• Content Descriptors for Protection, Accessibility, Rating, etc.
  – Enables common encryption, but different DRM (DECE-like)
DEPLOYMENT CONSIDERATIONS
Common Uses Cases

• MPEG-DASH supports simple and advanced use cases:
  – On-Demand, Live and time-shift (nPVR) streaming
  – Dynamic ad-insertion
  – Dynamic update of program
  – Delivery of same content on three screens
  – Delivery of any multimedia content (2D, 3D, animation, graphics, multiview, subtitles, text, etc.), not just AV
  – Support of multiple languages and different audio configuration
  – etc.

• Simple use cases can be gradually extended to more complex and advanced ones
Migration Scenarios

• Most generated content/production equipment for legacy Adaptive Bitrate Streaming systems can be used for MPEG-DASH:
  – generic encoders can be reused, DASH adds descriptive metadata for better client operations
  – HLS Content suitable for DASH M2TS Main profile.
  – Smooth Streaming Content suitable for DASH ISOBMFF Live profile.
• Manifest files can be easily converted to MPD format
  – XML conversion from m3u8 and Smooth Streaming manifests.
  – Deployment of two manifest files (legacy and DASH MPD) in parallel (low overhead)
• Documentation in preparation ...
• It’s not a competition
Next steps

• Complete standardization work
  – Formal approval of all specifications
  – Conformance, interoperability and reference software

• Towards deployments
  – Generate end-to-end system specs based on DASH including codecs, DRM, profiles, etc. (OIPF, 3GPP, HbbTV, HD Forum, etc.)
  – Generate guidelines, white papers, test content and software
  – Promotional efforts: Licensing, interoperability, plug-fests, etc.
  – Combine it with browsers, the web and HTML-5

• Everyone is invited to contribute
More Information

• Draft Specifications
  – 14496-12:2008/FDAM-3:
    http://www.3gpp.org/ftp/Inbox/LSs_from_external_bodies/
    ISO_IEC_JTC1_SG29_WG11/29n12310.zip
  – 23001-7: http://www.3gpp.org/ftp/Inbox/LSs_from_external_bodies/
    ISO_IEC_JTC1_SG29_WG11/29n12313.zip
  – 23009-1: http://www.3gpp.org/ftp/Inbox/LSs_from_external_bodies/
    ISO_IEC_JTC1_SG29_WG11/29n12316.zip

• More information from Qualcomm including Qualcomm‘s licensing position

• Several other companies have declared or expressed willingness to declare favorable licensing conditions
THANK YOU

Comments – Questions – Feedback