Short Tutorial on MPEG-DASH

Dynamic Adaptive Streaming Over Internet

ISO/IEC 23009-1

MPEG DASH Subgroup
MPEG-DASH ISO/IEC 23009-1

- Provides a superset for system specifications
  - Open IPTV Forum HTTP Adaptive Streaming (HAS)
  - System specifications may define more: codecs, DRM, etc.

- Timeline and Activities
  - Draft International Standard (DIS) 23009-1 available publicly
  - 2 months balloting period until October 2011
  - Parallel approval process for extensions to ISO base media FF to support DASH
  - Continuous coordination with 3GPP and other organizations (DECE, OIPF, etc.)
  - Conformance and Reference Software activities kicked off (see WD 23009-2)
MPEG-DASH Standard Development

<table>
<thead>
<tr>
<th>Stage</th>
<th>Steps</th>
<th>Approval Process</th>
<th>Timeline (date of issue)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploration</td>
<td>1st Workshop on MMT</td>
<td>-</td>
<td>July 2009</td>
</tr>
<tr>
<td></td>
<td>1st Draft CFP</td>
<td>-</td>
<td>October 2009</td>
</tr>
<tr>
<td></td>
<td>2nd Workshop on MMT</td>
<td>-</td>
<td>January 2010</td>
</tr>
<tr>
<td>Committee Stage</td>
<td>Call for Proposals</td>
<td>MPEG consensus</td>
<td>April 2010</td>
</tr>
<tr>
<td></td>
<td>Working Draft</td>
<td>MPEG consensus</td>
<td>July 2010</td>
</tr>
<tr>
<td></td>
<td>Committee Draft</td>
<td>SC29 Ballot</td>
<td>October 2010</td>
</tr>
<tr>
<td>Approval Stage</td>
<td>Draft International Standard</td>
<td>JTC1 Ballot</td>
<td>January 2011</td>
</tr>
<tr>
<td></td>
<td>2nd Draft International Standard</td>
<td>JTC1 Ballot</td>
<td>August 2011</td>
</tr>
<tr>
<td></td>
<td>International Standard (submit for publication)</td>
<td></td>
<td>Dec 2011 (expected)</td>
</tr>
</tbody>
</table>

DASH is an example of developing the standard in the shortest time possible.
Adaptive Dynamic Streaming Simple Use Case

Period 1
- 5M
- 2M
- 0.5M
- Trick Mode

Video

Audio 1
- English
- 128K AAC
- 48K AAC

Audio 2
- French
- 128K AAC
- 48K AAC

Period 2

Period 3

Surround

Surround (48K AAC)
MPEG-DASH Design Principles

- DASH is an enabler
  - provides formats to enable efficient and high-quality streaming over the Internet
  - System definition left to other organizations (SDOs, Fora, Companies, etc.)
- DASH is not:
  - system, protocol, presentation, codec, middleware, client specification
- 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?
Information Classification

- MPD and Index Information for DASH Access client
  - Core specification aspects of DASH

- Initialization 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 Description (MPD) Data Model

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

Period,
- start=100
- baseURL=http://www.e.com/
  • bandwidth=500kbit/s
  • width 640, height 480

Adaptation Set 1
- video

Adaptation Set 2
- audio

Representation 1
- duration=10s
- Template: ./ahs-5-$Index$.3gs

Segment Info
- Initialization Segment
  http://www.e.com/ahs-5.3gp

Segment Info
- Media Segment 1
  start=0s
  http://www.e.com/ahs-5-1.3gs

Segment Info
- Media Segment 2
  start=10s
  http://www.e.com/ahs-5-2.3gs

Segment Info
- Media Segment 3
  start=20s
  http://www.e.com/ahs-5-3.3gs

Segment Info
- Media Segment 20
  start=190s
  http://www.e.com/ahs-5-20.3gs

Splicing of arbitrary content
Selection of Components
Select/Switch of Bandwidth
MPD Schema Overview

- **Profile Identifier**
- **Type**: ‘Live’ or ‘On-Demand’
- **Adaptation Set**: Set of switchable Representations
- **Representation**: Encoded version of a media component
- **Period**: Time sequence of Media Presentation
MPD Schema - Representation

- Video/Audio Parameters
- Codecs, Container
- Descriptors
- Bandwidth
- URL Construction
- Template-based
- Playlist-based
MPD Information

- Information of Media Streams for the purpose to initially select or reject Representations
  - Examples: 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 media start time and duration of a Media Segment
  - for live service, instructions on starting playout such that media segments will be available in time for fluent playout in the future

- Switching and splicing relationships across Representations
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 by 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
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 (≈10 sec – 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 on-demand common with live • High switching granularity on segment level</td>
<td>• Large number of files • Large number of URLs • Fixed request size • Switching granularity on segment level</td>
</tr>
<tr>
<td></td>
<td>• Small number of files • Small number of URLs • High switching</td>
<td>• Need for Segment Index • Difference from Live</td>
</tr>
</tbody>
</table>

13
DASH Selected Features

- 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
  - Codec independent: guidelines for integrating any other format
- Support for server and client-side component synchronization
- Support for efficient trick mode
- Simple splicing and (targeted) ad insertion
- Definition of quality metrics for logging processes
- Content Descriptors for Protection, Accessibility, Rating, etc.
- Multi-DRM support.
- Profile: restriction of DASH and system features
Profiles

- Set of restrictions on the Media Presentation (MPD & Segments)
- Permission for DASH clients that only implement the required features
- 6 Profiles defined in ISO/IEC 23009:
Highlighted Features

- Supports adaptive on demand and live streaming such as MPEG-4 file format and MPEG-2TS.
- Efficient and ease of use of existing CDNs, proxies, caches, NATs and firewalls.
- Control of entire streaming session by the client.
- Support of seamless switching of tracks.
- The concept of switching and selectable streams.
- Signaling, delivery, utilization of multiple DRM schemes.
- Supports ad-insertion.
- Segments with variable durations.
- Sub-segment alignment indication to simplify switching and avoiding overlapping fragments.
Highlighted Features

- Manifest fragmentation and assembly for external inclusion of elements.
- Content Descriptors for accessibility, roles, rating and camera views.
- Multiple base URLs for the same content.
- Clock drift control for live sessions.
- Scalable Video Coding (SVC) and Multiview Video Coding (MVC).
- Subsetting of representation groups according to the content author’s guidance.
- Quality metrics for reporting the session experience.