An Introduction to DASH (Dynamic Adaptive Streaming over HTTP)

Presented by Ye-Kui Wang (slides mostly by Thomas Stockhammer)

April 2015, San Diego
Motivation: Consumer Internet Video Composition

Source: http://ciscovni.com, EB: 1e18 bytes
Global Mobile Data Traffic

Percentage Increase: 46%

Source: Ericsson (June 2013)

Source: http://www.ericsson.com/mobility-report
Adaptive Streaming over HTTP

• **Basic Approach: Adapt Video to Web rather than Changing the Web**

• **Streaming realized by continuous Short Downloads**
  – Downloads in small chunks to minimize bandwidth waste
  – Enables monitoring consumption and tracking clients

• **Adaptation to Dynamic Conditions and Device Capabilities**
  – Adapts to dynamic conditions anywhere on the path through the Internet or home network
  – Adapts to display resolution, CPU and memory resources of the client
  – Facilitates “any device, anywhere, anytime” paradigm

• **Improved Quality of Experience**
  – Enables faster start-up and seeking (compared to progressive download)
  – Reduces and may eliminate rebuffering, skips, freezes and stutters

• **Use of HTTP**
  – Well-understood naming/addressing approach
  – Provides easy traversal for all kinds of middleboxes (e.g., NATs, firewalls)
  – Enables cloud access, leverages existing HTTP caching infrastructure
  – Enables client-driven deployments
  – Enables reuse of existing web technologies: authentication, authorization, etc.
Scalability and Cost: Leveraging HTTP Caches

Send Segment 3

Segment 3

Send Segment 3

Segment 3

Send Segment 3

Segment 3

Time
Adaptive Streaming over HTTP – Common Understanding

1. Encode each video at multiple bitrates
2. Split the videos into small temporal segments
3. Encrypt each segment
4. Make each segment addressable via an HTTP-URL
5. Client makes decision on which segment to download
6. Client acquires a license for encrypted content
7. Client splices segments together and plays back
MPEG DASH
ISO/IEC 23009-1

Common Delivery Format

Confidence
Timeline and Standards in MPEG

Fastest time ever that a standard was developed in MPEG to address the demand of the market

- **ISO/IEC 23009 Parts**
  - Part 1: Media Presentation Description and Segment Formats
  - Part 2: Conformance and Reference Software
  - Part 3: Implementation Guidelines

- **Other Relevant MPEG Standards**
  - 14496-12: ISO Base Media File Format (ISOBMFF)
  - 14496-15: Carriage of NAL unit structured video in ISOBMFF
  - 23001-7: Common Encryption in ISOBMFF
  - ISO/IEC 14496-30: Timed Text and other visual overlays in ISOBMFF
  - Codec specs for AVC, HEVC, audio, etc.
MPEG – Dynamic Adaptive Streaming over HTTP

• **Goal**
  – Develop an international, standardized, efficient solution for HTTP-based streaming of MPEG media

• **Major Objectives and Design Principles**
  – Do the necessary, avoid the unnecessary
  – Be lazy: reuse what exists in terms of codecs, formats, content protection, protocols and signaling
  – Be backward-compatible (as much as possible) to enable deployments aligned with existing proprietary technologies
  – Be forward-looking to provide ability to include new codecs, media types, content protection, deployment models (ad insertion, trick modes, etc.) and other relevant (or essential) metadata
  – Enable efficient deployments for different use cases (live, VoD, time-shifted, etc.)
  – Focus on formats describing functional properties for adaptive streaming, not on protocols or end-to-end systems or implementations
  – Enable application standards and proprietary systems to create end-to-end systems based on DASH formats
  – Support deployments by conformance and reference software, implementation guidelines, etc.
Scope of MPEG DASH (Shown in Red)
Major Functional Components (1) – Data Model

- Provide information to a client, where and when to find the data that composes A/V experience → MPD
- Provide the ability to offer a service on the cloud and HTTP-CDNs → HTTP-URLs and MIME Types
- Provide service provider the ability to combine/splice content with different properties into a single media presentation → Periods
- Provide service provider to enable the client/user selection of media content components based on user preferences, user interaction device profiles and capabilities, using conditions or other metadata → Adaptation Sets
- Provide ability to provide the same content with different encodings (bitrate, resolution, codecs) → Representations
- Provide extensible syntax and semantics for describing Representation and Adaptation Set properties → Descriptors
- Provide ability to access content in small pieces and do proper scheduling of access → Segments and Subsegments
- Provide ability for efficient signaling and deployment optimized addressing → Playlist, Templates, Segment Index
- Provide ability to enable reuse of existing encapsulation and parsing tools → MPEG2-TS and ISO-BMFF
DASH Data Model

MPD
- Period id = 1
  start = 0 s
- Period id = 2
  start = 100 s
- Period id = 3
  start = 300 s
- Period id = 4
  start = 850 s

Period id = 2
start = 100 s
- Adaptation Set 0
  subtitle turkish
- Adaptation Set 1
  video
- Adaptation Set 2
  audio english
- Adaptation Set 3
  audio german

Adaptation Set 1
BaseUrl=http://abr.rocks.com/
- Representation 2
  Rate = 500 Kbps
- Representation 3
  Rate = 1 Mbps
- Representation 4
  Rate = 2 Mbps

Representation 3
Rate = 2 Mbps
Resolution = 720p
- Segment Info
  Duration = 10 s
  Template: 3/$Number$.mp4

Segment Access
- Initialization Segment
  http://abr.rocks.com/3/0.mp4
- Media Segment 1
  start = 0 s
  http://abr.rocks.com/3/1.mp4
- Media Segment 2
  start = 10 s
  http://abr.rocks.com/3/2.mp4

Splicing of arbitrary content, e.g. ad insertion
Selection of Components/Tracks based on properties
Selecting/Switching of Representation based on bandwidth, etc.
Well-defined media format, i.e. ISO BMFF or MPEG-2 TS
Media Delivery Format, chunks with unique addresses + associated timing
Representations in One Adaptation Set

➡ Enabling Seamless Switching

Role and Properties of each Media Component, e.g. language, main video, accessibility

“Original” Video

Common to all Reps

Processing A
(Subsampling, etc.)

Processing B
(Subsampling, etc.)

Processing C
(Subsampling, etc.)

Encoding 1

Encoding 2

Encoding 3

Encoding 4

Encoding 5

Decoding 1

Decoding 2

Decoding 3

Decoding 4

Decoding 5

HTTP Access Client

Processing⁻¹
A*

Processing⁻¹
B*

Processing⁻¹
C*

DASH Client

Codec, profile, level, bandwidth

frame rate, spatial resolution, field/frame coding
• Representations are conceptual, Segments are accessible and well defined data units
• For ISO BMFF based media formats, segments are based on fragmented movie files
• Signaling per Representation in MPD
  – The type of the segment, mostly an Initialization or Media Segment
  – The URL of each media segment through explicit list or template, or byte range
    • Number template: \texttt{http://abr.rocks.com/Rep1/$Number$.mp4} \rightarrow \texttt{http://abr.rocks.com/Rep1/1.mp4}
    • Time template: \texttt{http://abr.rocks.com/Rep1/$Time$.mp4} \rightarrow \texttt{http://abr.rocks.com/Rep1/3465.mp4}
    • Segment Index & byte ranges: \texttt{http://abr.rocks.com/Rep1.mp4} \rightarrow \texttt{http://abr.rocks.com/Rep1.mp4} range 3190-8959
  – The (approximate) start time and duration of each media segment (for seeking)
  – Start with SAP indication: Indicates SAP type at the start of a segment
  – The mapping of the internal movie time to the global media presentation timeline
  – The segment availability times for dynamic services
Major Functional Components (2) - Timing

• **Common Media Presentation Time**
  – Provide ability to present content from different adaptation sets synchronously
  – Provide ability to support seamless switching across different representations

• **Switching Support features**
  – Signalling of Stream Access Points
  – Segment Alignment to avoid overlap downloading and decoding

• **Play-out and decode times per Segment and Track fragment**
  – Provide ability to randomly access and seek in the content

• **Segment Availability Time**
  – Mapped to wall-clock time
  – Expresses when a segment becomes available on the server and when ceases it to be available
  – Provide ability to support live and time-shift buffer services with content generated/removed on the fly
Major Functional Components (3) - Operations

• Provide ability for personalized access to media presentation, e.g. targeted advertisement → **MPD Assembly with xlink**
• Provide ability to provide redundant content offering → **Multiple Base URLs**
• Provide ability to announce unforeseen/unpredictable events in live services → **MPD Updates**
• Provide ability to send events associated with media times → **Inband and MPD-based Event Messages**
• Provide the ability to log and report client actions → **DASH metrics**
• Provide ability to efficiently support trick modes → **Dedicated IDR-frame Representations and Sub-representations**
• Provide ability to signal collection of a subset/extension of tools → **Profiles and Interoperability Points**
Status of MPEG DASH in Industry

• DASH Industry Forum (http://dashif.org)
  – Founded in 2012 to promote and catalyze market adoption of MPEG DASH (70 members)
  – Established a mediator role among different communities: standardization organizations, interoperability groups, larger and smaller business entities, researchers, open source community, public and press
  – Successful demonstrations and events at IBC’12, MWC’13, NAB’13, …
  – Published DASH-AVC/264 Interoperability Guidelines
  – Published draft versions of test vectors, test services, conformance software and open source reference client based HTML-5 extensions and Javascript

• Many SDOs adopt MPEG-DASH as THE technology for Internet TV
  – HbbTV, DTG, 3GPP, DLNA, ATSC, OIPF, CableLabs and many more …

• Internet Streaming Services based on DASH
  – Netflix continues to stream DASH compatible content (but does not use MPDs)
  – NAGRA and Abertis telecom launch hosted multi-screen service
  – DASH is the transport format for LTE broadcast systems
  – Samsung and Orange announced to deploy DASH and HEVC
  – …
DASH Related Standardization Activities in Different SDOs

• MPEG ➔ bug fixing & some core experiments
• 3GPP ➔ lots of work on DASH over eMBMS/LTE Broadcast
• DVB ➔ defining a toolset for broadcasters to deploy their services with DASH over the top (mostly HbbTV context)
• ATSC, SCTE ➔ deploying OTT/Hybrid based on DASH
• DASH-IF ➔ create and demonstrate functionalities and interoperability
• W3C ➔ browser integration
• IETF ➔ delivery optimizations: CDNI, HTTP2.0, Multicast
Ongoing Work on DASH in MPEG

- Recurring work: Corrigenda, Conformance, Guidelines, integration of new codecs
- Amendment of Second Edition on the way to support
  - Improved live services and robustness
  - Profile for Ad Insertion
- Specific topics currently under work
  - Descriptor for Spatial Relationship in DASH (SRD)
  - Server and Network assisted DASH Operation (SAND)
  - DASH Client Authentication, Content Access Authorization, Controlling the Client Behavior
  - Quality-based streaming
  - DASH over emerging protocols HTTP/2.0 and web sockets
  - SAP-Independent Segment Signalling (SISSI)
  - Content aggregation and playback control (CAPCO)
  - ...