Guidelines for Implementation: DASH264 Interoperability Points

January 14, 2013

DASH Industry Forum

Version 0.9 (For Community Review)

Note: This document is no specification.

1 Scope

2 The scope of the interoperability point defined in this document is to provide basic support for
3 high-quality video distribution over the top. Both live and on-demand services are supported.

Disclaimer

This document is not yet final. It is provided for public review until March 15th, 2013. If you have comments on the document, please mail the comments to: iop-track@dashpg.org with a detailed description of the problem and the comment. Based on the received comments a final document will be published by March 31st, 2013.

This is a draft document made available by DASH-IF. The technology embodied in this draft specification may be covered under patents, including patents owned by such companies. No patent license, either implied or express, is granted to you by this document. This draft specification is provided on an as-is basis without any warranty whatsoever.

In addition, this document may include references to documents and/or technologies controlled by third parties. Those documents and technologies may be subject to third party rules and licensing terms. No intellectual property license, either implied or ex-press, to any third party material is granted to you by this document or DASH-IF. DASH-IF makes no any warranty whatsoever for such third party material.

Contents

List of Figures

Acronyms, abbreviations and definitions

For acronyms, abbreviations and definitions refer to ISO/IEC 23009-1 [1].

In addition, the following abbreviations and acronyms are used in this document:

AAC	Advanced Audio Coding
AVC	Advanced Video Coding
DRM	Digital Rights Management
DTV	Digital Television
FCC	Federal Communications Commission
GOP	Group-of-Pictures
KID	common Key IDentification
IDR	Instantaneous Decoder Refresh

1	PPS	Picture Parameter Set
2	SD	Standard Definition
3	SMPTE	Society of Motion Picture and Television Engineers
4	SPS	Sequence Parameter Set
5	TT	Timed Text
6	TTML	Timed Text Markup Language

1. Introduction

The scope of the DASH264 interoperability point is the basic support high-quality video distribution over the top. Both live and on-demand services are supported. It is expected that the client supports at least

- presentation of high-definition video up to 720p (based on H.264/AVC [5][6] Progressive High Profile)
- presentation of stereo audio
- support of basic subtitles
- basic support for encryption/DRM.

In addition, it is recognized that certain clients may only be capable to operate with H.264/AVC Main Profile. Therefore content authors may provide and signal a specific subset of DASH264 by providing a specific profile identifier referring to a standard definition presentation. This interoperability point is defined as DASH264 SD.

The test cases and test vectors for DASH264 Interoperability Points are defined in [24]. The conformance software for DASH264 Interoperability Points is defined in [25] which itself is based on the MPEG conformance software [2].

This version of the document defines the following Interoperability Points:

Interoperability Point	Identifier	Section
DASH264	`http://dashif.org/guidelines/dash264`	5.3
DASH264 SD	`http://dashif.org/guidelines/dash264#sd`	6.3

Beyond this, this document is expected to define further extensions in future versions.

2. DASH-Related Aspects

2.1. Scope

DASH264 is based on the ISO base media file format and has much in common with a superset of the ISO BMFF On-Demand and the ISO BMFF Live profile [1], sections 8.3 and 8.4, respectively. DASH264 is intended to provide basic support for on-demand and live content. The primary constraints imposed by this profile are the requirement that each Representation is provided in one of the following two ways

- as a single Segment, where Subsegments are aligned across Representations within an Adaptation Set and where Subsegments must begin with Stream Access Points (SAPs). This permits scalable and efficient use of HTTP servers and simplifies seamless switching. This is mainly for on-demand use cases.
- as a sequence of Segments where each Segment is addressable by a template-generated URL. This is mainly for live use case.

In addition, (Sub)Segments are constrained so that for switching video Representations within one Adaptation Set the boundaries are aligned without gaps or overlaps in the media data. Furthermore, switching is possible by a DASH client that downloads, decodes and presents the media stream of the come-from Representation and then switches to go-to Representation by downloading, decoding and presenting the new media stream. No overlap in downloading, decoding and presentation is required for seamless switching of Representations in one Adaptation Set.

2.2. DASH features

2.2.1. Introduction

This section introduces the detailed constraints of the MPD and the DASH segments in a descriptive way as done for the MPEG-DASH [1] standard. The interoperability point has significant commonality with the ISO BMFF Live and On-Demand profile from the DASH specification.

Specifically:

- Segment formats are based on ISO BMFF with fragmented movie files
- Alignment with ISO BMFF Live & On-Demand Profiles, i.e. within each Adaptation Set the following applies
  - Fragmented movie files are used for encapsulation of media data
  - (Sub)Segments are aligned to enable seamless switching

Beyond the constraints provided in the ISO BMFF profiles, the following additional restrictions are applied.

- IDR-like SAPs (i.e., SAPs type 2 or below) at the start of each (sub)segment for simple switching.
- Segments have almost equal duration. The maximum tolerance of segment duration is ±25% and the maximum accumulated deviation over multiple segments is ±25% of the signaled segment duration (i.e. the `@duration` attribute or the **`S`**`@d` in the **`SegmentTimeline`**). Such fluctuations in actual segment duration may be caused by for example ad replacement or specific IDR frame placement. Note that the last segment in a representation may be shorter according to ISO/IEC 23009-1.
- The **`SegmentTimeline`** is only used in order to signal occasional shorter Segments (possibly caused by encoder processes) or to signal gaps in the time line. It is not used for providing Segments with significantly varying duration.

- only unmultiplexed Representations are supported, i.e. each Representation only contains a single media component.
- Addressing schemes are restricted to
    - templates with number-based addressing
    - templates with time-based addressing
    - subsegments with segment index. In this case either the `@indexRange` attribute is expected to be present or the first Segment Index box is contained in the first 1450 byte of the Segment. Note: 1450 bytes are chosen to fit into typical Maximum Transmission Unit sizes for Ethernet of 1500 byte.
- the '`lmsg`' brand for signaling the last segment is applied
- In case multiple Adaptation Sets with `@contentType='video'` are offered, exactly one video Adaptation Set is signaled as the main one unless different Adaptation Sets contain the same content with different quality or different codecs. In the latter case, all Adaptation Sets with the same content shall be signaled as the main content.
- Restrictions on the presence of certain elements and attributes as defined section 2.2.4.

It is expected that a DASH264 client is able to process content offered under these constraints. More details on expected client procedures are provided in section 2.3.

2.2.2. Media Presentation Description constraints

DISCLAIMER: This section serves for the definition of the interoperability point, but is not intended as a normative specification.

NOTE: The term "ignored" in the following description means, that if an MPD is provided and a client that complies with DASH264 removes the element that may be ignored, then the MPD is still complying with the constraints of the MPD as defined in ISO/IEC 23001-9, section 5.

The Media Presentation Description shall conform to the following constraints:

— The rules for the MPD and the segments as defined in ISO/IEC 23001-9, section 7.3, shall apply.

— Representations with value of the `@mimeType` attribute other than `video/mp4`, or `audio/mp4` may be ignored. Additional profile or codec specific parameters may be added to the value of the MIME type attribute. For details refer to specific parameters below.

— The **`Subset`** element may be ignored.

— the **`Period.SegmentList`** element shall not be present.

— if the **`AdaptationSet.SegmentList`** is present in an **`AdaptationSet`** element then this **`AdaptationSet`** element may be ignored.

- if the **Representation.SegmentList** is present in a **Representation** element then this **Representation** element may be ignored.

- Elements using the `@xlink:href` attribute may be ignored from the MPD. The Representations conforming to this profile are those not accessed through an Adaptation Set that uses an `@xlink:href`.

- An **AdaptationSet** containing a **ContentComponent** element may be ignored, i.e. an Adaptation Set with multiplexed media streams may be ignored.

- An **AdaptationSet** element not including the parameters as mandated in section 2.2.4 for an Adaptation Set may be ignored.

- A **Representation** element not including the parameters as mandated in section 2.2.4 may be ignored.

- The **MPD**`@profile` attribute includes the URN "`urn:com:dashif:dash264`"

- If the **MPD**`@type` is equal to "`static`" and the **MPD**`@profile` attribute includes "`urn:mpeg:dash:profile:isoff-on-demand:2011`" then

  - **AdaptationSet** elements with **AdaptationSet**`@subsegmentAlignment` not present, or set to '`false`' may be ignored.

  - **Representation** elements with a `@subsegmentStartsWithSAP` value absent, zero or greater than 2 may be ignored.

  - if the **Representation** element does not contain a **BaseURL** element then this **Representation** element may be ignored.

- If the **MPD**`@type` is equal to "`dynamic`", then

  - the **MPD**`@profile` attribute shall include the signaling for the "`urn:mpeg:dash:profile:isoff-live:2011`"

- if the **MPD**`@profile` **MPD**`@profile` attribute includes "`urn:mpeg:dash:profile:isoff-live:2011`", then

  - **AdaptationSet** elements with **AdaptationSet**`@segmentAlignment` not present, or set to '`false`' may be ignored.

  - **Representation** elements with a `@segmentStartsWithSAP` value absent, zero or greater than 2 may be ignored.

  - The attribute **MPD**`@maxSegmentDuration` shall be present.

— If a Period contains multiple Adaptation Sets with value of the `@contentType="video"` then at least one Adaptation Set shall contain a Role element `<Role scheme="urn:mpeg:dash:role:2011" value="main">` and each Adaptation Set containing such a **`Role`** element shall provide perceptually equivalent media streams.

2.2.3. Segment format constraints

Representations and Segments referred to by the Representations in the profile-specific MPD for this profile, the following constraints shall be met:

— Representations shall comply with the formats defined in ISO/IEC 23009-1, section 7.3.

— In Media Segments, all Segment Index ('`sidx`') and Subsegment Index ('`ssix`') boxes, if present, shall be placed before any Movie Fragment ('`moof`') boxes.

— If the **`MPD`**`@type` is equal to "`static`" and the **`MPD`**`@profile` attribute includes "`urn:mpeg:dash:profile:isoff-on-demand:2011`", then

  — Each Representation shall have one Segment that complies with the Self-Initializing Media Segment as defined in section 6.3.5.2 in ISO/IEC 23009-1.

— If the **`MPD`**`@type` is equal to "`dynamic`" or if it includes **`MPD`**`@profile` attribute includes "`urn:mpeg:dash:profile:isoff-live:2011`", then

  — if the Media Segment is the last Media Segment in the Representation, this Media Segment shall carry the '`lmsg`' compatibility brand.

2.2.4. Presence of Attributes and Elements

Elements and attributes are expected to be presented for certain Adaptation Sets and Representations to enable suitable initial selection and switching.

Specifically the following applies:

- For any Adaptation Sets with value of the `@contentType="video"` the following attributes shall be present
  - `@maxWidth` (or `@width` if all Representations have the same width)
  - `@maxHeight` (or `@height` if all Representations have the same height)
  - `@maxFrameRate` (or `@frameRate` if all Representations have the same frame rate)
  - `@par`
- For any Representation within an Adaptation Set with value of the `@contentType="video"` the following attributes shall be present:

- `@width`, if not present in **`AdaptationSet`** element
- `@height`, if not present in **`AdaptationSet`** element
- `@frameRate`, if not present in **`AdaptationSet`** element
- `@sar`

- For Adaptation Set or for any Representation within an Adaptation Set with value of the `@contentType="video"` the attribute `@scanType` must not be present.
- For any Adaptation Sets with value of the `@contentType="audio"` the following attributes shall be present
  - `@lang`
- For any Representation within an Adaptation Set with value of the `@contentType="audio"` the following elements and attributes shall be present:
  - `@audioSamplingRate`, if not present in **`AdaptationSet`** element
  - **`AudioChannelConfiguration`**, if not present in **`AdaptationSet`** element

2.2.5. Dimension Restrictions

No restrictions are defined on MPD size, or on number of elements.

2.3. Client Implementation Guidelines (Informative)

The interoperability point as defined in section 3.2 can also be understood as permission for DASH clients that only implement the features required by the description to process the Media Presentation (MPD document and Segments). However, the detailed DASH264 DASH-related client operation is not specified. Therefore, it is also unspecified how a DASH client exactly conforms. This document however provides guidelines on what is expected for conformance to this interoperability point.

The DASH-related aspects in DASH264 as well as for the ISO BMFF based On-Demand and Live profiles of ISO/IEC 23009-1 are designed such that a client implementation can rely on relatively easy processes to provide an adaptive streaming service, namely:

- selection of the appropriate Adaptation Sets based on descriptors and other attributes
- initial selection of one Representation within each adaptation set
- download of (Sub)Segments at the appropriate time
- synchronization of different media components from different Adaptation Sets
- seamless switching of representations within one Adaptation Set

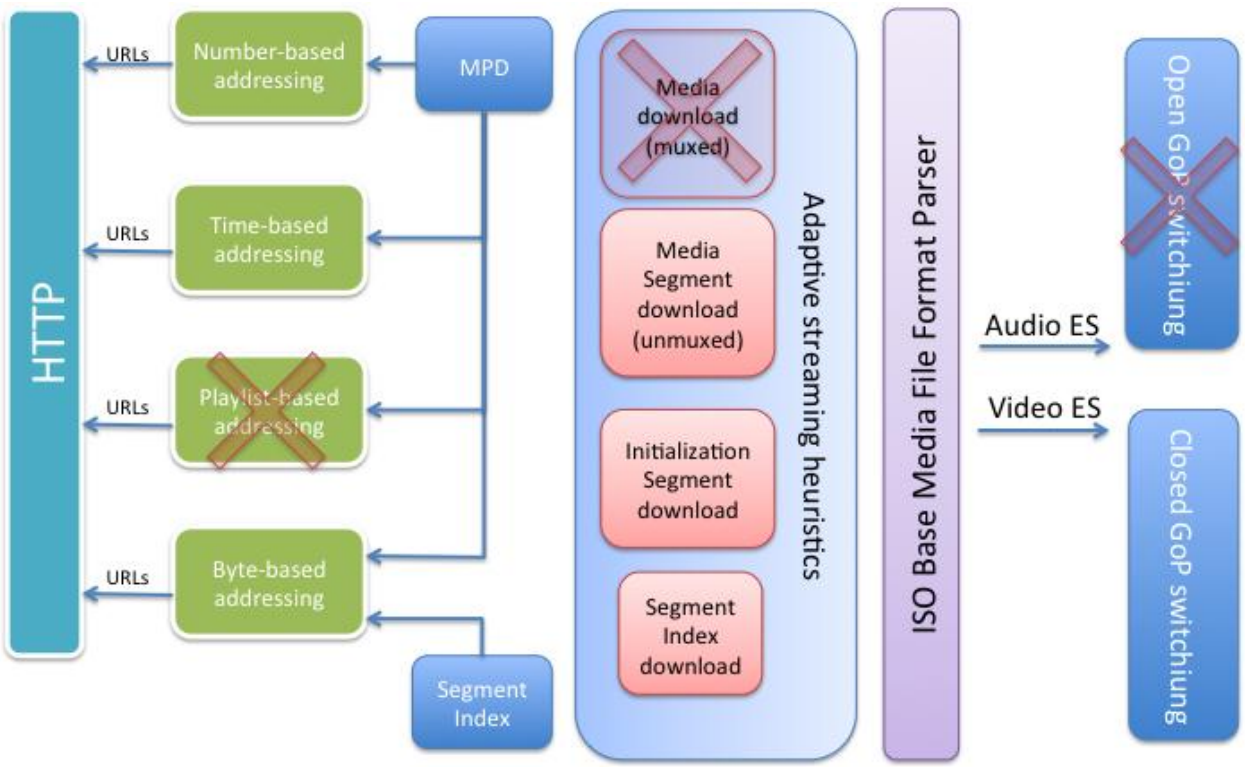


Figure 1 DASH aspects of a DASH264 client compared to a client supporting the union of DASH ISO BMFF live and on-demand profile.

Figure 1 shows the DASH aspects of a DASH264 client compared to a client supporting all features of the DASH ISO BMFF Live and On-Demand profile. The main supported features are:

- support of HTTP GET and partial GET requests to download Segments and Subsegments
- three different addressing schemes, number and time-based templating as well as byte range based requests.
- support of metadata as provided in the MPD and Segment Index
- download of Media Segments, Initialization Segments and Segment Index
- ISO BMFF parsing
- synchronized presentation of media components from different adaptation sets
- switching of video streams at closed GOP boundaries

The features of the client are expected be tested in an interoperability effort to ensure full support. Based on test cases, more detailed client requirements may be added.

2.4. Transport-Related Issues

2.4.1. General

It is expected that servers and clients operating in DASH264 environments support the normative parts of HTTP/1.1 as defined in RFC2616 [18]. Specifically

- Client are expected to support byte range requests, i.e. issue partial GETs to subsegments. Range requests may also be issued by using Annex E of 23009-1 using the syntax of the second example in Annex E.3,

  BaseURL@byteRange="$base$?$query$&range=$first$-$last$"

- HTTP Servers serving segments are expected to support suitable responses to byte range requests (partial GETs). HTTP Servers may also support the syntax using Annex E of 23009-1 using the syntax of the second example in Annex E.3,

  BaseURL@byteRange="$base$?$query$&range=$first$-$last$"

- Clients are expected to follow the reaction to HTTP status and error codes as defined in section A.7 of ISO/IEC 23009-1.

3. Media Coding Technologies

3.1. Introduction

In addition to DASH-specific constraints, DASH264 also adds restrictions on media codecs and other technologies. This section provides an overview on technologies for different media components and how they fit into the DASH-related aspects of DASH264.

3.2. Video

3.2.1. General

The codec that were considered for basic video support up to 1280 x 720p at 30 fps is H.264 (AVC) Progressive High Profile Level 3.1 decoder [6]. This choice is based on the tradeoff of content availability, support of in existing devices and compression efficiency.

Further, it is recognized that certain clients may only be capable to operate with H.264/AVC "Progressive" Main Profile Level 3.0 and therefore content authors may provide and signal a specific subset of DASH264 by providing a specific profile identifier referring to a standard definition presentation as well as to the codec.

For the integration in the context of DASH, the following applies for H.264 (AVC)

- The encapsulation of H.264/MPEG-4 AVC video data in ISO BMFF is defined ISO/IEC 14496-15 [7].
- Clients are expected to support Inband Storage for SPS/PPS based on Draft Amendment for ISO/IEC 14496-15 as issued from MPEG#101 [23].
- SAP types 1 and 2 correspond to IDR-frames in [6].

- The signaling of the different video codec profile and levels for the codecs parameters is according to RFC6381 [8] is documented in Table 1. Note that any of the codecs present in Table 1 conforms to the profile level combination that is supported in DASH264.

Table 1 Codecs parameter according to RFC6381 [8]

Profile	Level	Codec Parameter
H.264 (AVC) Constrained Baseline Profile X=?1????00	1.1	`avc[1..4].42X00B`
	1.2	`avc[1..4].42X00C`
	1.3	`avc[1..4].42X00D`
	3.0	`avc[1..4].42X01E`
H.264 (AVC) "Progressive" Main Profile Y=??????00	1.1	`avc[1..4].4DY00B`
	1.2	`avc[1..4].4DY00C`
	1.3	`avc[1..4].4DY00D`
	3.0	`avc[1..4].4DY01E`
H.264 (AVC) Progressive High Profile Y=??????00	1.1	`avc[1..4].64Y00B`
	1.2	`avc[1..4].64Y00C`
	1.3	`avc[1..4].64Y00D`
	3.0	`avc[1..4].64Y01E`
	3.1	`avc[1..4].64Y01F`

3.2.2. Video Metadata

No specific metadata beyond the one defined in 23009-1 are identified. With respect to the video metadata, the following elements and attributes from ISO/IEC 23009-1 are relevant:

- the picture aspect ratio as described by the `@par` attribute as defined in section 5.3.3 of ISO/IEC 23009-1
- the `@width` and `@height` attributes for signaling the horizontal and vertical visual presentation in section 5.3.7 of ISO/IEC 23009-1
- the `@sar` attribute for signaling sample aspect ratio of the video media component type in section 5.3.7 of ISO/IEC 23009-1

- the `@framerate` attribute for signaling the output frame rate of the video media component type in section 5.3.7 of ISO/IEC 23009-1
- the `@maximumSAPPeriod` attribute for signaling the maximum SAP interval of the video media component type in section 5.3.7 of ISO/IEC 23009-1
- the `@codingDependency` attribute for signaling if coding dependencies across access units exist of the video media component type in section 5.3.7 of ISO/IEC 23009-1

The provisioning of video metadata in the MPD is discussed in section 2.2.4.

3.3. Audio

3.3.1. Introduction

The support of audio in media streaming is essential. All devices considered for DASH264-based are expected to support stereo audio.

3.3.2. DASH-specific aspects for audio codecs

The following functionalities are required to support an audio codec in DASH264:

- encapsulation into fragmented ISO base media file format
- signaling of profiles and level in the `@codec` parameter
- definition of stream access points of type 1 and 2
- audio metadata

3.3.3. MPEG-4 High Efficiency AAC v2 Profile, level 2

3.3.3.1. Overview

The only candidate codec that was considered for basic stereo audio support is:

- MPEG-4 High Efficiency AAC v2 Profile, level 2 [9]

  Note:

  - HE-AACv2 is also standardized as Enhanced aacPlus in 3GPP TS 26.401 [11]
  - The HE-AACv2 Profile includes support of the MPEG-4 AAC Profile [9] and the MPEG-4 HE-AAC Profile
  - Dynamic Range Control metadata is defined in DVB [17], integrated in the audio bitstream.

3.3.3.2. DASH-related Issues

In the context of DASH, the following applies for the High Efficiency AAC v2 Profile

- The content is expected to be prepared according to the MPEG-DASH Implementation Guidelines [3] to make sure each (sub-)segment starts with a SAP of type 1.
- The signaling of MPEG-4 High Efficiency AAC v2 for the codecs parameters is according to RFC6381 [8] and is documented in Table 2. Table 2 also provides information on the ISO BMFF encapsulation.

Table 2 Codecs parameter according to RFC6381 [8] and ISO BMFF encapsulation

Codec	Codec Parameter	ISO BMFF Encapsulation	SAP type
MPEG-4 AAC Profile [9]	`mp4a.40.2`	ISO/IEC 14496-14 [10]	1
MPEG-4 HE-AAC Profile [9]	`mp4a.40.5`	ISO/IEC 14496-14 [10]	1
MPEG-4 HE-AAC v2 Profile [9]	`mp4a.40.29`	ISO/IEC 14496-14 [10]	1

3.3.4. Audio Metadata

3.3.4.1. General

Metadata for audio services is defined in ISO/IEC 23009-1.

3.3.4.2. ISO/IEC 23009-1 audio data

With respect to the audio metadata, the following elements and attributes from ISO/IEC 23009-1 are relevant:

- the `@audioSamplingRate` attribute for signaling the sampling rate of the audio media component type in section 5.3.7 of ISO/IEC 23009-1
- the **`AudioChannelConfiguration`** attribute for signaling audio channel configuration of the audio media component type.in section 5.3.7 of ISO/IEC 23009-1.

3.4. Auxiliary Components

3.4.1. Introduction

Beyond regular audio and video support, TV programs typically also require support for auxiliary components such as subtitles and closed captioning. For example, a Federal Communications Commission (FCC) Advisory Committee has recommended that a standard for the closed-captioning of online video content developed by the Society of Motion Picture and Television Engineers (SMPTE).

3.4.2. Basic Subtitles and Closed Captioning

The chosen technology for basic subtitles and closed captioning is W3C TTML [14] and the SMPTE profile on SMPTE Timed Text [15]. Graphics-based subtitles and closed captioning are also supported by SMPTE Timed Text [15].

Support for other technologies such as

- CEA-708 Digital Television (DTV) Closed Captioning [12]

- 3GPP Timed Text [13]
- Web VTT [16]

are not expected in DASH264, but may be required in certain environments. Conversion of CEA-608 and CEA-708 into SMPTE TT may be done according to SPMTE 2052-10 [19].

3.4.3. DASH-specific aspects of Auxiliary components

In the context of DASH, the following applies for text/subtitling:

- All graphics type samples are SAP type 1.
- The signalling of the different text/subtitling codecs for the codecs parameters is according to RFC6381 [8] is documented in Table 3. Table 3 also provides information on ISO BMFF encapsulation.
- For live services, encapsulation in ISO BMFF is definitely necessary. However, for On-Demand cases, the full file of subtitles may be provided as XML data.

Table 3 Codecs parameter according to RFC6381 [8] and ISO BMFF encapsulation

Codec	MIME type	Codec Parameter @codecs	ISO BMFF Encapsulation
SMPTE Timed Text [15] without encapsulation	`application/xml+ttml`	`not present`	n/a
SMPTE Timed Text [15] with ISO BMFF encapsulation	`application/mp4`	`stpp`	Text of ISO/IEC CD 14496-30 Timed Text and Associated Images in ISO Base Media File Format [20]

4. DRM-Related Aspects

4.1. Introduction

DASH264 does not intend to specify a full end-to-end DRM system. However DASH264 provides a framework for multiple DRMs to protect a content file by adding instructions or *Protection System Specific*, proprietary information in predetermined locations to a file that is encrypted with Common Encryption as defined in IDO/IEC 23001-7 [21] and [22].

The Common Encryption ('cenc') protection scheme specifies encryption parameters that can be applied by a scrambling system and key mapping methods using a common key identification

(KID) to be used by different DRM systems such that the same encrypted version of a file can be combined with different DRM systems that can store proprietary secure information for licensing and key retrieval in the Protection System Specific Header box (‘pssh’). The DRM scheme for each pssh box is identified by the SystemID in that box.

The recommendations in this document reduce the encryption parameters and use of the encryption metadata to specific use cases for VOD and live content with key rotation.

4.2. Base Technologies

The base standards to support common encryption in combination with ISO BMFF are

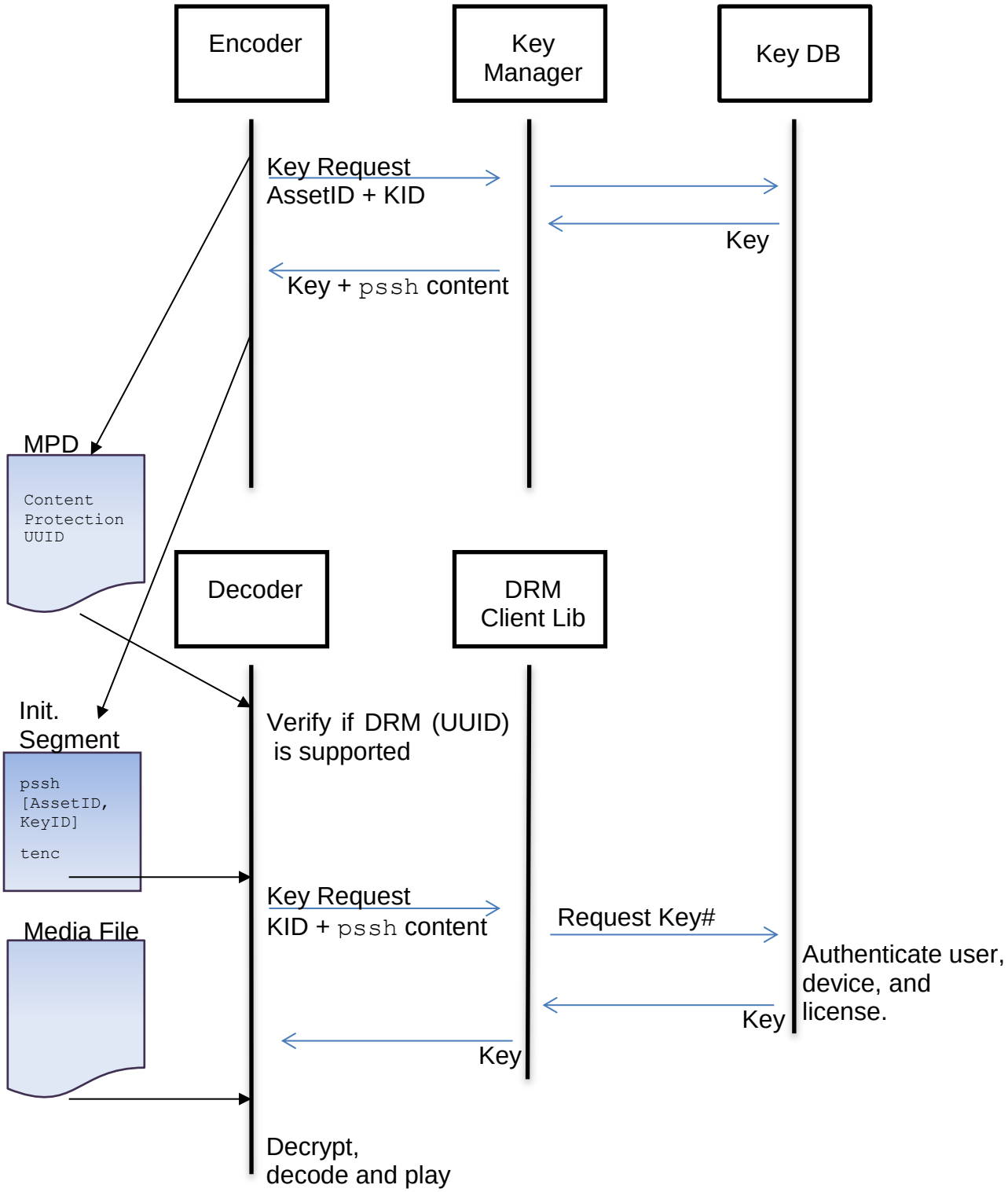
- Common Encryption as defined in ISO/IEC 23001-7 with CTR mode [21].
- Key rotation as defined in ISO/IEC 23001-7 23001-7 DAmd. 1 [22].

The main DRM elements are:

1. The **`ContentProtection`** descriptor (see [1] 5.3.7.2-Table 9, 5.8.5.2 and [1] 5.8.4.1) that contains the URN for signaling of the Common Encryption Scheme as well as the specific DRM
2. '`tenc`' parameters that specify encryption parameters and KID (see [21] 8.2.1). The '`tenc`' information is in the Initialization Segment and may also be in the MPD (see [21] 8.2.1).
3. '`pssh`' parameters that are "Protection System Specific" (see [21] 8.1), The pssh information is in the content file or initialization segment (See [21] 8.1 and 8.2) or in the MPD (see [1] 5.8.4.1). Information in the MPD increases the MPD size but may allow faster parsing, earlier access and addition of DRMs without content modification.

4.3. Workflow Overview

The diagram below shows a simple workflow with pssh information in the initialization segment for informational purpose.



4.4. Integration and Mapping to DASH

4.4.1. MP4 Structure Overview

Table 4 provides pointers to relevant information in the specifications to understand the standard DRM components and where the relevant information is located. The table is for informational purpose only.

Table 4 Boxes relevant for DRM systems

Box	Full Name / *Usage*	Info at
`moof`	movie fragment *One 'moof' box for each fragment in each stream*	ISO BMFF [4], 8.32 + [1]
`moov`	movie box, container for all the metadata *One 'moov' box per elementary stream. 1 x for each video stream, + 1 x for the audio stream*	ISO BMFF [4] , 8.1
`pssh`	protection system specific header box	[21], 8.1.1
`saio`	SampleAuxiliaryInformationOffsetsBox *Contains the offset of the IVs & encryption data.*	[21], 5
`saiz`	SampleAuxiliaryInformationSizesBox *Contains the size of the IVs & encryption data.*	[21], 5
`schi`	scheme information box	[21], 4
`seig`	CencSampleEncryptionInformation GroupEntry *Contains tenc information in sample in segments for key rotation.*	[21], 6
`sinf`	protection scheme information box	[21], 4
`stsd`	sample descriptions (codec types, initialization etc.)	ISO BMFF [4], 8.16
`tenc`	track encryption box contains tenc parameters	[21], 8.2.1

4.4.2. Box Hierarchy

The following shows the box hierarchy and composition:

- In the 'moov' box:
  - one or more 'pssh' boxes
  - in 'trak::mdia::minf::stbl::stsd':
    - the 'sinf' box that contains:
      - the 'frma' box

        - the ‘schm’ box
        - the ‘schi’ box that contains:
            - the ‘tenc’ box
- In the ‘moof’ box:
    - in the ‘traf’ box:
        - the ‘saiz’ box
        - the ‘saio’ box
        - if using key rotation, the ‘sbgp’ box
        - if using key rotation, the ‘sgpd’ box that contains:
            - the ‘seig’ box

4.5. DASH264-Specific DRM Aspects

To enable signaling of a specific DRM scheme in DASH using the Base Technologies as presented in section 4.2 one of the following two options can be applied:

4.5.1. `pssh` and `tenc` Parameters in Movie or Movie Fragment Box

The `pssh` and `tenc` parameters are exclusively provided in the movie or movie fragment box, i.e. in the Initialization Segment (and possibly in the movie fragment box for key rotation) for the live profile or in the movie box for the On-Demand profile:

- **`A ContentProtection`** descriptor in MPD specifies the specific DRM scheme.
    - An example is provided below

      ```
      <ContentProtection schemeIdUri="urn:uuid:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx"
      value="DRMNAME version">
      ```

    - The URN (marked by x) is generated and provided by single DRM provider and uniquely identifies this DRM system.
    - The `@value` attribute describes the DRM system and version in a human readable form.
- ‘`tenc`’ parameters are provided by `tenc` box in the content file that specify encryption parameters and KID as specified in [21], section 8.2.1.
- ‘`pssh`’ parameters are provided by `pssh` box in the content file as specified in [21], section 8.1.

4.5.2. pssh and tenc Parameters in MPD

An extension namespace is defined in order to enable inclusion of certain DRM parameters in the **`ContentProtection`** element. Note that this proposal is submitted to MPEG for consideration to define an extension namespace in the context of ISO/IEC 23001-7 and may therefore be subject to changes.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema targetNamespace="urn:dashif:iop:drm:2012"
    attributeFormDefault="unqualified"
    elementFormDefault="qualified"
    xmlns:xs="http://www.w3.org/2001/XMLSchema"
    xmlns="urn:dashif:iop:drm:2012" >
```

```
    <xs:simpleType name="KeyIdType">
        <xs:restriction base="xs:hexBinary">
            <xs:length value="32"/>
        </xs:restriction>
    </xs:simpleType>

    <xs:simpleType name="KeyIdListType">
        <xs:list itemType="KeyIdType" />
    </xs:simpleType>

    <xs:complexType name="Tenc">
        <xs:sequence>
            <xs:any namespace="##other" processContents="lax" minOccurs="0"
maxOccurs="unbounded"/>
        </xs:sequence>

        <xs:attribute name="isEncrypted" type="xs:unsignedInt" use="required"/>
        <xs:attribute name="ivSize" type="xs:unsignedInt" default="16"/>
        <xs:attribute name="keyId" type="KeyIdType" use="required"/>

        <xs:anyAttribute namespace="##other" processContents="lax"/>
    </xs:complexType>

    <xs:complexType name="Pssh">
        <xs:sequence>
            <xs:any namespace="##other" processContents="lax" minOccurs="0"
maxOccurs="unbounded"/>
        </xs:sequence>

        <!-- base64-encoded content of the `pssh` box -->
        <xs:attribute name="data" type="xs:base64Binary" use="required"/>
        <xs:attribute name="keyIdList" type="KeyIdListType"/>

        <xs:anyAttribute namespace="##other" processContents="lax"/>
    </xs:complexType>

</xs:schema>
```

With this extension namespace `dash264drm`, the following applies.

- A **`ContentProtection`** descriptor in MPD specifies the specific DRM scheme and may add the 'tenc' and 'pssh' information in the extension namespace.
- The **`ContentProtection`** descriptor in MPD specifies the DRM. The URN (marked by x) is generated and provided by single DRM provider and uniquely identifies this DRM system.
- The `@value` element describes the DRM system and version in a human readable form.
- The KID is a 32 Hex character string that uniquely identifies the encryption key.
- The **`PsshData`** is base64 encoded.
- An example if provided below:

```
<ContentProtection schemeIdUri="urn:uuid:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx" val-
ue="DRMNAME version">
   <dash264drm:Tenc isEncrypted="1" ivSize="8" KID="KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK" />
   <dash264drm:PsshData="BASE64 encoded DRM specific pssh data"/>
</ContentProtection>
```

The ‘tenc’ parameter that specifies encryption parameters and KID is also present in the movie box, as specified in [21], section 8.2.1.

The tenc and pssh parameters in the MPD are useful to allow license evaluation, key identification and retrieval before availability of the initialization segment, which may distribute client requests and is relevant to allow an early decision by the client if this key is already available and to otherwise retrieve it before or during download of the initialization segment.

If the pssh parameters are also present in the movie box, the MPD information takes precedence, assuming that the parameters in the MPD are easier to update and contain the most recent version.

The schema is applied to test cases and may be used by individual DRMs but it is not required for individual DRMs to follow this structure.

4.6. Key Rotation

4.6.1. Introduction

Key rotation is mainly used to allow changes in entitlement for continuous live content. It is used as defined in [22] with the following requirements:

- In the initialization segment, the movie box ‘moov’ contains ‘tenc’ parameters and may contain a ‘pssh’ box for each DRM to store root license information for authentication and authorization.
- In addition, each Movie Fragment may contain at most one ‘pssh’ in each ‘moof’ box per SystemID that contains sufficient information to acquire keys for this movie fragment, when combined with:
    - information from ‘pssh’ in ‘moov’
    - KID from ‘seig’ box

  (This will likely result in some redundant pssh boxes but will facilitate processing and trick play, of linear content that is later made available as VOD assets)

- Any KIDs in Movie Fragments override the ‘tenc’ parameter of the ‘default_KID’, as well as the ‘not encrypted’ parameter.

4.6.2. Encryption of Different Representations

Generally, different Representations of one Adaptation Set are protected by the same license, i.e. encrypted with the same key. That means all Representations have the same value of ‘default_KID’ in their ‘tenc’ boxes in their Initialization Segments.

In the case of key rotation, that applies to the root license (one per DRM) and the same value of KID in each leaf license contained in each Media Segment.

In cases where HD and SD content is contained in one asset, different license rights may be required for each quality level. It then is often advisable to create individual Adaptation Sets for each quality level, each with a different **ContentProtection** descriptor in the Adaptation

Set. While there may be some Representations that are equivalent in both Adaptation Set and therefore increase the content size, their size typically relatively small and switching between an HD and SD Adaptation Set is difficult to be applied seamlessly because these quality levels typically vary in DRM output controls, use different decryption licenses and keys and use different decoding parameters for e.g. subsampling, entropy coding, aspect ratios and color spaces.

The test vectors are limited to a single license (per DRM) per Adaptation Set but this does not explicitly exclude the viability of different licenses within one Adaptation Set.

4.7. Signaling

The DRM system is signaled with a URN as described in [1] 5.8.5.2. The list below contains some URNs. The Reference column contains sources for vendor specific information about the DRM system.

DRM System	@schemeIDURI	Reference
Marlin	5E629AF5-38DA-4063-8977-97FFBD9902D4	Marlin Adaptive Streaming Specification - Simple Profile, V1.0
Nagra Media PRM	adb41c24-2dbf-4a6d-958b-4457c0d27b95	
NDS	A68129D3-575B-4F1A-9CBA-3223846CF7C3	
Playready	79f0049a-4098-8642-ab92-e65be0885f95	"MPEG DASH Content Protection using Microsoft PlayReady", accessible here http://www.microsoft.com/playready/documents/
Verimatrix ViewRight Web / DASH	9a27dd82-fde2-4725-8cbc-4234aa06ec09	ViewRight Web / DASH Integration Manual
Adobe	F239E769-EFA3-4850-9C16-A903C6932EFB	
OMA, Widevine, TBD		

4.8. Common Encryption Test DRM

4.8.1. Introduction

In order to test common encryption without the necessity to do tests for a specific DRM, or all supported DRMs, a common encryption test DRM is defined.

Specifically the following aspects are defined:

- To test the encryption with common encryption scheme parameters, the key is provided in a separate file.
- To test the parsing of DRM relevant fields, two different test scenarios are defined to communicate the encryption parameters in the MPD (see section 4.5.2) and in the movie box (see section 4.5.1). The latter case also includes key rotation.

In the interest of testing independently of a specific DRM system, the keys are provided directly in lieu of the DRM information that is otherwise used to obtain the keys.

The use of an external file allows flexible referencing of the same key from different locations, to e.g. use the same key for audio, video or different Representations.

4.8.2. Test of Common Encryption

The key file location is the MPD directory or configurable in the player to avoid OS dependent path references. Its file name is the KID in 32 Hex lower case digits with .txt extension. The content is the decryption key in lower case Hex digits e.g.

`bdff1a347bd8e9f523f5ee6b16273d6e.txt` contains:
`050526bf6d3c386ffe5fc17c93506eca`

The key file name can be stored in the pssh to verify the creation and parsing of pssh information. If the pssh information is not present, the file name can also be derived directly with the knowledge of the KID.

In the test vectors 3 different test values for `@schemeIdUri` are defined to represent multi DRMs:

```
00000000-0000-0000-0000-000000000000
00000000-0000-0000-0000-000000000001
00000000-0000-0000-0000-000000000002
```

The test of common decryption is included in the successful decryption in the above cases.

4.8.3. Test Scenarios

4.8.3.1. Introduction

Different test scenarios are defined which are then mapped to specific test cases in [24]. The first test scenario uses a single key with

1. `pssh` and `tenc` parameters in the movie box

2. pssh and tenc parameters in the MPD.

Another test scenario implements key rotation with tenc and pssh information in the MPD. Finally, a use case for interleaving of unencrypted content is added.

4.8.3.2. Test Scenario 1: pssh and tenc Parameters in Movie Box

The simulation verifies the signaling of the DRM in the MPD, specifically the pssh and tenc information as it must be exercised to access the keys.

The signaling of encryption scheme(s) in MPD:

```
<ContentProtection schemeIdUri="urn:uuid:00000000-0000-0000-0000-000000000000">
<ContentProtection schemeIdUri="urn:uuid:00000000-0000-0000-0000-000000000001">
<ContentProtection schemeIdUri="urn:uuid:00000000-0000-0000-0000-000000000002">
```

The pssh box, if present, contains the base64 encoded filename of the key file. .

4.8.3.3. Test Scenario 2: pssh and tenc Parameters in MPD

The simulation verifies the encoding of the parameters in the MPD as described in 4.5.2. The key file is indicated in the **Pssh**@data attribute as base64 encoded KID in lower case with .txt extension. For example, for a KID of bdff1a347bd8e9f523f5ee6b16273d6, the key will be in the file bdff1a347bd8e9f523f5ee6b16273d6e.txt.

Full Pssh data with required base64 encoding in this case is:

```
<dash264iop:Pssh data=
"YmRmZjFhMzQ3YmQ4ZTlmNTIzZjVlZTZiMTYyNzNkNmUudHh0"/>
```

A separate key file is used for each key when key rotation is used.

4.8.3.4. Test Scenario 3: pssh and tenc Parameters in MPD with Key Rotation

In this case, the pssh information may contain root license information. For the test scenario, the pssh information does not contain relevant key information but is present as a place holder. The static place holder is the base64 encoding of the string: "possible root pssh license info", i.e.:

```
<dash264iop:Pssh license="
cG9zc2libGUgcm9vdCBwc3NoIGxpY2Vuc2UgaW5mbw=="/>
```

A separate key file with different $KeyId$ value is used for each new key.

4.8.3.5. Test Scenario 4: pssh and tenc Parameters in MPD with Key Rotation and unencrypted elements

This extends the previous test scenario with segments that are signaled as unencrypted that are combined with encrypted segments.

5. Interoperability point DASH264

5.1. Introduction

The scope of the DASH264 interoperability point is the basic support high-quality video distribution over the top. Both, live and on-demand services are supported. It is expected that the client supports at least the presentation of

- high-definition video up to 720p (Progressive HP 3.1)
- stereo audio
- basic subtitle support
- basic support for encryption/DRM

The compliance to DASH264 may be signalled by an `@profile` attribute with the value "`http://dashif.org/guidelines/dash264`"

5.2. Supporters

This interop points are supported by the following DASH IF members: Akamai, bitmovin, CastLabs, Cisco, Dolby, Digital Primates, DTS, Elemental Technologies, Envivio, Ericsson, Fraunhofer, Harmonic, Imagine Communications, Intel, InterDigital, Media Excel, Microsoft, Netflix, Path1, Qualcomm, RealNetworks, RGB Networks, Sony, Sorenson Media, Thomson Video Networks, Verimatrix.

5.3. Definition

A client that conforms to consume content based on such a profile support the following features:

- All DASH-related features as defined in section 2 of this document. It is recommended to use the HTTP-URL construction as defined in section A.3.
- H.264/MPEG AVC Progressive High Profile at level 3.1 as defined in section 3.2.
- MPEG-4 HE-AAC v2 level 2 profile audio codec as defined in section 3.3. Dynamic Range Control is not expected to be supported.
- subtitle and closed captioning support using SMPTE-TT as defined in section 3.4.2
    - For On-Demand single file download is sufficient.
    - For live services and/or if key rotation is to be supported, the encapsulation into ISO BMFF is necessary.
- content protection based on common encryption and key rotation as defined in section 4. No specific DRM is defined, but at the DRMs listed in section 4 are enabled. Furthermore, the following holds:
    - each Initialization Segment within one Adaptation Set shall contain an equivalent `pssh` box, i.e. license acquisition for one Representation is sufficient to ensure switching within Adaptation Set.

- in case of inband key delivery, the `pssh` box version 2 (as defined in [22]) shall be equivalent for all Representations within one Adaptation Set, i.e. license acquisition for one Representation is sufficient to ensure switching within Adaptation Set.
- The client is expected to support MPD-based parsing and movie box based parsing of DRM related parameters for common encryption.

6. Interoperability Point DASH264 SD

6.1. Introduction

it is recognized that certain clients may only be capable to operate with H.264/AVC Main Profile. Therefore content authors may provide and signal a specific subset of DASH264 by providing a specific profile identifier referring to a standard definition presentation. This interoperability point is defined as DASH264 SD.

The compliance to DASH264 SD may be signaled by an `@profile` attribute with the value "`http://dashif.org/guidelines/dash264#sd`"

6.2. Supporters

This interop points are supported by the following DASH IF members: Akamai, bitmovin, CastLabs, Cisco, Dolby, Digital Primates, DTS, Elemental Technologies, Envivio, Ericsson, Fraunhofer, Harmonic, Imagine Communications, Intel, InterDigital, Media Excel, Microsoft, Netflix, Path1, Qualcomm, RealNetworks, RGB Networks, Sony, Sorenson Media, Thomson Video Networks, Verimatrix.

6.3. Definition

A client that conforms to consume content based on such a profile support the following features:

- All features as defined in section 6, except:
  - Instead of H.264/MPEG AVC Progressive High Profile at level 3.1, the highest video codec configuration is H.264/MPEG AVC Progressive Main Profile at level 3.0 as defined in section 3.2.

7. References

[1] ISO/IEC 23009-1:2012 Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats.

[2] ISO/IEC 23009-2:2012 Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 2: Conformance and Reference Software (Study of DIS).

[3] ISO/IEC 23009-3:2012 Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 3: Implementation Guidelines (Working Draft).

[4] ISO/IEC 14496-12:2012 Information technology -- Coding of audio-visual objects -- Part 12: ISO base media file format

[5] ISO/IEC 14496-10: Information technology -- Coding of audio-visual objects -- Part 10: Advanced Video Coding

[6] ITU-T Recommendation H.264 (01/2012): "Advanced video coding for generic audio-visual services" | ISO/IEC 14496-10:2010: "Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding".

[7] ISO/IEC 14496-15:2010 Information technology -- Coding of audio-visual objects -- Part 15: Advanced Video Coding (AVC) file format

[8] IETF RFC 6381, The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types, August 2011.

[9] ISO/IEC 14496-3:2009 - Information technology -- Coding of audio-visual objects -- Part 3: Audio

[10] ISO/IEC 14496-12:2012 Information technology -- Coding of audio-visual objects -- Part 14: The MP4 File Format

[11] 3GPP (2005-01-04). "ETSI TS 126 401 V6.1.0 (2004-12) - Universal Mobile Telecommunications System (UMTS); General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description (3GPP TS 26.401 version 6.1.0 Release 6)"

[12] CEA-708-D: Digital Television (DTV) Closed Captioning, August 2008

[13] 3GPP TS 26.245: "Transparent end-to-end Packet switched Streaming Service (PSS); Timed text format"

[14] W3C Timed Text Markup Language (TTML) 1.0, November 2010.

[15] SMPTE ST 2052: "Timed Text"

[16] W3C WebVTT - W3C Web Video Text Tracks, Living Standard — Last Updated 15 February 2012

[17] ETSI TS 101 154 v1.11.1 (2012), “Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream”, Section 6.4.3, Annex C.5

[18] IETF RFC2616, Hypertext Transfer Protocol -- HTTP/1.1, June 1999.

[19] Recommended Practice (Conversion from CEA 608 to SMPTE-TT) RP 2052-10-2012 https://www.smpte.org/sites/default/files/rp2052-10-2012.pdf

[20] MPEG M13040, "Text of ISO/IEC CD 14496-30 Timed Text and Associated Images in ISO Base Media File Format", October 2012.

[21] ISO/IEC 23001-7:2012: "Information technology -- MPEG systems technologies -- Part 7: Common encryption in ISO base media file format files",

[22] ISO/IEC 23001-7:2012/DAM: "AES-CBC-128 and key rotation",

[23] MPEG M12846: "Study of ISO/IEC 14496-15:2010/PDAM 2 Carriage of HEVC", July 2012.

[24] DASH Industry Forum, "Guidelines for Implementation: DASH264 Test Cases and Vectors", under development.

[25] DASH Industry Forum, "Guidelines for Implementation: DASH264 Conformance Software", under development.