Guidelines for Implementation: DASH-IF Interoperability Point for ATSC 3.0

January 31, 2016

DASH Industry Forum

Version 1.0

1 Scope

2 The scope of this document is to provide a DASH interoperability point according to MPEG-DASH
3 [2] that is based on DASH-IF IOPs [1] and provides extensions to address use cases and require-
4 ments of ATSC 3.0 [3].

Disclaimer

This is a document made available by DASH-IF. The technology embodied in this document may involve the use of intellectual property rights, including patents and patent applications owned or controlled by any of the authors or developers of this document. No patent license, either implied or express, is granted to you by this document. DASH-IF has made no search or investigation for such rights and DASH-IF disclaims any duty to do so. The rights and obligations which apply to DASH-IF documents, as such rights and obligations are set forth and defined in the DASH-IF Bylaws and IPR Policy including, but not limited to, patent and other intellectual property license rights and obligations. A copy of the DASH-IF Bylaws and IPR Policy can be obtained at http://dashif.org/.

The material contained herein is provided on an "AS IS" basis and to the maximum extent permitted by applicable law, this material is provided AS IS, and the authors and developers of this material and DASH-IF hereby disclaim all other warranties and conditions, either express, implied or statutory, including, but not limited to, any (if any) implied warranties, duties or conditions of merchantability, of fitness for a particular purpose, of accuracy or completeness of responses, of workmanlike effort, and of lack of negligence.

In addition, this document may include references to documents and/or technologies controlled by third parties. Those third party documents and technologies may be subject to third party rules and licensing terms. No intellectual property license, either implied or express, 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.

If you have comments on the document or identify and bugs or problems, please submit comments as follows:

- at the github repository https://github.com/Dash-Industry-Forum/ATSC/issues
- at the public repository https://gitreports.com/issue/Dash-Industry-Forum/ATSC

Note that technologies included in this document and for which no test and conformance material is provided, are only published as a candidate technology, and may be removed if no test material is provided before releasing a new version of this guidelines document. The status of the test material can be verified on http://testassests.dashif.org.

Contents

List of Figures

List of Tables

Acronyms, abbreviations and definitions

For acronyms, abbreviations and definitions refer to ISO/IEC 23009-1 [2] and DASH-IF IOP [1].

References

[1] DASH-IF Interoperability Points: Guidelines for Implementation, version 4.0.

[2] ISO/IEC 23009-1:2014 Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats. Including:

ISO/IEC 23009-1:2014/Cor 1:2015

ISO/IEC 23009-1:2014/Cor 2:2015

ISO/IEC 23009-1:2014/Cor 3:2016 [Note: Expected to be published by Q1 of 2017. The Final Cor is available in the MPEG output document w16463.]

ISO/IEC 23009-1:2014/Amd 1:2015 High Profile and Availability Time Synchronization

ISO/IEC 23009-1:2014/Amd 2:2015 Spatial relationship description, generalized URL parameters and other extensions

ISO/IEC 23009-1:2014/Amd 3:2016 Authentication, MPD linking, Callback Event, Period Continuity and other Extensions

ISO/IEC 23009-1:2014/Amd 4:2016 Segment Independent SAP Signaling (SISSI), MPD chaining, MPD reset and other extensions [Note: Expected to be published by Q1 of 2017. The FDAM is available in the MPEG output document w16461.]

All the above is expected to be rolled into a third edition of ISO/IEC 23009-1 as:

ISO/IEC 23009-1:2017 Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats. [Note: Expected to be published by mid of 2017. The draft third edition is available in the MPEG output document w16467.]

[3] ATSC Candidate Standard: A/300 “ATSC3.0”,

[4] ATSC Candidate Standard: A/331, Signaling, Delivery, Synchronization, and Error Protection, June 2016, http://atsc.org/candidate-standard/a331-atsc-candidate-standard-signaling-delivery-synchronization-and-error-protection/

[5] ATSC Candidate Standard: A/337, Application Signaling.

[6] ATSC Candidate Standard: A/341, Video, January 2017, http://atsc.org/candidate-standard/a341-atsc-candidate-standard-video/

[7] ATSC Candidate Standard: A/342-1, Audio Common Elements, December 2016, http://atsc.org/candidate-standard/a342-part-1-atsc-candidate-standard-audio-common-elements/

[8] ATSC Candidate Standard: A/342-2, AC-4 System. December 2016, http://atsc.org/candidate-standard/a342-part-2-atsc-candidate-standard-ac-4-system/

[9] ATSC Candidate Standard: A/342-3, MPEG-H System, December 2016, http://atsc.org/candidate-standard/a342-part-3-atsc-candidate-standard-mpeg-h-system/

[10] ATSC Candidate Standard: A/343, Captions and Subtitles, December 2016, http://atsc.org/candidate-standard/a343-atsc-candidate-standard-captions-and-subtitles/

[11] ATSC Candidate Standard: A/344, Application Runtime Environment.

[12] ETSI TS 126.346, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs (Release 13)

[13] ETSI TR 126.946, Digital cellular telecommunication system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Multimedia Broadcast/Multicast Service (MBMS) user service guidelines (Release 13)

[14] ETSI TS 126.247, Universal Mobile Telecommunications System (UMTS); LTE; Transparent end-to-end Packet-switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH) (Release 13)

[15] W3C Candidate Recommendation Media Source Extensions, 3 May 2016 http://www.w3.org/TR/media-source/

[16] ISO/IEC 14496-15:2017 (4th edition) Information technology -- Coding of audio-visual objects -- Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format.

[17] ISO/IEC 23001-8:2013 Information technology -- MPEG systems technologies -- Part 8: Coding-independent code points

[18] ISO/IEC 23008-3:2015 Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 3: 3D audio. Including:

ISO/IEC 23008-3:2015/Amd 1 MPEG-H, 3D Audio Profile and Levels

ISO/IEC 23008-3:2015/Amd 2 MPEG-H 3D Audio File Format Support.

ISO/IEC 23008-3:2015/Amd 3 MPEG-H 3D Audio Phase 2.

ISO/IEC 23008-3:2015/Amd 4 Carriage of system data.

[19] ISO/IEC 14496-30:2014 Information technology -- Coding of audio-visual objects -- Part 30: Timed text and other visual overlays in ISO base media file format. Including:

ISO/IEC 14496-30:2014, Cor 1:2015

ISO/IEC 14496-30:2014, Cor 2:2016 [Note: 14496-30:2014, Cor 2:2016 is expected to be published by end of 2016. The latest document is available in w15933.]

[20] ISO/IEC 23009-5:2016 Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 5: Server and Network Assisted DASH (SAND).

[21] ETSI TS 103 190-2 v1.1.1 2015-09 Digital Audio Compression (AC-4) Standard Part 2: Immersive and personalized audio

[22] IETF RFC 6381 The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types

[23] IETF RFC 5646 (BCP 47) Tags for Identifying Languages

[24] SMPTE: "Digital Object Identifier (DOI) Name and Entertainment ID Registry (EIDR) Identifier Representations," RP 2079-2013, Society of Motion Picture and Television Engineers, 2013.

[25] SMPTE: "Advertising Digital Identifier (Ad-ID®) Representations," RP 2092-1, Society of Motion Picture and Television Engineers, 2015.

[26] ITU: ITU-R Recommendation BT.709-5 (2002), "Parameter values for the HDTV standards for production and international programme exchange," International Telecommunications Union, Geneva

[27] ITU: ITU-R Recommendation BT.2020-1 (2014), "Parameter values for ultra-high definition television systems for production and international programme exchange," International Telecommunications Union, Geneva.

[28] IETF RFC 3986: Uniform Resource Identifier (URI): Generic Syntax, January 2005.

1. Introduction

This document provides a DASH interoperability point that is based on DASH-IF IOPs and provides extensions to address use cases and requirements of ATSC 3.0.
The documents minimizes references to ATSC specifications; it is expected that ATSC will reference this document in order to enable a full ATSC 3.0 service.
The usage of this Interoperability Point is not restricted to ATSC3.0.
This specification defines the identifiers in
Table 1.

Table 1 Identifiers defined in this Document

Identifier	Semantics	Type	Section
http://dashif.org/guidelines/dash-atsc-main	Main DASH Interoperability Point for ATSC	IOP	3.1
http://dashif.org/guidelines/dash-atsc-cgcompatibility	Color gamut capability	Video	5.3.2.7
http://dashif.org/guidelines/dash-atsc-videoposition	View position for stereoscopic content	Video	5.3.2.7
http://dashif.org/guidelines/dash-atsc-scenedisparity	Scene disparity signaling	Video	5.3.2.7
http://dashif.org/guidelines/dash-atsc-temporalsub-layering	Temporal Sub-Layering	Video	5.3.2.7
http://dashif.org/guidelines/dash-atsc-staggercast	Staggercast signaling	Audio	5.4.3.5
http://dashif.org/guidelines/dash-atsc-program	Program Signaling	Function	5.7.2
http://dashif.org/guidelines/dash-atsc-closedcaption	Closed Caption	Subtitle	5.5.3.2
http://dashif.org/guidelines/dash-atsc-RRTrating:1	Rating	Rating	5.7.3

2. Background and Assumptions (Informative)

2.1. Introduction

To set the context, this section provides background and assumptions, primarily shared by ATSC with DASH-IF. For a detailed overview on ATSC3.0, please refer to ATSC A/300 [3]. The ATSC A/300 standard [3] is the initial entry point to the ATSC 3.0 system. It provides both an overview

of the system and a guiding structure to the pertinent ATSC component standards that are to be followed depending on how the system is configured, as indicated by the system signaling.

2.2. ATSC3.0 Protocol Stack

According to the ATSC A/331 [4] the protocol stack as presented in Figure 1 expresses the major components of the ATSC delivery system. In particular, DASH formats play a central role as the encapsulation and delivery format in the context of ATSC 3.0 for broadcast, broadband and hybrid delivery.

In case of broadcast delivery, the interface between the underlying delivery system and the DASH Player is at least conceptually based on an HTTP proxy that is included in the end point of the delivery system. In addition to the interfaces to the transport system, the DASH Player as shown in Figure 1 also provides the functionality to play media properly and to interface with native or downloadable applications, typically in a browser-centric runtime environment.

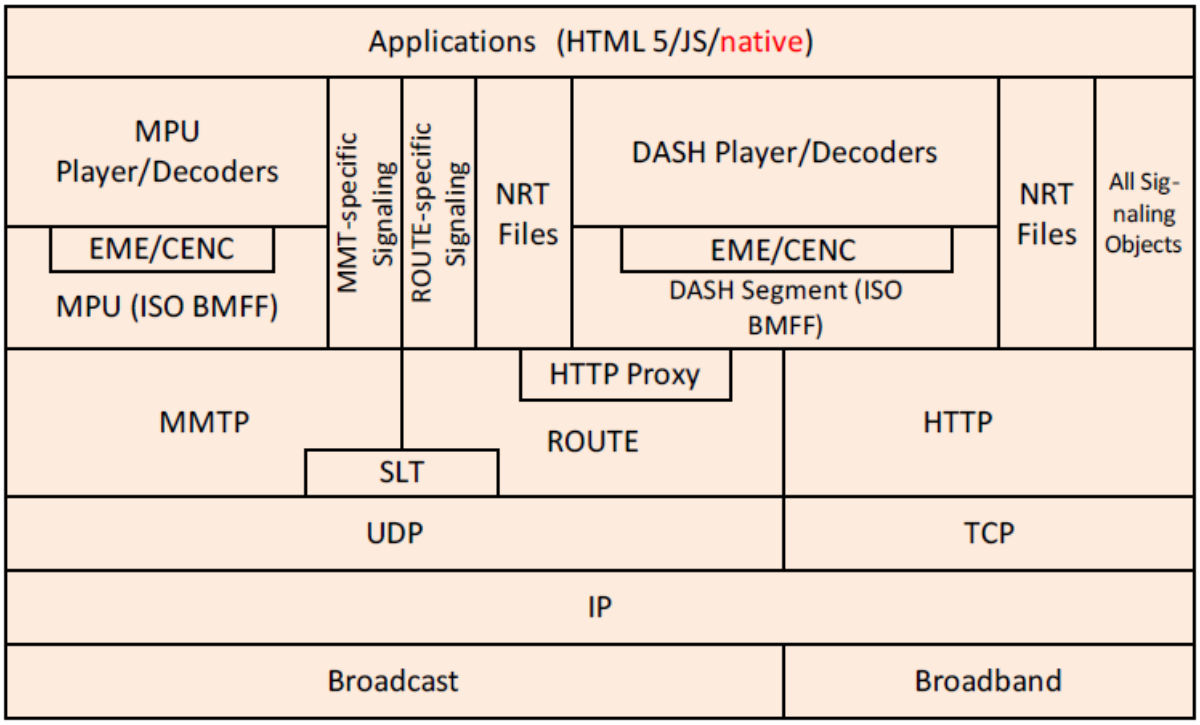

Figure 1 ATSC Protocol Stack[1]

2.3. Client Reference Architecture

2.3.1. Introduction

ATSC 3.0 as well as MPEG-DASH are defining emission standards. In addition, DASH formats terminate (at least primarily) in the DASH Player and it is assumed that the DASH Player controls

[1] Reproduced with permission.

the streaming session by issuing HTTP requests scheduled at appropriate times to download Segments from an HTTP server (possibly a distributed architecture using a CDN). In order to map DASH formats on top of ATSC delivery and create the appropriate service and user experience, it is considered useful to specify a reference architecture of an ATSC 3.0 receiver (or "client") device, referred to in this document as the Client Reference Model (CRM), in order to define and/or verify the proper emission specifications.

A decomposition of the functions and interfaces in the client enables the definition of proper emission formats in order to verify that the distribution formats result in expected functionality to fulfill the ATSC 3.0 system requirements.

By no means would such a reference client imply a normative implementation, as it would only provide an example implementation to verify the adequacy of the delivery specification.

The CRM is expected to decompose the ATSC 3.0 receiver device into the relevant network interfaces, device internal functions, interfaces to the application and interfaces to the media playout pipeline.

2.3.2. Overview: Functions and Interfaces

Figure 2 provides an overview of relevant functions and interfaces (IF) in the decomposition of the signaling and processing routines of the DASH Player. The DASH Player acts as a component in the ATSC 3.0 receiver client device.

The functions in the client are informative and do not imply a specific implementation. For example, Cache and HTTP Proxy may be implemented differently, but serve as a conceptual model and logical endpoint for service delivery.

The following functions are identified in the client reference model:

- ATSC 3.0 Physical Layer connections (possibly comprising multiple RF channels) and broadband connections provide the connectivity, via broadcast and broadband networks, to broadcasters/content providers to receive service signaling and data.
- ROUTE/UDP/IP and HTTP/TCP/IP that provide an object oriented transport protocol running on top of IP in order to receive DASH resources as well as other objects and files that are relevant for the ATSC 3.0 service, or an application associated with the service.
- HTTP proxy: A local (i.e. device-resident) HTTP proxy that may be used to abstract the underlying physical and transport layer to a client application, in particular the DASH player, but may also be a broadcaster application. Application specific data, transient service objects and NRT content may be provided through the HTTP proxy.
- Low-Level Signaling: Signaling delivered over UDP/IP that provides channel scanning and basic service description and entry point information to enable service selection and acquisition by the Basic TV Function.
- Service Signaling: A function that picks up service-related signaling for the selected service which provides information to the receiver and DASH Player on IP-level service acquisition, as well as static and dynamic configuration of the service.
- Cache: Temporary storage and handling of the MPD, Initialization Segment and Media Segments whose reception are facilitated by service signaling.
- Basic TV Function: A platform that provides at the minimum rendering capabilities for A/V services as well as simple means for interactivity, typically via a remote control.

- Application/Interactive Presentation: A native or downloaded application that makes use of broadcast or broadband delivered data in order to provide a potentially richer and interactive presentation to the end user.
- ATSC Events: A function that operates as a sink for ATSC events as defined in [5].
- DASH Player: A function that consumes MPDs and Segments, and communicates with other components in the CRM to which it interfaces to personalize the media experience based on platform capabilities, user preferences and user interaction. The DASH player also provides information to a DRM engine and media player in order to decrypt and decode media.
- Persistent Objects: Persistent storage of typically non-real time objects. This function may provide the media resources for a DASH Media Presentations through the HTTP Proxy.

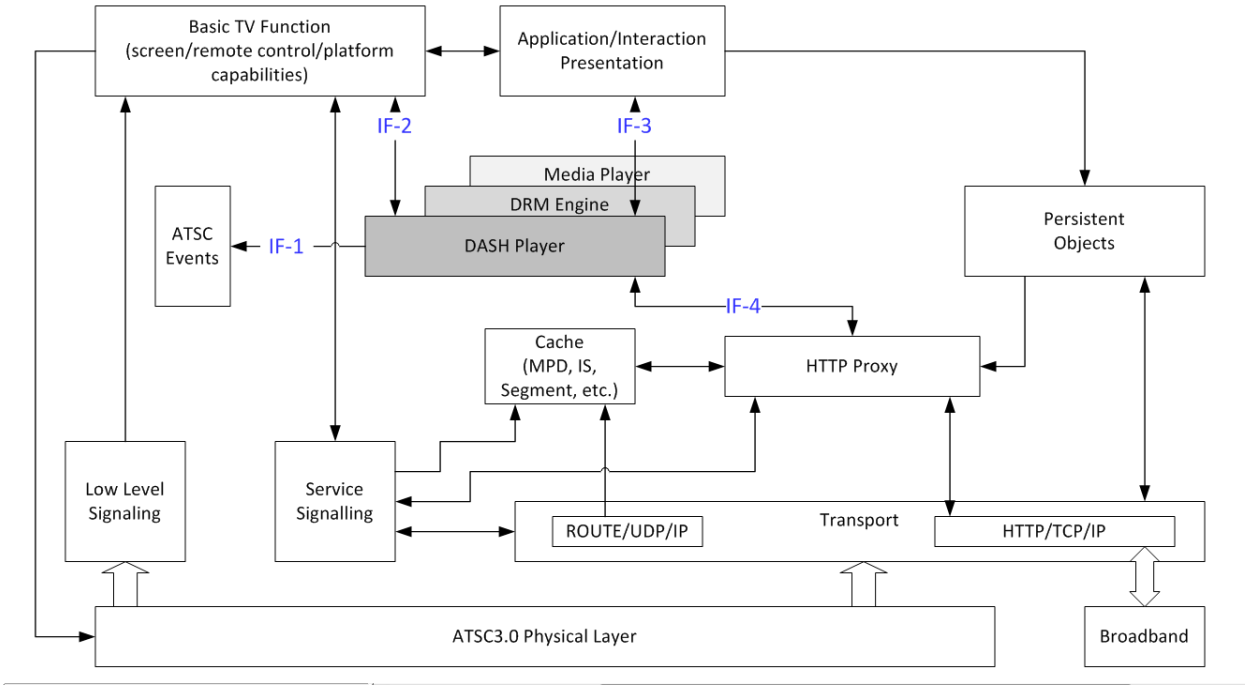

Figure 2 Client Reference Model

2.3.3. Relevant Interfaces

The logical functions in the CRM exchange information via the defined interfaces as described in this section to support the processing and playout of media data. Although the documented interfaces are conceptual, some of them may exchange information in a more formalized manner using well-defined APIs.

- IF-1: The ATSC specific events received by the DASH Player are dispatched to the ATSC event application through this interface.
- IF-2: If the service metadata includes an MPD, the MPD is handed to the DASH player and the DASH player is activated. In addition, the DASH player may exchange capability information with the Basic TV Function, for example on rendering and DRM capabilities, as well as on user preferences and settings.

- IF-3: For an app-enhanced linear service, or an app-based service, the app and the DASH player may exchange over IF-3 information regarding capabilities, personalization, app-specific events, targeting, etc.
- IF-4: A regular HTTP interface between the DASH player and the proxy. The interface follows HTTP methods, and may support extensions pertaining to error robustness and network information.

Other interfaces are conceptual and out of scope of this specification. More details on interfaces and the messages exchanged on the interface are provided in the remainder.

2.3.4. Typical Bootstrap and Service Signaling

A typical bootstrapping sequence is presented in the following:

1. The Basic TV Platform requests a pre-configured Service List Table (SLT) in Low Level Signaling (LLS). SLT is delivered to the Basic TV Function, which then provides a user interface for ATSC 3.0 Service selection. User chooses a particular ATSC 3.0 Service for rendering.
2. By using the SLT, the user selects the service to consume, and the Basic TV Function uses the Service Layer Signaling (SLS) entry point information carried in the SLT for the selected service to provide access information to the ROUTE/UDP/IP stack to retrieve the SLS. SLS is delivered to the Basic TV Function, but certain elements are added as transient service objects to be available directly for the application, i.e. the DASH player.
3. By using the SLS, the Basic TV Function provides access information to the ROUTE/UDP/IP stack for downloading the DASH-formatted media components of the selected Service, which can be in turn sent to the HTTP proxy/cache to be temporarily stored. Assuming that the selected Service is a linear service that includes a targeted ad insertion broadcaster application, the receiver platform provides access information to the ROUTE/UDP/IP stack for downloading the broadcaster application. Ad files can be downloaded as NRT content and passed to and cached in persistent storage (as Persistent Objects).
4. The broadcaster application may be automatically launched upon reception, or launched under the control of the receiver platform.
5. Via IF-2, the DASH Player exchanges service capability information with the Basic TV Function, for example on rendering and DRM capabilities, as well as on user preferences and settings.
6. Upon the selection of a service, the Basic TV Function activates the DASH Player via IF-2, causing the DASH Player to request Media Segments from the HTTP proxy, via IF-4, at or after the Media Segment availability start times. Media Segments delivered via broadcast will be sent by the ROUTE/UDP/IP stack to the Cache, for subsequent forwarding to the HTTP Proxy. Media Segments delivered via broadband will be directly sent by the HTTP/TCP/IP stack to the HTTP Proxy.
7. DASH Player sends Segment request/receives Segments to/from the HTTP proxy/cache over IF-4. In an alternative implementation, the ROUTE receiver, i.e. the ROUTE/UDP/IP stack in the Transport block, may stream MDE(s) to the DASH Player as described in Annex A of A/331 [4]

8. Upon reception of Media Segments or MDE, the composite function comprising the DASH Player, DRM Engine and Media Player decodes the received media content, and the decoded media is returned to the Basic TV Function for screen display.
9. During Service reception there may be the occurrence of an ad avail. The DASH Player will pass a remote Period element with XLink for resolution by the broadcaster application. The broadcaster application may provide the DASH Player a replacement Period which points to, for example, an Ad in the Persistent Objects store or other location.
10. After the ad avail, playout of the main program resumes based on repetition of steps 6-8.

2.4. Client and Service Types

2.4.1. Introduction

The service that includes a DASH Media Presentation may support different types and receiver models, with different levels of involvement of the application or browser in the DASH media consumption. Different service types are discussed in this sub-clause.

2.4.2. Client Type 1: Stand-alone

Client Type 1 is considered as a standalone without any interface to an app or browser, i.e. IF-3 in Figure 2 is not present and the client obtains all information primarily from IF-2.

2.4.3. Client Type 2: App-based Enhancement

In client type 2 as shown in Figure 3, the DASH player still acts as a stand-alone player, but through IF-3 in Figure 2 the DASH and media player may be partially controlled or at least some amount of interaction applies. The initial presentation is still launched through the DASH Player.

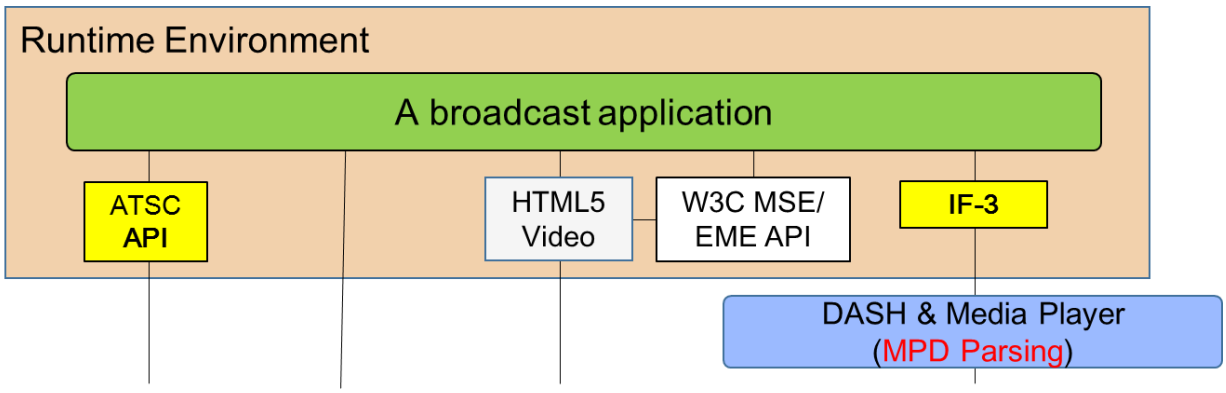

Figure 3 App-based Enhancement

2.4.4. Client Type 3: DASH Player in Video Element

In this case the app launches a DASH player through a <video> element that is provided with a URL to an MPD.

2.4.5. Client Type 4: App-based

In client type 4 as shown in Figure 4, the initial MPD is consumed in the app and all control is done in the application. In order to enable such a deployment, the content needs to be offered conforming to Media Source Extensions (MSE) [15].

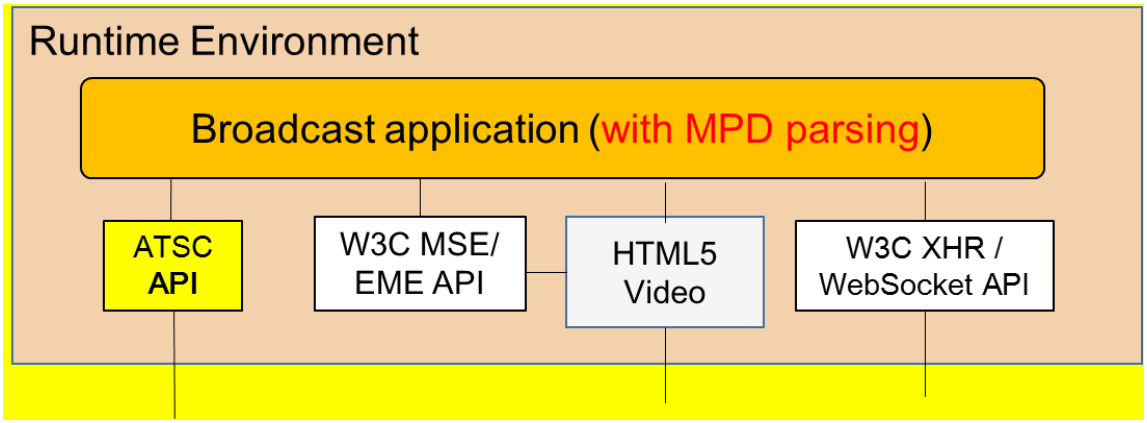

Figure 4 App-based Client

2.5. IF-1: Application Interface

The Application Interface enables communication of the DASH client with the application. An implementation of this interface is expected to be provided by a JSON RPC API defined in A/344 [11].

As an example, non ATSC-specific event streams may be supported. In addition, personalization information may be exchanged over this interface.

2.6. IF-2: Capabilities and User Settings/Interface

2.6.1. General

The MPD contains signaling on the property of the delivered media streams. These properties are also provided such that a Receiver can use this information to check if the stream matches platform capabilities. If the platform capabilities are not sufficient, the media stream is not considered for decoding and presentation. If the service contains more than one media stream of the same media type, then additional information needs to be provided to differentiate the media streams with the same media type and the DASH player typically needs to select one. In addition annotation can be provided that is used by the system to map against user preferences and presets (e.g. language or accessibility settings). Also signaling may be provided that supports the player in selecting a media stream when joining as well in the absence of other information. IF-2 is used by the DASH player to gather information from the platform on supported capabilities and user preferences and settings. Such a selection process needs to be done at join time and in case new content is spliced, i.e. DASH when a new Period is signaled.

The conceptual interface IF-2 expects that the DASH client can use the information in the MPD to query the platform for supported capabilities. The implementation of this interface is out of scope for this document. However, if for example an HTML-5 based user agent would be used to support track selection, parts of the interface may be implemented accordingly.

2.6.2. Video Specific Capabilities in context of ATSC 3.0

In the case of ATSC 3.0, typical differentiation of receiver capabilities for the video decoding and rendering pipeline may use one or multiple of the following properties:

- Codec capabilities

    - Single Layer Codec, Profile and Level
    - Scalable Codec
    - Temporal Sub-Layering
- Display/rendering capabilities
    - spatial and temporal resolution
    - Scan Format, interlace or progressive
    - HDR capabilities
    - 3D capabilities
    - Color space capabilities

2.6.3. Audio Specific Capabilities in context of ATSC 3.0

In the case of ATSC 3.0, typical differentiation of receiver capabilities for the audio decoding and rendering pipeline may use one or multiple of the following properties:

- Codec capabilities:
    - Codec, Profile and Level
- Rendering capabilities/environment
- User preferences and settings (accessibility, language, role)
- User interaction and Personalization

2.6.4. Subtitle/Caption Specific Capabilities in context of ATSC 3.0

In the case of ATSC 3.0, typical differentiation of receiver capabilities for the subtitle and caption decoding and rendering pipeline may use one or multiple of the following properties:

- User preferences and settings (e.g., accessibility, language)
- Rendering capabilities (e.g., text profile, image profile)

2.6.5. Transport Specific Capabilities in context of ATSC

In the case of ATSC 3.0, typical differentiation of receiver capabilities for the transport are:

- Broadcast-reception only
- Broadcast & Broadband
- Broadband only (no ATSC use case for broadband only, but media may primarily arrive through broadband, signaling always through broadcast)
- Maximum available broadband bandwidth
- Reception conditions, for example due to different robustness on the transport certain resources may or may not be available depending on the reception conditions.

2.6.6. DRM Specific Capabilities in context of ATSC

In the case of ATSC 3.0, typical differentiation of receiver capabilities for the DRM are:

- available DRM systems

2.7. IF-3: Application Interfaces

2.7.1. Introduction

The runtime environment is a relevant concept in ATSC 3.0. This section looks into possible interfaces between the DASH Player and an application.

2.7.2. Parental Control

Content advisories, in ATSC, are metadata associated with Programs, and not with individual components in contrast to the Rating descriptor in DASH. Each Program in the broadcast schedule may be associated with a content advisory rating. In the ATSC system, content advisory ratings shall be signaled as described in Section 5.7.3. The DASH client may communicate with the platform to understand the content rating associated with platform and apply this on Program level.

2.7.3. Personalization and Ad Insertion

Personalized content may be distributed. If done, then the content is differentiated through a RESTful architecture, i.e. personalization is achieved using personalized HTTP URLs and other HTTP methods that enable targeted content. The logic on how to personalize requests is outside the DASH Player, but the DASH Player communicates through IF-3 with the application for personalization information.

2.7.4. Media Control

The application may control the media playout, potentially in a dynamic fashion. Examples for media control may include scaling and positioning the video, muting audio, trick modes such as pause and resume or other aspects. The DASH Player may get information on how the media is controlled and may use the information to optimize its processing, e.g. selection of Adaptation Sets and Representations. For example, if audio is muted, download of audio may be dispensed. If the video is consumed in a thumbnail version with no audio then only a low resolution video may be downloaded. Details on how such information is exchanged between the DASH Player and application are out of scope, but a DASH MPD is expected to provide information in order to react to such dynamic information from the application.

2.8. IF-4: Transport Interfaces

2.8.1. Introduction

Figure 5 provides an overview on the transport interfaces. A DASH Player can communicate with a local proxy and cache that has intelligence to receive content from broadcast through ROUTE and broadband through HTTP/TCP/IP.

> Note: This description is only one possible implementation in order to show the use of a DASH Player in the ATSC 3.0 receiver model.

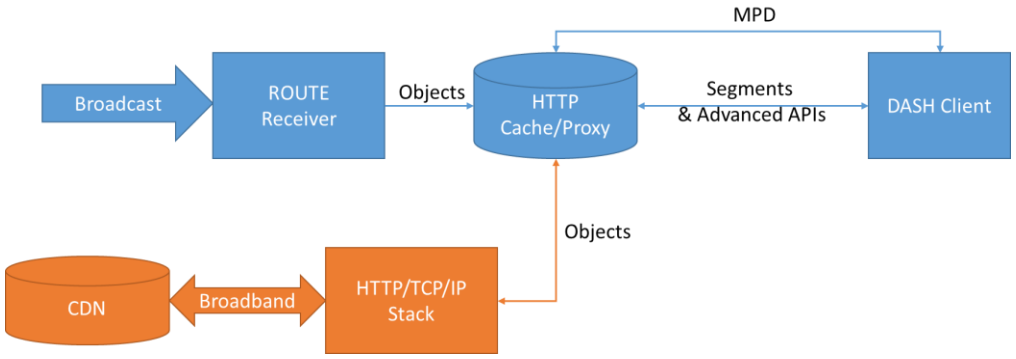

Figure 5 Receiver model for Broadcast and Broadband Reception

2.8.2. MPD and Segment-based – Regular File Delivery

In the regular file or Media Segment delivery mode, the DASH Player makes a content request for an entire Segment as the delivery object from the HTTP Proxy over IF-4. It uses the MPD to construct the Segment URLs for the requests. The corresponding media stream(s) is(are) delivered via broadcast and/or broadband, and forwarded by the Transport block as shown in Figure 2 to the HTTP Proxy, as an example implementation method depicted in the diagram. In this implementation method, the HTTP Proxy acts as a local HTTP server to return the requested Segments to the DASH Player over IF-4.

2.8.3. MDE-based for reduced startup delay

Figure 6 provides a possible implementation of the receiver in case the timing of the playout is controlled by the broadcast network and not the availability times in the MPD. DASH formats are distributed over broadband or broadcast. The MPD may be used as entry point or for example only when the broadband components are added. However, the timing of the broadcast/ROUTE distribution is determined by the broadcast transport and all relevant information may be provided through broadcast metadata. Startup may happen prior to reception of MPD and/or full segment. The MPD/DASH Player is still necessary for any hybrid aspects and to describe service details.

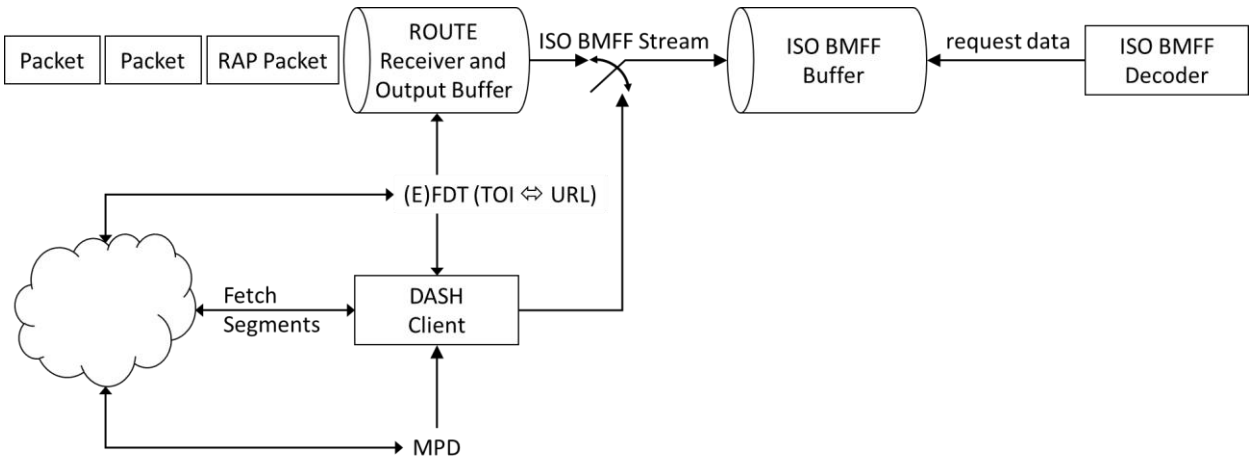

Figure 6 Model for MDE-based receptions

MDE-based delivery may be implemented by a regular DASH client using HTTP requests prior to full reception of segments and the proxy/cache provides the data with HTTP Chunked Transfer. By this, a progressive media consumption is enabled. If HTTP Chunked Transfer is not supported, then other means may be used to enable early consumption of Media Segments, e.g. using the WebSocket API to directly feed the MSE source buffer. For more details refer to Annex A.

2.8.4. Specific Methods for ATSC 3.0 beyond regular HTTP

2.8.4.1. Status Codes

Guidelines for handling request responses according to case 4 from above are provided in MPEG-DASH, Annex A.7 [2].

2.8.4.2. Robustness

Typical problems affecting robustness are documented in DASH-IF IOP, Annex B. The HTTP proxy and DASH Player may communicate using the tools defined in DASH-IF IOP, clause 4.8.

2.8.4.3. Network redirection

Suitable methods for communication between the HTTP Proxy/Cache and the DASH Player are provided in TR 26.946 [13], clause 7.2.4.

> Note: It is expected that updates will be provided once MPEG SAND [18] is fully defined and 3GPP has aligned as well.

2.8.4.4. Partial File Handling

Suitable handling of partial files is defined in clause 7.9.2 of TS 26.346 [12].

Guidelines for handling request responses with 200 OK with the Content-Type set to application/3gpp-partial and 416 Requested Range Not Satisfiable are provided in Annex A.9 of TS 26.247 [14].

2.9. Scope of this Specification

The scope of this specification is the definition of the DASH formats that conform to MPEG-DASH, but provide additional restrictions and extensions to fulfill the use cases and requirements documented by ATSC. The extensions include signaling for specific functionalities from ATSC including broadcast and hybrid services, specific media formats and codecs, subtitles, events, metadata, security and ad insertion functions.

In order to enable a complete end-to-end system, it is expected that receivers/DASH Players implement certain functions and processes, but this is outside of the scope of the specification. Nevertheless, expected receiver behavior is added in order to explain the assumptions when documenting the signaling requirements. It is expected that this information may be used to define more detailed receiver requirements in the context of receiver specification for the ATSC 3.0 emission standard.

3. DASH MPD and Segment Constraints

3.1. Interoperability Points Signaling

The conformance to *DASH-IF ATSC Main* may be signaled by a `@profiles` attribute with the value http://dashif.org/guidelines/dash-atsc-main.

A Media Presentation (MPD and Segment formats) conform to the IOP by offering content following the requirement and recommendations in the following sections:

- Clause 3.2: The requirements and recommendations from MPEG-DASH
- Clause 3.3: Requirements and recommendations related to DASH-IF IOPs
- Clause 4: Restrictions and Extensions on the Distribution Formats
- Clause 5: The Media Profiles and metadata as well as their mapping to DASH
- Clause 6 Ad Insertion requirements and recommendations
- Clause 7: DRM and Security Related requirements and recommendations

It is expected that with the combination of the ATSC specification and a usage of the DASH client following the CRM in clause 2.3, the ATSC use cases and requirements can be fulfilled.

3.2. Relation to MPEG-DASH

A *DASH-IF ATSC Main* Media Presentation shall conform to the ISO BMFF Broadcast TV Profile as defined in ISO/IEC 23009-1:2017, clause 8.11 [2].

> Note: As this profile is not yet fully defined and published, the key principles are included in clause 4 and Annex B.

3.3. Relation to DASH-IF IOP

The Media Presentation is built on the features from DASH-IF IOP v3.3 [1]. However, the DASH+ATSC Media Presentation is not expected to be conforming to DASH-IF IOP taking into account that certain features and requirements for ATSC need to be enabled, that had not been included in the requirements for DASH-IF IOP.

A *DASH-IF ATSC* Media Presentation shall follow the requirements and recommendations from DASH-IF IOP of the following features and sections:

- The DASH formats in clause 3.2.1, including segment formats and only non-multiplexed Representations.
- The DASH timing model in clause 3.2.7
- The Recommendations on Bandwidth and Minimum Buffer Time in clause 3.2.8
- The Trick mode support in clause 3.2.9
- The Adaptation Set Constraints in clause 3.2.10
- The Segment-based Media Time Information in clause 3.2.11
- The Content Offering within a Period in clause 3.2.12
- The Switching across Adaptation Sets in clause 3.8
- The Simple Live Operation as defined in clause 4.9.2

Note that the main live operation as defined in clause 4.9.3 may be used as well.

4. Distribution Formats

4.1. Introduction

4.1.1. Broadcast Distribution

In Broadcast Distribution, the broadcast channel is the only communication channel available to the DASH Player. Therefore, the DASH Player can only receive MPD and media segments through the broadcast channel. No return channel capability is available, but the client reference model as defined in clause 2 permits interfacing between the broadcast distribution and the DASH client.

Key aspects for linear TV services, in particular, broadcast services, are end-to-end latency and rapid channel change times. The distribution format should be easily integrated into ATSC delivery protocols, in particular ROUTE/UDP/IP for broadcast according to the CRM as introduced in clause 2. The distribution format is expected to support synchronization of supplemental content, such as accessibility components, supplementary languages, etc. with primary A/V content; both the supplemental content and the primary content may be delivered via Broadcast.

4.1.2. Hybrid Distribution

In addition to the broadcast channel, a broadband channel may also available to the DASH Player. While AV services may be pure broadcast, or hybrid broadcast/broadband, service signaling always starts on the broadcast channel. According to the ATSC A/331 specification [4], only a single MPD is used to signal content offerings on broadcast and broadband, the DASH Player may receive one MPD and Media Segments through the broadcast channel and/or the broadband channel.

The broadband channel may for example be used to:

- send additional service information,
- send Media Segments as part of a pure broadband service (on-demand content, catch-up content, time-shift services, etc.),
- send Media Segments as part of additional service components to a broadcast service,
- send additional Media Segments as an enhancement to broadcast Media Segments (using scalable coding),
- send Media Segments as a temporary replacement to broadcast Media Segments (for error recovery purposes (retransmission) or fast channel change purposes).

The formats should be easily integrated into ATSC delivery protocols, in particular, HTTP/TCP and ROUTE/UDP/IP. The same service may be offered through broadcast and broadband (with different quality), seamless transition from broadcast to broadband and back to broadcast is expected. The system is expected to support synchronization of supplemental content with primary content; both the supplemental content and the primary Content may be delivered via broadcast or broadband. The system is expected to provide the means for coping with variable content delivery latency.

4.2. Distribution Format

4.2.1. DASH Profile

This distribution format provides a restricted subset of MPEG-DASH primarily for distributing broadcast TV over broadcast and broadband services, including service offerings for combined broadcast and broadband services.

A *DASH-IF ATSC Main* Media Presentation shall conform to the ISO BMFF Broadcast TV Profile as defined in ISO/IEC 23009-1:2017, clause 8.11 [2].

Note: As the profile is not yet published, the profile is documented in Annex B.

In addition, the following constraints apply to the profile:

- The **`MPD`**`@type` shall be set to `dynamic`
- All Representations in one Adaptation Set shall have equal timescale values in all `@timescale` attributes and 'tkhd' `timescale` fields in Initialization Segments.
- The random access type as defined in ISO/IEC 23009-1:2017 clause 5.3.3.5, shall either be "`closed`" or "`open`".

Note that "publishing a new MPD" for broadcast distribution is equivalent of sending an MPD such that the new MPD is available on the local cache in the device.

The MPD Base URL's for broadcast resources are identified by using a relative reference per RFC3986 [28], where the first character in the URI is a "/".

4.2.2. ROUTE protocol constraints

In order for the ROUTE receiver to properly identify DASH segments, the following options are possible:

- If `$Number$` based addressing is used, the TOI field of a given ROUTE packet should be set to the `$Number$` value of the DASH segment it contains and the File mode with EFDT templating should be used.
- If `$Time$` is used and no segment sequences, the length of `$Time$` value should not exceed 32bits. The value of '0' and '1' shall not be used.
- If segment sequences are used with hierarchical addressing, then the entity mode ROUTE is expected to be applied in order to properly signal the Segments.

4.2.3. Segments, Random Access and Switching Points

Constraints on segmentation, random access and switching points follows the ISO BMFF Broadcast TV Profile as defined in ISO/IEC 23009-1:2017, clause 8.11. More details on requirements for random access and switching points may be provided for each codec.

Note: More details will be added in the next revision of this document.

4.3. Basic Use Cases and Recommendations

4.3.1. Broadcast Distribution

For broadcast distribution, the following recommendations apply:

- Only a single Representation per Adaptation Set should be present for broadcast distribution.
- the `@minimumUpdatePeriod` shall be set to 0. This permits to update the MPD with every new Segment.
- The open ended Segment Timeline with `@r=-1` should be used.

4.3.2. Hybrid Distribution

For hybrid distribution, the following recommendations apply:

- Representations that are expected to be seamlessly switchable (regardless whether they are distributed through broadcast or broadband) shall either be in the same Adaptation Set or the Representations shall be linked by using the Adaptation Set Switching signaling.
- If there are differences on the availability times between broadcast and broadband Representations, the `@availabilityTimeOffset` should be used.

4.4. Client Recommendations

The DASH client should check MPDs regularly for changes on the local cache, but should avoid parsing MPDs that have not changed.

Broadcast only clients are expected to support the simple live operation as defined in 4.9.2 of DASH-IF IOPs.

Hybrid clients are recommended to support the main live operation as defined in 4.9.3 of DASH-IF IOPs.

Access gain for applications to events carried in the event stream (which may be either signaled in the MPD, or carried in the Segments of a Representation) is relevant. Broadcaster-supplied applications can register for events of interest using a JSON RPC API defined in A/344 [11]. The application identifies events of interest by specifying their `schemeIdUri` and (optionally) `value`. For each event associated with a registered event, the receiver's DASH Player is expected to pass the associated data to the application over interface IF-3. Both "static" Events, whose timing is known well in advance, as well as "dynamic" Events, the timing of which can only be determined in real time as the program unfolds, are e expected to be supported by the receiver's DASH Player if the Runtime Application Environment specified in A/344 [11] is supported.

If an event is signaled as an inband event, the client is expected to parse each random access segment at least up to the first `'moof'` box. The DASH client parses the segment information and extract the earliest presentation time of the media segment.

If an `'emsg'` is detected that is set to the value defined in the MPD, the DASH client is expected to parse the segment information and extract the following values:

- `emsg.ptd` the presentation time delta as documented in the `emsg`.
- `emsg.ed` the event duration as documented in the `emsg`
- `emsg.message_data`

After parsing, the Segment is typically forwarded to the media pipeline if it also used for rendering, but it may either be dumped (if the Representation is only used to access the DASH event, such as muted audio).

5. Mapping of ATSC Media to DASH

5.1. Introduction

The media profile focusses on mapping ATSC media, in particular video, audio and subtitles/CC to MPEG DASH. This includes issues for MPD signaling as well as Representation/File Format constraints.

In addition, this section provides also the signaling of other media related information, such as the content model or media-time related events.

5.2. Content Model and Metadata

5.2.1. Introduction

The ATSC program or content played out by the user may be tracked for usage reporting. Content Identifiers are utilized for this tracking. Content identifier labeling is expected to be supported for broadcast and broadband content (including advertisements). As a minimum Content identifier values of type EIDR and Ad-ID, along with broadcaster-defined IDs (e.g., house numbers), are expected to be supported.

- "EIDR" indicates a content identification per the EIDR registry (http://eidr.org).
- "Ad-ID" indicates a content identifier per the Ad-ID registry (http://ad-id.org).

Extensibility should be provided for adding other content identifier types in future. Support for multiple content identifier values for the same content should be considered. Static (e.g. list of future scheduled content related content identifier values) and dynamic (e.g. unscheduled dynamically inserted advertisement related content identifier values) content identifiers signaling associated with content should be considered.

Programs and associated Ratings are defined in clause 5.7.

5.2.2. MPD Signaling

In order to annotate content, the DASH+ATSC Media Presentation author may use the Asset Identifier descriptor on Period level as defined in ISO/IEC 23009-1, clause 5.8.4.10 .

Two schemes are defined here:

- the value of `@schemeIdUri` set to "`urn:eidr`" and then the value of `@value` attribute descriptor shall be a valid canonical EIDR entry as defined in [24].
- the value of `@schemeIdUri` set to the "Designator" for either the "full" or "compact" encoding as defined in SMPTE 2092-1 [25] and then the value of `@value` attribute descriptor shall be a valid Ad-ID entry as defined in [25].

Other schemes may be used, including user private schemes, by using appropriately unique values of `@schemeIdUri`.

5.3. Video

5.3.1. Background and Use Cases (Informative)

ATSC A/300 mandates that when HEVC video compression is used with ATSC 3.0, the ATSC A/341 standard [6] is followed. When HEVC is used, support is provided up to 3840 x 2160p at 120 fps is HEVC Main 10 or Scalable Main 10 Profile, Level 5.2, Main Tier. The HEVC coded video includes legacy SD video and Interlaced HD video for support of existing content as well as Progressive Video. The progressive video allows the full range of advanced features including high dynamic range (HDR), wide color gamut (WCG), 3D, and temporal layering.

AFD and Bar Data are considered such that the active area of the picture does not necessarily need to fill the entire coded area.

When Spatial Scalable Coding is employed, both HD and UHD videos are encoded where HD video is coded in a base layer and UHD video is coded in enhancement and base layers.

When Temporal sub-Layering is applied, one video stream shall include two temporal video substeams. The video stream can be decoded with different frame rates according to the decoder's capabilities.

5.3.2. Service Offering Requirements and Recommendations

5.3.2.1. Constraints on HEVC Adaptation Sets and Bitstreams

The HEVC Adaptation Sets and bitstreams shall conform to DASH-IF IOP, Section 6.2 [1].

Switching type shall either be set to media switching or to bitstream switching.

5.3.2.2. MPD Signaling

5.3.2.2.1. IOP Constraints

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

All constraints of DASH-IF IOP, section 3.2.4 [1] on any Video Adaptation Set are applied except the constraint on `@scanType`.

For this IOP:

- For any Adaptation Set or for any Representation within an Adaptation Set with `@contentType="video"` the attribute `@scanType` need not be present, or if present, shall be set to `"progressive"` or `"interlaced"`.

  Note: default `@scanType` value is `"progressive"`.

5.3.2.3. DASH-specific aspects for H.265/HEVC video

For any Adaptation Set or for any Representation within an Adaptation Set with `@contentType="video"`, all constraints of DASH-IF IOP, section 6.2.3 [1] are applied.

The ATSC 3.0 video profiles are defined in A/341 [6].

Additionally, DASH-IF IOP, table 16 [1] is extended with the following entries from Table 2.

Table 2 Codecs parameter according to ISO/IEC 14496-15 [10]

Profile	Level	Tier	Constraints	The `@codecs` parameter	The `lhevcptl` parameter

HEVC Main 10	3.1	Main	progressive_source, non_packed, frame_only	hev1.2.4.L93.B0	n\a
			interlaced source, non_packed	hev1.2.4.L93.60	n\a
	4.1	Main	progressive source, non_packed, frame_only	hev1.2.4.L123.B0	n\a
			interlaced_source, non_packed	hev1.2.4.L123.60	n\a
	5.0	Main	progressive_source, non_packed, frame_only	hev1.2.4.L150.B0	n\a
	5.1	Main	progressive_source, non_packed, frame_only	hev1.2.4.L153.B0	n\a
	5.2	Main	progressive_source, non_packed, frame_only	hev1.2.4.L156.B0	n\a
HEVC Scal-able Main 10	5.1	Main	progressive_source, non_packed, frame_only, non temporal layering	lhe1	0, 1.0.7.80.L153.BD.88
	5.2	Main	progressive_source, non_packed, frame_only, non_temporal_layering	lhe1	0, 1.0.7.80.L156.BD.88
			progressive source, non_packed, frame_only, temporal_layering	lhe1	0, 1.0.7.80.L153.BD.88, 2.1.7.80.L156.BD.88

Note: The 'hev1', 'hev2' and 'lhe1' sample entry ensures convenient random access and switching without the need of searching and fetching parameter sets from earlier samples. The other sample entries ('hvc1', 'hvc2', and 'lhv1') do not guarantee such convenient random access and switching. Part 15 mandates parameter sets presence for 'hev1', 'hev2', and 'lhe1' types to randomly access at any IRAP picture and rely only on parameter sets from either the sample description (i.e., the IS) or from that sample onwards.

Note: The 'hev2' sample entry is only used for a representation exclusively containing the higher sub-layer of the base layer.

Note: When an HEVC Main 10 Profile or HEVC Scalable Main 10 Profile bitstream has a constant picture rate equal to 120, 120/1.001, or 100 pictures per second, temporal sub-layering with two temporal sub-layers may be applied.

When temporal sub-layering with two temporal sub-layers is applied, the bitstream shall contain exactly two sub-layers, with TemporalId equal to 0 and 1, respectively. Each sub-layer can be the output layer set.

Additionally, all relevant constraints to HEVC codec of DASH-IF-IOP, section 6.2.5 [1] are applied.

Note: The Codecs parameter signals the profile and level of the entire bitstream. For instance, when Temporal Layering is used, the Codecs parameter indicates the profile and level of the entire bitstream.

5.3.2.4. ATSC Legacy SD

This section defines the DASH related constraints required for Legacy SD in DASH-IF IOP Section 6.2.1 [1].

Any Adaptation Set signaling Legacy SD shall contain only one Representation.

5.3.2.5. ATSC Interlaced HD video

This section defines the DASH related constraints required for Interlaced HD in DASH-IF IOP Section 6.2.2 [1].

Any Adaptation Set signaling Interlaced HD shall contain only one Representation.

5.3.2.6. ATSC progressive video

This section defines the DASH related constraints required for ATSC progressive video in DASH-IF IOP Section 6.2.3 [1].

If the content is encoded using HEVC Scalable Main 10 Profile, the base layer Representation of each enhancement layer Representation shall be identified using `@dependencyId`.

5.3.2.7. Adaptation Sets constraints

All constraints of DASH-IF IOP, section 6.2.5 [1] on any Adaptation Set are applied except the following constraints:

- Only the active video area shall be encoded so that devices can frame the height and width of the encoded video to the size and shape of their currently selected display area without extraneous padding in the decoded video, such as "letterbox bars" or "pillar-box bars".

The additional following constraints are applied to the Adaptation Sets:

- Color space of all representations within one Adaptation Set shall be the same. The color space shall be one of the followings: Rec. 709 [26] or Rec. 2020[27].
- If the color space of the content of an Adaptation Set is Rec. 2020, then an Essential or Supplemental Descriptor shall be present at that Adaptation Set element, with `@schemeIdUri` of `urn:mpeg:mpegB:cicp:colourprimaries` URI and `@value` of `"9"` [17].
- If the color space of the content of an Adaptation Set is compatible with Rec. 709, then an Essential or Supplemental Descriptor shall be present at the Adaptation Set element, with `@schemeIdUri` of `http://dashif.org/guidelines/dash-atsc-cgcompatibility` URI and `@value` of `"1"`.
- For stereoscopic video content, the view position shall be signaled using an Essential or a Supplemental Descriptor at the Adaptation Set element of the "left" video, with `@schemeIdUri` of `http://dashif.org/guidelines/dash-atsc-videoposition` URI and `@value` equal to the value of `@id` of the "right" Adaptation Set. The scene disparity range shall be signaled using a Supplemental Descriptor at the Adaptation Set element of either left or right video, with `@schemeIdUri` of `http://dashif.org/guidelines/dash-atsc-scenedisparity` URI and `@value` of comma separated of two parameters. The first parameter represents the minimum disparity, and shall be an integer between `-1024` and `1023`. The second parameter represents the maximum disparity and shall be an integer between `0` and `2047`.
- When Temporal Sub-Layering with constraints defined in section 6.3.4 of A/341 [7] is used in a Representation, then a Supplemental Descriptor shall be present at that Representation, with `@schemeIdUri` of `http://dashif.org/guidelines/dash-atsc-temporalsub-layering` URI. The value of the `@value` attribute shall consist of two parts separated by a delimiter ',' with second part optionally present:
  - The first part will be an 8-bit unsigned integer with value equal to the Level for temporal sub-layer zero of the Representation. This will be equal to the value of syntax element `sub_layer_level_idc[ 0 ]` of the Representation.
  - The second part if present will be coded as a string using process defined for Codecs MIME type specification in Annex E section E.3 of ISO/ IEC 14496-15 for single

layer HEVC with syntax element sub_layer_profile_space[0], sub_layer_tier_flag[0], sub_layer_profile_idc[0], sub_layer_profile_compatibility_flag[0][j] for j in the range of 0 to 31, inclusive, and each of 6 bytes of the constraint flags starting from sub_layer_progressive_source_flag[0] respectively substituted for element general_profile_space, general_tier_flag, general_profile_idc, general_profile_compatibility_flag[j] for j in the range of 0 to 31, inclusive, and each of 6 bytes of the constraint flags starting from general_progressive_source_flag. If the second part is absent then all other profile_tier_level() parameters for the temporal sub-layer zero besides the sub_layer_level_idc[0] parameter which is signalled in the first part shall be inferred to be same as the value of those parameters signalled in Codecs parameter for the Representation. If all Representations of an Adaptation Set contain Temporal Sub-Layering with constraints defined in section 6.3.4 of A/341 [6] and all Representations have the same profile, tier, level and flags information for temporal sub-layer zero, then the above descriptor may be used at the Adaptation Set element.

- When temporal sub-layering with two temporal sub-layers is used in two Representations, each temporal sub-layer is carried in a Representation respectively, @codecs values shall be present at the Representation to signal the profile/level/tier described in Sample Description of the track contained in each Representation (see 5.3.2.8 for details). When the first containing VCL NAL units with TemporalId greater than 0 only and the second containing VCL NAL units with TemporalId equal to 0 only, the first Representation shall be associated to the second Representation by using @dependencyId attribute in the MPD.

5.3.2.8. Segment Format and Encapsulation Requirements for H.265/HEVC video

The encapsulation of HEVC single-layer bitstream in a file shall be according to Clause 8 and Clause 9 of ISO/IEC 14496-15 [16] with the following constraints applied:

- Each track shall carry only one layer or a subset of one layer, and the HEVC bitstream shall be carried in at most two tracks.
- Each track shall be encapsulated in one DASH Representation.
- Extractors and aggregators shall not be included in any track.
- If a track carries a subset containing VCL NAL unit with TemporalId greater than 0 only, the sample entry type shall be 'hev2'. Otherwise, the sample entry type shall be 'hev1' as defined in [16].
- When temporal sub-layering is applied and all samples (for both TemporalId=0 and 1) are carried in a single track, the track shall contain sample group description box containing sample group entry type 'tscl' and corresponding sample-to-group box which assigns a sample group for each sample within that track.
- When temporal sub-layering is used and sub-layers are carried in separate tracks, the following requirements apply.

— The 'hev1' sample entry of the track (carrying VCL NAL unit with TemporalId equal to 0 only) shall indicate the level of the substream, i.e. the value of sub_layer_level_idc[0] in the SPS if the value of sub_layer_level_present_flag[0] equal to 1.

— The 'hev2' sample entry of the track (carrying VCL NAL unit with TemporalId greater than 0 only) shall indicate the level of entire stream (including both temporal sub-layers).

— In the track with sample entry type of 'hev2', the decoding time of each sample containing VCL NAL units shall be equal as in the case when both temporal sub-layers are stored in a single track.

- The encapsulation rules for HEVC as defined in DASH-IF IOP v3.3 [1] apply.

The encapsulation of an SHVC bitstream in a file shall be according to Clause 9 of ISO/IEC 14496-15 [16] with the following constraints applied:

- Each track shall carry only one layer or a subset of one layer, and the SHVC bitstream shall be carried in at most two tracks.

  Note: With this constraint in place, a sample entry cannot contain both the HEVC and L-HEVC configurations, and the two layers of an SHVC bitstream have to be carried in two tracks, one for each layer.

- Each track shall be encapsulated in one DASH Representation.
- Extractors and aggregators shall not be included in any track.
- The base track (i.e., the track containing the base layer) shall use the sample entry type 'hev1' as defined in [16].
- For each track that carries a layer for which the VCL NAL unit has nuh_layer_id greater than 0 or a subset of such a layer, the sample entry type shall be 'lhe1'.
- The external base layer sample group shall not be included in any track.
- When temporal sub-layering is applied and all samples (for both TemporalId=0 and 1) of a layer are carried in a single track, the track shall contain sample group description box containing sample group entry type 'tscl' and corresponding sample-to-group box which assigns a sample group for each sample within that track.

No additional constraint on Segments other than imposed by the DASH profile is specified.

Note: Switching from the base layer (BL) to the enhancement layer (EL) can only occur at a segment or subsegment of the EL Representation starting with a sample containing an IRAP picture at the EL. Switching from the EL to the BL can occur at the start of any segment or subsegment of the BL Representation, regardless of whether that segment or subsegment starts with a sample containing an IRAP picture at the EL.

5.3.2.9. Multiple Frame Rate Temporal Filtering Information Signaling

The Multiple Frame Rate Temporal Filtering allows efficient delivery of video with independent effective shutter intervals. When the Multiple Frame Rate Temporal Filtering described in A/341 Section 6.3.4.1 and Annex D [6] is used the constraints described in section A/341 6.3.4 regarding High Frame Rate Temporal Sub-Layering also apply. When Multiple Frame Rate Temporal Filtering as described in A/341 Section 6.3.4.1 [6] is used in a Representation, then a Essential Descriptor shall be present at that Representation, with @schemeIdUri set equal to http://dashif.org/guidelines/dash-atsc-multiframerate-temporal-

filtering. The value of the @value attribute shall indicate a parameter which indicates a 2 bit field expressed as a 2 character string representing 2 binary bits which shall indicate the values of temporal filtering parameters temporal_filter_w1 and temporal_filter_w2. The temporal_filter_w1 and temporal_filter_w2 parameters are used in the recovery process as described in the Annex D, section D.1.1 in A/341 [6]. In this case temporal_filter_w1 parameter shall indicate the weight of the temporally preceding temporal sub-layer 1 picture that contributes to the current temporal sub-layer 0 picture and temporal_filter_w2 parameter shall indicate the weight of the high frame rate picture (not provided in the raw stream) in the current temporal position that contributes to the current temporal sub-layer 0 picture. The values of **temporal_filter_w1** and temporal_filter_w2 are inferred based on the signaled @value as shown in Table 3. The value of temporal_filter_w1 plus temporal_filter_w2 shall equal 1.

Note that this technology is expected to require specific APIs from the DASH client to the media decoder implementation and video display pipeline and may therefore not be usable to systems where the such APIs are not available.

Table 3 Values of Multiple Frame Rate Temporal Filtering parameters

@value parameter	temporal_filter_w2	temporal_filter_w1
'00'	4/5	1/5
'01'	2/3	1/3
'10'	4/7	3/7
'11'	1/2	1/2

A receiver capable of High Frame Rate playback but not capable of recovery process as described in A/341 Section 6.3.4.1 [6] should select a Representation (if available) without a Essential Descriptor with @schemeIdUri set equal to http://dashif.org/guidelines/dash-atsc-multiframerate-temporal-filtering.

If all Representations of an Adaptation Set use Multi-Frame Rate Temporal Filtering with same temporal filter weights then the above descriptor may be used at the Adaptation Set element.

Regarding switching, the following is supported:

a) A receiver capable of only Standard Frame Rate playback as defined in A/341 Section 6.3.4.1 [6] may switch between a Standard Frame Rate Representation and a Representation utilizing High Frame Rate Temporal Sub-Layering as defined in A/341 Section 6.3.4 [6] with Multiple Frame Rate Temporal Filtering as defined in A/341 Section 6.3.4.1 [6]. If multiple Representations with Multiple Frame Rate Temporal Filtering with different weighting factors are available, the one with the highest available value for temporal_filter_w1 minimizes temporal aliasing (strobing) and may be preferred.
b) A receiver capable of only Standard Frame Rate playback as defined in A/341 Section 6.3.4.1 [6] may switch between a Standard Frame Rate Representation and a Representation utilizing High Frame Rate Temporal Sub-Layering as defined in A/341 Section 6.3.4 [6] but not utilizing Multiple Frame Rate Temporal Filtering.

c) A receiver capable of High Frame Rate playback as defined in A/341 Section 6.3.4.1 [6] may switch between any Representations utilizing High Frame Rate Temporal Sub-Layering as defined in A/341 Section 6.3.4 [6] with or without Multiple Frame Rate Temporal Filtering as defined in A/341 Section 6.3.4.1 [6].

d) A receiver capable of High Frame Rate playback as defined in A/341 Section 6.3.4.1 [6] may switch between a Standard Frame Rate Representation and a Representation utilizing High Frame Rate Temporal Sub-Layering as defined in A/341 Section 6.3.4 [6] with or without Multiple Frame Rate Temporal Filtering as defined in A/341 Section 6.3.4.1 [6].

5.4. Audio

5.4.1. Background and Basic Use Cases (Informative)

The use cases provided by ATSC to DASH-IF are expected to be supported by the client reference model. The client can select audio components based on e.g.:

- the audio language preference setting of the receiver
- the accessibility settings of the receiver
- the codec capabilities of the receiver
- the output preference of the receiver (e.g. stereo vs. multichannel output)
- new parameters or methods for signaling of next generation audio defined by DASH-IF in order to signal immersive and personalized content
- the network connectivity, if applicable (access to hybrid content via Ethernet or WiFi). This may for example include that certain languages are only available if the receiver provides broadband connectivity.
- the usage of impairment techniques which rely on additional audio streams

Audio that consists of multiple components that contribute to an experience is expected to be supported. Personalization based on multi-component audio is expected to be supported. Multi-component audio is able to coexist with single-component audio. Signaling is defined to be agnostic to the underlying format of the audio stream. Signaling of availability of audio tracks to provide for user selection is expected. Signaling of Next Generation Audio (NGA) on systems level as well as evaluation of related content signaling by the decoder is expected to be enabled in order to address requirements of different client architectures. NGA codecs introduce the concept of Preselections which cannot be described sufficiently by today's collection of DASH parameters. The audio and DASH signaling experts extended parameters as required to enable NGA Preselections. ATSC 3.0 also expects the availability of signaling for accessibility services. The signaling is also expected to enable utilization of NGA codec features i.e. coding of audio elements. The signaling should enable delivery of audio elements for impairment services via broadcast as well as via broadband.

5.4.2. Assumptions and Definitions

5.4.2.1. Introduction

The Preselection element as defined in ISO/IEC23009-1:2014/Amd.4:2016 [2] is used for audio signaling in the context of ATSC 3.0. It is specifically adapted to address the next generation audio concepts. For common concepts of ATSC 3.0 audio, see A/342-1 [7].

Note: As ISO/IEC23009-1:2014/Amd.4:2016 [2] is not yet published, the relevant concepts are provided in Annex C.

5.4.2.2. Bundle

In the context of ATSC 3.0 audio, a Bundle is a closed set of audio elements that can contribute to the playout of one NGA audio decoder. Examples for audio elements are an English dialogue, German dialogue, or Music & Effects. The referred audio elements can be carried in one or separate tracks or in one or separate Adaptation Sets. Typically, not all audio elements of one bundle are played out at the same time. The set of audio elements of one audio Bundle can provide multiple personalization options like different languages, flexible gain or spatial location of audio elements, typically exposed through a user interface. A Bundle typically contains several Preselections.

5.4.2.3. Preselection

A Preselection is a personalization option to produce a complete audio experience. It is associated with one or more audio elements from one Bundle plus additional parameters like gain or spatial location. A Preselection can be considered the NGA equivalent of alternative audio tracks containing complete mixes using traditional audio codecs. Multiple Preselection instances can refer to the same set of elements in a Bundle for example with different settings for gain and spatial location. Only audio elements of the same Bundle can contribute to the decoding and rendering of a Preselection.

The Preselection concept is common to both NGA codecs referenced by ATSC 3.0 and is mapped to the systems layer to provide a basic selection mechanism, e.g. for user preferred languages, accessibility, etc.

5.4.2.4. Compound Stream

One audio elementary stream comprising more than one audio element.

5.4.2.5. Full-Compound Stream

One audio elementary stream comprising all audio elements belonging to one audio Bundle.

5.4.3. Codec-Independent Mapping to DASH

5.4.3.1. Additional Attributes

The following attributes are available in Adaptation Sets and Media Content Components for ATSC 3.0 as given in ISO/IEC23009-1:2014/Amd.4:2016 [2].

Table 4 MPD Adaptation Set

Element or Attribute Name	Use	Description
Adaptation Set		
@tag	O	Tag to be used to identify this adaptation set towards an external scope (e.g. decoder)

Table 5 MPD Media Content Component

Element or Attribute Name	Use	Description
Media Content Component		
@tag	O	Tag to be used to identify this content component towards an external scope (e.g. decoder)

5.4.3.2. Preselection

A Preselection is a personalization option to produce a complete audio experience as defined above in clause 5.4.2.3. By using a Preselection as a starting point, the client can avoid unnecessary consumption of network resources by selecting only those Adaptation Sets necessary for a specific Preselection and only downloading one Representation of each selected Adaptation Set.

Two different methods are defined how to signal Preselections in the MPD: The Preselection Descriptor and the Preselection Element.

The Preselection descriptor is defined in 5.3.11.2 of ISO/IEC23009-1:2014/Amd.4:2016 [2]. It enables simple setups and backward compatibility, but may not be suitable for advanced use cases. The usage of the Preselection descriptor in ATSC 3.0 is provided in clause 5.4.3.4.

The Preselection element is defined in 5.3.11.3 and 5.3.11.4 of ISO/IEC23009-1:2014/Amd.4:2016 [2]. More refinements for NGA in ATSC 3.0 on Preselection Elements are defined in clause 5.4.3.3.

5.4.3.3. Preselection Element

The concept of Preselection Elements is orthogonal to the concept of Adaptation Sets. The Preselection element is provided on Period level.

A subset and constrained usage of the Preselection element is shown in Table 6. Note that the "Use" column may be different from what is defined in ISO/IEC23009-1:2014/Amd.4:2016 [2] and provide specific constraints when using the Preselection element for NGA in ATSC 3.0. Other elements and attributes than provided in Table 6 should only be present if needed for backward-compatibility and may be ignored by the DASH client. The detailed semantics can be found in ISO/IEC23009-1:2014/Amd.4:2016 [2].

Table 6 MPD Preselection for NGA in ATSC

Element or Attribute Name	Use	Description
Preselection		
@id	OD De-fault=1	See ISO/IEC23009-1:2014/Amd.4:2016 [2].
@audioSamplingRate	O	See ISO/IEC23009-1:2014/Amd.4:2016 [2].
@codecs	M	See ISO/IEC23009-1:2014/Amd.4:2016 [2].
@selectionPriority	OD	See ISO/IEC23009-1:2014/Amd.4:2016 [2].

Element or Attribute Name	Use	Description
	de-fault=1	
`@preselectionComponents`	M	See ISO/IEC23009-1:2014/Amd.4:2016 [2].
`@tag`	M	See ISO/IEC23009-1:2014/Amd.4:2016 [2]. Note that the tag is mandatory ATSC Audio and provides a unique binding of the Preselection to the decoder.
`Language`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2]. Note that the `@lang` attribute should not be present. If present, at least one **`Language`** element shall be present that expresses the language of `@lang` redundantly.
`Role`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2]. The usage should be restricted to the Role scheme defined in ISO/IEC 23009-1 [2] and the following values: `main`, `alternate`, `supplementary`, `commentary`, `dub`, and `emergency`.
`Accessibility`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2]. The usage should be restricted to the Role scheme defined in ISO/IEC 23009-1 [2] and the following values: `descriptions`, `enhanced-audio-intelligibility`.
`Viewpoint`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2]. The view point descriptor may be used to annotate Adaptation Sets from different media types that are preferably played jointly, e.g. and audio and video presenting the view from the same view point.
`Rating`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2]. For usage, please refer to clause 5.7.3.
`Label`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2].
`AudioChannelConfiguration`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2].
`EssentialProperty`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2]. The following schemes and values are expected to be recognized by a receiver: - Content Interactivity descriptor as defined in ISO/IEC23009-1:2014/Amd.4:2016 [2], clause 5.8.5.11 with value set to 1. - Others defined by the codec specifically
`SupplementalProperty`	0 … N	See ISO/IEC23009-1:2014/Amd.4:2016 [2].

Element or Attribute Name	Use	Description
Legend: For attributes: M=Mandatory, O=Optional, OD=Optional with Default Value, CM=Conditionally Mandatory. For elements: <minOccurs>..<maxOccurs> (N=unbounded) Elements are **bold**; attributes are non-bold and preceded with an @.		

5.4.3.4. Preselection Descriptor

A scheme is defined to be used with an Essential or Supplemental Descriptor as "urn:mpeg:dash:preselection:2016". The value of the Descriptor provides two fields, separated by a comma:

- the tag of the Preselection
- the id of the contained elements/content components of this Preselection list as white space separated list in processing order. The first id defines the main element.

If the Adaptation Set includes the main element, then the Supplemental descriptor may be used to describe contained Preselections in the Adaptation Set.

If the Adaptation Set does not contain the main element the Essential Descriptor may be used instead.

The bundle is inherently defined by all elements that are included in all Preselections that include the same main element. Preselections are defined by the metadata that is assigned to each of the elements that are included in the Preselection.

> Note: This signaling may be simple for basic use cases, but is expected to not provide full coverage for all use cases.

> Note: The signaling constraints in Table 6 apply on Adaptation Set level if the Preselection property descriptor is used.

5.4.3.5. Staggercast Audio Descriptor

Staggercast is a robustness feature that can be optionally added to a program. It consists of delivering a redundant version of the audio possibly coded with lower quality (e.g. lower bitrate, number of channels, etc.) and with a significant lead ahead of the audio with which it is associated.

> Note: For live content, staggercast audio stream may be sent ahead of the main audio stream by, for instance, taking advantage of the internal delay of encoding a video GoP. "

Receivers that support the Staggercast feature can switch to the Staggercast stream should main audio become unavailable. The delivery offset (delay) between Staggercast audio and regular audio should be chosen high enough to provide robustness given the sufficient time diversity between both audio streams.

To explicitly signal that a Representation is only suitable for Staggercast, a scheme is defined to be used with an Essential Property Descriptor as "http://dashif.org/guidelines/dash-atsc-staggercast". The value of the Descriptor is a comma-separated list of the id attribute of the Adaptation Sets to which the Staggercast Representation belongs.

To enable staggercast audio impairment capability, the MPD shall be constructed as follows:

- Include an additional Adaptation Set that that contains one and only one Staggercast audio Representation.
- Annotate the Adaptation Set with a Staggercast Audio descriptor.

- Staggercast Representation shall be time-aligned with the Representation it belongs to in the main Adaptation Set.

If an Adaptation Set is annotated with a Staggercast Descriptor then the receiver is expected to not select such Representation for regular playout. If the receiver supports the Staggercast feature, it is expected to buffer both the main audio and the Staggercast audio in order to be able to switch to the Staggercast audio, should main audio become unavailable.

> Note: The amount of delay between main audio and Staggercast audio can be inferred from the MPD by comparing the value of the `@availabilityTimeOffset` information of the two Adaptation Sets.

5.4.4. Codec-specific Issues

5.4.4.1. Introduction

This section provides codec-specific issues that on how codecs can be mapped on the generic data structure defined in clause 5.4.3. This typically includes for each codec

- Codecs parameter settings
- Usage of the Preselection elements
- Random Access Point and Switching Point requirements
- The definition of bitstream switching or media level switching
- File format encapsulation requirements

5.4.4.2. Dolby AC-4 specific details

5.4.4.2.1. General

This section provides more details on Attributes and Elements used with AC-4. See ATSC A/342-2 [8].
ISO Base Media File Format Packaging Rules for AC-4 are described in ATSC A/342-2 [8], section 5.6. Random Access and Bitstream Switching is defined in ATSC A/342-2 [8], section 5.6.4.

Table 7 provides the element and attribute settings for AC-4.

Table 7 – AC-4 Elements and Attributes

Element or Attribute Name	Description
`@codecs`	For AC-4 the value of the codecs attribute shall be created according to the syntax described in RFC 6381 [22]. The value shall consist of the dot-separated list of the 4 following parts of which the latter three are represented by two-digit hexadecimal numbers: • The `fourCC` "ac-4" • The `bitstream_version` as indicated in the `ac4_dsi_v1` structure. • The `presentation_version` as indicated for the selected presentation in the `ac4_dsi_v1` structure. • The `mdcompat` parameter as indicated in the `ac4_presentation_v1_dsi` structure of the selected presentation. Example: "ac-4.02.01.03"

	The AC-4 ac4_dsi_v1 structure is described in Annex E of ETSI TS 103 190-2 [21].
Preselection@tag	This field shall correspond to the value of the presentation_group_index in the ac4_presentation_v1_dsi associated with an AC-4 presentation within the ac4_dsi_v1 structure.
AdaptationSet@tag	This field shall correspond to the value of the presentation_group_index in the ac4_presentation_v1_dsi associated with an AC-4 presentation within the ac4_dsi_v1 structure.
ContentComponent@tag	This field shall correspond to the value of the presentation_group_index in the ac4_presentation_v1_dsi associated with an AC-4 presentation within the ac4_dsi_v1 structure.
AudioChannelConfiguration	For AC-4 the Audio Channel Configuration descriptor shall use the "tag:dolby.com,2015:dash:audio_channel_configuration:2015" scheme URI. The value shall contain a six-digit hexadecimal representation of a 24-bit speaker group index bit field, which describes the channel assignment of the referenced AC-4 bit stream according to Table 27 in Annex A.3 of ETSI TS 103 190-2 [21]. This value is represented by the presentation_channel_mask_v1 parameter in the ac4_dsi_v1 structure. For example, for a stream with an 3/2/2 (5.1.2) Immersive Audio channel configuration using speakers L, R, C, Ls, Rs, TL, TR, LFE, the value shall be ”E30000” (hexadecimal equivalent of the binary value 1110 0011 0000 0000 0000 0000). The parameter b_presentation_channel_coded in the ac4_dsi_v1 structure indicates false if the audio contains objects. For content that conveys audio objects that may be rendered to positions/coordinates independent from speaker configurations, the hexadecimal value ”000000” should be indicated.
@audioSamplingRate	Example: "48000" for 48 kHz The indication shall correspond to the sampling frequency derived from the parameters fs_index and dsi_sf_multiplier inside the ac4_dsi_v1 structure described in Table E.4 in Annex E.9.3 of ETSI TS 103 190-2 [21].
@mimeType	The MIME type to be used with AC-4 shall be ”audio/mp4”.
RandomAccess	The type to be used with AC-4 shall be ”closed”, i.e. the SAP type is 1.
Language	The language indicated should correspond to the information conveyed in the language_tag_bytes of the ac4_substream_group_dsi structure (within the ac4_dsi_v1 structure) which is tagged as “dialog” or “complete main” in the corresponding content_classifier.
Role	The Role@value should be set by the content author. Note: The indication of the content_classifier from the ac4_substream_group_dsi structure is not sufficient to enable setting of an accurate indication for the Role descriptor in context of Preselections, describing entire experiences rather than individual audio elements.
Accessibility	The content_classifier field in the ac4_substream_group_dsi structure defined in ETSI TS 103 190-2 [21] describes the type of audio conveyed by audio elements.

	In case one or more audio elements related to an AC-4 Preselection indicate "visually impaired", an Accessibility descriptor shall indicate "descriptions" according to the Role scheme defined in ISO/IEC 23009-1 [2]. If one or more audio elements referenced by an AC-4 Preselection indicate a content type other than "music and effects" by means of the corresponding `content_classifier`, an Accessibility descriptor with the value "enhanced-audio-intelligibility" according to the Role scheme defined in ISO/IEC 23009-1 [2] may be used to indicate that the AC-4 Preselection enables the ability for a receiver to change the relative level of dialog to enhance dialog intelligibility. In case one or more audio elements related to an AC-4 Preselection indicate "Associated service: emergency (E)" by means of the value '110' in the corresponding content_classifier, an Accessibility descriptor shall indicate "emergency" according to the Role scheme defined in ISO/IEC 23009-1.
`Label`	The Label for a Representation should be set by the content author.

The value of the Preselection Property Descriptor provides two fields, separated by a comma:

- The first field shall correspond to the value of the `presentation_group_index` in the `ac4_presentation_v1_dsi` associated with an AC-4 presentation within the `ac4_dsi_v1` structure.
- The second field shall contain the whitespace separated list of **`AdaptationSet`** or **`ContentComponent`** ids which are included in the indicated Presentation.

5.4.4.2.2. Immersive Audio for Headphones Content Descriptor

If the content of an AC-4 Preselection has been tailored for headphones and therefore should be rendered on headphones, a Supplemental Property Descriptor should be used to indicate this property.

For AC-4 the Immersive Audio for Headphones Content Descriptor uses the "tag:dolby.com,2016:dash:virtualized_content:2016" scheme URI.

The value is set according to the `b_pre_virtualized flag` from the corresponding presentation_v1_dsi in the ac4_dsi_v1 defined in ETSI TS 103 190-2 [21].

5.4.4.3. MPEG-H Audio specific details

5.4.4.3.1. Packaging for ISOBMFF

5.4.4.3.1.1. MPEG-H Audio specific details

The storage of MPEG-H Audio is specified in ISO/IEC 23008-3:2015/Amd 2 [18]. Additional constraints on the audio elementary stream are specified in ISO/IEC 23008-3:2015 section 5.5.6 and section 5.7 [18]. See also ATSC A/342-3 section 5.2 [9] for constraints in the context of ATSC 3.0.

5.4.4.3.1.2. ISOBMFF sample entry

MPEG-H Audio supports both, storage of raw Access Units (AU) and storage of MHAS streams in the ISOBMFF. For this profile only MHAS streams shall be used. The sample entry in ISOBMFF shall be `'mhm1'` for single streams and `'mhm2'` when multiple streams are used. MHAS allows the in-band signaling of configuration information that can be used, e.g. for dynamic reconfigurations at Segment boundaries for easy ad-insertion as well as general purpose splicing and

trimming operations. MHAS is defined in 23008-3 section 14 [18]. Further, all rules and constraints specified in ATSC A/342-3 section 5.2.1 [9] apply.

5.4.4.3.1.3. Random Access and Bitstream Switching

Random Access and Stream Access Points for MPEG-H 3D Audio are described in section 5.7 of ISO/IEC 23008-3:2015 [18].

For delay-free priming of the decoder, the first AU of the audio stream shall contain an `AudioPreRoll()` element with `numPreRollFrames` set to 1 according to ISO/IEC 23008-3:2015 Amd 3 [18].

The MHASPacketLabel shall have different values for all representations of an adaptation set. Further, all rules and constraints specified in ATSC A/342-3 section 5.2.2 [9] apply.

In case of hybrid broadcast/broadband or multi-stream delivery the Random Access Points of all streams within a bundle shall be aligned.

For Stream Access Points that are supposed to be used for seamless switching, the same restrictions apply.

Table 8 MPEG-H Audio Elements and Attributes

Element or Attribute Name	Description
`@codecs`	For MPEG-H Audio, the value of the codecs attribute shall be created according to the syntax described in RFC 6381 [22]. The value consists of the following two parts separated by a dot: • The fourCC "`mhm1`" • The hex value of the profile-level-id starting with '`0x`' Example: "`mhm1.0x0D`" The profile-level-id is defined in ISO/IEC 23008-3 [18]
`AdaptationSet``@tag`	This field lists the `mae_groupIDs` as defined in ISO/IEC 23008-3 [18] that are contained in the Adaptation Set separated by white spaces.
`Preselection``@tag`	This field indicates the `mae_groupPresetID` as defined in ISO/IEC 23008-3 [18] that refers to a Preset in scope of MPEG-H Audio.
`ContentComponent``@tag`	This field indicates the `mae_groupID` as defined in ISO/IEC 23008-3 [18] which is contained in the Media Content Component.
`AudioChannelConfiguration`	For MPEG-H Audio, the Audio Channel Configuration descriptor shall use the "`urn:mpeg:mpegB:cicp:ChannelConfiguration`" scheme URI. The value shall be taken from the `ChannelConfiguration` table as defined in ISO/IEC 23001-8 [17]. Valid numbers for value are 1-7,9-12, 14-17 or 19
`@audioSamplingRate`	Example: "48000" for 48 kHz The indication shall correspond to the sampling frequency derived from the `usacSamplingFrequencyIndex` or `usacSamplingFrequency` as defined in ISO/IEC 23003-3.
`RandomAccess`	The type to be used with MPEG-H Audio shall be "`closed`", i.e. the SAP type is 1.
`@mimeType`	The MIME type to be used with MPEG-H Audio shall be "`audio/mp4`".

Language	The language indicated should correspond to the information conveyed in `mae_contentLanguage` of the default dialog element: The `maeGroup` which is marked as default in `mae_switchGroupDefaultGroupID` and is tagged in `mae_contentKind` as dialogue. This information is carried in the `AudioSceneInformation()` of the MPEG-H Audio stream as defined in ISO/IEC 23008-3.
Role	The Role for a Preselection should be set by the content author.
Accessibility	If the `mae_contentKind` value of at least one Audio Element is set to '9' ("audio-description/visually impaired"), an Accessibility descriptor shall indicate "descriptions" according to the Role scheme defined in ISO/IEC 23009-1 [2]. If at least the Audio Elements with a `mae_contentKind` value of '2' ("dialogue") have `mae_allowGainInteractivity` set to '1' and `mae_interactivityMaxGain` set to a non-zero value in the corresponding `mae_GroupDefinition()` structure, an Accessibility descriptor with the value "enhanced-audio-intelligibility" according to the Role scheme defined in ISO/IEC 23009-1 [2] may be used to indicate that the Preselection enables the ability for a receiver to change the relative level of dialog to enhance dialog intelligibility. the mae_contentKind value of at least one Audio Element is set to '12' ("emergency"), an Accessibility descriptor shall indicate "emergency" according to the Role scheme defined in ISO/IEC 23009-1. The accessibility information indicated for a Preselection should also correspond to the `mae_groupPresetKind`. The `mae_contentKind` field and all other fields mentioned above that start with a "`mae_`" prefix are carried in the AudioSceneInformation() of the MPEG-H Audio stream as defined in ISO/IEC 23008-3.
Label	The Label for a Preselection should be set by the content author.

The value of the Preselection Property Descriptor provides two fields, separated by a comma:

- The first field shall correspond to the value of the `mae_groupPresetID` as defined in ISO/IEC 23008-3 [18] that refers to a Preset in scope of MPEG-H Audio.
- The second field shall contain the whitespace separated list of `Adaptation Set` or `Content Component` ids which are included in the indicated Preset.

5.4.5. Service Offering Requirements and Recommendations

Note: this section will be provided in the next revision of this document following the multi-track work currently completed in DASH-IF including Accessibility use cases.

5.4.6. Expected Client Behavior

Note: this section will be provided in the next revision of this document following the multi-track work currently completed in DASH-IF.

5.5. Subtitling and Closed Captioning

5.5.1. Background and Use Cases (Informative)

ATSC 3.0 subtitles and closed captioning is defined in A/343 [10] which is based on W3C TTML IMSC1 as profiled in DASH-IF IOP [1]. Two profiles are included:

- Text Profile requiring a font rendering engine in the decoder
- Image Profile with PNG files

ATSC 3.0 Closed Captions are required to be carried as files and to be presented appropriately for ATSC 3.0 Video (e.g., 3D, HDR video). In order to provide the signaling of the presence of timed text-based data streams and closed captioning services on MPD level, descriptors on DASH level are defined.

5.5.2. Assumptions

The following closed caption metadata as provided in ATSC A/343, section 7.1 [10] is expected to be present for certain Adaptation Sets and Representations to enable suitable initial selection and switching:

- Language: the dominant language of the closed caption text
- Role: the purpose of the closed caption text, e.g., main, alternate, commentary.
- Display aspect ratio: the display aspect ratio assumed by the caption authoring in formatting the caption windows and contents.
- Easy reader: this metadata, when present, indicates that the closed caption text tailored to the needs of beginning readers
- Profile: this metadata indicates whether text or image profile is used.
- 3D support: this metadata, when present, indicates that the closed caption text is tailored for both 2D and 3D video.

5.5.3. Service Offering Requirements and Recommendations

5.5.3.1. DASH-specific aspects for Timed Text based Closed Caption

All constraints of DASH-IF IOP, section 6.4.4 [1] are applied; 14496-30 COR1 and COR2 [19] are applied.

- Mix of 2D and 3D closed captioning data per Period shall not be allowed.
- Only ISOBMFF encapsulation is permitted; and thus the only `@codecs` values are `"sbtt.ttml.im1t"` or `"stpp.ttml.im1i"`.

5.5.3.2. MPD-based Signaling of Timed Text based Closed Caption service metadata

This subsection provides methods MPD-based Signaling of Timed Text based Closed Caption services. Closed Caption metadata should be signaled properly using descriptors available in ISO/IEC 23009-1, specifically Role, Essential Property and Supplemental Property descriptors.

The language attribute shall be set on the Adaptation Set. Role element shall be used as necessary and the DASH role scheme may be used.

The Essential Property and/or Supplemental Property descriptors with the `@schemeIdURI` equal to `"http://dashif.org/guidelines/dash-atsc-closedcaption"`, and `@value` attribute to contain the Caption Service Metadata described in section 7.1 in [A/343] as a semicolon-separated string. The `@value` syntax shall be as described in the ABNF below.

```
@value = "ar" ":" aspect-ratio ["," easy-reader]["," profile] ["," 3d-support]
aspect-ratio = (%d1-%d99) "-" (%d1-%d99)
easy-reader = "er" ":" BIT; default value 0
profile = "profile" ":" BIT; default value 0 for text profile
3d-support = "3d" ":" BIT; default value 0
```

Based on the above ABNF, following parameters are defined for Timed Text Closed Caption metadata:

- `aspect-ratio` may be set to any value pairs, including: “4-3”, “16-9”, and “21-9”.
- `easy-reader` shall be set as a Boolean value; it is set as ‘1’ if present, otherwise the default is 0.
- `profile` shall be set as a Boolean value; it is set as ‘1’ for image profile if present, otherwise the default is 0 for text profile.
- `3d-support` shall be set as a Boolean value; it is set as ‘1’ if the 3D is supported, otherwise the default is 0.

5.6. Interactivity Events

5.6.1. Background and Basic Use Cases (Informative)

ATSC 3.0 Application Signaling specifies mechanisms for signaling app-based enhancements in both linear services containing app-based enhancements and standalone app-based services (which consist entirely of app-based enhancements), as well as mechanisms for delivering activation notifications, or “events” which activate or change the state of the associated applications at precise times in the media presentation timeline and can be mapped to wall-clock time. The details of application signaling are specified in A/337 [6]. Note that this section only deals with IF-1 of Figure 2, i.e. events and triggers as defined in A/337 [6]. Generic events may be used as well, and if so, they may be using IF-3 in Figure 2, as for example discussed in clause 6. Note also that the function “ATSC events” in Figure 2 may be part of the Application and therefore IF-1 and IF-3 coincide.

Some relevant feature for event signaling are summarized. The format is expected to support signaling of events with precise timing such that the action of the triggered application operations can be synchronized. The format is expected to support signaling of a series of events. The format is expected to support signaling of events using the MPD as well as part of Media Segments of Representations, e.g., using the `‘emsg’` box [2]. Both broadcast- and broadband-delivered content may support events.

5.6.2. Mapping to DASH

The existing MPEG-DASH Event Mechanism as defined in ISO/IEC 23009-1, clause 5.10, shall be used to carry ATSC events. The working draft of ATSC A/337[5], section 5.4 defines the ATSC events including a scheme ID URI as well as values for different events (a table update Event Stream used in the context of devices that have access to an ATSC 3.0 broadcast stream, and for a table update Event Stream used in a redistribution setting).
Application-specific Event Streams may be defined by application developers. The only constraints are that the `schemeIdUri/value` combination must be globally unique, such as by the use of a `schemeIdUri` controlled by the application developer, and by proper management of the value attributes. In order to get access to these Events, applications register callback routines for them, and the callback routines are called when such Events arrive.

5.6.3. Service Offering Requirements and Recommendations

Interactivity Events may be carried:

- As MPD Events as defined in ISO/IEC 23009-1, clause 5.10.2
- As Inband events as defined in ISO/IEC 23009-1, clause 5.10.3

If Inband events are used, then at least all Representations of all main audio Adaptation Sets shall contain an **InbandEventStream** element with `@schemeIdUri` set to the ATSC-defined value. In addition, all non-dependent Representations of at least one media type/group should contain an **InbandEventStream** element with `@schemeIdUri` set to the ATSC-defined value.

5.6.4. Expected Client Behavior

The DASH client shall download at least one Representation that contains `InbandEventStream` element set to the ATSC-defined value.

The process as defined in clause 4.4 is expected to be used.

The event information is handed to the ATSC event function.

5.7. Programs and Program Ratings

5.7.1. Program Definition in ATSC

According to ATSC, a Program is defined as follows:

> *Program — Content of a defined composition and scheduled duration intended by the broadcaster to be treated as a programming unit.*

Programs may map to a content fragment identified in the Electronic Service Guide (ESG).

5.7.2. Program Signaling

Program signaling is out of scope for this profile.

5.7.3. Program Rating Signaling in DASH

When using DASH, the ratings value shall be specified by the **MPD.Period.AdaptationSet.Rating** element. When the content advisory corresponds to a rating system defined by an RRT, the value of **Rating**`@schemeIdUri` shall be set equal to "`http://dashif.org/guidelines/dash-atsc-RRTrating:1`". The `@value` string shall be set equal to the content advisory ratings string specified in A/331 Section 7.3.1 [4]. Alternatively or in addition, content advisories corresponding to other rating systems may be included. For content advisories not corresponding to defined RRTs, different **Rating**`@schemeIdUri` values shall be used, as specified by appropriate regional authorities.

The **Rating** element is a child element of **AdaptationSet**, thus any or all Adaptation Sets in a Period could be labeled with a content advisory. When the entire Program is to be associated with one content advisory rating (the usual case), at least one instance of the Rating element with a value of "`http://dashif.org/guidelines/dash-atsc-RRTrating:1`" for **Rating**`@schemeIdUri` shall be included in the Period as an **MPD.Period.AdaptationSet.Rating** element. Multiple Rating elements with different values for **Rating**`@schemeIdUri` may be included in the Period as **MPD.Period.AdaptationSet.Rating** elements. In the DASH MPD no **ContentComponent** element shall include a **Rating** element.

The rules for placement of a **Rating** element with a value of "`http://dashif.org/guidelines/dash-atsc-RRTrating:1`"for **Rating**`@schemeIdUri` shall be as follows:

- When a Period includes only one Adaptation Set containing one or more video components (e.g. those with `@mimeType="video/mp4"`), the **`Rating`** element shall appear in that **`AdaptationSet`**.
- When a Period includes multiple Adaptation Sets each with `@mimeType="video/mp4"`containing video components, the Rating element shall appear in each Adaptation Set among these whose **`Role`**`@schemeIdUri` is equal to `"urn:mpeg:dash:role:2011"`and **`Role`**`@value` is equal to `"main"`.
- When a Period includes no Adaptation Sets describing video components, i.e. none of the **`AdaptationSet`** elements have `@mimeType="video/mp4"`, the Rating element shall appear in each **`AdaptationSet`** listed in the MPD for that Period.

6. Ad Insertion

6.1. Background (Informative)

An ATSC 3.0 receiver accesses broadcast signaling identifying the availability of streaming services delivered within the broadcast stream, by broadband, or by a combination of the two (hybrid services). An ATSC 3.0 receiver which supports the application runtime environment defined in A/344 [11] can, under the control of a broadcaster-supplied application, present personalized ads to the viewer. When a personalized ad is played, it replaces the content that is present in the regular stream (e.g. content that is played by receivers not supporting the runtime environment).

As described in the Client Reference Model in Section 2.3.2, receivers include a DASH Player that is responsible for managing the playout of DASH Media Segments. The locations of ad avails are defined as DASH Periods. The MPD delivered in the signaling can identify one or more ad avails by placing an XLink in a future Period. When the DASH Player sees an MPD update containing an XLink, it interacts with the broadcaster application over interface IF-3 to attempt to resolve it. If resolution is successful, one or more Period elements are returned to the DASH Player, which replaces the Period that had contained the XLink with the one or more new Period elements.

Personalized ad insertion requires making choices about which ad content is appropriate for a particular viewer. In the ATSC 3.0 receiver, such choices are made by the broadcaster application. Once an XLink to be resolved is received by the app, it can perform appropriate logical operations, using whatever personalization information it has access to, to choose the appropriate ad content. Alternatively, the app might pass the XLink, with appropriate query terms, to a broadcaster server which would perform the decision logic.

6.2. Use Cases (Informative)

6.2.1. Series Fan

The broadcaster wishes to target personalized ads to fans of a certain TV series. Based on Joe's recent viewing of six hours of the "marathon" for this series, he is presented with an ad for memorabilia, while others in his neighborhood view different advertising in that slot.

6.2.2. Swing Shift Viewer

Based on Ted's TV viewing hours being predominantly in the 11pm to 4am time period, he is presented with an ad for employment services, while others in his neighborhood view different advertising in that slot.

6.2.3. Young Cat Lover

Emily had interacted with her favorite cartoon show on Saturday to indicate her love of cats. On Sunday morning, she is presented with an ad for cat toys, while others in her neighborhood are presented with ads for different products.

6.2.4. Geographic Location

A broadcaster wishes to play an ad for a car dealership local to the west side of town to those living there, and an ad for a different dealership to those living on the east side of town

6.2.5. Generic Personalized Ads

A viewer watching TV is presented personalized ads during broadcast ad spots. Characterization of typical decisions for personalized ad insertion include:

- Demographics (age, gender, location, income, education)
- Interests (arts & entertainment, finance, autos, cooking, survival, sports, etc.)
- Viewing behavior (program/channel selection, time of day, channel surfer, ads watched vs. skipped, etc.)
- Device characteristics (make/model/vintage, capabilities, etc.)

6.2.6. Incidence of Breaking News during Replacement Ad Viewing

A TV viewer is watching a replacement ad which is interrupted with breaking news. The replacement ad stops playing and the breaking news is viewed.

6.2.7. Trick Mode Access associated with Replacement Ad Viewing

A TV viewer watches a replacement ad during a previously recorded show. He/she is able to pause and rewind during that replacement ad.

6.2.8. Replacement Ad Containing Interactivity Components

A TV viewer watches a replacement ad that also has interactive elements. The user uses the TV remote control to start the interaction by highlighting and selecting an icon that is on-screen. Types of interactive elements might include:

- The ability to receive a coupon for a product by typing in their mobile phone number.
- View the location of the nearest car dealer onscreen in an overlay that does not interfere with critical visual elements of the ad.
- Get more detailed product information on a registered companion device (tablet or smart phone).

6.3. Assumptions

The following system aspects are assumed:

- The receiver implements the Application Runtime Environment specified in A/344 [11].
- Ad avails are identified by the placement of Periods with XLinks in the MPD.
- The receiver's DASH Player resolves ATSC app-specific XLinks by interacting with the broadcaster-supplied application through the JSON WebSocket RPC API defined in A/344 [11].
- Non-personalized ads may be included in broadcast content, either not exposed as separate Periods or associated with Periods.
- XLink resolution may fail. In that case, the client is expected to delete the XLink and use the default Content.
- XLinks to be communicated to the broadcaster application are identified as such by a specified URI pattern in the href attribute. Xlinks not matching the pattern may appear, including for example http(s) URLs. Receivers not supporting a given form of XLink resolution are expected to delete the associated XLinks from the Period.
- The broadcaster app, at the discretion of the app designer and subject to the availability of broadband access, may append personalization data to an XLink and forward it to a broadcaster's resolution server for processing. Upon receiving a response from the broadcaster server, the replacement Period(s) may be returned to the receiver's DASH player using the XLink resolution API defined in A/344 [11].

6.4. Service Offering Requirements and Recommendations

6.4.1. General

Service offering should follow the server-driven ad insertion approach, as defined in DASH-IF IOP [1], clause 5.3, which uses remote periods to represent avails. Remote period resolution is performed by a broadcaster-supplied app.

The service offering may contain inband `'emsg'` boxes or/and **`EventStream`** elements, carrying payloads such as SCTE 35 cue messages. Treatment of specific event payloads is outside the scope of this document, and the client is expected to be able to play seamlessly irrespective of whether the above events were handled by an application.

6.4.2. Remote Periods

An avail is represented by one or more remote Period elements.

Each remote Period element shall contain "default content", i.e., it would be a playable non-empty Period would its **`Period`**`@xlink:href` attribute be deleted.

If the **`Period`**`@xlink:href` attribute is present, the `@xlink:actuate` attribute shall be present and have the value "`onLoad`".

6.4.3. XLink API

An XLink to be resolved by a broadcaster-supplied application is identified by a `Period@xlink:href` attribute containing a URI conforming to a format specified in A/344 [11]. Resolution of Remote Periods with such URIs is expected to be handled by applications and is outside the scope of this document.

6.5. Expected Client Behavior

6.5.1. XLink

MPD Periods with XLink URIs conforming to a format specified in A/344 [12] are resolved by local apps via the JSON-RPC API defined in A/344 [11].

If Remote Period dereferencing time exceeds 3 seconds, the client should assume that dereferencing failed. The consequence is no modification to the broadcast MPD and thus playback of the "default" content.

6.5.2. Events

Events are expected to be passed to apps using same mechanism as described in clause 5.6.4, however events with non-ATSC `@schemeIdUri` values should be expected. For more details on expected receiver handling, refer to clause 4.4.

7. DRM and Security

7.1. Introduction

The following describes the content protection and DRM solution using Common Encryption of media, and DASH MPD signaling of DRM licenses.

It is assumed that devices will connect to a DRM license server to receive a device or user specific license that will authorize access to protected content. The method and frequency of license server connection is a deployment choice and can range from one-time provisioning when a device is purchased, to unlimited on-demand online license downloads. Broadcast delivery of individualized licenses (cryptographically bound to a device or user) is not specified by DASH-IF.

Device independent "child" licenses that contain a Media Segment decryption key can be accessed by all authorized devices and users (with a "parent" license), and may be delivered in every Media Segment to facilitate random access and key rotation.

The model is based on a "parent/child" hierarchy of licenses and keys supporting "key rotation" and subscriptions at the content level. In addition to a common scrambling algorithm, the following steps are needed to authorize playback:

- Devices must be initialized and registered by an authorization server in order to identify the device or user to be authenticated and authorized, and must establish a cryptographic identity to a DRM client to allow the license server to generate cryptographically bound licenses. Note that different devices may use different DRM clients.

- Devices need to retrieve device or user bound licenses that authorize a set of content determined by the Operator, typically, a subscription to a service. A license may authorize limited permissions, such as a time limit, resolution limit, geographical limit, etc.
- Optionally, enforcement of authorization may be repeated per program segment or time interval by changing the key used to encrypt corresponding Media Segments, thus requiring the DRM system to verify authorization for that device or user in order to extract the key delivered by a child license within the Broadcast stream.

Note that the system does not support broadcast-only distribution of individual licenses.

7.2. Device Initialization

DRM specific protocols are used for enabling the device in the operator network. It is a one-time operation requiring connections to the operator head-end for uniquely identifying and authenticating the device. For example, the DRM system may perform an operation with a hardware embedded DRM client key, or may install a domain certificate on each authorized device belonging to a particular user so that a single license can authorize all the devices in that domain.

7.3. License Delivery

Licenses are retrieved by the device using DRM specific protocols. It requires connecting to authentication, authorization, and licensing servers. How often this connection is required depends on the validity period of the licenses that are delivered. This can be a one-time operation if the license has an infinite life time (some years) or this can be on a regular basis (e.g. every month) for renewing a subscription for example.

7.4. Key Rotation

Section of 7.5 of DASH-IF IOP [1] defines different mechanisms for key rotation. In ATSC 3.0, the key hierarchy as described in subsection 7.5.3.4 is to be used.

How often keys are rotated is a deployment choice. Typically, parent licenses at the Entitlement Management Level (EML or "parent license") are expired every month for a subscription service so authorization will fail if a user stops their subscription. At the Entitlement Control Level (ECL), child licenses change more frequently, typically per show or time interval. Each change requires an authorization check because a valid parent license must be present in order to extract the new key from the child license in the Media Segment, so authorization limitations (location, expiration, resolution, etc.) will be checked by the DRM system. Historically, key rotation was used to prevent key factoring and distribution when 8-bit keys were used and factoring took minutes, later seconds. With Common Encryption and 128-bit keys, key factoring is no longer a reason to use frequent key rotation.

7.5. Content Encryption

Common ENCryption (CENC) of NAL structure video and other media data with AES-128 CTR mode is used. The use of the cenc scheme follows guidelines defined in Section of 7.4 of DASH-IF IOP [1].

7.6. Manifest Signaling

DASH IF specifies the use of **`ContentProtection`** Descriptors in the MPD to identify:

1. **Adaptation Sets encrypted using a default_KID**.
   The **`ContentProtection`**`@schemeIdUri="urn:mpeg:dash:mp4protection:2011"` contains the attribute `cenc:default_KID`, which equals the `default_KID` field in the Track Encryption Box (‘`tenc`’) of the Initialization Segments.
2. **DRM licenses that are available and necessary.**
   There should be a ContentProtection Descriptor for each DRM system supported, identified by a UUID, and containing any information defined by that DRM system. These **`ContentProtection`** Descriptors have `@schemeIdUri="urn:uuid:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx"`, where the UUID value is registered at http://www.dashif.org/identifiers/protection.

A DASH player can make a license request or verify the presence of a license for the `default_KID` indicated and any of the DRM systems that it supports. That license can either provide the key to decrypt the content, or if a parent license, the key to access child licenses broadcast in Media Segments that contain the keys to decrypt the content. Protection System Specific Header Boxes (‘`pssh`’) SHALL NOT be used in Initialization Segments to signal encryption or DRM licenses. Players SHOULD pass any ‘`pssh`’ boxes present in Media Segments to the DRM system (“Content Decryption Module”). MPD signaling follows guidelines defined in Section of 7.6 of DASH-IF IOP [1].

8. Relevant Use Cases and Content Offering Guidelines

Note: This section will be provided in the next revision of this document.

Annex A MDE Delivery Methods

A.1 HTTP Media Segment Delivery

In conventional, HTTP file playback DASH client fetches Media Segments shortly after they become available in the Cache as shown above in Figure A1.1.

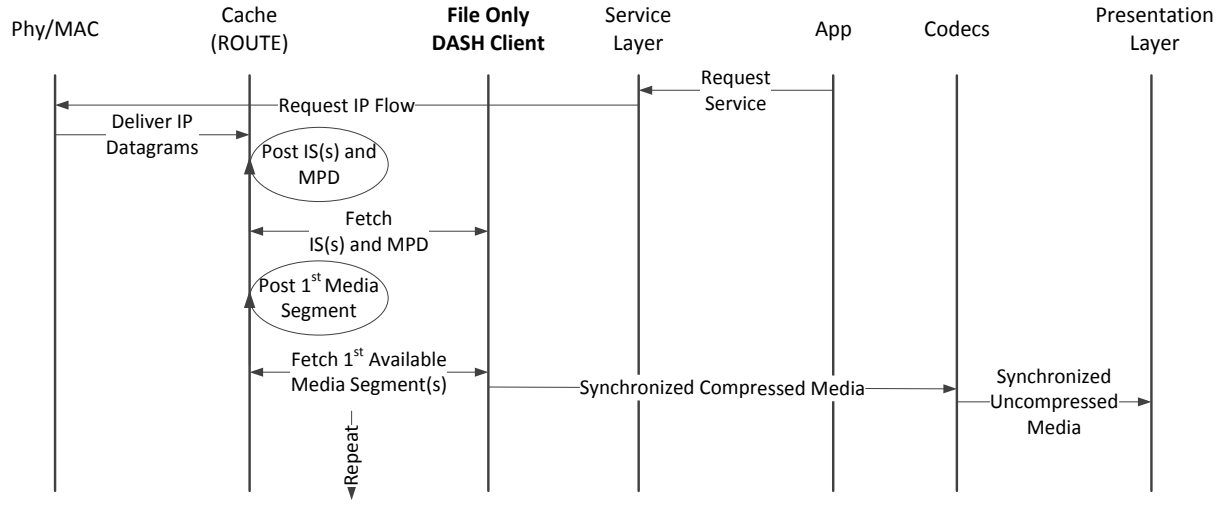

Figure A1.1: Call Flow for HTTP File Delivery to DASH Client

For MDE delivery as shown below in Figure A1.2, the MDE aware DASH Client requests the desired Media Segment prior to the MPD defined availability time and the Cache streams MDEs to the DASH client upon expiry of the ROUTE timer for the requested Media Segment.

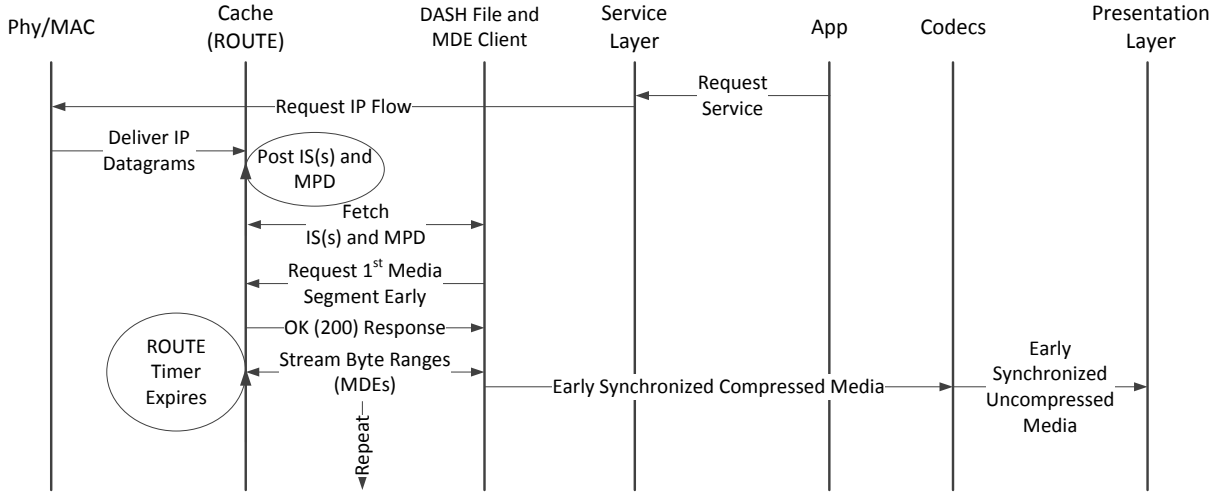

Figure A1.2: Call Flow for HTTP MDE Delivery to MDE Aware DASH Client

A.2 Websocket Delivery of MDE

Figure A2.1 above depicts a typical call flow for Websocket delivery of MDE to a client. The DASH client establishes a Websocket connection to the HTTP proxy via a well-known URL or address (e.g. ws://127.0.0.1:8080). In the drawing above, the DASH client can optionally receive notification of a channel change and immediately start receiving MDE's upon service acquisition. The MPD in this example is delivered in-band apriori as per the description in Section 2.7.3, which allows for hybrid use cases.

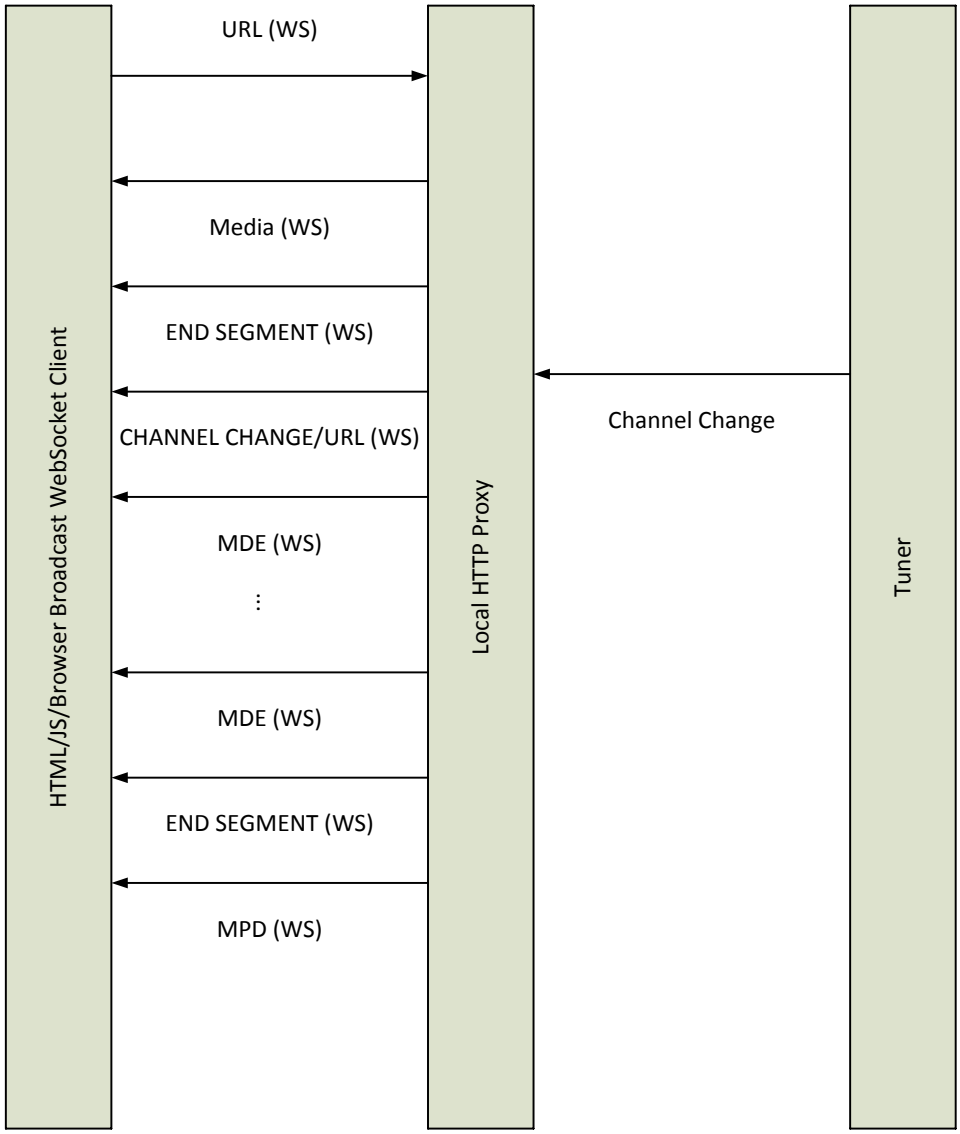

Figure A2.1: Call Flow for Websocket Delivery of MDE

Annex B Broadcast TV Profile and Related Information from ISO/IEC 23009-1 Amd.4

Note: This Annex will be removed once ISO/IEC 23009-1:2017 [2] is available. The section numbers replicate the numbers in ISO/IEC 23009-1.

5.3.3.4 Switching within Adaptation Sets

Switching refers to the presentation of decoded data from one Representation up to a certain time t, and presentation of decoded data of another Representation from time t onwards, for details refer to 4.3.

The **`Switching`** element as defined in Table AAA provides instructions of switch points within an Adaptation Set and the permitted switching options as defined in Table BBB. When this element is present, it signals opportunities for simple switching across Representations in one Adaptation Set. This element may be used instead of the attributes `@segmentAlignment` or `@bitstreamSwitching`.

Table BBB defines different switching strategies that provide instructions to the client on the procedures to switch appropriately within an Adaptation Set.

Table AAA — Switch Point Signalling

Element or Attribute Name	Use	Description
`Switching`		Switching logic description for the associated Representation
`@interval`	M	specifies the interval between two switching points in the scale of the `@timescale` on Representation level. Any Segment for which the earliest presentation time minus the `@t` value of the **`S`** element describing the segment is an integer multiple of the product of `@timescale` and `@interval` is a switch-to opportunity, i.e. it enables to switch to this Representation with the switching strategy as defined by the `@type` value. The value should be chosen such that the resulting time matches MPD start time of segments, otherwise no switching will be described

Element or Attribute Name	Use	Description
@type	OD default: 'media'	specifies the switching strategy for the switch points identified in by the @interval attribute. Switching strategies are defined in Table BBB.

Table BBB — Switching Strategies

Type	Description
media	Media level switching: In this case switching is possible at the switch point by decoding and presenting switch-from Representation up to switch point t, initializing the switch-to Representation with the associated Initialization Segment and continue decoding and presenting the switch-to Representation from time t onwards.
bitstream	Bitstream switching: In this case switching is possible at the switch point by decoding and presenting switch-from Representation up to switch point t, and continue decoding and presenting the switch-to Representation from time t onwards. More specifically, the concatenation of two Representations at the switch point results in a results in a "conforming Segment sequence" as defined in 4.5.4 with the media format as specified in the @mimeType attribute. Initialization of the switch-to Representation is not necessary and is not recommended. In order to enable this feature, it is recommended to use the same Initialization Segment for all Representations in the Adaptation Set, i.e. the highest profile/level is signaled in the Initialization Segment.

The XML schema snippet is as follows:

```
<!-- Switching  -->
<xs:complexType name="SwitchingType">
  <xs:attribute name="interval" type="xs:unsignedInt" use="required"/>
  <xs:attribute name="type" type="SwitchingTypeType"/>
  <xs:anyAttribute namespace="##other" processContents="lax"/>
</xs:complexType>

<!—Switching Type type enumeration -->
<xs:simpleType name="SwitchingTypeType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="media"/>
    <xs:enumeration value="bitstream"/>
  </xs:restriction>
</xs:simpleType>
```

5.3.3.5 Switching across Adaptation Sets

Representations in two or more Adaptation Sets may provide the same content. In addition, the content may be time-aligned and may be offered such that seamless switching across Representations in different Adaptation Sets is simplified. Typical examples are the offering of the same content with different codecs, for example H.264/AVC and H.265/HEVC and the content author wants to provide such information to the receiver in order to seamlessly switch Representations (as defined in 4.5.1) across different Adaptation Sets.

A content author may signal such seamless switching property across Adaptation Sets by providing a Supplemental Descriptor along with an Adaptation Set with `@schemeIdURI` set to `urn:mpeg:dash:adaptation-set-switching:2016` and the `@value` is a comma-separated list of Adaptation Set IDs that may be seamlessly switched to from this Adaptation Set.

If the content author signals the ability of Adaptation Set switching and as `@segmentAlignment` or `@subsegmentAlignment` are set to `TRUE`, the (Sub)Segment alignment element shall be valid for all Representations in all Adaptation Sets for which the `@id` value is included in the `@value` attribute of the Supplemental descriptor.

If the content author signals the ability of Adaptation Set switching and **`Switching`** element is provided, the signaled switch points apply for all Representations in all Adaptation Sets for which the `@id` value is included in the `@value` attribute of the Supplemental descriptor.

As an example, a content author may signal that seamless switching across an H.264/AVC Adaptation Set with **`AdaptationSet`**`@id="4"` and an HEVC Adaptation Set with **`AdaptationSet`**`@id="5"` is possible by adding a Supplemental Descriptor to the H.264/AVC Adaptation Set with `@schemeIdURI` set to `urn:mpeg:dash:adaptation-set-switching:2016` and the `@value="5"` and by adding a Supplemental Descriptor to the HEVC Adaptation Set with `@schemeIdURI` set to `urn:mpeg:dash:adaptation-set-switching:2016` and the `@value="4"`.

In addition, if the content author signals the ability of Adaptation Set switching for any Adaptation Sets then the parameters as defined for an Adaption Set shall also hold for all Adaptation Sets that are included in the `@value` attribute. Note that this constraint may result that the switching may only be signaled with one Adaptation Set, but not with both as for example one Adaptation Set signaling may include all spatial resolutions of another one, whereas it is not the case the other way round.

5.3.5.5 Random Access to Representations

Random Access refers to start processing, decoding and presenting the Representation from the random access point at time t onwards by initializing the Representation with the Initialization Segment, if present and decoding and presenting the Representation from the signaled Segment onwards. Random Access point may be signaled with the **`RandomAccess`** element as defined in Table CCC.

Table DDD provides different random access point types.

Table CCC — Random Access Signalling

Element or Attribute Name	Use	Description
`RandomAccess`		Random Access Information
`@interval`	M	specifies the position of the random access points in the Representations. The information is specified in the scale of the

Element or Attribute Name	Use	Description
		@timescale on Representation level. Any Segment for which the MPD start time minus the @t value of the **S** element describing the segment is an integer multiple of the product of @timescale and @interval is a random access opportunity, i.e. it enables randomly access to this Representation with the random access strategy as defined by the @type value. The value should be chosen such that the resulting time matches MPD start time of segments, otherwise no random access will be described.
@type	OD default: "closed"	specifies the random access strategy for the random access points in by the @interval attribute. The value shall use a type present in Table DDD. If the value of the type is unknown, the DASH client is expected to ignore the containing Random Access element.
@minBufferTime	O	specifies a common duration used in the definition of the Representation data rate (see @bandwidth attribute in 5.3.5.2 and 5.3.5.4). If not present, then the value of the MPD level is inherited.
@bandwidth	O	Consider a hypothetical constant bitrate channel of bandwidth with the value of this attribute in bits per second (bps). Then, if the Representation is continuously delivered at this bitrate, starting at any RAP indicated in this element a client can be assured of having enough data for continuous playout providing playout begins after @minBufferTime * @bandwidth bits have been received (i.e. at time @minBufferTime after the first bit is received). For dependent Representations this value specifies the bandwidth according to the above definition for the aggregation of this Representation and all complementary Representations. For details see 5.3.5.4. If not present, the value of the Representation is inherited.

1

Table DDD — Random Access Strategies

Type	Informative description
`closed`	Closed GOP random access. This implies that the segment is a Random Access Segment as well as the segment starts with a SAP type of 1 or 2. Note that SAP type 1 or 2 is a necessary condition, but not sufficient. In addition, all requirements of a Random Access Segment need to be fulfilled.
`open`	Open GOP random access. This implies that the segment is a Random Access Segment as well as the segment starts with a SAP type of 1, 2 or 3. Note that SAP type 1, 2 or 3 is a necessary condition, but not sufficient. In addition, all requirements of a Random Access Segment need to be fulfilled.
`gradual`	Gradual decoder refresh random access. This implies that the segment is a Random Access Segment as well as the segment starts with a SAP type of 1, 2, 3 or 4. Note that SAP type 1 2, 3 or 4 is a necessary condition, but not sufficient. In addition, all requirements of a Random Access Segment need to be fulfilled.

The XML schema snippet is as follows:

```
<!-- Random Access  -->
<xs:complexType name="RandomAccessType">
  <xs:attribute name="interval" type="xs:unsignedInt" use="required"/>
  <xs:attribute name="type" type="RandomAccessTypeType"/>

  <xs:attribute name="minBufferTime" type="xs:duration"/>

  <xs:attribute name="bandwidth" type="xs:unsignedInt"/>
  <xs:anyAttribute namespace="##other" processContents="lax"/>
</xs:complexType>

<!– Random Access Type type enumeration -->
<xs:simpleType name="RandomAccessTypeType">
  <xs:restriction base="xs:string">
    <xs:enumeration value="closed"/>
    <xs:enumeration value="open"/>

    <xs:enumeration value="gradual"/>
  </xs:restriction>
</xs:simpleType>
```

8.11.1 General

This profile provides a restricted profile primarily for distributing broadcast TV over broadcast and broadband services, including service offerings for combined unicast and broadcast services. The profile is based on ISO-BMFF. In order to enable those advanced use cases, this profile introduces the main restrictions that follows compared to the extended live profile:.

- Use a single `@timescale` for all Representations in one Adaptation Set
- Use Segment Timeline for signaling of segment durations
    - The timing of the segments in the MPD is accurate
    - The Segment Timeline may be on Representation level to allow different segment durations in different Representations. However, it may be defaulted on Adaptation Set level

    - The Segment Timeline may use open ended `@r` (-1) or closed `@r` (>=0)
    - The Segment Timeline may use Segment sequences and Hierarchical Templating
- Each Representation shall provide at least one **`RandomAccess`** element.
- If an Adaptation contains more than one Representation, then at least one **`Switching`** element shall be present.
- Segment alignment and start with SAP signalling may be used for backward compatible deployments, but should generally not be used.
- Data URLs as defined in RFC2397 may be used for Initialization Segments.

The ISO-Base Media File Format Broadcast TV profile is identified by the following URN: `"urn:mpeg:dash:profile:isoff-broadcast:2015"`.

8.11.2 Media Presentation Description constraints

8.11.2.1 General

The Media Presentation Description shall conform to the following constraints:

— The rules for the MPD as defined in ISO/IEC 23009-1 7.3, shall apply.

— The rules for the Segments as defined in 7.3.5 of ISO/IEC 23009-1 shall apply.

— Periods which do not conform to the constraints in 8.11.2.2 may not be presented

— Representations not inferred to have `@profiles` equal to the profile identifier as defined in 8.11.1 may be ignored

8.11.2.2 Constraints on Period elements

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

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

— **`AdaptationSet`** elements that do not conform to 8.11.2.3 may be ignored

8.11.2.3 Constraints on AdaptationSet elements

— **`AdaptationSet`** element may be ignored unless **`AdaptationSet.SegmentTemplate`** is present and/or for each Representation within this Adaptation Set **`Representation.SegmentTemplate`** element is present;

— **`AdaptationSet`** element may be ignored unless **`AdaptationSet.RandomAccess`** is present and/or for each Representation within this Adaptation Set **`Representation.RandomAccess`** element is present;

— **AdaptationSet** element that contains more than one Representation may be ignored unless **AdaptationSet.Switching** is present and/or for each Representation within this Adaptation Set **Representation.Switching** element is present and all the **SegmentTemplate** elements conform to 8.11.2.5;

— **InBandEventStream** shall only be used on Adaptation Set level.

— **Representation** elements that do not conform to 8.11.2.4 may be ignored

8.11.2.4 Constraints on Representation elements

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

— **Representation** elements may be ignored if **Representation.RandomAccess** element is not present and also no **AdaptationSet.RandomAccess** element is present.

— **InBandEventStream** shall not be present on Representation level.

— Segment Timeline shall be used for signaling of segment durations and the following restrictions shall apply:

  - The timing of the segments in the MPD shall be accurate.
  - The Segment Timeline may be open ended `@r` (-1) or may closed `@r` (>=0).
  - The Segment Timeline may contain Segment Sequences as defined in 5.3.9.6.4 and Hierarchical Templating as defined in 5.3.9.6.5.

— The Segment Timeline may be on Representation level to allow different segment durations in different Representations. However, it may be defaulted on Adaptation Set level.

8.11.2.5 Constraints on SegmentTemplate elements

— `@initialization` attribute may include data URLs as defined in RFC 2397.

8.11.3 Segment format constraints

Representations and Segments complying with this profile shall meet the following constraints:

— Representations shall comply with the formats defined in section 7.3.5.

- If Segment Sequences as defined in 5.3.9.6.4 and Hierarchical Templating as defined in 5.3.9.6.5 are used, then the first Segment of a Segment Sequence shall not carry 'dums' brand in the Segment Type box ('styp') as major brand and all other Segments of the Segment Sequence shall carry 'dums' brand in the Segment Type box ('styp') as major brand.

8.11.4 MPD Updates and Inband Event Streams

In order for a DASH client to operate without frequent MPD requests and use the information contained in Inband Event Streams, the content authoring needs to obey certain rules.

In case of **MPD**@type="dynamic" and the MPD indicates that one or several Representation(s) contain an inband event stream in order to signal MPD validity expirations, then the following applies:

- The **MPD**@publishTime shall be present.
- The **MPD**@minimumUpdatePeriod should be set to a small number, preferably 0.
- for each newly published MPD, that includes changes that are not restricted to any of the following (e.g. a new Period):
  - The value of the **MPD**@minimumUpdatePeriod is changed,
  - The value of a **SegmentTimeline.S**@r has changed,
  - A new **SegmentTimeline.S** element is added.
    - Any information that has been fallen outside the timeshift buffer. .

  the following shall be done
  - a new MPD shall be published with a new publish time **MPD**@publishTime
  - an 'emsg' box shall be added to each segment of each Representation that contains an **InbandEventStream** element with
    - scheme_id_uri = "urn:mpeg:dash:event:2012"
    - @value either set to 1 or set to 3
    - the value of the **MPD**@publishTime of the previous MPD as the message_data

Annex C Preselections for Audio from ISO/IEC 23009-1:2014/Amd.4

Note: This will be removed once ISO/IEC 23009-1:2017 [2] is available. The section numbers replicate the numbers in ISO/IEC 23009-1.

5.3.11 Preselection

5.3.11.1 Overview

The concept of Preselection is primarily motivated for the purpose of Next Generation Audio (NGA) codecs in order to signal suitable combinations of audio elements that are offered in different Adaptation Sets. However, the Preselection concept is introduced in a generic manner such that it can be extended and be used also for other media types and codecs.

Each Preselection is associated to a bundle. A bundle is a set of media components which may be consumed jointly by a single decoder instance. Elements are addressable and separable components of a bundle and may be selected or deselected dynamically by the application, either directly or indirectly by the use of Preselections. Media components are mapped to Adaptation Sets by either a one-to-one mapping or by the inclusion of multiple media components in a single Adaptation Sets. Furthermore, Representations in one Adaptation Set may contain multiple media components that are multiplexed on elementary stream level or on file container level. In the multiplexing case each media component is mapped to a Media Content component as defined in 5.3.4. Each media component in the bundle is therefore identified and referenced by the `@id` of a Media Content component, or, if only a single media component is contained in the Adaptation Set, by the `@id` of an Adaptation Set.

Each bundle includes a main media component that contains the decoder specific information and bootstraps the decoder. The Adaptation Set that contains the main media component is referred to as main Adaptation Set. The main media component shall always be included in any Preselection that is associated to a bundle. In addition, each bundle may include one or multiple partial Adaptation Sets. Partial Adaptation Sets may only be processed in combination with the main Adaptation Set.

A Preselection defines a subset of media component in a bundle that are expected to be consumed jointly. A Preselection is identified by a unique tag towards the decoder. Multiple Preselection instances can refer to the same set of streams in a bundle. Only media components of the same bundle can contribute to the decoding and rendering of a Preselection.

In the case of next generation audio, a Preselection is a personalization option that is associated with one or more audio components from one plus additional parameters like gain, spatial location to produce a complete audio experience. A Preselection can be considered the NGA-equivalent of alternative audio tracks containing complete mixes using traditional audio codecs.

A bundle, Preselection, main media component, main Adaptation Set and partial Adaptation Sets may be defined by one of the two means:

— A preselection descriptor is defined in 5.3.11.2. Such a descriptor enables simple setups and backward compatibility, but may not be suitable for advanced use cases.

— A preselection element as defined in 5.3.11.3 and 5.3.11.4. The semantics of the Preselection element is provided in Table 17c in 5.3.11.3, the XML syntax is provided in 5.3.11.4.

The instantiation of the introduced concepts using both methods is provided in the following.

In both cases, if the Adaptation Set is not including the main Adaptation Set, then the Essential descriptor shall be used together with the `@schemeIdURI` as defined in 5.3.11.2.

5.3.11.2 Preselection Descriptor

A scheme is defined to be used with an Essential Descriptor as "`urn:mpeg:dash:preselection:2016`". The value of the Descriptor provides two fields, separated by a comma

— the tag of the Preselection

— the id of the contained content components of this Preselection list as white space separated list in processing order. The first id defines the main media component.

If the Adaptation Set contains the main media component, then the Supplemental descriptor may be used to describe contained Preselections in the Adaptation Set.

If the Adaptation Set does not contain the main media component then the Essential Descriptor shall be used.

The bundle is inherently defined by all media components that are included in all Preselections that include the same main media component. Preselections are defined by the metadata that is assigned to each of the media components that are included in the Preselection. Note that this signalling may be simple for basic use cases, but is expected to not provide a full coverage for all use cases. Therefore, the Preselection element is introduced in 5.3.11.3 to cover more advanced use cases.

5.3.11.3 Semantics of Preselection element

As an alternative to the Preselection descriptor, Preselections may also be defined through the Preselection element as provided in Table 17d. The selection of Preselections is based on the contained attributes and elements in the Preselection element.

Table 17d — Semantics of `PreSelection` element

Element or Attribute Name	Use	Description
`Preselection`		
`@id`	OD default=1	specifies the id of the Preselection. This shall be unique within one Period.
`@preselectionComponents`	M	specifies the ids of the contained Adaptation Sets or Content Components that belong to this Preselection as white space separated list in processing order. The first tag defines the main media component.
`@lang`	O	same semantics as in Table 5 for `@lang` attribute
`Accessibility`	0 … N	specifies information about accessibility scheme For more details, refer to 5.8.1 and 5.8.4.3.
`Role`	0 … N	specifies information on role annotation scheme For more details, refer to 5.8.1 and 5.8.4.2.
`Rating`	0 … N	specifies information on rating scheme. For more details, refer to 5.8.1 and 5.8.4.4.
`Viewpoint`	0 … N	specifies information on viewpoint annotation scheme. For more details, refer to 5.8.1 and 5.8.4.5.
`CommonAttributesElements`	-	specifies the common attributes and elements (attributes and elements from base type **RepresentationBaseType**). For details see 5.3.7.
Legend: For attributes: M=Mandatory, O=Optional, OD=Optional with Default Value, CM=Conditionally Mandatory. For elements: <minOccurs>..<maxOccurs> (N=unbounded) Elements are **`bold`**; attributes are non-bold and preceded with an @.		

5.3.11.4 XML Syntax for Preselection element

```
<!-- Preselection -->
<xs:complexType name="PreselectionType">
  <xs:complexContent>
    <xs:extension base="RepresentationBaseType">
      <xs:sequence>
        <xs:element name="Language" type="xs:language" minOccurs="0" maxOccurs="unbounded"/>
      </xs:sequence>
      <xs:attribute name="id" type="StringNoWhitespaceType" use="required"/>
      <xs:attribute name="preselectionComponents" type="StringVectorType" use="required"/>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

5.8.5.11 Audio Interactivity Descriptor

A scheme is defined to be used with an Essential Property or Supplemental Property Descriptor as "`urn:mpeg:dash:audio-interactivity:2016`".

This descriptor indicates if the associated audio content (Adaptation Set, Preselection or Representation) contains media components that are enabled for user interactivity through associated metadata. The descriptor is used e.g. to facilitate user interface (UI) resource management in the receiving client. Interactivity involves user interaction with elements, i.e. the user can modify dynamically for example the gain, spatial position or mute/unmute status of audio elements. Therefore, a UI is required to enable this kind of personalization during playback. A supplemental descriptor should be used if a UI is not mandatory to select and play the corresponding audio elements. An essential descriptor should be used if a UI is mandatory in order to play the corresponding audio elements. The `@value` attribute is owned by the codec in use. The detailed semantics of the descriptor are also owned by the codec in use.

Document History

Version	Additions	Date
0.01	Initial Draft	Nov 19, 2015
0.10	Initial Version shown to ATSC	Jan 19, 2016
0.30	Initial Version sent to ATSC 3.0 for review	Feb 11, 2016
0.35	Commented Version from ATSC 3.0 with initial resolutions	Mar 15, 2016
0.50	Intermediate Version after MPEG#115	June 1st, 2016
0.60	Version after Call July 8th	July 11th, 2016
0.65	Version shared with ATSC on July 12th	July 12th, 2016
0.80	Version sent to DASH-IF IOP for Community Review approval	August 1st, 2016
0.90	Version published for Community Review	August 3rd, 2016
0.93	Updated Version prior to call September 15	September 15th, 2016
0.95	Version created for ATSC final review	September 20th, 2016

0.97	Version created based on comments from ATSC for IOP approval	December 6th, 2016
0.98	Version created after IOP call on December 6th.	December 7th, 2016
0.99	Version sent for IPR Review	Dec 15th, 2016
0.991	Version sent for Board Approval	Jan 30th, 2017

1