DASH-IF IOP-8 V5.0.0 (2021-11)

DASH-IF Interoperability Points;
Part 8: Audio

DASH Industry Forum

3855 SW 153rd Dr.
Beaverton, OR 97003 - USA

Email : admin@dashif.org

Important notice

The present document can be downloaded from:
http://www.dashif.org/guidelines

Contents

Tables

Intellectual Property Rights

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 pe mitted 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 warranty whatsoever for such third-party material.

Note that technologies included in this document and for which no test and conformance material is provided, are only published as candidate technologies, and may be removed if no test material is provided before releasing a new version of this guidelines document. For the availability of test material, please check https://www.dashif.org.

Foreword

This Technical Specification (TS) has been produced by the DASH-IF Technical Working Group.

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in deliverables except when used in direct citation.

Executive Summary

Part 8 of DASH-IOP v5 describes the audio interoperability points for the DASH-IF ecosystem. Audio coding profiles, ISOBMFF packaging and MPD parameters are defined for these points of interoperability. Also defined is the application of audio Preselections as applied to object-based audio. Preselections may include audio components from one or more streams.

Introduction

The present document is Part 8 of a multipart set of documents, collectively called "DASH-IF Interoperability Points, V5.0" (IOP V5) and effectively extends IOP v5 Part 2[2] with the CMAF audio extensions. Version 5 of the DASH-IF IOP is defined for usage with the MPEG DASH specification (ISO/IEC 23009-1 [1]) and further constrained to deliver media formatted according to the MPEG CMAF Specification (ISO/IEC 2300-19 [16].

It is worth noting here that the DASH-IF IOP V4.3 is still available for DASH delivery of media tracks that are not constrained to CMAF.

The goal of the DASH-IF IOP specifications is facilitating a common interoperability of media and media services at key points of the delivery chain from the server to the client.

The following is a list of the parts of IOP V5 at the time of publication of the present document:

1. Overview, architecture and interfaces
2. Core principles and CMAF mapping
3. On-demand services
4. Live and low-latency live services
5. Ad insertion
6. Content protection
7. Video
8. Audio (this document)
9. Text
10. Events
11. Additional functionalities
12. Conformance and reference tools

1 Scope

The present document describes the audio interoperability points for the DASH-IF ecosystem. Audio coding profiles, ISOBMFF packaging and MPD parameters are defined for these points of interoperability. Also defined is the application of audio Preselections as applied to object-based audio. Preselections may include audio components from one or more streams.

The audio interoperability points include formats from MPEG, Dolby and DTS as follows:

- The MPEG profiles defined include MPEG-4 AAC profiles: AAC-LC, HE-AAC, HE-AACv2 and xHE-AAC. The MPEG-H Audio profiles include both LC and BL profiles for both single stream and multi-stream delivery.

- The Dolby profiles defined include Enhanced AC-3, Enhanced AC-3 with Joint Object Coding, and AC-4 for both single stream and multi-stream delivery.

- The DTS profiles defined include DTS core, DTS-HD core + extension, DTS-HD LBR, DTS-UHD Profile 2 and DTS-UHD Profile 3 for both single and multi-stream delivery.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

The following referenced documents are necessary for the application of the present document.

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

[2] DASH-IF IOP-2, V5.0.0 "DASH-IF Interoperability Points: Guidelines for Implementation; Part 2: Core principles and CMAF mapping"

[3] ISO/IEC 23091-3:2018, "Information technology - Coding-independent code points - Part 3: Audio"

[4] ETSI TS 102 366 v1.4.1, "Digital Audio Compression (AC-3, Enhanced AC-3) Standard"

[5] ETSI TS 103 190-2 V1.2.1, "Digital Audio Compression (AC-4) Standard; Part 2: Immersive and personalized audio"

[6] ETSI TS 102 114 v1.6.1, "DTS Coherent Acoustics; Core and Extensions with Additional Profiles"

[7] ETSI TS 103 491 v 1.2.1, "DTS-UHD Audio Format; Delivery of Channels, Objects and Ambisonic Sound Fields"

[8] ISO/IEC 23008-3:2019, "Information technology - High efficiency coding and media delivery in heterogeneous environments - Part 3: 3D Audio, Second Edition"

[9] ETSI TS 103 420 v1.2.1, "Backwards-compatible object audio carriage using Enhanced AC-3"

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

[11] ISO/IEC 23003-1:2007, " Information technology - MPEG audio technologies - Part 1: MPEG Surround"

[12] ISO/IEC TR 23009-3:2015, "Information technology - Dynamic adaptive streaming over HTTP (DASH) - Part 3: Implementation Guidelines"

[13] ISO/IEC 14496-3:2019, " Information technology -- Coding of audio-visual objects - Part 3: Audio"

[14] ISO/IEC 14496-14:2020, " Information technology - Coding of audio-visual objects - Part 14: MP4 file format"

[15] ISO/IEC 23003-4:2020, " Information technology - MPEG audio technologies - Part 4: Dynamic range control"

[16] ISO/IEC 23000-19, "Information technology - Multimedia application format (MPEG-A) - Part 19: Common media application format (CMAF) for segmented media"

[17] DASH-IF, "Guidelines for Implementation: DASH-IF Interoperability Points", Version 4.3

[18] DASH-IF, "Audio Amendment to Guidelines for Implementation: DASH-IF Interoperability Points", Version 1.1

2.2   Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, DASH-IF cannot guarantee their long-term validity.

The following referenced documents are not necessary for the application of the present document, but they assist the user regarding a particular subject area.

[i.1] ETSI TS 103 285 v.1.3.1: "Digital Video Broadcasting (DVB); MPEG-DASH Profile for Transport of ISO BMFF Based DVB Services over IP Based Networks"

3   Definition of terms, symbols and abbreviations

3.1   Terms

For the purposes of the present document, the following terms apply:

Audio Component - Media Component as defined in ISO/IEC 23009-1 [1]

Preselection – A subset of Audio Components as defined in ISO/IEC 23009-1 [1]

3.2   Symbols

For the purposes of the present document, the following symbols apply:

3.3   Abbreviations

For the purposes of the present document, the following abbreviations apply:

AAC	Advanced Audio Coding
AC-3	Dolby AC-3 audio coding system
AC-4	Dolby AC-4 audio coding system
DRM	Digital Rights Management
DTS-HD	Extended DTS audio coding system
DTS-UHD	DTS-UHD Audio coding system
E-AC-3	Enhanced AC-3
DVB	Digital Video Broadcasting
HE-AAC	High Efficiency AAC
HEVC	High-Efficiency Video Coding
HTTP	HyperText Transport Protocol
IOP	InterOperability Point
ISO	International Standards Organization
JOC	Joint Object Coding
MHA	Encapsulation of raw MPEG-H 3D Audio frames into ISO BMFF
MHAS	MPEG-H 3D Audio Stream
MHM	Encapsulation of MHAS packets into ISO BMFF
MPEG	Moving Pictures Experts Group

PCM	Pulse Code Modulation
PS	Parametric Stereo
SAP	Stream Access Point
SBR	Spectral Band Replication
URL	Universal Resource Location
USAC	Unified Speech and Audio Coding
xHE-AAC	Extended High Efficiency AAC: a specific USAC profile

4 CMAF Media Profiles

4.1 Media Profiles

The CMAF Audio Profiles described in the present specification and their properties are shown in Table 4-1. It should be noted that the present specification does not deprecate Version 4.3 of the DASH Industry Forum Guidelines for Implementation [17] nor the Audio Amendment [18] that accompanies it.

Table 4-1 is a comprehensive list of the audio media profiles that support CMAF packaging along with some important parameters. These codecs, along with any additional codecs added since this publication can be found in the online DASH-IF specification repository found here: https://dashif.org/codecs/audio/.

Table 4-1 CMAF Audio Profiles

Audio Profiles	compatibility brand	@codec	Media Profile	Profile Details (clause)
MPEG-4 AAC Profile	caac	mp4a.40.2	ISO/IEC 23000-19 [16] Annex A.3	5.1 (stereo) 5.2 (multichannel)
MPEG-4 HE-AAC Profile		mp4a.40.5		
MPEG-4 HE-AAC v2 Profile		mp4a.40.29		
Enhanced AC-3	ceac	ec-3	ETSI TS 102 366 [4] Annex J	5.3
AC-4 single stream	ca4s	ac-4.x.y.z	ETSI TS 103 190-2 [5] Annex H	
AC-4 main	ca4m			
DTS core	dts1	dtsc	ETSI TS 102 114 [6] Annex H	5.4
DTS-HD core + extension		dtsh		
DTS-HD LBR		dtse		
DTS-UHD Profile 2	dts2	dtsx	ETSI TS 103 491[7] Annex E	
DTS-UHD Profile 3	dts3	dtsy		
MPEG-H Audio LC Profile Level [1, 2, 3]	cmhs	mhm1.[0x0B, 0x0C, 0x0D]	ISO/IEC 23000-19 [16] Annex J	5.5
MPEG-H Audio BL Profile Level [1, 2, 3]		mhm1.[0x10, 0x11, 0x12]		
MPEG-H Audio LC Profile Level [1, 2, 3] multistream	cmhm	mhm2.[0x0B, 0x0C, 0x0D]		
MPEG-H Audio BL Profile Level [1, 2, 3] multistream		mhm2.[0x10, 0x11, 0x12]		
MPEG-D USAC (xHE-AAC)	casu	mp4a.40.42	ISO/IEC 23000-19 [16] Annex K	5.6

Table 4-2 lists the additional audio formats defined in Version 4.3 of the DASH-IF Guidelines [17]. Depending on the use case, these profiles may still be viable for CMAF adaptation sets. Such suitability is left to the implementer to determine. Please refer to the DASH-IF Audio Addendum [18] on specific guidance regarding the DASH specific issues.

Table 4-2 Additional Audio Profiles

Media Profiles	@codec	Profile Details
Dolby TrueHD	`mlpa`	Please refer to DASH-IF Audio Amendment [18]
DTS-HD Lossless (no core)	`dtsl`	
MPEG Surround	`mp4a.40.30`	
MPEG-H 3D Audio LC Profile Level 1	`mha[1, 2].0x0B`	
MPEG-H 3D Audio LC Profile Level 2	`mha[1, 2].0x0C`	
MPEG-H 3D Audio LC Profile Level 3	`mha[1, 2].0x0D`	
MPEG-H 3D Audio BL Profile Level 1	`mha[1, 2].0x10`	
MPEG-H 3D Audio BL Profile Level 2	`mha[1, 2].0x11`	
MPEG-H 3D Audio BL Profile Level 3	`mha[1, 2].0x12`	

4.2   CMAF Audio Adaptation Sets

In addition to the general provisions defined in 23009-1 [1] Adaptation Sets shall comply with the provisions of Table 4-3. Additional information on the usage if the listed Elements and Attributes may be provided in the codec specific clauses.

Table 4-3 Audio MPD elements and attributes

Element or Attribute	Use	Value
`@mimeType`	M	`audio/mp4`
`@codecs`	M	The value of the codecs attribute shall be created according to the syntax described in RFC 6381 [10]. If `@mimeType` is set to `"audio/mp4"`, then this element shall be present and set to one of the `@codecs` values defined in Table 4-1.
Role	M	shall use the schema `"urn:mpeg:dash:role:2011"` which is defined in 23009-1 [1]
AudioChannelConfiguration	0..N	the default schema for all audio codecs is `urn:mpeg:mpegB:cicp:ChannelConfiguration` as defined in ISO/IEC 23091-3[3]. Alternate namespaces are described in the codec specific clauses
`@audioSamplingRate`	O	use as defined in 23009-1 [1]
`@lang`	M, O	use as defined in 23009-1 [1], Optional for AdaptationSets referenced by Preselections
`@startWithSAP`	O	If this attribute is present it shall be set to 1
Accessibility	O	shall use the schema `"urn:mpeg:dash:role:2011"` which is defined in 23009-1 [1]

4.3   NGA and Preselections

4.3.1   Overview

Next Generation Audio (NGA) codecs include AC-4, defined in ETSI TS 103 190-2 [5], MPEG-H 3D Audio, defined in ISO/IEC 23008-3 [8], and DTS-UHD defined in ETSI TS 103 491 [7].

New concepts and signalling mechanisms have been introduced as part of ISO/IEC 23009-1 [1] in order to support the variety of new use-cases that Next Generation Audio enables. In particular, the concept of Preselections has been introduced as a method for defining and signalling audio personalization options.

4.3.2   Signalling of Preselections

The NGA codecs support the concept of component-based audio, i.e. the audio program can be constructed from a set of separate Audio Components (i.e. media content components containing audio). Examples of Audio Components are dialogues (in potentially different languages), ambient background sound, music or effects. The Audio Components may be delivered in a single stream or in multiple streams.

The concept of Preselections as defined in ISO/IEC 23009-1 [1], allows the codec to offer different combinations of those Audio Components, either for automatic selection based on user preferences or for manual selection by the user.

Two different methods are defined to signal Preselections in the MPD: The Preselection Descriptor and the Preselection Element. The Preselection descriptor, defined in ISO/IEC 23009-1 [1] clause 5.3.11.2, enables simple setups and backward compatibility but may not be suitable for advanced use cases. See codec-specific details below.

The Preselection Element is defined in ISO/IEC 23009-1 [1] clauses 5.3.11.3 and 5.3.11.4. The `@lang` attribute, the Role and Accessibility descriptors in the Preselection Element, as well as other parameters, such as a profile & level indication on the `@codecs` attribute are related only to that Preselection and not to the stream(s) referenced by the Preselection element.

Additional MPD Elements and Attributes are recommended for use when describing an NGA Preselection. These are listed in Table 4-4. Note that additional guidance in applying these parameters to specific NGA Audio codecs may be included in their respective clauses.

Table 4-4 Elements and attributes for NGA Preselections

Element or Attribute	Use	Usage or value
`Accessibility`	0..N	Indicate whether a Preselection or a Component has accessibility considerations.
`AudioChannelConfiguration`	0..N	the default schema for all audio codecs is `urn:mpeg:mpegB:cicp:ChannelConfiguration` as defined in ISO/IEC 23091-3[3]. Alternate namespaces are described in the codec specific clauses
`EssentialProperty`	M	An **AdaptationSet** referring to Auxiliary audio streams shall include a **Preselection EssentialProperty** descriptor as specified in MPEG-DASH
`Label`	0..N	If there are multiple Preselections, this label should be set by the content author.
`Role`	O	DASH role scheme, `urn:mpeg:dash:role:2011`, may be used to indicate a key attribute of the Preselection.
`SupplementalProperty`	O	If an **AdaptationSet** referring to the Main Audio stream is referenced by one or more **Preselection** elements, the **AdaptationSet** should include a Preselection **SupplementalProperty** descriptor as specified in MPEG-DASH
`@lang`	M	as defined in 23009-1 [1] when dialog is present
`@preselectionComponents`	O	as defined in 23009-1 [1]
`@selectionPriority`	O	as defined in 23009-1 [1]
`@tag`	O	as defined in 23009-1 [1]

For NGA Preselections, the type of audio experience should be signalled by utilizing the descriptors as provided in Table 4-5. Note that additional guidance in applying these parameters to specific NGA Audio codecs may be included in their respective clauses.

Table 4-5 NGA Preselection experience signalling

Experience Description	Role Descriptor @value	Accessibility descriptor @value
Normal Audio	main	**Accessibility** is not present in this use case
Secondary Language (dubbed audio)	dub	**Accessibility** is not present in this use case
Audio Description (Visual Description Service)	alternate	description
Enhanced Intelligibility	alternate	enhanced-audio-intelligibility
Emergency Information (for the hard of viewing)	alternate	emergency

5   Media Profile Specific Information

5.1   MPEG High Efficiency AAC, Stereo

5.1.1   General

To support interoperability, presentations conforming to the DASH-IF Interoperability Guidelines that contain audio shall contain at least one basic stereo audio adaptation set.

The codec for basic stereo audio support shall conform to MPEG-4 High Efficiency AAC v2 Profile, level 2 [17].

For all HE-AAC and HE-AACv2 bitstreams, explicit backwards compatible signalling should be used to indicate the use of the SBR and PS coding tools.

> Note: Explicit backwards compatible signalling is mandatory in DVB DASH [i.1] and is necessary to remain compatible.

5.1.2   DASH-specific aspects for HE-AACv2 audio Level 2

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

- The content should be prepared according to ISO/IEC TR 23009-3 [12] to make sure each (Sub)Segment starts with a SAP of type 1.

- The signalling of MPEG-4 High Efficiency AAC v2 for the codecs parameters is according to IETF RFC6381 [10] and is documented in Table 4-1. Table 5-1 provides information on the ISO BMFF encapsulation.

- For content with SBR, i.e. @codecs = mp4a.40.5 or @codecs=mp4a.40.29, @audioSamplingRate signals the resulting sampling rate after SBR is applied, e.g. 48 kHz even if the AAC-LC core operates at 24 kHz. For content with PS, i.e. @codecs=mp4a.40.29, **AudioChannelConfiguration** signals the resulting channel configuration after PS is applied, e.g. stereo even if the AAC-LC core operates at mono.

Table 5-1 MPEG 4 AAC stereo profiles and ISO BMFF encapsulation

Codec	Codec Defined	ISO BMFF Encapsulation
MPEG-4 AAC Profile		
MPEG-4 HE-AAC Profile	ISO/IEC 14496-3[13]	ISO/IEC 14496-14 [12]
MPEG-4 HE-AAC v2 Profile		

> Note: Since both, HE-AAC and HE-AACv2 are based on AAC-LC, for the above-mentioned "Codec Parameter" the following is implied:
>
> - mp4a.40.5 = mp4a.40.2 + mp4a.40.5
> - mp4a.40.29 = mp4a.40.2 + mp4a.40.5 + mp4a.40.29

5.1.3   AAC Audio Metadata

5.1.3.1   General

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

5.1.3.2   ISO/IEC 23009-1 audio data

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

• the @audioSamplingRate attribute for signalling the sampling rate of the audio media component type in section 5.3.7 of ISO/IEC 23009-1 [1]

• the **AudioChannelConfiguration** element for signalling audio channel configuration of the audio media component type.in section 5.3.7 of ISO/IEC 23009-1 [1]. For this element the scheme and values defined in ISO/IEC 23091-3[3] for the **ChannelConfiguration** should be used.

5.2 MPEG-4 High Efficiency AAC Profile v2, Multichannel

5.2.1 Overview

Support for multichannel content is available in the HE-AACv2 Profile, starting with level 4 for 5.1 and level 6 for 7.1. All MPEG-4 HE-AAC multichannel profiles are fully compatible with the DASH-AVC/264 baseline interoperability point for stereo audio, i.e. all multichannel decoders can decode DASH-IF IOPS stereo content.

5.2.2 DASH-specific issues

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

• The content shall be prepared according to ISO/IEC TR 23009-3 [12] to make sure each (sub)segment starts with a SAP of type 1.

• Signalling of profile levels is not supported in RFC 6381 but the channel configuration shall be signalled by means of the **ChannelConfiguration** element in the MPD.

• The signalling of MPEG-4 High Efficiency AAC v2 for the codecs parameters is according to RFC6381 [11] and is documented in Table 4-1. Table 5-2 provides information on the ISO BMFF encapsulation.

• For all HE-AAC bitstreams, explicit backward-compatible signalling of SBR shall be used.

• The content should be prepared incorporating loudness and dynamic range information into the bitstream also considering DRC Presentation Mode in ISO/IEC 14496-3 [13].

• Decoders shall support decoding of loudness and dynamic range related information, i.e. dynamic_range_info() and MPEG4_ancillary_data() in the bitstream.

Table 5-2 MPEG-4 AAC multichannel profiles and ISO BMFF encapsulation

Codec	Codec Defined	ISO BMFF Encapsulation
MPEG-4 AAC Profile [13]		
MPEG-4 HE-AAC Profile [13]	ISO/IEC 14496-3[13]	ISO/IEC 14496-14 [14]
MPEG-4 HE-AAC v2 Profile [13]		

Note: Since both, HE-AAC and HE-AACv2 are based on AAC-LC, for the above mentioned "Codec Parameters" the following is implied:
mp4a.40.5 = AOT 2 + AOT 5

5.3 Dolby Multichannel Technologies

5.3.1 Overview

The considered technologies from Dolby for advanced audio support are:

• Enhanced AC-3 (Dolby Digital Plus), ETSI TS 102 366 [4]

• AC-4, ETSI TS 103 190-2 [5]

The references to the codec definition and file encapsulation are listed in Table 5-3.

Table 5-3 Dolby profiles and ISO BMFF encapsulation

Codec	Codec Defined	ISO BMFF Encapsulation
Enhanced AC-3	ETSI TS 102 366 [4]	ETSI TS 102 366 Annex F [4]
AC-4	ETSI TS 103 190-2 [5]	ETSI TS 103 190-2 Annex E [5]

5.3.2   DASH-specific issues

In the context of DASH, the following applies:

- The signalling of the different audio codecs for the codecs parameters is documented in ETSI TS 102 366 7[4] and ETSI TS 103 190-2 which also provides information on ISO BMFF encapsulation.

- For E-AC-3 the Audio Channel Configuration may use any of the following:

  - `urn:mpeg:mpegB:cicp:ChannelConfiguration` as defined by **ChannelConfiguration** in ISO/IEC 23091-3 [3]

  - `tag:dolby.com,2014:dash:audio_channel_configuration:2011` as defined in TS 102 366 [4] clause I.1.2.1.

- For AC-4, the Audio Channel Configuration may use any of the following used as further specified in TS 103 190-2 [5] clause G.3.1:

  - `urn:mpeg:mpegB:cicp:ChannelConfiguration` as defined by **ChannelConfiguration** in ISO/IEC 23091-3 [3]

  - `tag:dolby.com,2015:dash:audio_channel_configuration:2015` as defined in TS 103 190-2 [5] clause G.3.1

- For E-AC-3, the presence of JOC enhanced AC-3 extension information is signalled using the following supplemental descriptors as specified in ETSI TS 103 420 [9] clause D.2

  - `tag:dolby.com,2018:dash:EC3_ExtensionType:2018`

  - `tag:dolby.com,2018:dash:EC3_ExtensionComplexityIndex:2018`

5.3.3   Dolby Enhanced AC-3 specific issues

Dolby Enhanced AC-3 tracks shall be constrained according to the CMAF specific constraints as provided in ETSI TS 102 366 [4] Annex J.

If the backward-compatible object audio carriage using Enhanced AC-3 according to ETSI TS 103 420 [9] is used, these tracks shall be constrained according to the CMAF specific requirements as provided in ETSI TS 103 420 [9] Annex E. Additionally, a compatibility brand of `'ceao'` should be used.

5.3.4   Dolby AC-4 specific issues

5.3.4.1   AC-4 CMAF Media Profiles

Constraints on CMAF tracks carrying Dolby AC-4 are specified in ETSI TS 103 190-2 [5] Annex H.

CMAF audio adaptation sets carrying Dolby AC-4 shall comply with the AC-4 single-stream Media Profile or, if advanced multi-stream presentations are used, the AC-4 main Media Profile, as further constrained in the following sections.

5.3.4.2   AC-4 Elementary stream constraints for the use of DASH

For the use of DASH with the ISOBMFF, the following additional constraints shall apply when packaging AC-4 audio into DASH Representations:

- The following parameters shall remain constant within each Representation:

  - frame_rate_index

  - fs_index and, if present, b_sf_multiplier and sf_multiplier

  - presentation_config

  - channel_mode

- content_classifier

- - b_language_indicator and, if present, the language_tag_bytes or the concatenated version of data from all language_tag_chunk.

• AC-4 access units should be encoded with the same frame rate as the associated video frame rate.

• AC-4 access units should be encoded temporally aligned with the video access units from the corresponding video to ensure continuous alignment of video and audio access units in order to utilize the features of A/V alignment.

• AC-4 I-Frames should be placed temporally aligned with the I-Frames of the video to enable synchronous switching.

5.3.4.3   ISO Based Media File Format Packaging

ISO Base Media File Format Packaging Rules for Dolby AC-4 are described in ETSI TS 103 190-2 [5] Annex E.

5.3.4.4   Random Access Points, Stream Access Points and CMAF Switching Sets

Random Access and Stream Access Points for AC-4 Audio are described in ETSI TS 103 190-2 [5] Annex H.2, referring to the AC-4 sync sample requirements in ETSI TS 103 190-2 [5] Annex E.3.

If a CMAF switching set contains more than one CMAF track, the requirements given in ETSI TS 103 190-2 [5] Annex H.3 shall apply.

5.3.4.5   Dynamic Range Control and Loudness

The requirements from ETSI TS 103 190-2 [5] Annex H.1.5 shall be applied with parameters set in accordance with the applicable regional loudness regulation.

5.3.5   AC-4 MPD Element and Attribute Settings

Table 5-4 summarizes the mapping of relevant MPD elements and attributes to AC-4 Audio.

Most of the elements can be derived from the AC-4 `ac4_dsi_v1` structure as described in ETSI TS 103 190-2 [5] Annex E.

Note that usage of the elements and attributes listed below depends on the employed DASH profile.

Table 5-4 AC-4 element and attribute settings

Element or Attribute	Usage or value
`@codecs`	For AC-4 the value of the codecs attribute shall be created according to the syntax described in RFC 6381 [10]. 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 referenced presentation in the **ac4_dsi_v1** structure. The mdcompat parameter, indicating the compatibility level for the referenced presentation. Example: `ac-4.02.01.03,` signalling AC_4 audio with `bitstream_version=2`, `presentation_version=1` and `md_compat=3`. In case of **AdaptationsSets**, the term referenced presentation shall refer to that presentation with the lowest mdcompat value amongst all presentations with `presentation_version` < 2 and that are fully contained in this **AdaptationSet**.
Preselection@tag	This field shall correspond to the value of the **presentation_id** in the **ac4_presentation_v1_dsi** associated with the referenced AC-4 presentation.

Element or Attribute	Usage or value

`AudioChannelConfiguration`	For AC-4 the Audio Channel Configuration descriptor shall use one of the following schemes
	    • `urn:mpeg:mpegB:cicp:ChannelConfiguration` as defined by **`ChannelConfiguration`** in ISO/IEC 23091-3 [3]
	
	    • `tag:dolby.com,2015:dash:audio_channel_configuration:2015` as defined in TS 103 190-2 Annex G.3.1 [5].
	
	`urn:mpeg:mpegB:cicp:ChannelConfiguration` is the preferred scheme.

`@audioSamplingRate`	The value shall be set to the sampling frequency as specified in TS 103 190-2 [5] Annex G.2.6. Example: For `fs_index = 1` and `dsi_fs_multiplier = 0`, the value is 48000.
`@lang`	The language indicated by the lang attribute should correspond to that language signalled in the `language_tag_bytes`, which is tagged as "`dialogue`" or "`complete main`" in the corresponding **`content_classifier`**. NOTE: The `language_tag_bytes` are contained in the **`ac4_substream_group_dsi`** structure, within the **`ac4_dsi_v1 structure`**. For **`AdaptationSets`** that are referenced by Preselection elements, the lang attribute should not be present on the **`AdaptationSet`** element. In cases where it is present, the indicated language should correspond to that presentation with the lowest `mdcompat` value amongst all presentations with `presentation_version` < 2 and that are fully contained in this **`AdaptationSet`**.
`Role`	The **`Role`** for a **`Preselection`** 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`	In case one or more audio elements contained in a Presentation indicate a content type visually impaired, an Accessibility descriptor shall indicate description according to the Role scheme defined in ISO/IEC 23009-1 [1]. If one or more audio elements contained in a Presentation indicate a content type other than music and effects, an Accessibility descriptor indicating enhanced-audio-intelligibility according to the Role scheme defined in ISO/IEC 23009-1 [1] may be used. In case one or more audio elements contained in a Presentation indicate Associated service: emergency (E), an Accessibility descriptor may indicate emergency according to the Role scheme defined in ISO/IEC 23009-1 [1].
`SupplementalProperty`	If the content of a Presentation has been tailored for consumption via headphones, an Immersive Audio for Headphones SupplementalProperty descriptor should be used as specified in ETSI TS 103 190-2 [5] clause G.2.12.1. An audio framerate SupplementalProperty descriptor should be used as specified in ETSI TS 103 190-2 [5] clause G.2.12.2.

5.4   DTS Audio Technologies

5.4.1   Overview

The considered technologies from Xperi (DTS) for advanced audio support are:

- DTS-HD, TS 102 114 [6]

- DTS-UHD, TS 103 491 [7]

5.4.2   DASH specific issues

Table 5-5 provides a list of the relevant codecs and their reference for ISOBMFF encapsulation.

Table 5-5 DTS profiles and ISO BMFF encapsulation

Codec	Codec Defined	ISO BMFF Encapsulation
DTS core	ETSI TS 102 114 [6]	ETSI TS 102 114 [6] Annex E
DTS-HD core + extension		
DTS-HD LBR extension		
DTS-UHD Profile 2	ETSI TS 103 491 [7]	ETSI TS 103 491 [7] Annex B
DTS-UHD Profile 3		

A summary of MPD elements and attributes specific to DTS-HD are found in Table 5-6.

Table 5-6 DTS-HD element and attribute settings

Element or Attribute	Description
`@codecs`	This attribute specifies the codecs used to encode all representations within the adaptation set and the value shall be one of "dtsc", "dtsh", or "dtse" corresponding to the composition of the elementary stream. This value shall match the AudioSampleEntry
AudioChannelConfiguration	`urn:mpeg:mpegB:cicp:ChannelConfiguration` as defined by **ChannelConfiguration** in ISO/IEC 23091-3 [3] is the preferred schema for **AudioChannelConfiguration**. `tag:dts.com,2014:dash:audio_channel_configuration:2012` where the `@value` is set to the number of output channels, as defined in ETSI TS 102 114 [6] Annex G can also be used.

Additional requirements for delivering DTS-HD using DASH are discussed in TS 102 114 [6] Annex G.

Additional information for MPD elements and attributes specific to DTS-UHD are found in Table 5-7.

Table 5-7 DTS-UHD element and attribute settings

Element or Attribute	Description
`@codecs`	For DTS-UHD, @codecs is the associated 4cc with no additional suffix. if DecoderProfile = 2 then `@codecs = 'dtsx'` if DecoderProfile = 3 then `@codecs = 'dtsy'` This value shall match that used for the AudioSampleEntry
preselection`@tag`	DTS-UHD bitstreams carry two levels of organization that can be signalled with the `preselectionComponents` value. When a Component is identified by an Audio Presentation Index, the tag value shall be formatted as "Px" where x is the presentation ID within the elementary stream. When discrete audio objects are being selected from the elementary stream, the `@tag` parameter is a space delimited string of the desired Object IDs. (See TS 103 491 [7])
AudioChannelConfiguration	When RepresentationType (defined in DTS-UHD Annex B) is set to 0, 1 or 2, then DTS-UHD shall use one of the following schema to describe the channel layout: `urn:mpeg:mpegB:cicp:ChannelConfiguration` as defined by **ChannelConfiguration** in ISO/IEC 23091-3 [3]. `tag:dts.com,2018:uhd:audio_channel_configuration` where the value is according to TS 103 491 Annex D [7]
Accessibility	A DTS-UHD elementary stream carries accessibility information in the object property `m_ucAssociatedAssetType` as described in ETSI TS 103 491 [7].

Additional requirements for delivering DTS-UHD using DASH are discussed in TS 103 491 Annex D [7].

Other mandatory and recommended elements and attributes for delivery of audio tracks are according to this specification and ISO/IEC 23009-1 [1].

5.4.3   DTS-UHD specific issues

5.4.3.1   Sink frames and non-sync frames

DTS-UHD elementary streams are organized in groups of frames (GoF) that begin with a sync frame. Following the sync frame are non-sync frames. Playback can on start on a sync frame, therefore any random-access point requires sync frame alignment.

5.4.3.2   DTS-UHD Profiles

Two DTS-UHD profiles are currently defined, described in TS 103 491 [7] Annex F.

5.4.3.3   Multi-stream support

DTS-UHD can be used with the Preselection Descriptor to facilitate multi-stream playback. Additional considerations for multi-stream are discussed in TS 103 491 [7] Annex G.

5.5   MPEG-H 3D Audio

5.5.1   Overview

MPEG-H 3D Audio is defined in ISO/IEC 23008-3 [8] and is a Next Generation Audio (NGA) codec. MPEG-H 3D Audio encoded content shall comply with Level 1, 2 or 3 of the MPEG-H Low Complexity (LC) Profile or the MPEG-H Baseline (BL) profile as defined in ISO/IEC 23008-3 [8] clause 4.8.

MPEG-H 3D Audio content might comply with both Profiles (BL and LC) as defined in Annex P1 of ISO/IEC 23008-3 [7]. If a bitstream is compatible to both profiles, the service provider can choose the Codecs Parameter based on the service requirements. The same audio content might be referenced from different Adaptation Sets with different Codec Parameters.

The clauses to follow clarify DASH specific requirements for MPEG-H 3D Audio, such as:

- Codec parameters settings and signalling

- Usage of MPD elements and attributes

- File format encapsulation modes and requirements

- Loudness and Dynamic Range Control requirements

5.5.2   DASH-specific Issues

The carriage of MPEG-H 3D Audio in the ISO BMFF is specified in ISO/IEC 23008-3 [8] clause 20. Storage of MHAS streams is specified in ISO/IEC 23008-3 [8] clause 20.6. The MPEG-H Audio Stream (MHAS) format is defined in ISO/IEC 23008-3 [8] clause 14. Clause 5.5.4 provides more information on this encapsulation.

5.5.3   Element and Attribute Settings

Table 5-8 summarizes the mapping of relevant MPD elements and attributes to MPEG-H Audio.

Table 5-8 MPEG-H Audio elements and attributes settings

Element Name or Attribute	Description
@codecs	The signalling of the codecs parameters is according to RFC6381 [10] and ISO/IEC 23008-3 [8] clause 21. The value consists of the following two parts separated by a dot: • the sample entry 4CC code ('mhm1','mhm2') • '0x' followed by the hex value of the profile-level-id, as defined in in ISO/IEC 23008-3 [8] See Table 5-9 for more details.

`AdaptationSet`@tag	This field lists the `mae_groupIDs` as defined in ISO/IEC 23008-3 [8] 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 [8] that refers to a Preset in the scope of MPEG-H Audio.
`ContentComponent`@tag	This field indicates the `mae_groupID` as defined in ISO/IEC 23008-3 [8] which contains the Media Content Component.
`AudioChannelConfiguration`	For MPEG-H Audio, the Audio Channel Configuration descriptor shall use the scheme URI `"urn:mpeg:mpegB:cicp:ChannelConfiguration"`. The value shall be taken from the **ChannelConfiguration** table as defined in ISO/IEC 23091-3 [3]. Valid numbers for value are 0-7, 9-12, 14-17 or 19. The value 0 should only be used if the exact Audio Channel Configuration cannot be determined, e.g. a live service with in-band configuration changes within a period, or for object only audio scenes.
`@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 [8].
`@lang`	The language indicated should correspond to the information conveyed in `mae_contentLanguage` of the default dialog element. The default dialog corresponds to the Group (`mae_groupDefinition()`) which is marked as default in `mae_switchGroupDefaultGroupID` and is tagged in `mae_contentKind` as dialogue. This information is carried in the `mae_audioSceneInfo()` of the MPEG-H Audio stream as defined in ISO/IEC 23008-3 [8]. The language of a Preselection should correspond to the information conveyed in `mae_contentLanguage` of the selected dialog. The selected dialog corresponds to the Group (`mae_groupDefinition()`) which is marked as on (`mae_groupPresetConditionOnOff == 1`) for the given **Preselection**@tag and is tagged in `mae_contentKind` as dialogue. This information is carried in the `mae_audioSceneInfo()` of the MPEG-H Audio stream as defined in ISO/IEC 23008-3 [8].
`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 `"description"` according to the Role scheme defined in ISO/IEC 23009-1 [1]. 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 [1] 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. If the `mae_contentKind` value of at least one Audio Element is set to '12' ("emergency"), an Accessibility descriptor may indicate

	"emergency" according to the Role scheme defined in ISO/IEC 23009-1 [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 [8].

If present, the Preselection Tag value of the Preselection Descriptor, shall be set to the corresponding mae_groupPresetID as defined in ISO/IEC 23008-3 [8].

5.5.4   MHM Encapsulation

Storage of MHAS into ISO BMFF shall be according to ISO/IEC 23008-3 [8] clause 20 with the additional constraints described below.

The sample entry 'mhm1' shall be used for encapsulation of MHAS packets into ISO BMFF files, according to ISO/IEC 23008-3 [8] clause 20.6. The sample entry 'mhm2' shall be used in cases of multi-stream delivery, i.e. the MPEG H Audio Scene is split into two or more streams for delivery as described in ISO/IEC 23008-3 [8] clause 14.6.

All MHAS packet types defined in ISO/IEC 23008-3 [8] clause 14, may be present in the stream, except the following packet types that shall not be present:

- PACTYP_CRC16

- PACTYP_CRC32

- PACTYP_GLOBAL_CRC16

- PACTYP_GLOBAL_CRC32

If Audio Scene Information, defined in ISO/IEC 23008-3 [8] clause 15 is present, it shall be always encapsulated in an MHAS packet of type PACTYP_AUDIOSCENEINFO. Audio Scene Information shall not be included in the mpegh3daConfig() structure carried in the MHAS packet of type PACTYP_MPEGH3DACFG.

The MHAConfigurationBox(), defined in ISO/IEC 23008-3 [8] clause 20.6, is optional for MHM.

> Note: In case of in-band configuration changes within a period, the MHAConfigurationBox() should not be present in the corresponding sample entry.

The content is expected to be prepared according to the ISO/IEC TR 23009-3 [12] to make sure each (sub-)segment starts with a Stream Access Points (SAP) of type 1 (i.e. a sync sample). For MHM encapsulation in particular, a sync sample shall consist of the following MHAS packets, in the following order:

- PACTYP_MPEGH3DACFG

- PACTYP_AUDIOSCENEINFO (if Audio Scene Information is present)

- PACTYP_BUFFERINFO

- PACTYP_MPEGH3DAFRAME

MPEG-H Audio sync samples contain Immediate Playout Frames (IPFs), as specified in ISO/IEC 23008-3 [8] clause 20.2, thus the audio data encapsulated in the MHAS packet PACTYP_MPEGH3DAFRAME shall contain the AudioPreRoll() syntax element, as defined in ISO/IEC 23008-3 [8] clause 5.5.6, and shall follow the requirements for stream access points as defined in ISO/IEC 23008-3 [8] clause 5.7.

The audio configuration is delivered as part of the MHAS packet PACTYP_MPEGH3DACFG and, therefore, the AudioPreRoll() structure carried in the MHAS packet PACTYP_MPEGH3DAFRAME shall not contain the

`Config()` structure, i.e. the `configLen` field of the `AudioPreRoll()` shall be 0. All rules defined in ISO/IEC 23008-3 [8] clause 20.6.1 regarding sync samples shall also apply.

The `MHASPacketLabel` shall have different values for all representations that comprise an experience.

Additional MHAS packets may be present between the MHAS packets listed above or after the MHAS packet `PACTYP_MPEGH3DAFRAME`, with one exception: when present, the `PACTYP_AUDIOSCENEINFO` packet shall directly follow the `PACTYP_MPEGH3DACFG` packet, as defined in ISO/IEC 23008-3 [8] clause 14.4.

Table 5-9 MPEG-H profiles and ISO BMFF encapsulation

Codec	Codec Defined	ISO BMFF Encapsulation
MPEG-H 3D Audio LC Profile Level [1, 2, 3]	ISO/IEC 23008-3 [8]	ISO/IEC 23008-3 [8]
MPEG-H 3D Audio BL Profile Level [1, 2, 3]		

5.5.5   MHM Configuration Change Constraints

A configuration change takes place in an audio stream when the content setup or the Audio Scene Information changes (e.g., when changes occur in the channel layout, the number of objects etc.) and, therefore, new `PACTYP_MPEGH3DACFG` and `PACTYP_AUDIOSCENEINFO` packets are required upon such occurrences. A configuration change usually happens at program boundaries, but it may also occur within a program.

The following constraints apply:

- At each configuration change, the `MHASPacketLabel` shall be changed to a different value from the `MHASPacketLabel` in use before the configuration change occurred. A configuration change may happen at the beginning of a new ISO BMFF file or at any position within the file. In the latter case, the File Format sample that contains a configuration change shall be encoded as a sync sample (i.e. an IPF) as defined above. Note, that also after a configuration change the `MHASPacketLabel` needs to have different values for all representations comprising an experience.

- A sync sample that contains a configuration change and the last sample before such a sync sample may contain a truncation message (i.e., a `PACTYP_AUDIOTRUNCATION` packet in the MHAS stream) as defined in ISO/IEC 23008-3 clause 14.4 [8]. If MHAS packets of type `PACTYP_AUDIOTRUNCATION` are present, they shall be used as described in ISO/IEC 23008-3 [8] clause 14.4.

5.5.6   MPEG-H Audio Multi-Stream Constraints

The multi-stream-enabled MPEG-H Audio System is capable of handling Audio Components delivered in several different elementary streams (e.g., the main MHAS stream containing one complete audio main, and one or more auxiliary MHAS streams, containing different languages and audio description). The MPEG-H Audio Metadata information (MAE) allows the MPEG-H Audio Decoder to correctly decode several MHAS streams.

The following constraints apply when using the sample entry `'mhm2'`:

- One MHAS stream shall be the main stream (Main Adaptation Set), i.e., in exactly one MHAS stream the Audio Scene Information shall have the `mae_isMainStream` field set to 1. In all other MHAS streams the `mae_isMainStream` shall be set to 0. All Representations of an MHAS stream with `mae_isMainStream` set to 1 form the Main Adaptation Set.

- In each auxiliary MHAS stream (i.e., streams with `mae_isMainStream` field set to 0) the `mae_bsMetaDataElementIDoffset` field in the Audio Scene Information shall be set to the index of the first metadata element in the auxiliary MHAS stream minus one. Each auxiliary Stream (and all its representations) form(s) a Partial Adaptation Set.

- For the main and the auxiliary MHAS stream(s), the `MHASPacketLabel` shall be set according to ISO/IEC 23008-3 [8] clause 14.6.

- All MHAS elementary streams that carry Audio Components of one complete experience shall be time aligned.

- In each auxiliary MHAS elementary stream (i.e., streams with `mae_isMainStream` field set to 0), IPFs shall be aligned to the IPFs present in the main stream (i.e., the stream with `mae_isMainStream` field set to 1).

5.5.7   Loudness and Dynamic Range Control

Loudness metadata shall be embedded within the `mpegh3daLoudnessInfoSet()` structure as defined in ISO/IEC 23008-3 clause 6.3 [8]. Such loudness metadata shall include at least the loudness of the content rendered to the default rendering layout as indicated by the `referenceLayout` field (see ISO/IEC 23008-3 [8] clause 5.3.2). More precisely, the `mpegh3daLoudnessInfoSet()` structure shall include at least one `loudnessInfo()` structure with `loudnessInfoType` set to 0, whose `drcSetId` and `downmixId` fields are set to 0 and which includes at least one `methodValue` field with `methodDefinition` set to 1 or 2 (see ISO/IEC 23008-3 [8] clause 6.3.1 and ISO/IEC 23003-4 [15] clause 7.3). The indicated loudness value shall be measured according to applicable regional loudness regulations.

DRC metadata shall be embedded in the `mpegh3daUniDrcConfig()` and `uniDrcGain()` structures as defined in ISO/IEC 23008-3 [8] clause 6.3. For each included DRC set the `drcSetTargetLoudnessPresent` field as defined in ISO/IEC 23003-4 clause 7 [15] shall be set to 1. The `bsDrcSetTargetLoudnessValueUpper` and `bsDrcSetTargetLoudnessValueLower` fields shall be configured to continuously cover the range of target loudness levels between -31 dB and 0 dB. The embedded DRC metadata should allow for a decoder output loudness of at least -16 LKFS.

Loudness compensation information (`mae_LoudnessCompensationData()`), as defined in ISO/IEC 23008-3 [8] clause 15.5, shall be present in the Audio Scene Information if the `mae_allowGainInteractivity` field (according to ISO/IEC 23008-3 [8] clause 15.3) is set to 1 for at least one group of audio elements.

5.6   MPEG-D Unified Speech and Audio Coding

5.6.1   Overview

MPEG-D Unified Speech and Audio Coding (USAC) has been designed to provide consistently high audio quality with a variety of content that comprises a mixture of audio and speech signals. Using such a codec in a DASH streaming environment enables adaptive switching capability from 12 kbps stereo up to transparency.

ISO/IEC 23000-19 [16] defines a media profile for MPEG-D USAC that is suitable for streaming applications and therefore can be referenced here.

5.6.2   DASH-specific issues

In the context of DASH-IF IOPs, the following applies to the xHE-AAC profile:

- Content representations encoded with MPEG-D USAC shall comply with the Extended High Efficiency AAC (xHE-AAC) CMAF media profile `'cxha'`, as defined in ISO/IEC 23000-19 [16], providing support up to 5.1 multichannel coding.

- All representations of an adaptation set shall conform to the CMAF switching set constraints.

- The codec signalling is according to RFC6381 [10] and documented in Table 4-1.

- The profiles mime sub- parameter of the `@mimetype` attribute should include `'cxha'`.

- If the `ChannelConfiguration` parameter is present in the Movie header, then the identical channel configuration shall be signalled by means of the **`AudioChannelConfiguration`** element in the MPD, according to the values specified in ISO/IEC 23000-19 [16] Table K.2.

- The CMAF xHE-AAC media profile defined in ISO/IEC 23000-19 [16] requires each CMAF Fragment to start with an SAP of type 1.

Table 5-10 MPEG-D USAC profile and ISO BMFF encapsulation

Codec	Codec Defined	ISO BMFF Encapsulation
MPEG-D USAC	ISO/IEC 23008-3 [8]	ISO/IEC 23000-19 [16]

Annex A (Informative): Legacy DASH-IF interoperability points for audio

Extension	Identifier
DASH-IF multichannel audio extension with Enhanced AC-3	`http://dashif.org/guidelines/dashif#ec-3`
DASH-IF multichannel audio extension with Dolby TrueHD	`http://dashif.org/guidelines/dashif#mlpa`
DASH-IF multichannel audio extension with AC-4	`http://dashif.org/guidelines/dashif#ac-4`
DASH-IF multichannel audio extension with DTS Digital Surround	`http://dashif.org/guidelines/dashif#dtsc`
DASH-IF multichannel audio extension with DTS-HD High Resolution Audio and DTS-HD Master Audio	`http://dashif.org/guidelines/dashif#dtsh`
DASH-IF multichannel audio extension with DTS Express	`http://dashif.org/guidelines/dashif#dtse`
DASH-IF multichannel audio extension with DTS-HD Lossless (no core)	`http://dashif.org/guidelines/dashif#dtsl`
DASH-IF multichannel audio extension with DTS-UHD Profile 2	`http://dashif.org/guidelines/dashif#dtsx`
DASH-IF multichannel audio extension with DTS-UHD Profile 3	`http://dashif.org/guidelines/dashif#dtsy`
DASH-IF multichannel audio extension with MPEG Surround	`http://dashif.org/guidelines/dashif#mps`
DASH-IF multichannel audio extension with HE-AACv2 level 4	`http://dashif.org/guidelines/dashif#heaac-mc51`
DASH-IF multichannel audio extension with HE-AACv2 level 6	`http://dashif.org/guidelines/dashif#heaac-mc71`
DASH-IF multichannel audio extension with MPEG-H 3D Audio	`http://dashif.org/guidelines/dashif#mpeg-h-3da`
DASH-IF audio extension with USAC	`http://dashif.org/guidelines/dashif#cxha`

Annex B (Informative): Change History

Date	Version	Information about changes
2021-12	5.0.0	Initial publication