**CHANGE REQUEST**

<table>
<thead>
<tr>
<th>DASH-IF IOP</th>
<th>CR</th>
<th>rev</th>
<th>Current version:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>xxxx</td>
<td>0</td>
<td>4.3</td>
</tr>
</tbody>
</table>

**Status:**
- Draft
- Internal Review
- Community Review
- Editor's Proposal
- Agreed

**Title:**
Alignment of the MPEG-H IOP with the IOP for ATSC 3.0 services

**Source:**
Fraunhofer IIS, Dolby Labs, and Xperi

**Supporting Companies:**
Fraunhofer IIS, Dolby Labs, Xperi

**Category:**
A (addition of feature)

**Date:**
2020-04-20

**Reason for change:**
AC-4 and MPEG-H have diverged over time from the specifications in the DASH-IF spec for ATSC.

For MPEG-H: Due to the strict timeline of the ATSC 3.0 specification, the MPEG-H Audio efforts have been mainly focused on the DASH-IF IOP for ATSC 3.0 services. As a result, the main MPEG-H IOP is lacking additional ISO BMFF encapsulation, as specified in the profile for ATSC 3.0 services. This misalignment is at all effects not beneficial for DASH-IF. This document proposes an update to the main DASH-IF IOP such that it aligns with the IOP for ATSC 3.0. In particular, meeting the following objectives is desired:

- The general DASH MPEG-H IOP (this document) is a superset of the IOP for ATSC 3.0 with regard to ISO BMFF encapsulation.
- The document includes the correct references, a short description of the general MPD signalling and a small restructuring of the sections to better fit the changes.

For AC-4: Updates to AC-4 have been introduced to other important specifications, including ATSC 3.0. These specifications cover use cases not previously covered by DASH-IF IOP. This document proposes an update that aligns with ATSC 3.0 IOP, and includes signalling for the advanced use cases enabled by those services.

For DTS: Updates to general sections in include DTS-UHD

**Summary of change:**
The proposed changes aim at aligning the general MPEG-H and AC-4 IOP with the IOP for ATSC 3.0 services with regard to ISO BMFF encapsulation.

**Consequences if not approved:**
AC-4 and MPEG-H bit streams for DASH-IF ATSC profile are not valid for the DASH-IF IOP, making the convergence of broadcast and OTT delivery harder.

**Sections affected:**
9.2.5, 9.4.6

**Other comments:**

**Disclaimer:**
This document is not yet final. It is provided for public review until the deadline mentioned below. If you have comments on the document, please submit comments by one of the following means:
- dashif-iop@groupspaces.com with a subject tag [Audio], or
Please add a detailed description of the problem and the comment.
Based on the received comments a final document will be published latest by the expected publication date below, integrated in a new version of DASH-IF IOP if the following additional criteria are fulfilled:
  - All comments from community review are addressed
  - The relevant aspects for the Conformance Software are provided
  - Verified IOP test vectors are provided

<table>
<thead>
<tr>
<th>Commenting Deadline:</th>
<th>June 15th, 2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Publication:</td>
<td>Q3/2020</td>
</tr>
</tbody>
</table>
Change #1

Add the following definitions:

- **MHAS**  MPEG-H 3D Audio Stream
- **MHA**  Encapsulation of raw MPEG-H 3D Audio frames into ISO BMFF
- **MHM**  Encapsulation of MHAS packets into ISO BMFF
- **Audio Component**  An Audio Component is a Media Component as defined in ISO/IEC 23009-1 Amd. 4 Clause 5.3.11.

Change #2

Update references as follows:


Add the following references as follows:


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


Change #3

Update Table 2 as indicated below:

<table>
<thead>
<tr>
<th>Extension</th>
<th>Identifier</th>
<th>Version</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>DASH-IF multichannel audio extension with MPEG-H 3D Audio</td>
<td><a href="http://dashif.org/guidelines/dashif#mha">http://dashif.org/guidelines/dashif#mha</a></td>
<td>4.x</td>
<td>9.4.6.3</td>
</tr>
</tbody>
</table>

Change #4

Change section 9.2.5.1 as follows:
9.2.5.1. Overview

MPEG-H 3D Audio is defined in ISO/IEC 23008-3 [64] 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, as defined in ISO/IEC 23008-3, clause 4.8 [64].

The sections 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

Change #5

Add Section 9.1.1 and 9.1.2

9.1.1 NGA Overview

Next Generation Audio (NGA) codecs include AC-4, defined in ETSI TS 103 190-2 [x] and MPEG-H 3D Audio, defined in ISO/IEC 23008-3 [64], and DTS-UHD defined in ETSI TS 103 491 [AB].

New concepts and signalling mechanisms have been introduced as part of ISO/IEC 23009-1:2019 [4] in order to support the variety of new use-cases that AC-4 and MPEG-H 3D Audio enables. In particular, the concept of Preselections has been introduced as a method for defining and signalling audio personalization options.

9.1.2 Signalling of Preselections

The NGA codecs support the concept of component based audio, i.e. the audio program can be constructed from set of separate Audio Components (i.e. media content components containing audio). Examples of Audio Components are: dialogs (in potentially different languages), ambient background sound, music or effects.

The concept of “Preselections” as defined in ISO/IEC 23009-1:2019 [4], 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. The Audio Components may be delivered in a single stream or in multiple streams.

Two different methods are defined 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/IEC 23009-1:2019 [4]. It 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 5.3.11.3 and 5.3.11.4 of ISO/IEC 23009-1:2019 [4]. The Role and Accessibility descriptors on 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.

For all AdaptationSets referenced by Preselection elements, the following rules presented in Table 2 shall apply:
Table 2 NGA Element and Attribute Settings

<table>
<thead>
<tr>
<th>Element or Attribute Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SupplementalProperty</td>
<td>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</td>
</tr>
<tr>
<td>EssentialProperty</td>
<td>An AdaptationSet referring to Auxiliary audio streams shall include a Preselection EssentialProperty descriptor as specified in MPEG-DASH</td>
</tr>
<tr>
<td>@preselectionComponents</td>
<td>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</td>
</tr>
</tbody>
</table>

Additional MPD Elements and Attributes are recommended for use when describing an NGA Preselection. These are listed in Table 3. Note that additional guidance in applying these parameters to specific NGA Audio codecs may be included in their respective sections.

Table 3 Additional recommended MPD Elements for NGA Audio

<table>
<thead>
<tr>
<th>Element or Attribute Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>@lang</td>
<td>If a Preselection includes a dialog component, this attribute should be used to indicate the language.</td>
</tr>
<tr>
<td>Role</td>
<td>DASH role scheme, urn:mpeg:dash:role:2011, to indicate a key attribute of the Preselection.</td>
</tr>
<tr>
<td>Accessibility</td>
<td>Indicate whether a Preselection or a Component has accessibility considerations.</td>
</tr>
<tr>
<td>Label</td>
<td>If there are multiple Preselections, this label should be set by the content author.</td>
</tr>
</tbody>
</table>

**Change #6**

Add section 9.2.1.3 as follows:

9.2.1.3. Dolby AC-4 specific issues

9.2.1.3.1 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 ETSI TS 103 190-2 [x], Annex E.

9.2.1.3.2 AC-4 Element and Attribute Settings

Table XX 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 Annex E of ETSI TS 103 190-2 [YY].

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

Table 3 AC-4 Element and Attribute Settings

<table>
<thead>
<tr>
<th>Element or Attribute Name</th>
<th>Description</th>
</tr>
</thead>
</table>
| @codecs                   | For AC-4 the value of the codecs attribute shall be created according to the syntax described in RFC 6381 [ZZ]. 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  
  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. |
| AudioChannelConfiguration  | For AC-4 the Audio Channel Configuration descriptor shall use one of the following schemes  
  - urn:mpeg:mpegB:cicp:ChannelConfiguration as defined in ISO/IEC 23001-8 [49]  
  - tag:dolby.com,2015:dash:audio_channel_configuration:2015 as defined in TS 103 190-2 [YY], Annex G.3.1  
  urn:mpeg:mpegB:cicp:ChannelConfiguration is the preferred scheme. |
<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>@audioSamplingRate</td>
<td>The value shall be set to the sampling frequency as specified in TS 103 190-2 [YY], Annex G.2.6.</td>
</tr>
<tr>
<td>Example:</td>
<td>For $fs_index = 1$ and $dsi_fs_multiplier = 0$, the value is 48000.</td>
</tr>
<tr>
<td>RandomAccess@type</td>
<td>The type attribute of the RandomAccess element shall be set to “closed”.</td>
</tr>
<tr>
<td>@mimeType</td>
<td>The MIME type attribute shall be set to ”audio/mp4”.</td>
</tr>
<tr>
<td>@startWithSAP</td>
<td>The startWithSAP attribute shall be set to 1.</td>
</tr>
<tr>
<td>@lang</td>
<td>The language indicated by the lang attribute should correspond to that language signalled in the language_tag_bytes, which is tagged as “dialog” or “complete main” in the corresponding content_classifier.</td>
</tr>
<tr>
<td>NOTE:</td>
<td>The language_tag_bytes are contained in the ac4_substream_group_dsi structure, within the ac4_dsi_v1 structure.</td>
</tr>
<tr>
<td></td>
<td>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 &lt; 2 and that are fully contained in this AdaptationSet.</td>
</tr>
<tr>
<td>Role</td>
<td>The Role for a Preselection should be set by the content author.</td>
</tr>
<tr>
<td>Note:</td>
<td>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.</td>
</tr>
<tr>
<td>Accessibility</td>
<td>In case one or more audio elements contained in a Presentation indicate a content type visually impaired, an Accessibility descriptor shall indicate descriptions according to the Role scheme defined in ISO/IEC 23009-1 [xx].</td>
</tr>
<tr>
<td></td>
<td>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 [xx] may be used.</td>
</tr>
<tr>
<td></td>
<td>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 [xx].</td>
</tr>
</tbody>
</table>
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 [33], clause G.2.12.1.

An audio framerate SupplementalProperty descriptor should be used as specified in ETSI TS 103 190-2 [33], clause G.2.12.2.

**Change #7**

Rename the current section 9.2.5.2 “DASH-specific issues” as section 9.2.5.3 “MHA encapsulation” and update Table 29 as follows:

<table>
<thead>
<tr>
<th>Codec</th>
<th>Codecs Parameter</th>
<th>ISO BMFF Encapsulation</th>
<th>SAP Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEG-H 3D Audio LC Profile Level 1</td>
<td>mha[1, 2].0x0B</td>
<td>ISO/IEC 23008-3</td>
<td>1</td>
</tr>
<tr>
<td>MPEG-H 3D Audio LC Profile Level 2</td>
<td>mha[1, 2].0x0C</td>
<td>ISO/IEC 23008-3</td>
<td>1</td>
</tr>
<tr>
<td>MPEG-H 3D Audio LC Profile Level 3</td>
<td>mha[1, 2].0x0D</td>
<td>ISO/IEC 23008-3</td>
<td>1</td>
</tr>
</tbody>
</table>

**Change #8**

Add the following as section 9.2.5.2 “DASH-specific issues”.

### 9.2.5.2. DASH-specific issues

The carriage of MPEG-H 3D Audio in the ISO BMFF is specified in ISO/IEC 23008-3, clause 20 [64]. Two different packaging methods are specified:

- **MHA**: storage of raw audio frames, as specified in ISO/IEC 23008-3, clause 20.5 [64]. Section 9.2.5.3. provides more information on this encapsulation.
- **MHM**: storage of MHAS streams, as specified in ISO/IEC 23008-3, clause 20.6 [64]. The MPEG-H Audio Stream (MHAS) format is defined in ISO/IEC 23008-3, clause 14 [64]. Section 9.2.5.4. provides more information on this encapsulation.

#### 9.2.5.2.1 Element and Attribute Settings

Table 26 summarizes the mapping of relevant MPD elements and attributes to MPEG-H Audio.

<table>
<thead>
<tr>
<th>Element or Attribute Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>@codecs</td>
<td>The signalling of the codecs parameters is according to RFC6381 [10] and ISO/IEC 23008-3, clause 21 [64]. The value consists of the following two parts separated by a dot:</td>
</tr>
</tbody>
</table>
• the sample entry 4CC code ('mha1', 'mha2', 'mhml', 'mhm2')
• '0x' followed by the hex value of the profile-level-id, as defined in ISO/IEC 23008-3 [64]

See Table 27 and Table 28 for more details.

**AdaptationSet**

This field lists the mae_groupIDs as defined in ISO/IEC 23008-3 [64] that are contained in the Adaptation Set separated by white spaces.

**Preselection**

This field indicates the mae_groupPresetID as defined in ISO/IEC 23008-3 [64] that refers to a Preset in the scope of MPEG-H Audio.

**ContentComponent**

This field indicates the mae_groupID as defined in ISO/IEC 23008-3 [64] 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 23001-8:2016 [49]. 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 [64].

**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".

@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 [64].

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 [64].

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 [4].

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 [4] 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.

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 [64].

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 [64].

Change #9

Add the following as section 9.2.5.4 “MHM encapsulation”.

9.2.5.4. MHM encapsulation

Storage of MHAS into ISO BMFF shall be according to ISO/IEC 23008-3 clause 20 [64] 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, clause 20.6 [64]. 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, clause 14.6 [64].

All MHAS packet types defined in ISO/IEC 23008-3, clause 14 [64], 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 clause 15 of ISO/IEC 23008-3 [64], is present, it shall be always encapsulated in an MHAS packet of type PACTYP_AUDIOSCENEINFO [64]. Audio Scene Information shall not be included in the mpegh3daConfig() structure carried in the MHAS packet of type PACTYP_MPEGH3DACFG.

As defined in ISO/IEC 23008-3 clause 20.6 [64], the MHAConfigurationBox() is optional for MHM, but if present it shall also be consistent with the configuration of the audio bitstream, as described in 9.2.5.3.

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 MPEG-DASH Implementation Guidelines [6] 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, clause 20.2 [64], thus the audio data encapsulated in the MHAS packet PACTYP_MPEGH3DAFRAME shall contain the AudioPreRoll() syntax element, as defined in sub-clause 5.5.6 of ISO/IEC 23008-3 [64], and shall follow the requirements for stream access points as defined in clause 5.7 of ISO/IEC 23008-3 [64].

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, clause 20.6.1 [64] 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, clause 14.4 [64].

<table>
<thead>
<tr>
<th>Codec</th>
<th>Codec Parameter</th>
<th>ISO BMFF Encapsulation</th>
<th>SAP Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEG-H 3D Audio LC Profile Level 1</td>
<td>mhm[1, 2].0x0B</td>
<td>ISO/IEC 23008-3</td>
<td>1</td>
</tr>
<tr>
<td>MPEG-H 3D Audio LC Profile Level 2</td>
<td>mhm[1, 2].0x0C</td>
<td>ISO/IEC 23008-3</td>
<td>1</td>
</tr>
</tbody>
</table>
9.2.5.4.1. 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 [61]. If MHAS packets of type PACTYP_AUDIOTRUNCATION are present, they shall be used as described in ISO/IEC 23008-3, clause 14.4 [64].

9.2.5.4.2. 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) an Partial Adaptation Set.

- For the main and the auxiliary MHAS stream(s), the MHASPacketLabel shall be set according to ISO/IEC 23008-3, clause 14.6 [64].

- 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).

**Change #10**

Add the following as section 9.2.5.5 “Loudness and Dynamic Range Control”.

**9.2.5.5. Loudness and Dynamic Range Control**

Loudness metadata shall be embedded within the `mpeg3daLoudnessInfoSet()` structure as defined in ISO/IEC 23008-3, clause 6.3 [64]. 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, clause 5.3.2 [64]). More precisely, the `mpeg3daLoudnessInfoSet()` 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, clause 6.3.1 [64] and ISO/IEC 23003-4, clause 7.3 [88]). The indicated loudness value shall be measured according to applicable regional loudness regulations.

DRC metadata shall be embedded in the `mpeg3daUniDrcConfig()` and `uniDrcGain()` structures as defined in ISO/IEC 23008-3, clause 6.3 [64]. For each included DRC set the `drcSetTargetLoudnessPresent` field as defined in ISO/IEC 23003-4, clause 7 [88] 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, clause 15.5 [64] shall be present in the Audio Scene Information if the `mae_allowGainInteractivity` field (according to ISO/IEC 23008-3, clause 15.3 [64]) is set to 1 for at least one group of audio elements.

**Change #11**

Update section 9.4.6 accordingly:

**9.4.6.1. Introduction**

Compliance to DASH-IF multichannel audio extension with MPEG-H 3D Audio [x] may be signaled by a `@profile` attribute with the value `http://dashif.org/guidelines/dashif#mpeg-h-3da`.

**9.4.6.3. Definition**

Content may be authored claiming conformance to DASH-IF multichannel audio extension with MPEG-H 3D Audio,

- if the content is multichannel audio content as defined in section 9.4.1, and
- if a client can properly play the content by supporting at least the following features:
  - all multichannel audio client features as defined in section 9.4.1,
MHA, MHM and the DASH-specific features defined in section 9.2.5.