Exploration on alignment of ISOBMFF/DASH/CMAF terminology, concepts and solutionsINTERNATIONAL STANDARD© ISO/IEC 2015 – All rights reservedISO/IEC 15444-12:2015(E) 63Part 12: ISO base media file formatInformation technology — JPEG 2000 image coding systemTechnologies de l'information — Codage des objets audiovisuels — Partie 12: Format ISO de base pour les fichiers médiasInformation technology — JPEG 2000 image coding system — Part 12: ISO base media file formatE2015-02-20(60) PublicationISO/IECISO/IEC J   International Standard 2015ISO/IEC 15444‑ISO/IEC 15444‑12ISO/IEC 15444-12  Coding of audio, picture, multimedia and hypermedia informationInformation technology 291 2見出し 2見出し 1    02 STD Version 2.1c260   4C:\Users\shinji\_w\AppData\Roaming\Microsoft\Templates\STD\w15177\_14496\_5th.-restyle-R1.doc

Contents Page

1 General Considerations 2

1.1 Processing Diagram 5

1.2 Requirements and Scenarios 5

2 Tentative Mapping 5

2.1.1 Movie = Period = Presentation 6

2.1.2 Track = Representation =Track 6

2.1.3 Switch Group = AdaptationSet = Switching Set 7

2.1.4 Alternate Group = Group = Selection Set 7

2.1.5 Track Reference = Dependent Representation = Dependent CMAF Track 7

2.1.6 TBD = Associated Representation = TBD 8

2.1.7 TBD =Switchable AdaptationSet = TBD 8

2.1.8 TBD = Preselection = TBD 8

2.1.9 Track Group = TBD = TBD 8

2.1.10 Subtracks = Sub-Representation = TBD 8

2.1.11 TBD = ContentComponents = TBD 9

2.2 ISOBMFF clarifications/changes 9

2.2.1 Alternate groups 9

2.2.2 Track Groups 10

3 Use Case 1: Bitrate switching 10

3.1 Alternate groups and track selection box 10

3.1.1 Request 10

3.1.2 Proposal – Preselection related 11

3.2 Track groups and track group description 12

3.2.1 Request 1 12

3.2.2 Proposal 1 12

3.2.3 Request 2 13

3.2.4 Proposal 2 13

3.2.5 Request 3 14

3.2.6 Proposal 3 15

4 Use Case 2: Switching across AdaptationSets 16

4.1 Example 16

4.2 Request 18

4.3 Proposal 18

5 Use Case 3: Dependent Tracks 20

5.1 Request 20

5.2 Proposal 21

6 Preselections 21

6.1 Preselections and alternate groups 21

6.1.1 Request 21

6.1.2 Proposal 21

6.2 Preselection Main track signaling 22

6.2.1 Request 22

6.2.2 Proposal 22

7 Other topics 23

7.1 Terminology in ISOBMFF 8th Edition 23

7.1.1 Request 23

7.1.2 Proposal 23

7.2 New CMAF Profile 23

7.3 Additional considerations 24

7.3.1 Editorial changes in ISOBMFF 8th Edition 24

7.3.2 On segment\_order 24

8 Related documents 24

# General Considerations

## Purpose of the document

This exploration document and/or its technologies may eventually evolve into one or more of the following outcomes:

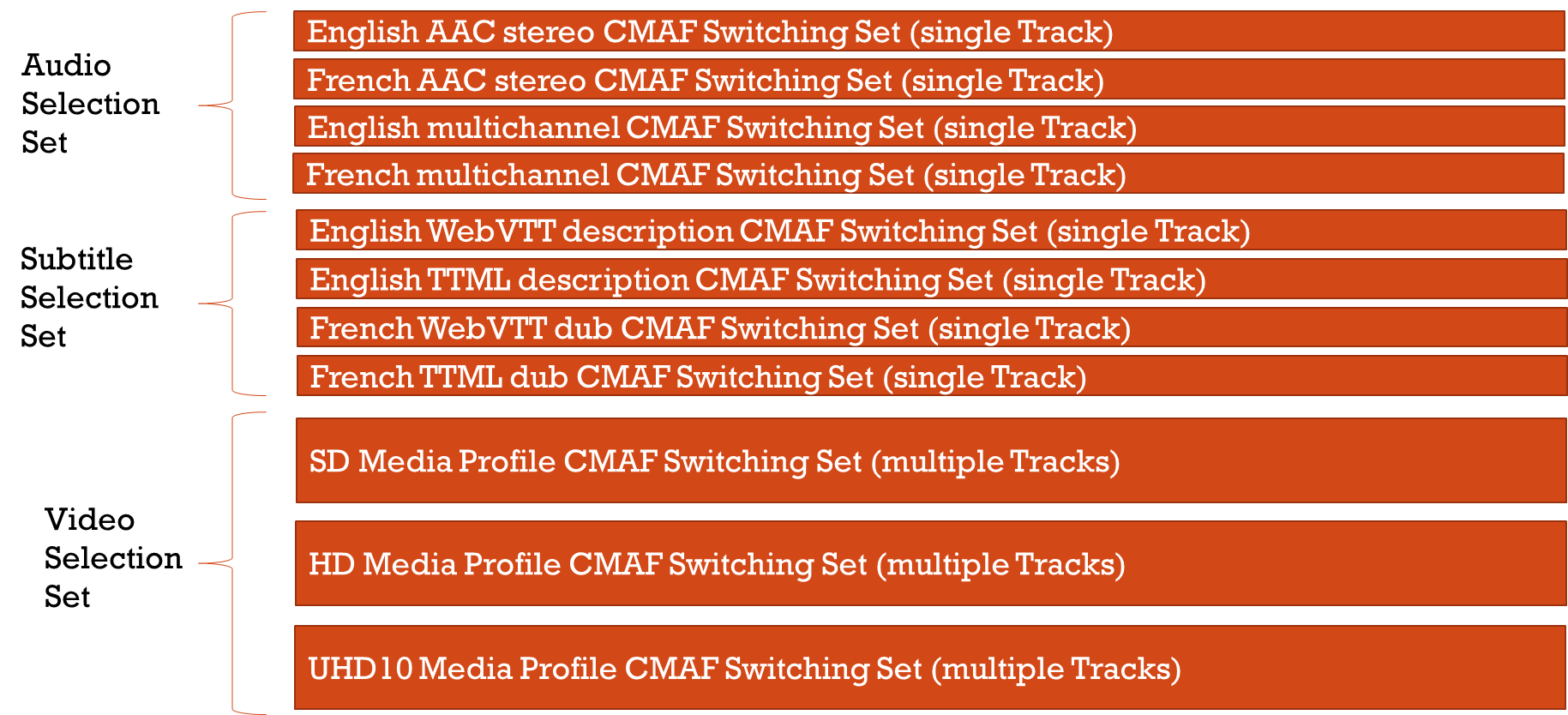
* Technical Report that describes the mapping between ISOBMFF, CMAF, and DASH terms and concepts.
* ISOBMFF extensions to indicate a track grouping that is equivalent to a CMAF Switching Set and a DASH Adaptation Set
* ISOBMFF extensions equivalent to indicating switching between DASH Adaptation Sets

## Background

Some structural functionalities in DASH and ISO BMFF have not been fully harmonized. The file format typically assumes that all tracks are contained in a single file and a single movie header documents all tracks included in the file, but also provides the relationship of these tracks for presentation. A few file format functionality that are relevant include

* Each media stream is contained in a track specialized for that media type (audio, video etc.), and is further parameterized by a sample entry.
* The sample entry
  + contains the ‘name’ of the exact media type (i.e., the type of the decoder needed to decode the stream) and any parameterization of that decoder needed.
  + The name also takes the form of a four-character code.
  + There are defined sample entry formats not only for MPEG-4 media, but also for the media types used by other organizations using this file format family.
  + They are registered at the MP4 registration authority.
* Tracks (or sub tracks) may be identified as alternatives to each other, and there is support for declarations to identify what aspect of the track can be used to determine which alternative to present, in the form of track selection data.
* Tracks may also be linked
* Tracks may be grouped
* Tracks may be encrypted
* Tracks have random access samples
* Tracks have assigned media specific properties/annotation (codec, width, height, etc.)
* Relationship in tracks are expressed as
  + Track references: track N uses or refers to track(s) K (hint, chap, scal)
  + Track groups: Tracks in the same group share a common feature
  + Track Selection: Provides selection information for alternate tracks
    - New features have been added or about to be added, such as dependencies, Pre-selection, etc.

Some of the file format principles do not carry forward to CMAF or DASH, because CMAF and DASH heavily rely on the concept of late binding, i.e. each track is stored in a separate file. This is shown in the figure below.



In a streaming environment this avoids combinatorial complexity or useless downloads, as clients only select the relevant tracks for the current situation, and only downloads these tracks and does synchronized playout. The HTML-5 MSE based playback exactly permits this.

However, by doing so, each track gets its own ISO BMFF/MP4 file and all of a sudden all concepts of the file format are no longer applicable because

1. The requirements for file to have unique track identifiers is not carried forward
2. Each file includes its own movie header and relationship about the tracks cannot be expressed

In a streaming environment, the information that is contained in a movie header for many tracks, needs to be populated to the manifest such that the client can early enough select the tracks based on this information.

In CMAF, some conceptual grouping was done, but not relying on file format information, but defining its own concepts

* CMAF Tracks
* CMAF Switching Sets
* CMAF Groups
* Depending switching sets
* Aligned switching sets
* Etc.

For example, in CMAF it is not specified that all tracks in a CMAF switching track shall conform to a track group according to the ISO BMFF.

## Processing Diagram

Packaging for Streaming   
(parameters: segment duration, low latency, …)

Packaging for local playback   
(losing delivery aspects: segmentation, availability, …)

*Should we keep mixing DASH and CMAF terms? Or just use DASH terms (e.g. Representation vs Track)?*

## Requirements and Scenarios

* It should be possible to generate MPD and CMAF tracks from a multi-track MP4. How easily?
* It should be possible to offer the same presentation options than the MPD offers in a local MP4, including bitrate variations (for debug purposes but also for battery efficiency use cases)

*We need to consider live cases, where you would only get the MP4 init segment (with multiple tracks in it), can you produce an MPD from that only? Not having the media segments you cannot know in advance if they are aligned and thus cannot produce the proper MPD grouping, without additional signaling.*

# Tentative Mapping

Table 1 shows commonly understood mappings between terminology from the different standards under this exploration.

Subsequent sections list definitions for terms from respective standards as tentatively mapped in Table 1.

*We could add diagrams of the hierarchical grouping in each spec.*

Figure 1: DASH High-Level Data Model

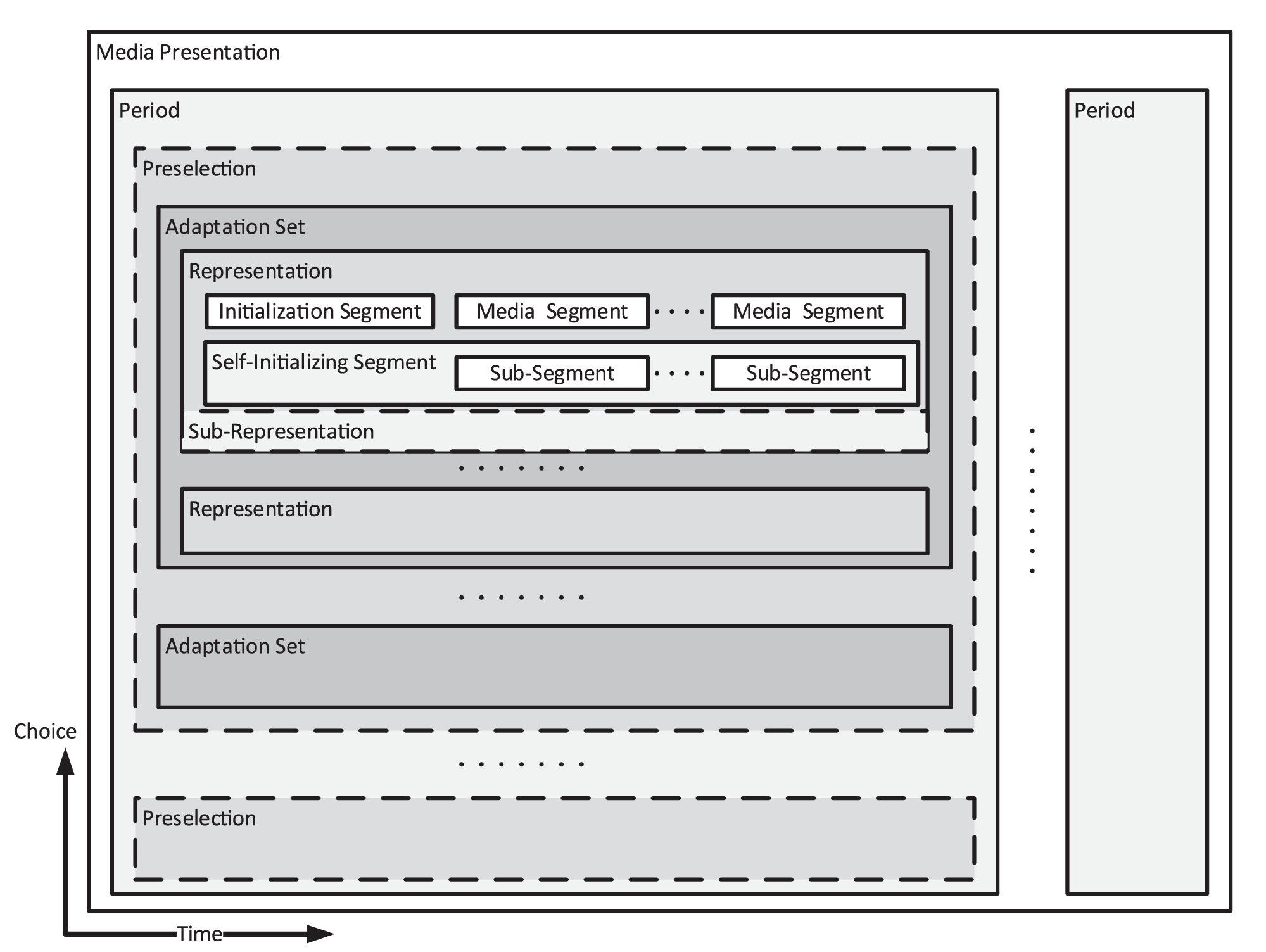


Figure 2: Media objects specified in CMAF and presented by externally specified applications,

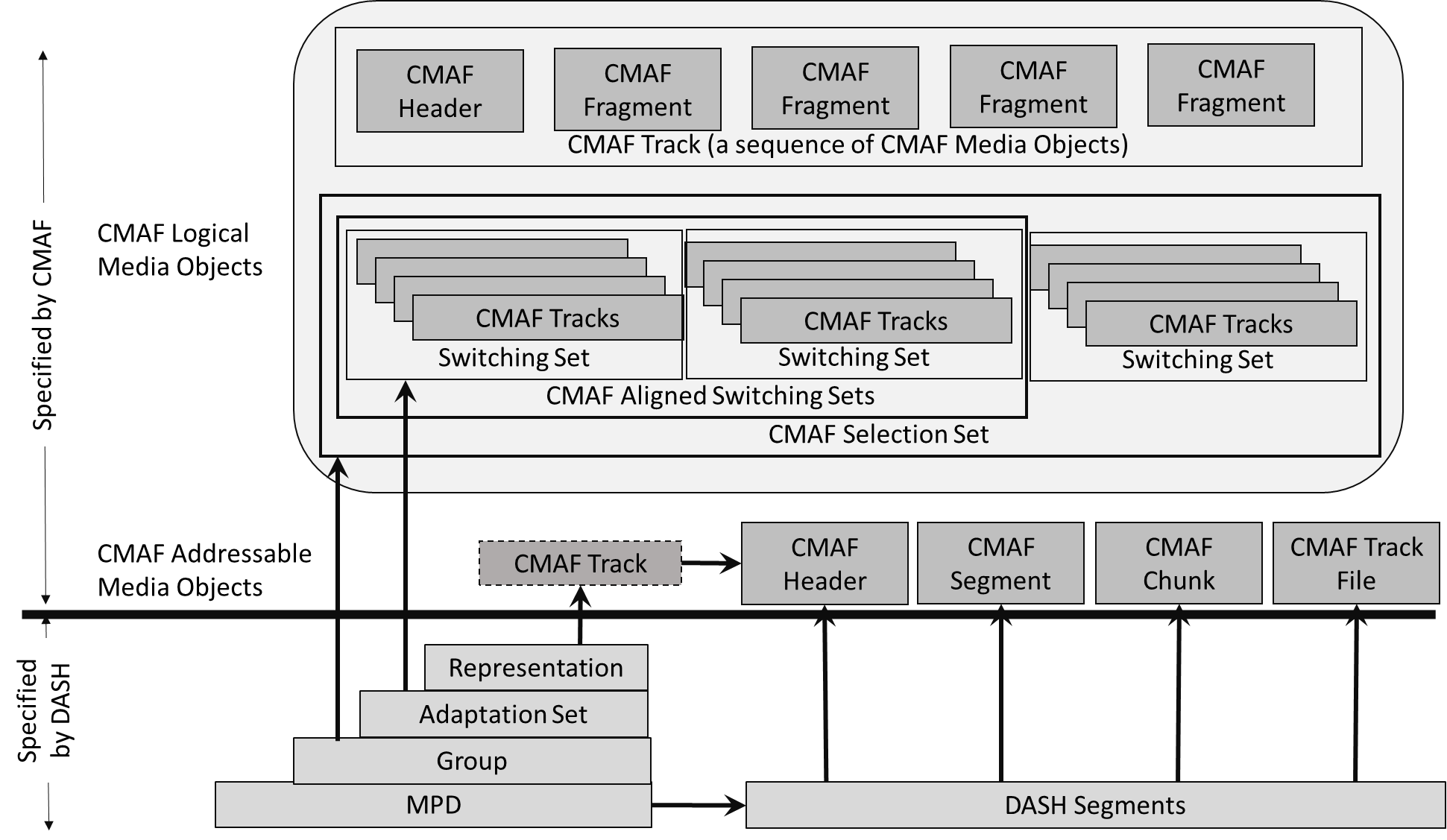


Table 1:Tentative Mappings

|  |  |  |
| --- | --- | --- |
| *ISOBMFF* | *DASH* | *CMAF* |
| Movie | Period | Presentation |
| Track | Representation | Track |
| Switch Group | AdaptationSet | Switching Set |
| Alternate Group | Group | Selection Set |
| Track Reference | Dependent Representation | Dependent CMAF Track |
|  | Associated Representation |  |
|  | Switchable AdaptationSet |  |
| (*see Ed.Note below*) | Preselection |  |
| Track Group |  |  |
| Subtracks | Sub-Representation |  |
|  | ContentComponent |  |
|  | Subset |  |

*Ed.Note on VR extractor tracks:*

*A simplified description: In OMAF, the Main Adaptation Set is mapped to an OMAF base track (which is an AVC/HEVC extractor track, an HEVC tile base track, or a VVC merge base track), and each Partial Adaptation Set is mapped to an 'alte' track group of AVC/HEVC tracks containing collocated slices/tiles or VVC subpicture tracks containing collocated subpictures. A track reference (e.g. 'scal') is included in the OMAF base track and lists all 'alte' track references.*

*This might be too detailed to be included in this document.*

### Movie = Period = Presentation

|  |  |
| --- | --- |
| ISOBMFF | **movie file** ISO base media file containing a MovieBox  **movie box** container box whose sub-boxes define the structure-data for a presentation ('moov')  **Presentation** one or more motion sequences, possibly combined with audio |
| DASH | **media content period** set of *media content components* that have a common timeline as well as relationships on how they can be presented  **Period** interval of the *Media Presentation*, where a contiguous sequence of all Periods constitutes the Media Presentation  *Is an ISOBMFF presentation equivalent to an entire Media Presentation with multiple periods? Or just one period? If an ISOBMFF presentation is an entire MP, how would you represent different number of tracks? How to handle timing? There could be some cases that would need special handling? Are there more cases not handled by this mapping?* |
| CMAF | **CMAF presentation** set of one or more *CMAF selection sets* that can be simultaneously decoded to produce a multimedia user experience, potentially including synchronized audio, video, and subtitles  *CMAF presentation assumes fixed set of CMAF tracks, this would correspond to a DASH period also. Is that true?* |

### Track = Representation =Track

|  |  |
| --- | --- |
| ISOBMFF | **track** timed sequence of related samples (q.v.) in an ISO base media file |
| DASH | **Representation** collection and encapsulation of one or more *media streams* in a delivery format and associated with descriptive metadata  *A DASH Representation can have multiple Content Components, be multiplexed. This document currently assumes “Single Component” Representations. Currently, CMAF and ISOBMFF do not support multiple media types in a track. ISOBMFF supports multiple codecs in a track, but in sequence not in parallel.* |
| CMAF | **CMAF track** sequence of *CMAF fragments* that are consecutive in presentation time, contain one media stream, conform to at least one structural CMAF brand, including an associated *CMAF header* that can initialize playback |

### Switch Group = AdaptationSet = Switching Set

|  |  |
| --- | --- |
| ISOBMFF | *No term definition* |
| DASH | **Adaptation Set** set of interchangeable encoded versions of one or several *media content components* |
| CMAF | **CMAF switching set** set of one or more *CMAF tracks*, where each track is an alternative encoding of the same source content, and are constrained to enable seamless track *switching* |

ISOBMFF alternate groups and switch groups are related concepts. An alternate group may contain 1 or more switch groups. Example: a single MP4 file containing AVC, HEVC and VVC tracks. Tracks of a single codec could be in a separate switching group within the same alternate group.

The signaling of alternate group and switching group could be also determined by looking at the movie header. The additional signaling could be used to provide author’s indications or the fact that the content has been prepared to enable switching.

DASH has bitstream switching, header switching.

The definition of Switching in DASH is very well defined but not clearly defined in ISOBMFF.

### Alternate Group = Group = Selection Set

|  |  |
| --- | --- |
| ISOBMFF | *No term definition* |
| DASH | **group** collection of *Adaptation Sets* that are not expected to be presented simultaneously |
| CMAF | **CMAF selection set** set of one or more *CMAF switching sets*, where each CMAF switching set encodes an alternative aspect of the same presentation over the same time period, only one of which is intended to be played at a time, e.g. an alternative language or codec |

### Track Reference = Dependent Representation = Dependent CMAF Track

|  |  |
| --- | --- |
| ISOBMFF | *No term definition*  *Some track references imply decoding dependencies, some don’t.* |
| DASH | **dependent Representation** *Representation* for which *Segments* from its *complementary Representations* are necessary for presentation and/or decoding of the contained *media content components*  **complementary Representation** *Representation* which complements at least one *dependent Representation*  *DASH puts restrictions (e.g. media type) to ensure that the same decoder can be initialized for both representations when fed with the dependent header. (to be checked)*  *We should look at output document w18291 (section 10).* |
| CMAF | *No term definition* |

### TBD = Associated Representation = TBD

|  |  |
| --- | --- |
| ISOBMFF | *t.b.d.*  *This could be a track reference, but with no decoding dependencies.* |
| DASH | **associated Representation** *Representation* which provides supplemental or descriptive information for at least one other Representation  *Metadata track contains information related to some other representation(s).* |
| CMAF | *t.b.d.* |

### TBD =Switchable AdaptationSet = TBD

|  |  |
| --- | --- |
| ISOBMFF | *t.b.d.* |
| DASH | *No term definition*  *Allowing switching from an Adaptation Set to another Adaptation Set, unidirectionally. (to be checked)* |
| CMAF | *t.b.d.* |

### TBD = Preselection = TBD

|  |  |
| --- | --- |
| ISOBMFF | *t.b.d. <How do you do VR? Extractor Tracks>* |
| DASH | **Preselection** set of *media content components* that are intended to be consumed jointly |
| CMAF | *t.b.d.* |

### Track Group = TBD = TBD

|  |  |
| --- | --- |
| ISOBMFF | *No term definition*  *ISOBMFF 8th edition introduces the track group definition as a way to optimize the storage of information common to tracks in a track group* |
| DASH | *t.b.d.* |
| CMAF | *t.b.d.* |

### Subtracks = Sub-Representation = TBD

|  |  |
| --- | --- |
| ISOBMFF | *No term definition*  *ISOBMFF subtracks is descriptive.* |
| DASH | **Sub-Representation** part of a *Representation* described in the *MPD* that is present in the entire *Period*  *Can be used to describe layers in the same representation.* |
| CMAF | *t.b.d.* |

### TBD = ContentComponents = TBD

|  |  |
| --- | --- |
| ISOBMFF | *t.b.d.* |
| DASH | **media content component** single continuous component of the *media content* with an assigned *media content component type*  **media content component type** single type of *media content* EXAMPLES Audio, video, or text.  *<Is this applicable to ContentComponent elements?>* |
| CMAF | *t.b.d.* |

## ISOBMFF clarifications/changes

Tracks in an Alternate Group are alternate!

Tracks in a Track Group are grouped semantically but they are not alternative. (maybe clarify/restrict that in ISOBMFF)

### Alternate groups

*Informative text from 14496-12, clause 8.10.3.1 on alternate groups:*

A typical presentation stored in a file contains one alternate group per media type: one for video, one for audio, etc. Such a file may include several video tracks, although, at any point in time, only one of them should be played or streamed. This is achieved by assigning all video tracks to the same alternate group. (See subclause 8.3.2 for the definition of alternate groups.)

All tracks in an alternate group are candidates for media selection, but it may not make sense to switch between some of those tracks during a session. One may for instance allow switching between video tracks at different bitrates and keep frame size but not allow switching between tracks of different frame size. In the same manner it may be desirable to enable selection – but not switching – between tracks of different video codecs or different audio languages.

The distinction between tracks for selection and switching is addressed by assigning tracks to switch groups in addition to alternate groups. One alternate group may contain one or more switch groups. All tracks in an alternate group are candidates for media selection, while tracks in a switch group are also available for switching during a session. Different switch groups represent different operation points, such as different frame size, high/low quality, etc.

For the case of non-scalable bitstreams, several tracks may be included in a switch group. The same also applies to non-layered scalable bitstreams, such as traditional AVC streams.

By labelling tracks with attributes it is possible to characterize them. Each track can be labelled with a list of attributes which can be used to describe tracks in a particular switch group or differentiate tracks that belong to different switch groups.

### Track Groups

*Informative text from 14496-12, clause 8.3.4 on track groups:*

This box (Track Group Box) enables indication of groups of tracks, where all tracks in a group share a particular characteristic or the tracks within a group have a particular relationship. The box contains zero or more boxes, and the particular characteristic or the relationship is indicated by the box type of the contained boxes. The contained boxes include an identifier, which can be used to conclude the tracks belonging to the same track group. The tracks that contain the same type of a contained box within the TrackGroupBox and have the same identifier value within these contained boxes belong to the same track group.

Track groups shall not be used to indicate dependency relationships between tracks. Instead, the TrackReferenceBox is used for such purposes.

# Use Case 1: Bitrate switching

Table 2: Terms for bitrate switching

|  |  |  |
| --- | --- | --- |
| *ISOBMFF* | *DASH* | *CMAF* |
| Switch Group | AdaptationSet | Switching Set |
| Track | Representation | Track |

Two alternative approaches for signalling the relation between multiple tracks enabling bitrate switching in an ISOBMFF are considered by contributions:

- alternate\_groups and switching\_group

- Track Groups

It needs to be decided which approach to follow.

## Alternate groups and track selection box

### Request

The MPEG-DASH specification defines the Preselection to enable the combination of different Adaptation Sets into a single decoding instance and user experience. The subclause 5.3.11.1 of MPEG DASH specifies the mapping of media components in Preselection to Adaptation Sets in 3 different ways, shown below.

*Media components can be mapped to Adaptation Sets in multiple ways:*

1. *by a one-to-one mapping between media components and Adaptation Sets;*
2. *by the inclusion of multiple media components in a single Adaptation Set where all encoded versions of the media components are multiplexed on the file-container level;*
3. *by the inclusion of multiple media components in a single Adaptation Set where all encoded versions of the media components are multiplexed on the elementary-stream level.*

The DIS text of ISOBMFF 8th edition (WG03 N0651) specifies signalling of Preselection related data structures in ISOBMFF. However, the current design of Preselection in ISOBMFF does not define any mechanism to signal tracks which are alternative of each other in the context of Preselection. In the DASH context this translates to Representations which are part of the same media Adaptation Set.

The ISOBMFF subclause 8.10.3.1 describes the relation of alternate groups (as indicated by non-zero values of alternate\_group in the TrackHeaderBox) and switch groups (indicated by TrackSelectionBox) described as follows:

*All tracks in an alternate group are candidates for media selection, but it may not make sense to switch between some of those tracks during a session. One may for instance allow switching between video tracks at different bitrates and keep frame size but not allow switching between tracks of different frame size. In the same manner it may be desirable to enable selection – but not switching – between tracks of different video codecs or different audio languages.*

*The distinction between tracks for selection and switching is addressed by assigning tracks to switch groups in addition to alternate groups. One alternate group may contain one or more switch groups. All tracks in an alternate group are candidates for media selection, while tracks in a switch group are also available for switching during a session. Different switch groups represent different operation points, such as different frame size, high/low quality, etc.*

A player can switch between Representations of an Adaptation Set. Since an alternate group can contain tracks for both selection and switching, a grouping mechanism that provides finer grouping than an alternate group is needed in Preselection for associating tracks to an Adaptation Set.

Two options for associating tracks to an Adaptation Set in the context of Preselection have been proposed:

1. A new track group mechanism which groups tracks that are alternative of each other in the context of Preselection. (see section 3.2.2)
2. A new attribute value for TrackSelectionBox that indicates a switch group of tracks that are alternative of each other in the context of Preselection. (see section 3.1.2)

### Proposal – Preselection related

*Option 2, change 1: In 8.10.3, add the last table row:*

|  |  |  |
| --- | --- | --- |
| *Name* | *Attribute* | *Pointer* |
| Codec | 'cdec' | Sample Entry (in SampleDescriptionBox of media track) |
| Screen size | 'scsz' | Width and height fields of VisualSampleEntry. |
| Max packet size | 'mpsz' | Maxpacketsize field in RtpHintSampleEntry |
| Media type | 'mtyp' | Handlertype in HandlerBox (of media track) |
| Media language | 'mela' | Language field in MediaHeaderBox |
| Bitrate | 'bitr' | Total size of the samples in the track divided by the duration in the TrackHeaderBox |
| Frame rate | 'frar' | Number of samples in the track divided by duration in the TrackHeaderBox |
| Number of views | 'nvws' | Number of views in the track |
| Alternative for preselection | 'adst' | Preselection track group |

*Option 2, change 2: in subclause 8.3.5.4.1.3 :*

The number of non-alternative tracks grouped by this preselection track group is the sum of the following:

* the number of tracks that have alternate\_group equal to 0 and are grouped by this preselection track group,
* the number of switch groups containing tracks ~~the number of unique non-zero alternate\_group values in all tracks~~ that are grouped by this preselection track group.

## Track groups and track group description

### Request 1

As per section 3.1.1.

### Proposal 1

*Option 1, change 1: add the following*

* + - * 1. Preselection media adaptation group box

Definition

A TrackGroupTypeBox with track\_group\_type equal to 'pmag', is also referred to as a PreselectionMediaAdaptationGroupBox. All the tracks that have a track group with track\_group\_type equal to 'pmag' and a particular value of track\_group\_id are alternative tracks in the context of a preselection. All the tracks of the same 'pmag' track group shall have the same non-zero alternate\_group value.

NOTE 1  In the context of DASH (ISO/IEC 23009-1), tracks with the same track\_group\_id are intended to be part of the same media Adaptation Set.

NOTE 2 When authoring a DASH MPD, the value of track\_group\_id can be used as the ID of the media Adaptation Set in Preselection.

Syntax

aligned(8) class PreselectionMediaAdaptationGroupBox extends TrackGroupTypeBox('pmag')  
{  
}

*Option 1, change 2: in subclause 8.3.5.4.1.3:*

*…*

The number of non-alternative tracks grouped by this preselection track group is the sum of the following:

* the number of tracks that have alternate\_group equal to 0 and are grouped by this preselection track group,
* the number of 'pmag' track groups containing tracks ~~the number of unique non-zero alternate\_group values in all tracks~~ that are grouped by this preselection track group.

…

### Request 2

One desired parameter to group representations of the same media is to combine all tracks encoded in different bitrates in one single ISOBMFF file.

To accomplish this use case and signal this kind of interrelation between tracks in ISOBMFF, it is proposed to define a new track\_group\_type to form “multirate” track groups.

Although unlikely to be distributed to end users due to its data overhead, such a file containing a “multirate” track group may serve as a mezzanine format provided as input to streaming packagers

This concept enables DASH Representations in AdaptationSets and likewise CMAF Tracks combined in CMAF Switching Sets.

### Proposal 2

*Add the following entries to the list of track\_group\_types:*

* + - 1. Semantics

track\_group\_type indicates the track grouping type and shall be set to one of the following values, or a value registered, or a value from a derived specification or registration:

'msrc' indicates that this track belongs to a multi-source presentation. Specified in 8.3.4.4.1.

'ster' indicates that this track is either the left or right view of a stereo pair suitable for playback on a stereoscopic display. Specified in 8.3.4.4.2.

'pres' indicates that this track contributes to a preselection. Specified in 8.3.4.4.3.

'mrtg' indicates that this track belongs to a group of track that provide the same content encoded in different bitrates

The pair of track\_group\_id and track\_group\_type identifies a track group. The tracks that contain a particular TrackGroupTypeBox having the same value of track\_group\_id and track\_group\_type belong to the same track group.

*Add the following new clauses:*

* + - 1. Track group definitions
         1. Multirate group box

Definition

Streaming applications typically provide their media assets in different qualities and bitrates to enable adaptation to network conditions. The multirate group allows to assign multiple tracks to one group to indicate a streaming server that each of the tracks in the group provide the content in different bitrates.

Optionally present boxes may provide a generic description of the content that is valid for all tracks.

Syntax

aligned(8) class MultirateGroupBox extends TrackGroupTypeBox('mrtg')  
{  
 Box boxes[];  
}

### Request 3

**ISOBMFF alternate groups**

One possible solution for signaling a CMAF switching set is to use the alternate\_group field to associate the tracks of a CMAF switching set. While the definition of this flag seems to be aligned with the CMAF switching set definition, it does not allow the definition of various attributes of a CMAF switching set.

The preselection track group considers the possible existence of alternate\_group in counting the number of tracks in a preselection track group.

**CMAF Switching Set Characteristics**

ISO/IEC 23000-19 defines various characteristics of a CMAF switching set. These characteristics are:

1. General CMAF switching sets constraints
2. Aligned CMAF switching sets
3. CMAF switching sets with single initialization
4. CMAF switching sets with principal header
5. CMAF structural and media profile brands

**Signaling the CMAF Switching set characteristics in ISOBMFF**

One possible way to signal the CMAF switching set characteristics is to use the track group box, similar to the way the preselection is done. In this case, a CMAF switching set track group entry box can be defined (such as ‘cmsg’) that includes the following parameters:

1. track\_group\_id value equal to alternate\_group of CMAF tracks in the switching set.
2. num\_tracks specifies the number of tracks in this switching sets
3. aligned specifies whether the tracks are time aligned.
4. single\_init specifies whether the tracks have single initialization
5. principal\_id indicates the track id that contains the principal header for this switching set.
6. structural\_brand indicates the CMAF structural brand that all the tracks in this switching set conform to.
7. mediaprofile\_brand indicates the CMAF file brand of a media profile that all tracks in this switching set conform to.

Each track of the CMAF switching set shall have its alternate\_group set equal to the track\_group\_id value of the CMAF track group entry box.

### Proposal 3

* + - * 1. CMAF Switching Group track group entry box

Definition

Box Type: 'cmss'  
Container: TrackGroupDescriptionBox  
Mandatory: No  
Quantity: Zero or More

CMAF Switching Group defines a set of tracks that are conforming the CMAF switching sets in ISO/IEC 23000-19. A CMAF Switching set can be qualified by its attributes.

CMAFSwitchingGroupEntryBox shall describe only track groups identified by track\_group\_type equal to 'cmsg'.

All CMAF Switching Groups with at least one contributing track having the track\_in\_movie flag set to 1 shall be qualified by CMAFSwitchingGroupEntryBoxes. Otherwise, the presence of the CMAFSwitchingGroupEntryBoxes is optional.

All attributes uniquely qualifying a CMAF Switching Group should be present in the CMAFSwitchingGroupEntryBox of the CMAF Switching Group.

Syntax

aligned(8) class CMAFSwtichingGroupEntryBox  
 extends TrackGroupEntryBox('cmsg', version=0, flags)  
{  
 unsigned int(8) num\_tracks;  
 unsigned int(8) aligned;  
 unsigned int(8) single\_init;  
 unsigned int(8) principal;  
 utf8string structural\_brand;  
 utf8string mediaprofile\_brand;  
}

Semantics

This box contains information about the CMAF switching group.

num\_tracks specifies the number of alternative tracks grouped by this CMAF switching group.

A track grouped by this CMAF switching track group is a track with alernate\_groupe equal to the ID of this switching group.

aligned indicates an aligned CMAF Switching set when it is set to a 1. Otherwise, it shall be set to 0. Other values are reserved.

Single\_init indicates the CMAF Switching set with a single initialisation segment for all tracks when it is set to a 1. Otherwise, it shall be set to 0. Other values are reserved.

principal specifies the track\_id of the track that its initialisation segment can be used for the initialisation of any track in this switching group.

structural\_brand specifies the CMAF structural brand of all tracks of this switching group.

mediaprofile\_brand specifies the media profile brand that all tracks of this switching group conform to.

# Use Case 2: Switching across AdaptationSets

Table 3: Terms for bitrate switching

|  |  |  |
| --- | --- | --- |
| *ISOBMFF* | *DASH* | *CMAF* |
| Switch Group | AdaptationSet | Switching Set |
| Track | Representation | Track |
|  | Switchable AdaptationSet |  |

Reference to DASH: ISO/IEC 23009-1, clause 5.3.3.5

## Example

To illustrate one simple use case, Code 1 provides a simple MPD excerpt with tow AdaptationSets with each containing two Representations. Both AdaptationSetds enable switching between each other.

Figure 1 shows how this manifest snippet may be translated into a box structure if alternate\_group are used. Note that the AdaptationSet swithcign descriptor is not translated in this case.

Figure 2 shows how this manifest snippet may be translated into a box structure if track grouping is used. In this case, the AdaptationSet switching descriptor is translated into a new box as provided in 4.2.

Code 1: Example MPD with AdaptationSet switching

<AdaptationSet id="1" lang="en">

    <SupplementalProperty schemeIdUri="urn:mpeg:dash:adaptation-set-switching:2016" value="2"/>

    <Representation id="11">

    <Representation id="12">

</AdaptationSet>

<AdaptationSet id="2" lang="en">

    <SupplementalProperty schemeIdUri="urn:mpeg:dash:adaptation-set-switching:2016" value="1"/>

    <Representation id="21">

    <Representation id="22">

</AdaptationSet>

Figure 3: Box-Strucutre for AdaptationSet switching using alternate\_group

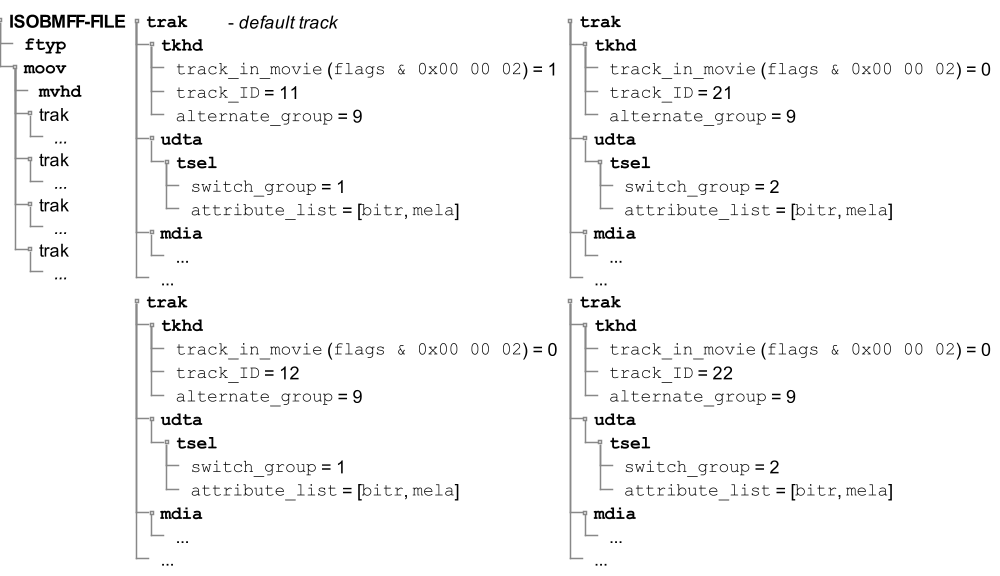
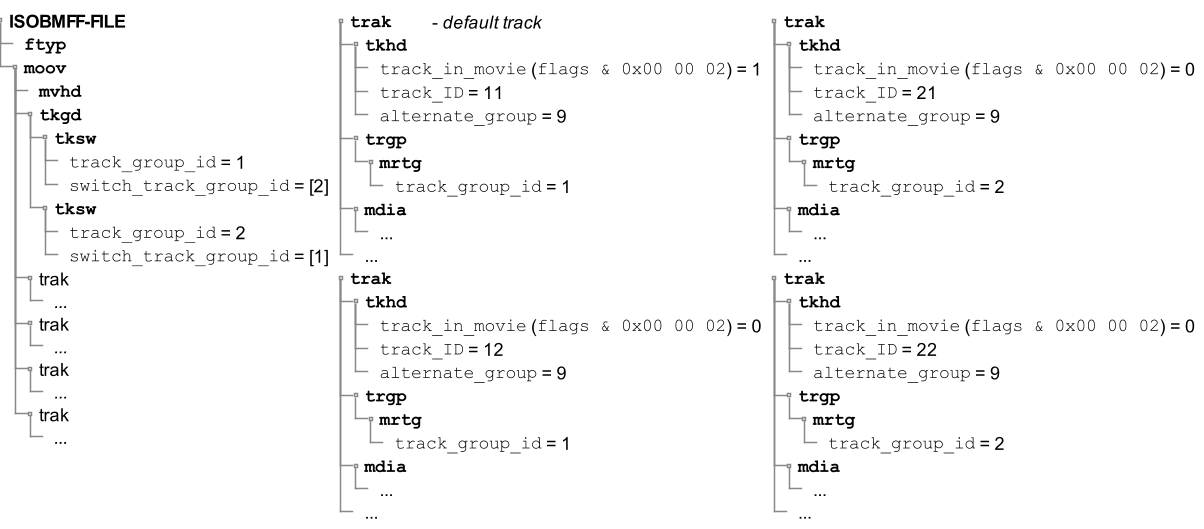


Figure 3: Box-Strucutre for AdaptationSet switching using track groups



## Request

It may be desirable to switch between content offerings provided in two or more multirate groups. One example could be two different viewpoints the content provider wants to offer to end user to select from seamlessly. This has to work on top of seamless switching amongst different bitrates within the multirate group.

Since this functionality applies to track groups, the newly introduced concept of TrackGroupEntryType boxes may be utilized to describe the switching options one (multirate) track group may have, with one of these derived boxes added as “description” to the source group and providing a list of ids of all destination groups a seamless switch is possible to.

In MPEG-DASH, this is implemented by a supplemental property descriptor with an assigned schemeIdUri in section 5.3.3.5 of ISO/IEC 23009-1.

To further generalize the track grouping concept, it may be desirable for applications to group not only tracks into track groups, but to form a superordinate group of track groups.

This can be easily accomplished by defining a group of track groups entry box as derivative of the track group entry box. While such a box may reference all containing track groups through their identifiers, this group should also announce its own new identifier, making itself a new group in the context of an ISOBMFF file. Further boxes may be used to describe the new group of track groups.

This concept enables, when enhanced with appropriate track\_group\_entry\_types, the definition of CMAF aligned switching sets and CMAF selection sets.

## Proposal

**8.3.5 Track group description box**

**8.3.5.4 Track group entry definitions**

**8.3.5.4.4 Switching track group entry box**

**8.3.5.4.4.1 Definition**

Box Type: 'tgsw'  
Container: TrackGroupDescriptionBox  
Mandatory: No  
Quantity: Zero or More

A descriptor indicating the possibility of (seamless) playback switching from one track to one or more other tracks.

[Ed.Note: For the previous paragraph, there seems to be an editing mistake in related contribution. Created relevant text by editor.]

Note: This enables the same functionality as per ISO/IEC 23009-1, section 5.3.3.5, “Switching across Adaptation Sets”.

**8.3.5.4.4.2 Syntax**

aligned(8) class SwitchingTrackGroupEntryBox  
 extends TrackGroupEntryBox('assw', version=0, flags)  
{  
 unsigned int(8) switch\_track\_group\_id[];  
}

**8.3.5.4.4.3 Semantics**

switch\_track\_group\_id is an array of track\_group\_ids referencing multirate track groups and indicating the switch destination track groups.

Using the switchingTrackGroupBox indicates that a player may switch playback from the multirate track group with the same track\_group\_id to any of those multirate track groups listed in the switch\_track\_group\_id array.

**8.3.5.4.5 Group of track group entry box**

**8.3.5.4.5.1 Definition**

Box Type: 'gtgr'  
Container: TrackGroupDescriptionBox  
Mandatory: No  
Quantity: Zero or More

When grouping of track groups is desired, the GroupOfTrackGroups box may be used.

The formed group of track groups is again considered to be a track group and can be referenced by its track\_group\_id.

**8.3.5.4.5.2 Syntax**

aligned(8) class GroupOfTrackGroupEntryBox  
 extends TrackGroupEntryBox('gtgr', version=0, flags)  
{  
 unsigned int(8) num\_tracks;  
 unsigned int(8) group\_track\_group\_id[num\_tracks];  
 // Boxes describing the group of track groups  
}

**8.3.5.4.5.3 Semantics**

group\_track\_group\_id is an array of track\_group\_ids referencing the track\_group\_ids of all track groups to be group by the new group of track groups.

The track\_group\_id of this Box shall not be set to a value of a track\_group\_id of any track in the file, but is used to define a new track group.

# Use Case 3: Dependent Tracks

Table 4: Terms for dependent tracks

|  |  |  |
| --- | --- | --- |
| *ISOBMFF* | *DASH* | *CMAF* |
| Track | Representation | Track |
| Alternate Group | Group | Selection Set |
| Track Reference | Dependent Representation | Dependent CMAF Track |

In DASH, dependent Representations are signalled using the @dependencyId attribute on Representations.

In ISOBMFF, dependencies between tracks are implemented using track references (TrackReferenceBox, section 8.3.3). The TrackReferenceBox, if present, is contained in the TrackBox of the track and references others tracks by their respective track\_ids. The type of the reference is provided by the reference\_type from a list of 12 types. Multiple references from one track are possible.

## Request

In MPEG DASH, dependencies between ~~tracks~~ Representations are only signalled on Representation level, not on AdaptationSets. This implies that such dependencies do not require to be implemented on track groups. Instead, the already existing mechanism provided by the TrackReferenceBox can be used together with a to-be-defined track\_reference\_type.

## Proposal

*Add the following entries to the list of track\_reference\_types:*

* + - 1. Semantics

The reference\_type shall be set to one of the following values, or a value registered or from a derived specification or registration:

* depc indicates that the referenced track(s) may contain media data required for decoding of the track containing the track reference, i.e., it should only be used if the referenced media track is used. The referenced tracks shall be tracks for the same type. The 'depc' dependency can, for example, be used for indicating the dependencies between “normal audio” and “associated audio” tracks.

# Preselections

## Preselections and alternate groups

### Request

* Specify the following:
  + When any two tracks belong to the same alternate group, they shall not both have a PreselectionGroupBox with the same value of track\_group\_id.
  + When processing a preselection to which a track contributes, that track can be replaced by any other track in the same alternate group that track belongs to.

### Proposal

*Proposed changes:*

* + - * 1. Preselection group box

Definition

The presence of a TrackGroupTypeBox with track\_group\_type equal to 'pres', which is also referred to as a PreselectionGroupBox, in a track indicates that this track contributes to a preselection.

All the tracks that have a track group with track\_group\_type equal to 'pres' and a particular value of track\_group\_id are part of the same preselection. The particular value of track\_group\_id is also referred to as the ID of the preselection.

NOTE 1 This means that a preselection is uniquely identified by the track\_group\_id of the track group.

When any two tracks belong to the same alternate group, they shall not both have a PreselectionGroupBox with the same value of track\_group\_id. However, when processing a preselection to which a track contributes, that track can be replaced by any other track in the same alternate group that track belongs to.

When multiple tracks contribute to a preselection, the optionally present PreselectionProcessingBox provides information on how to process the track containing this box in the context of the preselection and relative to other tracks. Consequently, the content of the PreselectionProcessingBox may differ for each track within a preselection.

NOTE 2 Preselections consisting of only one track do not require any track-related processing. In this case, the PreselectionProcessingBox is typically not present in the PreselectionGroupBox.

Semantics

This box contains information on what experience is available when this preselection is selected.

Boxes suitable to describe a preselection include but are not limited to the following list of boxes defined in this document:

* AudioElementBox (subclause 12.2.9)
* AudioElementSelectionBox (subclause 12.2.13)
* ExtendedLanguageBox (subclause 8.4.6)
* UserDataBox (subclause 8.10.1)
* KindBox (subclause 8.10.4)
* LabelBox (subclause 8.10.5)
* AudioRenderingIndicationBox (subclause 12.2.8)
* ChannelLayout (subclause 12.2.4)

If a UserDataBox is contained in a PreselectionTrackGroupEntryBox, then it shall not carry any of the above boxes.

num\_tracks specifies the number of ~~non-alternative~~ tracks grouped by this preselection track group.

A track grouped by this preselection track group is a track that has the 'pres' track group with track\_group\_id equal to the ID of this preselection.  
~~The number of non-alternative tracks grouped by this preselection track group is the sum of the following:~~

* ~~the number of tracks that have alternate\_group equal to 0 and are grouped by this preselection track group,~~
* ~~the number of unique non-zero alternate\_group values in all tracks that are grouped by this preselection track group.~~

The value of num\_tracks shall be greater than or equal to the number of ~~non-alternative~~ tracks grouped by this preselection track group in this file.  
A value equal to 0 indicates that the number of tracks grouped by this track group is unknown or not essential for processing the track group.  
NOTE 1 The value of num\_tracks can be greater than the number of ~~non-alternative~~ tracks containing a PreselectionGroupBox with the same track\_group\_id in this file when the preselection is split into multiple files.  
NOTE 2 When a player has access to fewer ~~non-alternative~~ tracks grouped by this preselection track group than indicated by num\_tracks, the player might need to omit the tracks grouped by this preselection track group.

...

## Preselection Main track signaling

### Request

Add a flag to the PreselectionGroupBox for signalling of the indication of the main track in a preselection.

### Proposal

* + - * 1. Preselection group box

Definition

Syntax

aligned(8) class PreselectionGroupBox extends TrackGroupTypeBox('pres')  
{  
 unsigned int(1) main\_track\_in\_preselection\_flag;  
 unsigned int(7) reserved;  
 PreselectionProcessingBox preselection\_processing; // optional  
}

Semantics

main\_track\_in\_preselection\_flag equal to 1 indicates that the track is the main track in the preselection. main\_track\_in\_preselection\_flag equal to 0 indicates that the track is not the main track in the preselection. One and only one track among all tracks contributing to a preselection shall have main\_track\_in\_preselection\_flag equal to 1.

NOTE When the media content is encapsulated as a DASH Media Presentation, the Adaptation Set containing the Representation for the main track in a preselection would be the Main Adaptation Set for the DASH Preselection.

preselection\_processing is an instance of the PreselectionProcessingBox, providing information needed for processing the containing track in the context of the preselection and relative to other tracks.

# Other topics

## Terminology in ISOBMFF 8th Edition

### Request

Regarding the use of the undefined term "media presentation" in the definition of "preselection"

**preselection**set of one or more tracks representing one version of the media presentation for simultaneous decoding or presentation

The term "media presentation" is used without being defined; note that this is a DASH term, and it probably sneaked into the ISOBMFF without being noticed. Maybe use "media content" instead? Although not defined either, "media content" appears to be more like a plain phrase that can probably be used without a definition, and it is indeed used at least six times in this DIS text.

### Proposal

* + 1. preselection

set of one or more tracks representing one version of the media presentation for simultaneous decoding or presentation [Ed. (YK): The term "media presentation" is used without being defined; note that this is a DASH term, and it probably sneaked into the ISOBMFF without being noticed. Maybe use "media content" instead? Although not defined either, "media content" appears to be more like a plain phrase that can probably be used without a definition, and it is indeed used at least six times in this DIS text.]

## New CMAF Profile

During MPEG#139, it was decided to collect new ideas around a new CMAF profile that allows to align ISO BMFF functionalities with CMAF and possibly carrying this forward to DASH, for example using the DASH Profile for CMAF content as defined in the 5th edition. It was proposed and agreed to:

1. Create a new brand ‘cmf3’ that requires  
   a. The track header fulfills all requirements as if the CMAF presentation would be included in a single file  
   b. that each track in the Switching Set has a unique track\_id (as if they would be included in a single file)  
   c. a switching set is treated as an alternate group and each switching set gets assigned a unique alternate\_group identifier.
2. Use the text in clause 7.8 as a baseline for the new brand,
3. Identify if additional functionalities need to be carried over from the file format.

The following change 3 was considered

Change 3

**7.8 The structural CMAF Brand 'cmf1'**

**7.8.1 General**

A CMAF track conforming to the CMAF structural brand 'cmf1' shall conform to constraints of the CMAF structural brand 'cmfc' and all remaining constraints in subclause 7.8.

These constraints introduced to signal that the CMAF tracks and CMAF switching set track headers are conforming as if all CMAF Tracks of the presentation also conforming to this brand would be included in a single ISO BMFF file.

**7.8.2 Track Header Box ('tkhd')**

CMAF TrackHeaderBoxes shall conform to subclause 7.5.4 with the following additional constraints.

— The track\_ID is set to a unique identifier over the entire lifetime of this CMAF presentation;

— All tracks within one CMAF Switching Set have alternate\_group set to the same value. Each CMAF Switching Set in the CMAF Presentation has a unique value for the alternate\_group

— For a video track, every decoder output signal shall have decoded and cropped image size in video spatial samples measured on a uniformly sampled square grid identical to the value of width and height defined in the Track Header.

End of Change 3

It was also considered to collect additional considerations and requirements for such a profile.

This document [i.e. the original contribution] provides a few design considerations that should support a more structured approach to introduce new features into modern MPEG-based streaming systems relying on ISO BMFF, CMAF and DASH.

## Additional considerations

When generating this document, the editor has discovered the following items not considered during MPEG#140, but deemed being still worth for consideration.

### Editorial changes in ISOBMFF 8th Edition

M60771 lists additional editorial edits in the “attachment” document. Worth considering them in DIS 8th Edition of 14496-12; not discussed at MPEG#140.

### On segment\_order

Per resolution form MPEG#140, segment\_order gets removed from DIS 8th Edition of 14496-12; mapping for this DASH attributes needs to get described from other ISOBMFF fields (track\_order,…)

# Related documents

WG03N0659 (Serial Number 21788):  
Technology under consideration for MPD writing guidelines based on ISOBMFF  
Online, July 2022

N17866:  
WD of ISO/IEC TR 23009-7 Delivery of CMAF contents with DASH  
Ljubljana, July 2018

[Ed.Note: Do we intent do merge these documents with this one?]