Dolby Digital Plus online delivery content creation

System development manual

Version 1.4.1
31 July 2018
Confidential information

Copyright

Dolby Laboratories, Inc.
1275 Market Street
San Francisco, CA 94103-1410 USA
Telephone 415-558-0200
Fax 415-645-4000
http://www.dolby.com

Trademarks

Dolby and the double-D symbol are registered trademarks of Dolby Laboratories

The following are trademarks of Dolby Laboratories:

Dialogue Intelligence™
Dolby®
Dolby Advanced Audio™
Dolby Atmos®
Dolby Audio™
Dolby Cinema™
Dolby Digital Plus™
Dolby Digital Plus Advanced Audio™
Dolby Digital Plus Home Theater™
Dolby Home Theater®
Dolby Theatre®
Dolby Vision™
Dolby Voice®
Feel Every Dimension™
Feel Every Dimension in Dolby™
Feel Every Dimension in Dolby Atmos™
MLP Lossless™
Pro Logic®
Surround EX™

All other trademarks remain the property of their respective owners.

Confidential information

Contents

1 Introduction to the content creation system development manual

This documentation supports the system designers and engineers who are building a content creation product. It lists the requirements for creating Dolby Digital Plus content in such a product and provides guidance to help in meeting the requirements.

1.1 Dolby Digital Plus Online Delivery Kit

This documentation is part of the Dolby Digital Plus Online Delivery Kit. For a Dolby Digital Plus online delivery product (content creation or playback system) to receive approval from Dolby Laboratories, it must meet the requirements listed in the corresponding system development manual for the technology or technologies the product supports. A summary of the requirements is available in *Requirements and recommendations*. Dolby Laboratories provides test materials to evaluate whether the product satisfies requirements and recommendations. Test materials are contained in the Dolby Digital Plus Online Delivery Kit.

1.2 New in this version

These updates have been made to the latest version of this documentation.

- In a Media Presentation Description (MPD) file for Dolby Digital Plus bitstreams carrying Dolby Atmos content, the codec attribute of an adaptation set must include the entry ec-3 instead of ec+3.
- In a HTTP Live Streaming (HLS) playlist for Dolby Digital Plus bitstreams carrying Dolby Atmos content, the codec attribute of the EXT-X-STREAM-INF playlist parameter must be set to ec-3 instead of ec+3. The CHANNELS attribute of the #EXT-X-MEDIA parameter must include the JOC identifier.
- Dolby Digital Plus bitstreams carrying Dolby Atmos content in MPEG-2 transport streams are not supported now. Descriptions about signaling Dolby Digital Plus bitstreams carrying Dolby Atmos content in an MPEG-2 transport stream have been removed.

1.3 Standards and Dolby documents

Standards and Dolby documents provide additional information to assist you in designing your product.

Standards

- ETSI TS 102 366 v1.4.1 (2017), *Digital Audio Compression (AC-3, Enhanced AC-3) Standard*, Annex E (normative), available from http://www.etsi.org. This document describes the Dolby Digital Plus (E-AC-3) bitstream syntax.
- ISO/IEC 14496-12:2012, *Information Technology—Coding of Audio-Visual Objects, Part 12: ISO Base Media File Format*, available from http://www.iso.org. This documentation is Part 12 of the MPEG-4 specification and describes storage of content in a media file.
- ISO/IEC 23009-1:2014, *Information Technology—Dynamic Adaptive Streaming over HTTP (DASH)—Part 1: Media Presentation Description and Segment Formats*, available from http://www.iso.org.
- ETSI TS 102 796 v1.4.1, *Hybrid Broadcast Broadband TV*, available from http://www.etsi.org.
- *HTTP Live Streaming—draft-pantos-http-live-streaming-23*, available from https://datatracker.ietf.org. (Search for "pantos".)
- ISO/IEC 13818-1:2013, *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information: Systems*, available from http://www.iso.org.
- RFC 6381, *The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types*, August 2011, available from http://tools.ietf.org/html.
- ISO 639-2:1998, *Codes for the Representation of Names of Languages, Part 2: Alpha-3 Code*, as maintained by the ISO 639/Joint Advisory Committee, available from http://www.iso.org.

1.4 Channel abbreviations

This table lists the channel notations used in this document.

Abbreviation	Channel
L	Left
R	Right
C	Center
Lc	Left Center
Rc	Right Center
LFE	Low-Frequency Effects
Lfh	Left Front Height
Rfh	Right Front Height
S	Mono Surround
Cs	Center Surround
Ls	Left Surround
Rs	Right Surround
Lscr	Left Screen
Rscr	Right Screen
Lsd	Left Surround Direct
Rsd	Right Surround Direct
Lb *	Left Back
Rb *	Right Back

Abbreviation	Channel
Lvh	Left Vertical Height
Cvh	Center Vertical Height
Rvh	Right Vertical Height
Lrs	Left Rear Surround
Rrs	Right Rear Surround
Lw	Left Wide
Rw	Right Wide
Tbl	Top Back Left
Tbr	Top Back Right
Tfl	Top Front Left
Tfr	Top Front Right
Tl	Top Left
Tr	Top Right
Ts	Top Surround
Ltm	Left Top Middle
Rtm	Right Top Middle
Lts	Left Top Surround
Rts	Right Top Surround
SW	Subwoofer speaker output

* Lb and Rb correspond to the SMPTE-defined Lrs and Rrs channels.

1.5 Contacting Dolby

Support services are available to address any questions and to provide advice about integrating Dolby technology into your product.

For product design or testing, contact Dolby at systemsupport@dolby.com. By utilizing Dolby expertise, especially during the design process, many problems that might require design revisions before a product is approved can be prevented.

Dolby is also available to review product plans, including preliminary design information, markings, displays, and control and menu layouts, with the goal of preventing problems early in the product development cycle.

If you have comments or feedback about this document, send us an email at documentation@dolby.com.

2 Product overview

A typical product built and tested using the resources in this kit delivers Dolby Digital Plus content in an online container that supports adaptive streaming.

2.1 Supported online containers

Requirements are provided for a product that delivers content using:

- Apple HTTP Live Streaming
- MPEG Dynamic Adaptive Streaming over HTTP (MPEG-DASH)
- ISO base media file format

2.2 Products that include other Dolby technologies

A content creation product may also include other functionality, such as Dolby Digital Plus encoding.

For an encoder product, the requirements in this documentation supplement (and are in addition to) the requirements in the kit for your specific product type. For example, the requirements in:

- Dolby Digital Plus professional encoder system development kit
- Dolby Digital Plus consumer encoder system development kit for software applications

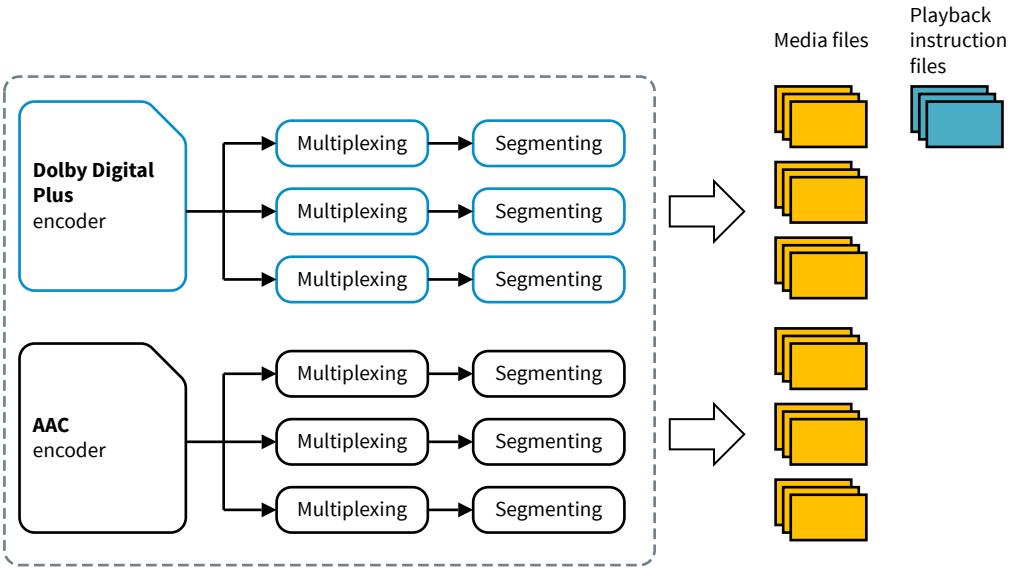
2.3 Adaptive streaming

The most common way to deliver online content is via adaptive streaming over HTTP. A multimedia presentation that is prepared for adaptive streaming is normally encoded in different ways (for example, at different bit rates). Additional or alternative content can also be encoded (for example, to provide audio in additional languages).

To prepare content for adaptive streaming, the content is encoded, packaged in a container format (such as an .mp4 file or an MPEG-2 transport stream), and segmented. One or more files that instruct a client application how to download and play back the content must also be created. These manifest or playlist files include instructions for the client to play back any of the available content that forms a media presentation.

Your product may perform all of these functions, or other products may perform some of the functions, depending on your work flow. For example, your product may encode and multiplex the content, but your work flow uses a third-party segmenter. This diagram gives an overview of the content preparation process.

Figure 1: Adaptive streaming content preparation overview



In an adaptive streaming delivery system, several versions of the content are made available on the server. For example, content is encoded at several quality levels to provide higher and lower bandwith-intensive versions. Content is divided into small segments (literally or virtually). The client checks the delivery conditions (such as available bandwidth and client capabilities) and requests the most appropriate version of the content.

During playback, the client monitors delivery conditions and can switch to a different presentation of the content. Small segments allow the client to switch to a different encoding of the content with little or no interruption to the playback.

The ability to switch between different content streams also allows delivery of alternative versions of the same content (for example, to switch to other camera angles for video, or to an alternative language for the audio).

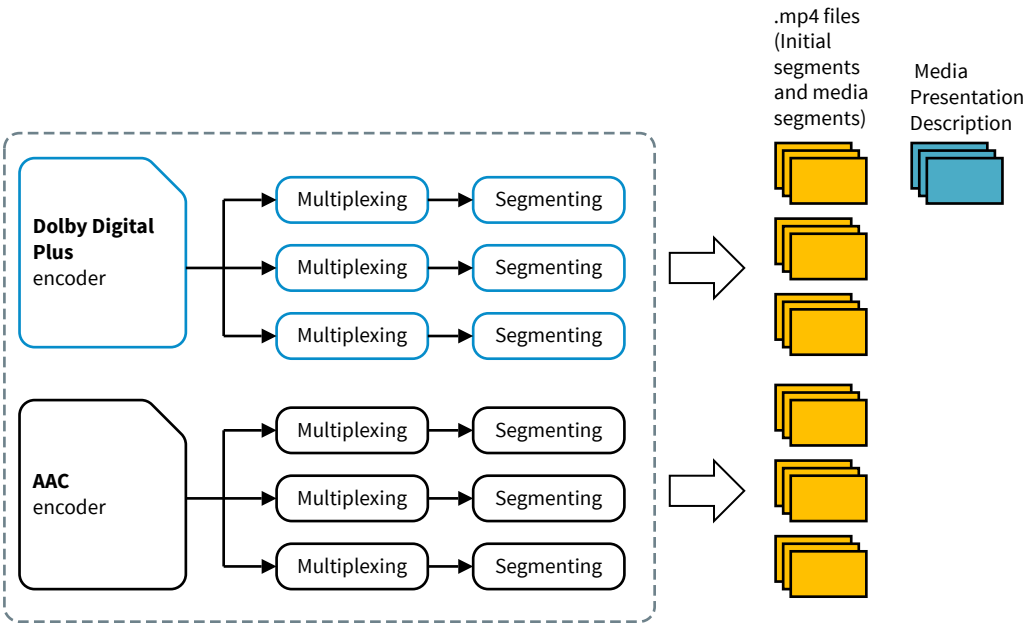
2.3.1 MPEG-DASH

MPEG-DASH is an adaptive bit-rate streaming technology where a media presentation is partitioned into segments and delivered from a server to a client via HTTP.

The content exists on the server in two parts: a Media Presentation Description (MPD), and segmented content. The MPD is a manifest of the available content, its various alternatives, their URL addresses, and other characteristics (including timing information and media characteristics such as video resolution and bit rates). The content fragments are the actual media bitstreams in the form of chunks, in single or multiple files.

The MPEG-DASH specification defines the MPD and the segment formats only. The delivery of the MPD and the media-encoding formats containing the segments—as well as the client behavior for segment retrieval, adaptation heuristics, and content playback—are not regulated by the MPEG-DASH standard. This diagram shows the content preparation process for MPEG-DASH.

Figure 2: Adaptive streaming content preparation for MPEG-DASH

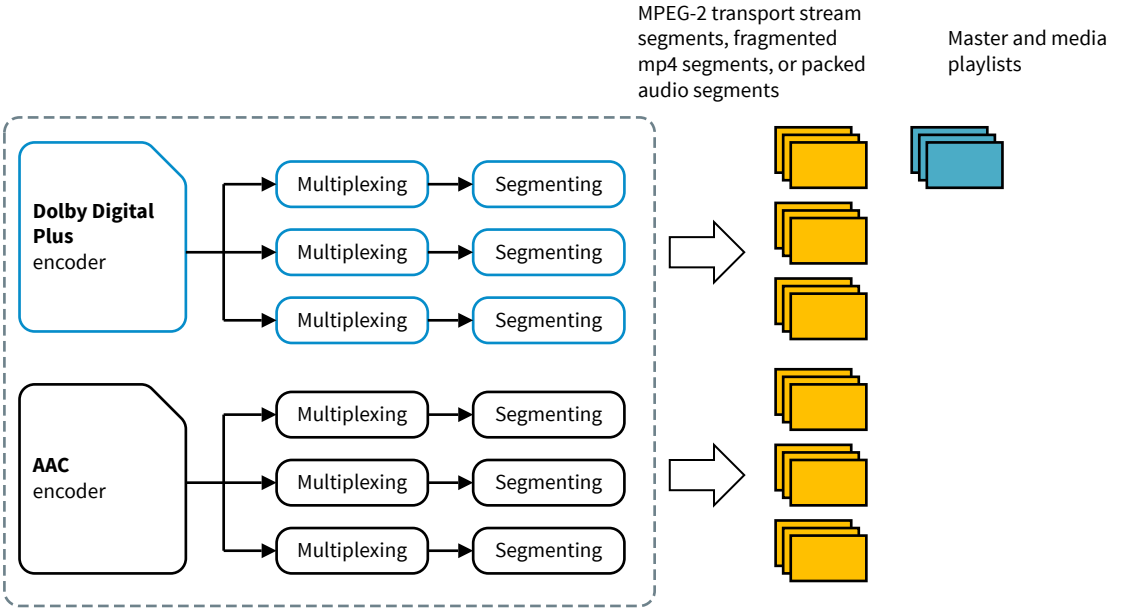


2.3.2 HTTP Live Streaming

HTTP Live Streaming is a protocol developed by Apple for transferring streams of media data. HTTP Live Streaming supports both live broadcasts and prerecorded content (video on demand).

HTTP Live Streaming uses segmented MPEG-2 transport streams or fragmented MP4 files for streaming a multimedia presentation that contains both audio and video content, and uses segmented packed audio or fragmented MP4 files for streaming audio-only content. These segmented streams are created from unbounded MPEG-2 transport streams, packed audio, or fragmented MP4 files by using a segmentation tool. This diagram shows the content preparation process for HTTP Live Streaming.

Figure 3: Adaptive streaming content preparation for HTTP Live Streaming



3 Requirements

An online delivery content creation product must meet all defined requirements for the online delivery formats it supports.

3.1 Dolby Digital Plus bitstream requirements

This section contains the requirements for a Dolby Digital Plus bitstream used in an online delivery format.

DR-5
Supports all valid data rates for Dolby Digital Plus, up to and including 3,024 kbps.

Mux-2
The `fscod` parameter is set to `0` (48 kHz) for all substreams present in the Dolby Digital Plus bitstream.

Mux-3
All substreams present in the Dolby Digital Plus bitstream have the same `numblkscod` value.

Mux-4
The `bsid` value of the Dolby Digital Plus bitstream remains constant.

Mux-5
No substreams within the Dolby Digital Plus bitstream have a `strmtyp` value of `2` (transcoded) or `3` (reserved).

Mux-6
The Dolby Digital Plus bitstream does not have an `acmod` value of `0` (1+1).

Mux-7
The number of independent substreams within the Dolby Digital Plus bitstream remains constant.

Mux-8
The values of `bsmod`, `acmod`, and `lfeon` remain constant within each independent substream of the Dolby Digital Plus bitstream.

Mux-9
The Dolby Digital Plus bitstream is encoded in big-endian format.

Mux-10

The number of dependent substreams within the Dolby Digital Plus bitstream remains constant.

Mux-11

The values of `acmod`, `lfeon`, and `chanmap` remain constant within each dependent substream of the Dolby Digital Plus bitstream.

3.2 MPEG-DASH requirements

A product that prepares Dolby Digital Plus bitstream for online delivery via MPEG-DASH must meet certain requirements.

Mux-37

Ensures that each MP4 sample consists of one Dolby Digital Plus access unit.

Mux-39

Uses the structure and values for the `EC3SampleEntry` and `EC3SpecificBox` boxes, as defined in Annex F of ETSI TS 102 366.

Mux-40

When creating an MP4 file with a Dolby Digital Plus track that contains Dolby Atmos content, uses the extended structure and values for `EC3SpecificBox`.

Mux-45

In the MPD file, the product uses these adaptation set attributes to describe Dolby Digital Plus content: `codecs` and `mimeType`.

Mux-46

In the MPD file, sets the `codecs` attribute value to `ec-3` for a Dolby Digital Plus bitstream.

Mux-47

In the MPD file, sets the `type` element of the `mimeType` attribute to `audio/mp4` for an MP4 file that contains one or more audio tracks, with no accompanying video.

Mux-48

In the MPD file, sets the `type` element of the `mimeType` attribute to `video/mp4` for an MP4 file that contains one or more video tracks.

Mux-49

In the MPD file, uses the `value` attribute of the `AudioChannelConfiguration` descriptor for Dolby Digital Plus bitstreams as specified in `schemeIdUri:tag:dolby.com, 2014:dash:audio_channel_configuration:2011`.

Mux-50
In the MPD file, sets the `AudioChannelConfiguration` descriptor value to the hexadecimal representation of the 16-bit field describing the channel assignment for the referenced bitstream.

Mux-51
A product that includes the `accessibility` descriptor in the MPD file sets the `schemeIdUri` attribute to `urn:tva:metadata:cs:AudioPurposeCS:2007`.

Mux-52
A product that includes the `accessibility` descriptor in the MPD file sets the `termID` attribute to the type of accessible audio service carried in the bitstream.

Mux-53
In the MPD file for Dolby Digital Plus streams that contain Dolby Atmos content, the product includes the `SupplementalProperty` descriptor to correctly describe attributes of the Dolby Atmos content.

3.3 HTTP Live Streaming requirements

A product that prepares Dolby Digital Plus bitstreams for online delivery via HTTP Live Streaming must meet certain requirements.

Mux-54
When creating more than one variant or rendition of a media presentation, produces content segments, corresponding to the same presentation period, that contain an identical number of video and audio access units.

Mux-55
Ensures that each packet payload contains at least one complete Dolby Digital Plus access unit.

Mux-56
Produces transport-stream segments that contain only complete packetized elementary stream (PES) packets.

Mux-57
Does not produce fragmented PES packets in transport-stream segments.

Mux-58
Places the bytes of a Dolby Digital Plus frame in big-endian format within the PES payload.

Mux-59
Byte aligns the elementary streams within the PES payload.

Mux-60
Assembles audio frames in the PES packet payload in the same sequence as they occur in the elementary stream.

Mux-61
Does not fragment access units within a payload, or across multiple payloads.

Mux-62
Sets the stream_type in the packet header to 0x87when the PES contains Dolby Digital Plus audio.

Mux-63
Sets the stream_id in the packet header to 0xBD (private_stream_1).

Mux-64
Includes the E-AC-3_audio_descriptor() in the Program Map Table (PMT) entry for the Dolby Digital Plus stream when the PES contains Dolby Digital Plus audio.

Mux-65
Includes the E-AC-3_audio_descriptor() in the PMT entry for the Dolby Digital Plus stream with additional parameters to indicate the presence of Dolby Atmos content when the PES contains Dolby Digital Plus audio with Dolby Atmos content.

Mux-66
Sets the descriptor_tag in the E-AC-3_audio_descriptor(), to 0xCC.

Mux-67
Sets the descriptor_length field in the E-AC-3_audio_descriptor() to the total number of bytes of the data portion of the E-AC-3_audio_descriptor() following this field.

Mux-68
Sets mainid_flag, asvc_flag, and mixinfoexists fields in the E-AC-3_audio_descriptor() to 0.

Mux-69
Sets the appropriate the substream flags (substream1_flag, substream2_flag, and substream3_flag) in the E-AC-3_audio_descriptor() when the PES contains Dolby Digital Plus audio with additional independent substreams containing associated audio services.

Mux-70
Sets the full_service_flag in the E-AC-3_audio_descriptor() to 1 to indicate whether the audio in substream 0 (and any dependent substreams associated with substream 0) is suitable for presentation alone (full-service audio) or to 0 to indicate that the audio is a partial service that needs to be combined with another audio service before presentation.

Mux-71
Sets the service_type in the E-AC-3_audio_descriptor() to indicate the type of audio service in substream 0 (and any dependent substreams associated with substream 0).

Mux-72
Sets the number_of_channels in the E-AC-3_audio_descriptor() to indicate the channels in substream 0 (and any dependent substreams associated with substream 0).

Mux-73
Sets the language_flag in the E-AC-3_audio_descriptor(), to 1 to indicate that the language field is present in the descriptor or to 0 to indicate that the language field is not present in the descriptor.

Mux-74
Sets the language_flag_2 in the E-AC-3_audio_descriptor() to 0.

Mux-75
Sets the bsid field (if present) in the E-AC-3_audio_descriptor(), to the same value as the bsid parameter in the Dolby Digital Plus bitstream.

Mux-76
Sets the appropriate substream fields (substream1, substream2, or substream3) in the E-AC-3_audio_descriptor(), the product to indicate the type of audio carried in any additional independent substreams.

Mux-77
Sets the language field for each independent substream in the E-AC-3_audio_descriptor() to indicate the language of the substream.

Mux-78
Produces transport stream segments where the presentation time stamp (PTS) of the first PES packet of each audio stream in the segment is equal to or greater than the PTS of the first PES packet of the video stream.

Mux-79
Produces transport stream segments where the time offset between the first video PTS and the first audio PTS is less than 2,880 PTS ticks.

Mux-80
Produces transport stream segments where the time offset between the Audio_In time and Video_In time of a segment (the A/V PTS offset) is identical to the time offset between the Audio_Out time and Video_Out time of the previous segment.

Mux-81
When creating a media presentation that contains audio only, packetizes the presentation as HLS packed audio with ID3 tags or as a fragmented MP4 file.

Mux-82
A product that uses the CODEC attribute of the EXT-X-STREAM-INF playlist parameter sets its value to ec-3 when creating a playlist for a media presentation that contains Dolby Digital Plus.

Mux-83
A product that includes the #EXT-X-MEDIA parameter includes a meaningful NAME attribute when creating a playlist for a media presentation that contains different renditions of Dolby Digital Plus content.

Mux-84
A product that includes the `#EXT-X-MEDIA` parameter in a playlist indicates the main and associated programs using the `<attribute-list>` of the parameter when creating a playlist for a media presentation that contains Dolby Digital Plus audio with one or more associated audio programs.

Mux-85
A product that includes the `#EXT-X-MEDIA` parameter in a playlist indicates the eight-bit substream information using the `DDP-PROGRAM0`, `DDP-PROGRAM1`, `DDP-PROGRAM2`, and `DDP-PROGRAM3` attributes for present substreams. when creating a playlist for a media presentation that contains Dolby Digital Plus audio with one or more associated audio programs.

4 Dolby Digital Plus streams in online delivery formats

The Dolby Digital Plus format provides a high level of flexibility and can support complex bitstream configurations. To minimize impact on decoding devices while ensuring that a wide range of applications can be supported, some constraints are applied to the permitted bitstream structure and to bitstream parameter settings.

4.1 Dolby Digital Plus bitstream structure

A Dolby Digital Plus bitstream is constructed from one or more substreams. Each substream is a sequence of frames that can carry up to 5.1 channels of audio.

The use of multiple substreams allows delivery of a single program with more than 5.1 channels, multiple independent programs, or a combination within a single Dolby Digital Plus bitstream. The substreams are time multiplexed. Section E.2.8 of ETSI TS 102 366 provides detailed information on the use and structure of multiple substreams within a Dolby Digital Plus bitstream.

The frames that make up each substream are constructed from smaller units called blocks, each representing 256 samples of audio from each channel carried in the substream. A frame may contain one, two, three, or six blocks of audio data (representing 256, 512, 768, or 1,536 samples of PCM audio).

If Dolby Atmos content is present in a Dolby Digital Plus bitstream, the `addbsi` field includes extensions specific to Dolby Atmos in the first independent substream (I0) of a 5.1-channel Dolby Digital Plus bitstream. The `complexity_index_type_a` field takes a value of 1 to 16 that indicates the decoding complexity of the Dolby Atmos bitstream. The syntax is shown in this table.

	Syntax	Word size in bits	Identifier	Value
`addbsie`				`1`
`addbsil`				`1`
`addbsi` includes:	`flag_ec3_extension_type_reserved`	`7`	`bslbf`	`0`
	`flag_ec3_extension_type_a`	`1`	`bslbf`	`1`
	`complexity_index_type_a`	`8`	`uimsbf`	`1 - 16`

4.2 Constraints on Dolby Digital Plus bitstreams in online containers

A Dolby Digital Plus bitstream must meet certain conditions to be delivered in any of the supported online containers.

If your product encodes Dolby Digital Plus bitstreams, ensure that your product uses the proper settings to meet these requirements. To prevent subsequent decoding problems, we recommend that a multiplexer reject a bitstream that is not compliant.

A Dolby Digital Plus bitstream must:

- Have a maximum data rate of 3,024 kbps.
- Have a sample rate of 48 kHz for all substreams present in the bitstream. (The value of the `fscod` parameter must be set to `0`.)
- Have a constant number of blocks per frame for all substreams in the bitstream. (The value of the `numblkscod` parameter must be the same for all substreams.)
- Have a constant number of independent substreams.
- Within each independent substream, these bitstream parameters remain constant:
  - `bsid`
  - `bsmod`
  - `acmod`
  - `lfeon`
  - `fscod`
- Have a constant number of dependent substreams.
- Within each dependent substream, these bitstream parameters remain constant:
  - `bsid`
  - `acmod`
  - `lfeon`
  - `fscod`
  - `chanmap`
- Have a `strmtyp` value of `0` or `1`. `strmtyp` values of `2` (transcoded) or `3` (reserved) are not supported.
- Have an `acmod` value of `1` or more. An `acmod` value of `0` (1+1) is not allowed.
- Be encoded in big-endian format.

Note: Dolby Digital Plus decoders must handle both little- and big-endian formats.

5 Dolby Digital Plus and MPEG-DASH

These topics provide guidance to assist you in meeting the requirements for delivering Dolby Digital Plus streams using the MPEG-DASH standard in conjunction with the ISO base media file format.

5.1 Packetization of Dolby Digital Plus bitstreams

This section describes the process of analyzing and preparing a Dolby Digital Plus bitstream for storage within an MPEG-DASH compliant ISO base media file.

As Dolby Digital Plus features a high level of flexibility and potentially complex bitstream configurations, this section defines constraints on Dolby Digital Plus that minimize impact on decoding devices while ensuring that a wide range of applications can be supported. To simplify the design of MPEG-DASH compliant demultiplexers, this section presents a specific method for grouping Dolby Digital Plus frames together for storage within an ISO base media file track sample.

5.1.1 Contents of the MP4 sample

Building an MP4 sample from a Dolby Digital Plus bitstream must meet certain conditions.

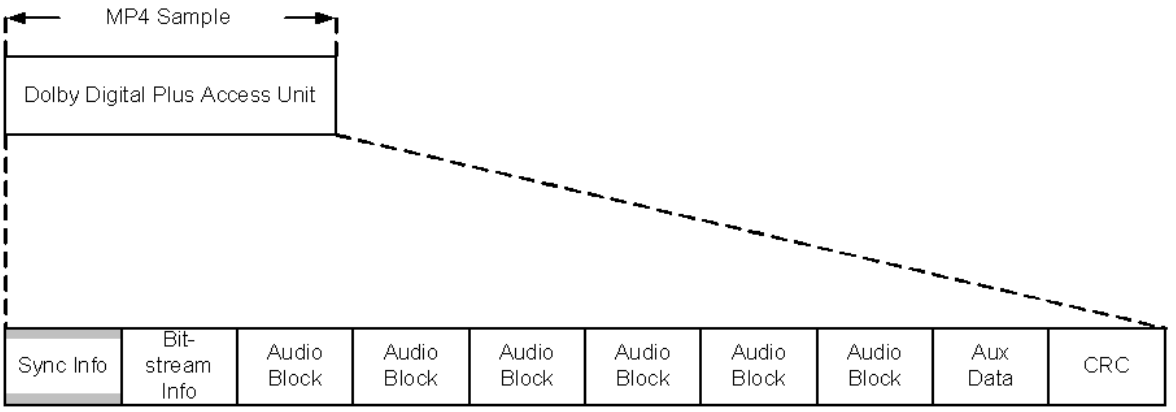
When multiplexing a Dolby Digital Plus bitstream into an MP4 sample, the product must:

- Ensure that each MP4 sample contains one access unit. One access unit consists of all of the parts of the bitstream required by the Dolby Digital Plus decoder to produce 1,536 samples of decoded audio for each audio channel present in the bitstream.
- Ensure that each MP4 sample contains only one complete Dolby Digital Plus access unit.
- Ensure that a Dolby Digital Plus access unit does not span multiple MP4 samples.

These constraints ensure the correct operation of a downstream Dolby Digital Plus decoding device, particularly when this device is capable of converting the Dolby Digital Plus stream to Dolby Digital. This conversion requires the correct set of six blocks of audio data to produce a Dolby Digital frame.

How Dolby Digital Plus data is structured within an MP4 sample depends on the configuration of the Dolby Digital Plus bitstream. This figure shows the construction of an MP4 sample that contains a single Dolby Digital Plus access unit consisting of six audio blocks.

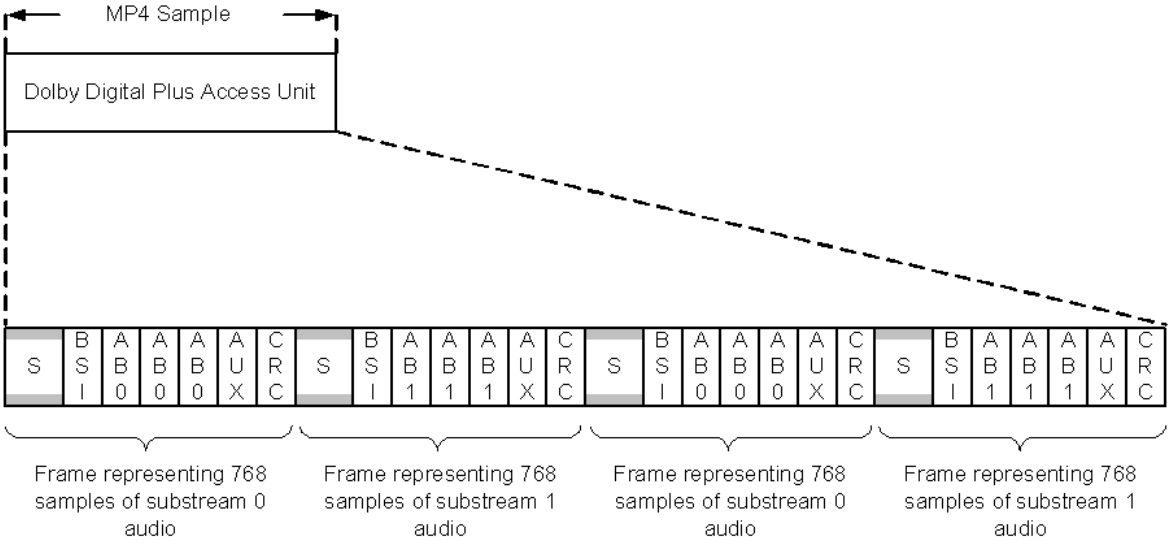
Figure 4: MP4 sample with a single substream with six blocks per frame



The six audio blocks represent 1,536 samples of audio from a single substream (substream 0).

This figure shows an MP4 sample that contains a single access unit of Dolby Digital Plus audio consisting of four frames.

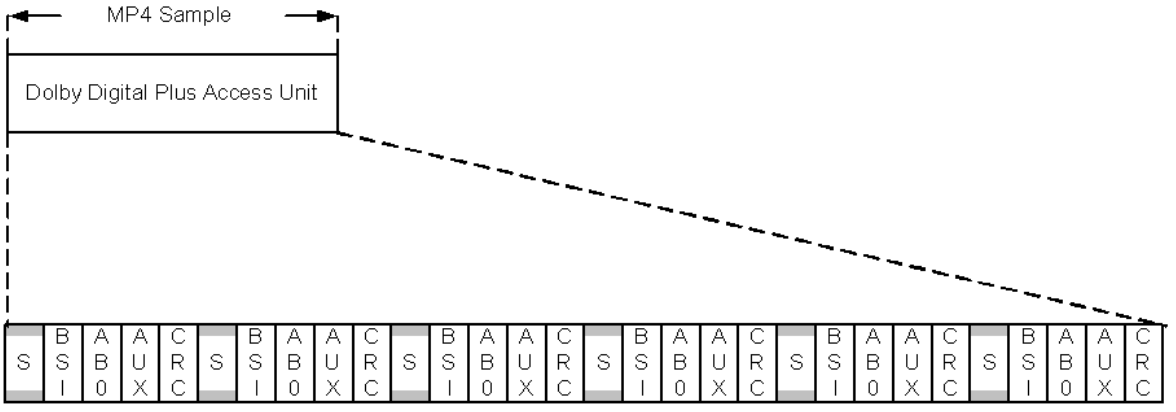
Figure 5: MP4 sample with two substreams with three blocks per frame



Each frame contains three audio blocks (denoted AB0 for substream 0 and AB1 for substream 1), each representing 256 samples of PCM audio from all channels in each substream.

This figure shows an MP4 sample that contains a single Dolby Digital Plus access unit consisting of six frames. Each frame contains one audio block, each representing 256 samples of PCM audio from every channel in the substream.

Figure 6: MP4 sample with a single substream with one block per frame



5.1.2 Parsing a Dolby Digital Plus bitstream

A product must use several bitstream information fields in the Dolby Digital Plus bitstream to identify access units and to properly signal the Dolby Digital Plus bitstream information in the MP4 file.

To locate the start of an access unit, the product checks the numblkscod, strmtyp, and substreamid values. The multiplexing processing logic depends on the number of blocks per frame (indicated by the numblkscod value) in the Dolby Digital Plus bitstream.

If the numblkscod value is 3:

1. Find the first frame with a strmtyp value of 0 and a substreamid value of 0. This indicates the start of an access unit.
2. Use the frame as the first frame of the MP4 sample.
3. Search for the next frame with a strmtyp value of 0 and substreamid value of 0. This indicates the start of the next access unit.
4. Use the frame immediately preceding the frame that starts the next access unit to complete the MP4 sample.
5. Use the frame that starts the next access unit as the first frame of the next MP4 sample.

If the numblkscod value is not 3:

1. Find the first frame with a strmtyp value of 0, a substreamid value of 0, and a convsync flag set to 1.

   This indicates the start of an access unit.
2. Use the frame as the first frame of the MP4 sample.
3. As frames are placed in the MP4 sample, keep track of the number of accumulated audio blocks.
4. Find the next frame with a strmtyp value of 0, a substreamid value of 0, and a convsync flag set to 1.

   If the frame immediately preceding the next frame completes six blocks (or a multiple of six blocks), the current sample is completed. Otherwise, keep searching until six blocks have been accumulated.

5.1.3 Signaling Dolby Digital Plus bitstreams in an MPEG-DASH compliant ISO base media file

The basic structures defined within ISO/IEC 14496-12 to identify audio tracks are used with specific extensions (defined in Annex F of ETSI TS 102 366) to provide detailed information on the characteristics of a Dolby Digital Plus stream.

The information is included in the stbl box of the moov box of the MPEG-DASH compliant ISO base media file. The locations and hierarchy of the boxes that must be included in stbl to identify a Dolby Digital Plus stream within an MPEG-DASH compliant ISO base media file are listed in the table.

Table 1: Sample table box hierarchy for Dolby Digital Plus audio tracks

Nesting Level				
4	5	6	7	Reference
stbl				ISO/IEC 14496-12
	stsd			
		EC3SampleEntry (header type set to ec-3)		Annex F of ETSI TS 102 366
			EC3SpecificBox (header type set to dec3)	Annex F of ETSI TS 102 366 and *Deriving the contents of the EC3Specific box* section of this information set
	stts			ISO/IEC 14496-12
	stsc			
	stsz			
	stz2			
	stco			
	co64			

In the table, the value of the nesting level provided for each box is based on the structure of the complete MPEG-DASH file, beginning with a nesting value of `0` for the ftyp and moov boxes. The table also includes a reference to where each box is defined.

For an unencrypted MPEG-DASH file, the EC3SampleEntry box header type value is ec-3. The value of the EC3SpecificBox header type is dec3.

Related information

Deriving the contents of the EC3SpecificBox on page 24

5.1.4 Signaling Dolby Digital Plus bitstreams with Dolby Atmos content in an ISO base media format file

In addition to the definitions in Annex F of ETSI TS 102 366, other information must be added to the EC3SpecificBox in a backward-compatible way to describe the characteristics of the Dolby Atmos content carried by a Dolby Digital Plus stream.

The extensions specific to audio objects are highlighted in the table. These extensions must be used to signal the presence of the Dolby Atmos content within an ISO base media file.

Syntax	Word size in bits	Identifier
EC3SpecificBox ()		
{		
BoxHeader.Size.....	32	uimsbf
BoxHeader.Type.....	32	uimsbf
data_rate.....	13	uimsbf
num_ind_sub.....	3	uimsbf
for (i = 0; i < num_ind_sub + 1; i++)		
{		
fscod.....	2	uimsbf

Syntax	Word size in bits	Identifier
bsid.....	5	uimsbf
reserved.....	1	bslbf
asvc.....	1	bslbf
bsmod.....	3	uimsbf
acmod.....	3	uimsbf
lfeon.....	1	bslbf
reserved.....	3	uimsbf
num_dep_sub.....	4	uimsbf
if (num_dep_sub > 0)		
{		
chan_loc.....	9	uimsbf
}		
else		
{		
reserved.....	1	bslbf
}		
}		
flag_ec3_extension_type_reserved.....	7	bslbf
flag_ec3_extension_type_a.....	1	bslbf
if (flag_ec3_extension_type_a == 1)		
{		
complexity_index_type_a.....	8	umisbf
}		
additional_data[ ].....	8	uimsbf
}		

These extensions are Dolby Atmos related.

flag_ec3_extension_type_a

The one-bit field must be set to 1 to indicate that Dolby Atmos content is carried in the first independent substream (I0) of a 5.1 Dolby Digital Plus stream.

flag_ec3_extension_type_reserved

This seven-bit field must be reserved for future application.

For content creation referring to this version of the specification, the reserved bits must be set to 0. For playback referring to this version of the specification, the reserved bits shall be ignored.

complexity_index_type_a

This eight-bit field indicates the decoding complexity of a Dolby Digital Plus bitstream carrying Dolby Atmos content. If the flag_ec3_extension_type_a bit in the addbsi has a value of 0, the complexity_index_type_a will not be available in the E-AC-3 descriptor. If the

flag_ec3_extension_type_a bit in the addbsi has a value of 1, the complexity_index_type_a in the E-AC-3 descriptor must take the value of the complexity_index_type_a field found in the addbsi of the Dolby Digital Plus bitstream.

A larger value indicates higher complexity.

additional_data[ ]

These optional bytes must be reserved for future use.

Other fields must conform to Annex F of ETSI TS 102 366.

5.1.5 Deriving the contents of the EC3SpecificBox

This topic provides additional details on the information specific to Dolby Digital Plus and how information is derived from the Dolby Digital Plus bitstream parameters.

The data inside the EC3SpecificBox can be calculated by using information in the Dolby Digital Plus bitstream parameters, including:

- frmsiz
- fscod
- numblkscod
- substreamid
- strmtyp
- bsmod
- chanmap
- addbsi

data_rate

The data_rate parameter indicates the data rate (in kbps) of the entire Dolby Digital Plus bitstream. The value is the sum of the data rates of all the substreams in the Dolby Digital Plus bitstream. When a bitstream uses variable data-rate encoding, data_rate indicates the maximum data rate of the bitstream.

The data rate of each substream is calculated using this equation:

$$\text{data\_rate\_sub} = \frac{(\text{frmsiz}+1)*\text{fs}}{\text{numblks}*16}$$

In this equation:

- frmsiz is the value of the frmsiz parameter in the Dolby Digital Plus frame.
- fs is the sampling frequency of the Dolby Digital Plus bitstream (in kHz). (The fs value is derived from the fscod parameter in the Dolby Digital Plus frame.)
- numblks is the number of audio blocks per frame. (The numblks value is derived from the numblkscod parameter in the Dolby Digital Plus frame.)

num_ind_sub

The num_ind_sub parameter indicates the number of independent substreams present in the Dolby Digital Plus bitstream. The value of num_ind_sub is equal to the value of the substreamid parameter found in the last independent substream of the bitstream. During bitstream parsing,

this will be the frame with a strmtyp value of 0 that precedes the frame with both a strmtyp value of 0 and a substreamid value of 0 (indicating that this frame belongs to the first independent substream of the bitstream).

bsmod

The bsmod parameter is optional in a Dolby Digital Plus bitstream. Detecting it requires additional parsing.

If the Dolby Digital Plus bitstream contains only one independent substream, inclusion of the value of bsmod in the EC3SpecificBox is optional. If parsing yields more than one independent substream, the value of bsmod must be included in the EC3SpecificBox for each substream to ensure that the system parsing the DASH-compliant file can quickly identify the audio services present in the bitstream.

This field itself is not optional. When there is no bsmod, the value must be set to zero.

num_dep_sub

The num_dep_sub parameter indicates the number of dependent substreams associated with an independent substream. Its value is equal to the value of the substreamid parameter found in the frame with a strmtyp value of 1 (that is, in the dependent substream) immediately preceding a frame with a strmtyp value of 0 (that is, in the independent substream).

chan_loc

The chan_loc field indicates channel locations (beyond the standard 5.1 channels) that are carried by dependent substreams associated with an independent substream. The contents of the chan_loc field are determined by parsing the chanmap bit field in every dependent substream associated with a particular independent substream, and setting the corresponding channel locations in the chan_loc field to a value of 1.

Because this field is used by the system only to indicate the unique channel locations present in the bitstream, it is not necessary to reflect replacement channels in this field. Therefore, duplicate channel locations in the chanmap field indicate replacement channels and can be ignored.

flag_ec3_extension_type_a

The flag_ec3_extension_type_a parameter indicates whether Dolby Atmos content is present in a bitstream. Its value is equal to the value of the flag_ec3_extension_type_a bit found in the addbsi.

complexity_index_type_a

The complexity_index_type_a parameter indicates the decoding complexity of a Dolby Digital Plus bitstream carrying Dolby Atmos content. If the flag_ec3_extension_type_a bit in the addbsi has a value of 0, the complexity_index_type_a will not be available in the EC3SpecificBox. If the flag_ec3_extension_type_a bit in the addbsi has a value of 1, the complexity_index_type_a in the EC3SpecificBox shall take the value of the complexity_index_type_a field found in the addbsi of the bitstream.

5.1.6 Packetization of Dolby Digital Plus bitstreams for the HbbTV environment

The HbbTV environment supports only one profile, the HbbTV ISO base media file format Live profile. This profile supports both live and on-demand streaming of ISO base media file format content.

The HbbTV ISO base media file format Live profile is based on the MPEG-DASH ISO base media file format Live profile, and all of the requirements and restrictions for the MPEG-DASH Live profile apply.

All Dolby Digital Plus bitstreams encapsulated in an HbbTV ISO base media file format file must follow the same requirements in the *Constraints on Dolby Digital bitstreams in online containers* section.

In addition, the HbbTV ISO base media file format file must meet these conditions:

- The size of the `moov` box must not exceed 2.5 MB (to prevent slow start-up times for broadband connections with a small bandwidth).
- The large size field may be used. The size of the box must not exceed 4 GB.

Related information

Constraints on Dolby Digital Plus bitstreams in online containers on page 17

5.2 Media Presentation Description

A Media Presentation Description (MPD) is a hierarchical XML document that provides information for an MPEG-DASH client about the available content for a media presentation.

An MPD includes:

- The sequence of periods that make up a media presentation.
- The available adaptation sets within each period that contain encoded versions of media content. For example, one adaptation set can contain video, another can contain audio, and a third can contain a different language audio track or an audio description.
- The representations contained in each adaptation set (for example, content rendered for different bandwidths).
- The URLs for media segments contained in each representation.

The media engine in your product uses the information in the MPD for accessing media segments containing the actual audio and video content and controlling playback.

5.2.1 Media Presentation Description with Dolby Digital Plus

This section defines the values that enable an MPD to properly describe a Dolby Digital Plus bitstream.

Although the syntax of the MPD is capable of using common XML elements to describe almost any media format, the encoding type and the configuration of a Dolby Digital Plus bitstream that is part of a content presentation constrain the parameter values of some of these elements.

The MPD supports these scenarios:

- Media presentations that consist of a single Dolby Digital Plus bitstream

- Media presentations that consist of multiple Dolby Digital Plus bitstreams, with each bitstream stored in a separate MP4 file or segment file

It is possible for the MPD to describe multiple audio services delivered using multiple Dolby Digital Plus bitstreams (for example, a main audio service and an associated audio service that are intended to be decoded and then mixed together).

Adaptation sets

An adaptation set describes the overall media presentation.

The media presentation typically consists of multiple instances of the same audio, video, or audio/video content, with each instance encoded at a different data rate. A representation describes the parameters of each individual encoding of an adaptation set:

- The codecs attribute is required. It specifies the codecs used to encode all representations within the adaptation set. For all Dolby Digital Plus bitstreams, the codecs attribute must include the entry ec-3[1].

  For Dolby Digital Plus bitstreams that carry Dolby Atmos content, the corresponding representations must include additional SupplementalProperty descriptors.

- The mimeType attribute describes the encapsulation format used to store the Dolby Digital Plus bitstreams present in the adaptation set. For adaptation sets that conform to ISO/IEC 14496-12, the mimeType attribute must be set to audio/mp4 (for ISO base media files that contain a Dolby Digital Plus audio track but no accompanying video track).

In some applications, multiple Dolby Digital Plus bitstreams may be used to simultaneously deliver different audio elements of the overall media presentation. For example, one bitstream carries a main audio service (the main audio), and a second bitstream carries an associated audio service (such as commentary) intended to be mixed with the main audio service before presentation to the listener.

If the content provider wishes to enable user selection of specific combinations of bitstreams in the playback device (allowing different renditions of the overall media presentation to be selected and delivered), separate adaptation sets must be defined for each bitstream. For example, one adaptation set is used to describe the main audio service on its own, and a second adaptation set describes the associated audio service that will be simultaneously delivered with the main audio service to the playback device, where both adaptation sets will be decoded and mixed together. Refer to the *MPD with associated audio services using Dolby Digital Plus* section for more details.

Related information

Deriving the contents of the EC3SpecificBox on page 24

Representations

Each adaptation set carries one or more representations. All representations in an adaptation set must be perceptually identical, meaning that the bit rate is the only major parameter that may differ across the Dolby Digital Plus bitstreams in one adaptation set.

AudioChannelConfiguration descriptor

A representation that includes Dolby Digital Plus must include an AudioChannelConfiguration descriptor, which unambiguously describes the channel configuration of the referenced Dolby Digital Plus bitstream. For a Dolby Digital Plus bitstream that carries Dolby Atmos content, this

[1] As specified in Annex E of ETSI TS 102 366, Dolby Digital Plus decoders are required to decode Dolby Digital Plus bitstreams with bsid values between 11 and 16, inclusive. Bitstreams that have a bsid value outside of this range must not be identified using the ec-3 .

descriptor describes the channel configuration of the backward-compatible channel coding of the presentation.

For Dolby Digital Plus bitstreams, the AudioChannelConfiguration descriptor must use the AudioChannelConfiguration scheme described in the schemeIdUri: tag:dolby.com, 2014:dash:audio_channel_configuration:2011.

The value element must contain a four-digit hexadecimal representation of the 16-bit bit field, which describes the channel assignment of the referenced Dolby Digital Plus bitstream according to the table.

Table 2: AudioChannelConfiguration descriptor

Bit	Location
0 (most-significant bit)	L
1	C
2	R
3	Ls
4	Rs
5	Lc/Rc pair
6	Lrs/Rrs pair
7	Cs
8	Ts
9	Lsd/Rsd pair
10	Lw/Rw pair
11	Lvh/Rvh pair
12	Cvh
13	Lts/Rts pair
14	LFE2
15	LFE

Bit 0, which indicates the presence of the L channel, is the most-significant bit of the AudioChannelConfiguration descriptor. For example, to indicate that the channel configuration of the Dolby Digital Plus bitstream is L, C, R, Ls, Rs, LFE, the value element must contain the value F801 (the hexadecimal equivalent of the binary value 1111 1000 0000 0001).

SupplementalProperty descriptor

If the referenced Dolby Digital Plus bitstream carries Dolby Atmos content, the representation must include two additional SupplementalProperty descriptors.

To signal the presence of audio objects, one descriptor must use the SupplementalProperty scheme described in the schemeIdUri: tag:dolby.com,2018:dash:EC3_ExtensionType:2018, and the value element must be set to JOC.

To describe the decoding complexity of the Dolby Digital Plus bitstream with Dolby Atmos content, the other descriptor must use the SupplementalProperty scheme described in the schemeIdUri: tag:dolby.com,2018:dash:EC3_ExtensionComplexityIndex:2018. The value element must contain a decimal representation of the eight-bit bit field for complexity_index_type_a in the EC3SpecificBox of a Dolby Digital Plus audio track. The complexity_index_type_a parameter has a value range of 1 to 16.

A playback client can use the information to select the suitable representation according to its decoding capability.

Accessibility descriptor

If the adaptation set provides for enhanced accessibility, the AdaptationSet may include an accessibility descriptor that describes the type of accessible audio service being provided. The required attribute schemeIdUri must be set to urn:tva:metadata:cs:AudioPurposeCS:2007, as defined in section B.1 of ETSI TS 102 822-3-1 v1.4.1, signaling the namespace for the accessibility descriptor.

The audio purpose classification scheme (AudioPurposeCS), which is used to describe the type of accessible audio service that is being delivered, is defined in section A.15 of ETSI TS 102 822-3-1 v1.4.1. The value of the termID attribute must be set to match the type of accessible audio service carried in the Dolby Digital Plus bitstream, which is indicated by the value of the bsmod parameter in independent substream 0 of the Dolby Digital Plus stream, or in the EC3SpecificBox of the Dolby Digital Plus audio track. The corresponding values of the termID attribute and bsmod parameter are listed in the table.

Table 3: Corresponding termID attribute and bsmod parameter values

termID attribute value	AudioPurposeCS name	bsmod parameter value
1	Audio description for the visually impaired	010
2	Audio description for the hearing impaired	011
3	Supplemental commentary	101
4	Director's commentary	101
5	Educational notes	101
6	Main program audio	000
7	Clean feed (no effects mix)	100

Role descriptor

As defined in the MPEG-DASH role scheme (urn:mpeg:dash:role:2011), the value attribute of the role descriptor must be set to describe the purpose of each adaptation set in the overall presentation, as listed here:

- If the adaptation set is delivering a full audio service intended for direct presentation to the listener, the value attribute must be main or alternate.
- If the adaptation set is delivering an audio service that must be decoded and mixed with a full audio service delivered in a different adaptation set before presentation to the listener (sometimes referred to as a receiver-mix service), the value attribute must be set to commentary.
- If the adaptation set is delivering a full audio service intended for direct presentation to the listener, but this audio service is intended as an alternative presentation to the main audio service (for example, when delivering a service that contains premixed main audio and audio elements for visually impaired listeners, sometimes referred to as a broadcast-mix service), the value attribute must be set to alternate.

5.2.2 Media Presentation Description with associated audio services using Dolby Digital Plus

It is useful in some scenarios to simultaneously deliver two audio services (one main and one associated) to a decoding device.

For example, the scenarios can be:

- The main audio service contains the majority of the audio program stream and is a self-contained presentation that can be decoded on its own.
- The associated audio service contains additional supplementary audio program elements intended to be decoded and mixed with the main audio service (for example, a director's commentary or a description of the program for a visually impaired listener).

The main and associated audio streams are stored in separate ISO base media files that are described in separate adaptation sets within the MPD. The `accessibility` and `role` descriptors describe the purpose of the audio streams.

dependencyID

An adaptation set that is delivering an associated audio service must not be decoded and presented to the listener on its own, but must always be mixed with the decoded audio from the adaptation set that is delivering the corresponding main audio service. Therefore, we recommend that the adaptation set that is delivering the associated audio service includes a `dependencyID` descriptor. This descriptor indicates the relationship of the associated audio service with the main audio service that it will be mixed with after decoding.

5.2.3 Media Presentation Description for the HbbTV environment

The MPD for an HbbTV presentation must follow the same requirements as an MPEG-DASH presentation. In addition, HbbTV-specific requirements apply.

- The `profiles` attribute must contain `urn:hbbtv:dash:profile:isoff-live:2012`, which identifies the HbbTV ISO base media file format Live profile.
- The size of an MPD conforming to the HbbTV profile must not exceed 100 KB.
- The HbbTV profile-specific MPD must provide this information for all audio representations:
  - `audioSamplingRate`
  - `AudioChannelConfiguration`
  - `lang`

Note: `lang` is an attribute of the `AdaptationSet` element and is inherited by its representations.

Restrictions on referenced content in a Media Presentation Description

There are several restrictions on referenced content.

For content referenced in an HbbTV MPD:

- The movie fragment box (`moof`) must contain only one track fragment box (`traf`).
- The track run box (`trun`) must allow negative composition offsets (as defined in ISO 14496-12) in order to maintain audio/visual presentation synchronization.

Restrictions on adaptation sets

There are several restrictions on adaptation sets.

For the set of representations in an adaptation set in an HbbTV profile-specific MPD:

- Each representation must contain only one media component; a single audio or video track. Other nonmedia components may be present if applicable.
- All representations must have the same `track_ID` in the track header box and track fragment header box.
- The initialization segment must be common for all representations. Additionally, all information necessary to decode any segment chosen from representations must be provided in the initialization segment.

  A common initialization segment means that all representations in an adaptation set have the same `stsd` box. There is one entry in the `stsd` box for each representation. Representations encoded with different parameters (for example, at different bit rates) use the same `sample_description_index` in the track fragment header to identify which one of the sample entries in the `stsd` box is applicable to them. Each segment must consist of a whole self-contained movie fragment.
- Segments must be at least one second long, except for the last segment in an MPD, which may be shorter.
- Each audio segment must have a duration of no more than 15 seconds. (Video segments have the same requirement.)

For a detailed list of nonaudio MPD requirements, including requirements for transitions between representations, see ETSI TS 102 796.

5.2.4 Media Presentation Description file examples

This section contains example MPD files for different media presentations.

Media Presentation Description for a single video component and single audio component

This MPD example describes a simple media presentation that consists of a single video component with a single 5.1-channel (L, C, R, Ls, Rs, LFE) Dolby Digital Plus audio component. Three representations of the video content and three representations of the audio content are provided, each at a different data rate.

The media presentation complies with the ISO base media file format live profile, as defined in ISO/IEC 23009-1:2012.

```
<?xml version="1.0" encoding="utf-8"?>
<MPD xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:dolby="http://www.dolby.com/ns/online/DASH" xmlns="urn:mpeg:dash:schema:mpd:2011"
xsi:schemaLocation="urn:mpeg:dash:schema:mpd:2011"
type="static" minimumUpdatePeriod="PT2S" timeShiftBufferDepth="PT30M"
availabilityStartTime="2011-12-25T12:30:01" minBufferTime="PT4S"
profiles="urn:mpeg:dash:profile:isoff-live:2011">
   <BaseURL>http://cdn1.example.com/</BaseURL>
   <BaseURL>http://cdn2.example.com/</BaseURL>
   <Period>
      <!-- Video -->
      <AdaptationSet mimeType="video/mp4" codecs="avc1.4D401F"
       frameRate="30000/1001" segmentAlignment="true" startWithSAP="1">
         <BaseURL>video/</BaseURL>
         <SegmentTemplate timescale="90000" media="$Bandwidth$/$Number$.mp4"
initialization="$Bandwidth$/0.mp4">
            <SegmentTimeline>
```

```
               <S t="0" d="180180" r="10"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <Representation id="v0" width="320" height="240" bandwidth="250000" />
         <Representation id="v1" width="640" height="480" bandwidth="500000" />
         <Representation id="v2" width="960" height="720" bandwidth="1000000" />
      </AdaptationSet>
      <!-- 5.1 channel English Audio -->
      <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="en" segmentAlignment="true"
startWithSAP="1">
      <SegmentTemplate timescale="48000" media="audio/en/$Bandwidth$/$Number$.mp4"
          initialization="audio/en/$Bandwidth$/0.mp4">
            <SegmentTimeline>
               <S t="0" d="96768" r="10"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <AudioChannelConfiguration schemeIdUri="tag:dolby.com,
2014:dash:audio_channel_configuration:2011" value="F801"/>
         <Representation id="a0" bandwidth="192000" />
         <Representation id="a1" bandwidth="256000" />
         <Representation id="a2" bandwidth="384000" />
      </AdaptationSet>
   </Period>
</MPD>
```

Media Presentation Description for main and associated audio services delivered in separate files

This is a simple example of a dynamic presentation, with multiple languages and multiple base URLs.

This MPD document describes content available from two sources (cdn1 and cdn2) with audio available in two different English-language presentations: main audio service only, or a visually impaired receiver-mix service. The visually impaired service is enabled by simultaneously delivering the Dolby Digital Plus bitstream containing the main audio service and an additional Dolby Digital Plus bitstream containing the associated audio service for visually impaired listeners.

Three versions of the video are provided at bit rates between 250 kbps and 1 Mbps in different spatial resolutions.

The media presentation complies with the ISO base media file format live profile, as defined in ISO/IEC 23009-1:2012.

```
<?xml version="1.0" encoding="utf-8"?>
<MPD xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:dolby="http://www.dolby.com/ns/online/DASH" xmlns="urn:mpeg:dash:schema:mpd:2011"
xsi:schemaLocation="urn:mpeg:dash:schema:mpd:2011"
type="dynamic"
minimumUpdatePeriod="PT2S"
timeShiftBufferDepth="PT30M"
availabilityStartTime="2011-12-25T12:30:00"
minBufferTime="PT4S"
profiles="urn:mpeg:dash:profile:isoff-live:2011">
   <BaseURL>http://cdn1.example.com/</BaseURL>
   <BaseURL>http://cdn2.example.com/</BaseURL>
```

```
   <Period>
      <!-- Video -->
      <AdaptationSet mimeType="video/mp4" codecs="avc1.4D401F" frameRate="30000/1001"
       segmentAlignment="true" startWithSAP="1">
         <BaseURL>video/</BaseURL>
         <SegmentTemplate timescale="90000" media="$Bandwidth$/$Number$.mp4"
initialization="$Bandwidth$/0.mp4">
            <SegmentTimeline>
               <S t="0" d="180180" r="12"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <Representation id="v0" width="320" height="240" bandwidth="250000" />
         <Representation id="v1" width="640" height="480" bandwidth="500000" />
         <Representation id="v2" width="960" height="720" bandwidth="1000000" />
      </AdaptationSet>
      <!-- English Audio -->
      <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="en" segmentAlignment="0"
startWithSAP="1">
         <Accessibility schemeIdUri="urn:tva:metadata:cs:AudioPurposeCS:2007" value="6"/>
         <Role schemeIdUri="urn:mpeg:dash:role:2011" value="main" />
         <SegmentTemplate timescale="48000" media="audio/en_main/$Bandwidth$/$Number$.mp4"
           initialization="audio/en_main/$Bandwidth$/0.mp4">
            <SegmentTimeline>
               <S t="0" d="96000" r="11"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <Representation id="a0" bandwidth="256000">
            <AudioChannelConfiguration schemeIdUri="tag:dolby.com,
2014:dash:audio_channel_configuration:2011" value="F801"/>
         </Representation>
      </AdaptationSet>
      <!-- English Audio for visually impaired listeners -->
      <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="en" segmentAlignment="true"
startWithSAP="1">
         <Accessibility schemeIdUri="urn:tva:metadata:cs:AudioPurposeCS:2007" value="1"/>
         <Role schemeIdUri="urn:mpeg:dash:role:2011" value="commentary" />
         <SegmentTemplate timescale="48000" media="audio/en_vi/$Bandwidth$/$Number$.mp4"
          initialization="audio/en_vi/$Bandwidth$/0.mp4">
            <SegmentTimeline>
               <S t="0" d="96000" r="11"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <Representation id="a1" dependencyId="a0" bandwidth="64000">
            <AudioChannelConfiguration schemeIdUri="tag:dolby.com,
2014:dash:audio_channel_configuration:2011" value="4000"/>
         </Representation>
      </AdaptationSet>
      <!-- French Audio -->
      <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="fr" segmentAlignment="0"
startWithSAP="1">
      <SegmentTemplate timescale="48000" media="audio/fr/$Bandwidth$/$Number$.mp4"
             initialization="audio/fr/$Bandwidth$/0.mp4">
               <SegmentTimeline>
                  <S t="0" d="96000" r="11"/>
               </SegmentTimeline>
```

```
            </SegmentTemplate>
         <Representation id="a2" bandwidth="192000">
            <AudioChannelConfiguration schemeIdUri="tag:dolby.com,
2014:dash:audio_channel_configuration:2011" value="F801"/>
         </Representation>
      </AdaptationSet>
   </Period>
</MPD>
```

Media Presentation Description for multiple-codec bitstreams

This MPD example describes a static media presentation with a single video component and two different audio presentations provided.

The first audio presentation is a 7.1-channel Dolby Digital Plus bitstream (with the language set to English). The second audio presentation is a two-channel AAC bitstream (with the language set to French).

```
<?xml version="1.0"?>
<MPD
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:dolby="http://www.dolby.com/ns/online/DASH"
 xmlns="urn:mpeg:dash:schema:mpd:2011"
 xsi:schemaLocation="urn:mpeg:dash:schema:mpd:2011"
 type="static"
 mediaPresentationDuration="PT0H0M30.080S"
 minBufferTime="PT4S"
 profiles="urn:mpeg:dash:profile:isoff-live:2011">
   <BaseURL>./</BaseURL>
   <Period start="PT0S">
      <!-- Video -->
      <AdaptationSet mimeType="video/mp4" codecs="avc1.42801f" segmentAlignment="true"
startWithSAP="1">
      <SegmentTemplate timescale="25000" media="ChID_voices_71_20_768_64_ddp_aac_V_$Number
$.mp4"
          initialization="ChID_voices_71_20_768_64_ddp_aac_V.mp4">
            <SegmentTimeline>
               <S t="0" d="66000" r="11"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <Representation id="1" bandwidth="462640" width="720" height="576"/>
      </AdaptationSet>
      <!-- Audio -->
      <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="en" segmentAlignment="true"
startWithSAP="1">
         <Role schemeIdUri="urn:mpeg:dash:role:2011" value="alternate" />
         <SegmentTemplate timescale="48000" media="ChID_voices_71_20_768_64_ddp_aac_A_2_
$Number$.mp4"
          initialization="ChID_voices_71_20_768_64_ddp_aac_A_2.mp4">
            <SegmentTimeline>
               <S t="0" d="142848" r="10"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <Representation id="2" bandwidth="768000">
            <AudioChannelConfiguration schemeIdUri="tag:dolby.com,
2014:dash:audio_channel_configuration:2011" value="FA01"/>
```

```
         </Representation>
      </AdaptationSet>
      <!-- Audio -->
      <AdaptationSet mimeType="audio/mp4" codecs="mp4a.40.02" lang="fr"
segmentAlignment="true" startWithSAP="1">
         <Role schemeIdUri="urn:mpeg:dash:role:2011" value="main" />
         <SegmentTemplate timescale="48000" media="ChID_voices_71_20_768_64_ddp_aac_A_3_
$Number$.mp4"
          initialization="ChID_voices_71_20_768_64_ddp_aac_A_3.mp4">
            <SegmentTimeline>
               <S t="0" d="143360" r="10"/>
            </SegmentTimeline>
         </SegmentTemplate>
         <Representation id="3" bandwidth="61384">
            <AudioChannelConfiguration schemeIdUri="urn:mpeg:dash:
23003:3:audio_channel_configuration:2011" value="2"/>
         </Representation>
      </AdaptationSet>
   </Period>
</MPD>
```

Media Presentation Description for a single video component and dual audio components with Dolby Atmos content

This MPD example describes a media presentation that consists of a single video component with three representations provided at bit rates between 1.5 Mbps and 5 Mbps in different spatial resolutions.

It contains a Dolby Digital Plus audio component with a 5.1-channel configuration at 192 kbps and a Dolby Atmos component at 448 kbps. The same stream can be used for both Dolby Digital Plus 5.1 and Dolby Atmos.

The media presentation complies with the ISO base media file format live profile, as defined in ISO/IEC 23009-1.

```
<?xml version="1.0"?>
<MPD xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:dolby="http://
www.dolby.com/ns/online/DASH" xmlns="urn:mpeg:dash:schema:mpd:2011"
xsi:schemaLocation="urn:mpeg:dash:schema:mpd:2011" type="static"
mediaPresentationDuration="PT0H30M0.064S" minBufferTime="PT4S"
profiles="urn:mpeg:dash:profile:isoff-live:2011">
    <BaseURL>./</BaseURL>
    <Period start="PT0S">
        <!-- Video -->
        <AdaptationSet mimeType="video/mp4" codecs="avc1.42c01e" segmentAlignment="true"
startWithSAP="1">
        <SegmentTemplate timescale="30000" media="example_V_$RepresentationID$_$Number
$.mp4" initialization="example_V_$RepresentationID$.mp4">
                <SegmentTimeline>
                    <S t="0" d="58058" r="930" />
                </SegmentTimeline>
            </SegmentTemplate>
            <Representation id="1" bandwidth="1501320" width="640" height="480" />
            <Representation id="2" bandwidth="2991576" width="1280" height="720" />
            <Representation id="3" bandwidth="4977520" width="1920" height="1080" />
        </AdaptationSet>
```

```
        <!-- Audio 5.1 -->
        <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="en"
segmentAlignment="true" startWithSAP="1">
        <SegmentTemplate timescale="48000" media="example_A_$RepresentationID$_$Number
$.mp4" initialization="example_A_$RepresentationID$.mp4">
                <SegmentTimeline>
                    <S t="0" d="142848" r="604" />
                </SegmentTimeline>
            </SegmentTemplate>
            <Representation id="4" bandwidth="192000">
                <AudioChannelConfiguration schemeIdUri="tag:dolby.com,
2014:dash:audio_channel_configuration:2011" value="F801" />
            </Representation>
        </AdaptationSet>

        <!-- Audio ATMOS -->
        <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="en"
segmentAlignment="true" startWithSAP="1">
        <SegmentTemplate timescale="48000" media="example_A_$RepresentationID$_$Number
$.mp4" initialization="example_A_$RepresentationID$.mp4">
                <SegmentTimeline>
                    <S t="0" d="142848" r="604" />
                </SegmentTimeline>
            </SegmentTemplate>
            <Representation id="5" bandwidth="448000">
                <AudioChannelConfiguration schemeIdUri="tag:dolby.com,
2014:dash:audio_channel_configuration:2011" value="F801" />
                <SupplementalProperty schemeIdUri="tag:dolby.com,
2018:dash:EC3_ExtensionType:2018" value="JOC"/>
                <SupplementalProperty schemeIdUri="tag:dolby.com,
2018:dash:EC3_ExtensionComplexityIndex:2018" value="16"/>
            </Representation>
        </AdaptationSet>
    </Period>
</MPD>
```

6 Dolby Digital Plus and HTTP Live Streaming

This section specifies the required data formatting and signaling between the server (sender) and the clients (receivers) to use Dolby Digital Plus audio bitstreams in Apple HTTP Live Streaming audio/video (A/V) delivery applications.

- HTTP Live Streaming transport stream segments
- Dolby Digital Plus packetization into an MPEG-2 transport stream
- MPEG-2 transport stream segment transition requirements
- Dolby Digital Plus packetization into an audio-only elementary stream
- Dolby Digital Plus packetization into a fragmented MP4 file
- HTTP Live Streaming playlist files

Because HTTP Live Streaming is codec agnostic, the implementation of Dolby Digital Plus support in HTTP Live Streaming is very similar to other audio codecs typically used in HTTP Live Streaming applications (for example, MPEG-4 AAC).

Note: Multiplexing and segmenting are described separately, although a product may incorporate both functions in a single operation.

6.1 HTTP Live Streaming transport stream segments

Generally, the HTTP Live Streaming application uses segmented MPEG-2 transport streams or fragmented MP4 files to deliver a multimedia presentation that contains both audio and video content. Apple recommends using the fragmented MP4 file solution.

For specific delivery and playback scenarios, such as music services and audio-only playback at low bit rate, delivery of audio-only content is also supported (in which case, either packed audio or fragmented MP4 file is used).

A transport stream segment, packed audio, and a fragmented MP4 file are each referred to as a content segment in this documentation.

From HTTP Live Streaming version 4, audio and video content can be packaged and delivered separately. This feature benefits situations where variations of a multimedia presentation must be delivered. For example, when delivering one video stream with four audio streams in four different languages, in order to avoid repeat packaging of the video four separate times, we recommend packaging each language in a separate packed audio or a fragmented MP4 file, and packaging the video in a video-only MPEG-2 transport stream or another fragmented MP4 file.

Typically, these media segments are created using segmentation tools from either a multiplexed MPEG-2 transport stream or a fragmented MP4 file, each of which provides the multimedia presentation for a certain variant.

Media stream segmentation should be performed based on the presentation time stamps (PTSs). The difference between the first (earliest) PTS in segment n and segment n + 1 is less than or equal to the segment duration indicated by the `#EXT-X-TARGETDURATION` attribute.

6.2 Dolby Digital Plus packetization into an MPEG-2 transport stream

This section describes the process of analyzing and preparing a Dolby Digital Plus bitstream for storage within an HTTP Live Streaming compliant MPEG-2 transport stream.

Note: Dolby Digital Plus bitstreams carrying Dolby Atmos content in an MPEG-2 transport stream are not supported.

6.2.1 Content of the PES packet payload

Building a PES packet payload from a Dolby Digital Plus bitstream must meet certain conditions.

When multiplexing a Dolby Digital Plus bitstream into a PES packet payload, the product must:

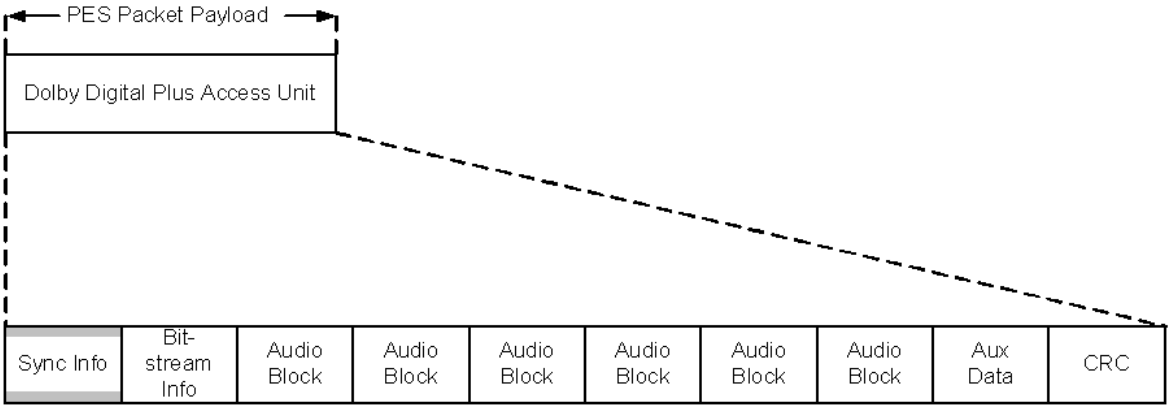
- Ensure that each PES packet payload contains complete access units. One access unit consists of all of the parts of the bitstream required by the Dolby Digital Plus decoder to produce 1,536 samples of decoded audio for each audio channel present in the bitstream.
- Ensure that a Dolby Digital Plus access unit does not span multiple PES packet payloads.

  Multiple complete Dolby Digital Plus access units may be placed within a single PES packet payload, but fragmentation of Dolby Digital Plus access units within a payload, or across multiple payloads, is not permitted. Multiplexing multiple, complete Dolby Digital Plus access units into a single PES packet payload can increase the overall efficiency of the transport stream, and in some cases provides bit-rate savings for the overall transport stream.
- Place Dolby Digital Plus frame bytes in the PES in big-endian format. (The first byte is `0x0B`.)
- Byte align the Dolby Digital Plus streams within the PES packet payload so that the initial eight bits of a Dolby Digital Plus frame reside in a single byte, placed at the start of the PES packet payload.
- Assemble frames in the same sequence in the PES packet payload as they occur in the Dolby Digital Plus stream.

These constraints ensure the correct operation of a downstream Dolby Digital Plus decoding device, particularly when this device is capable of converting the Dolby Digital Plus stream to Dolby Digital. This conversion requires the correct set of six blocks of audio data to produce a Dolby Digital frame.

How Dolby Digital Plus data is structured within a PES packet payload depends on the configuration of the Dolby Digital Plus bitstream. This figure shows the construction of a PES packet payload that contains a single Dolby Digital Plus access unit consisting of six audio blocks.

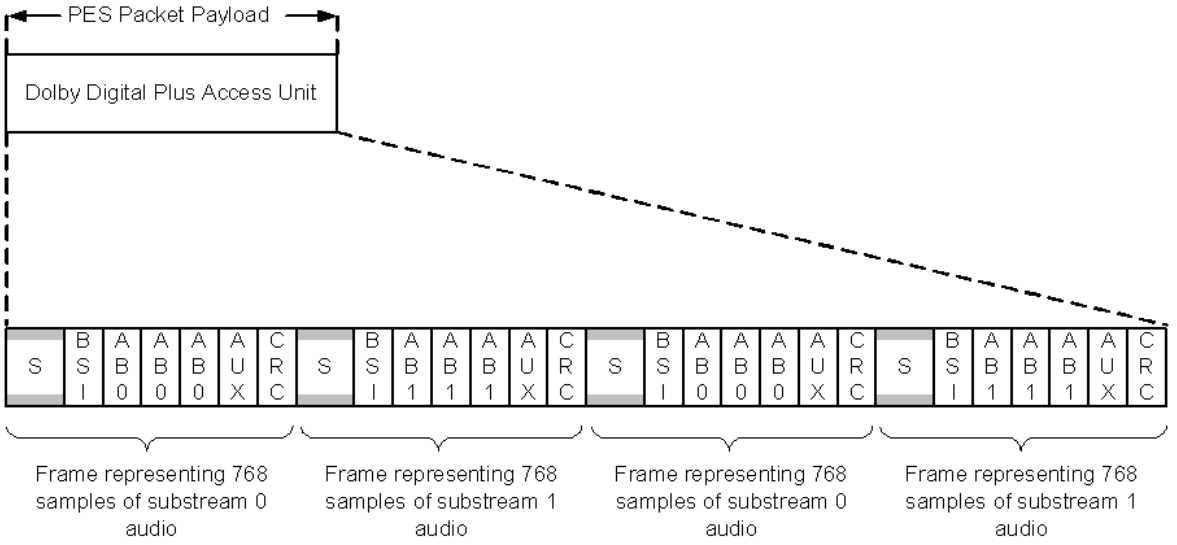
Figure 7: PES packet payload with a single substream with six blocks per frame



The six audio blocks represent 1,536 samples of audio from a single substream (substream 0).

This figure shows a PES packet payload that contains a single access unit of Dolby Digital Plus audio consisting of four frames.

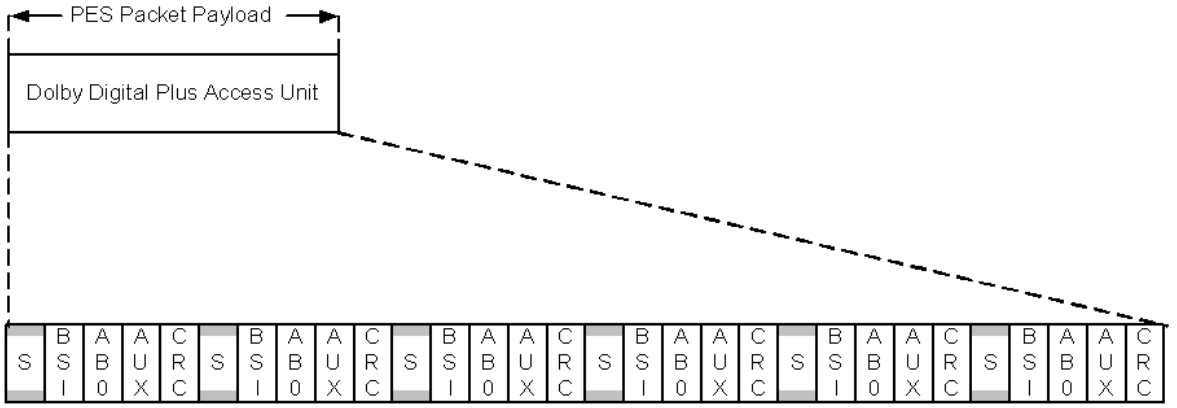
Figure 8: PES packet payload with two substreams with three blocks per frame



Each frame contains three audio blocks (denoted AB0 for substream 0 and AB1 for substream 1), each representing 256 samples of PCM audio from all channels in each substream.

This figure shows a PES packet payload that contains a single Dolby Digital Plus access unit consisting of six frames. Each frame contains one audio block, each representing 256 samples of PCM audio from every channel in the substream.

Figure 9: PES packet payload with a single substream with one block per frame



6.2.2 Parsing a Dolby Digital Plus bitstream

A product must use several bitstream information fields in the Dolby Digital Plus bitstream to identify access units and to properly signal the Dolby Digital Plus bitstream information in the PES packet payload.

To locate the start of an access unit, the product checks the `numblkscod`, `strmtyp`, and `substreamid` values. The multiplexing processing logic depends on the number of blocks per frame (indicated by the `numblkscod` value) in the Dolby Digital Plus bitstream.

If the `numblkscod` value is 3:

1. Find the first frame with a `strmtyp` value of `0` and a `substreamid` value of `0`. This indicates the start of an access unit.
2. Use the frame as the first frame of the program elementary stream packet payload.
3. Search for the next frame with a `strmtyp` value of `0` and `substreamid` value of `0`. This indicates the start of the next access unit.

4. If putting another access unit in the payload, continue placing frames in the PES packet payload and go back to step 3.

   Alternatively, use the frame immediately preceding the frame that starts the next access unit to complete the current PES packet payload.

5. Use the frame that starts the next access unit as the first frame of the next PES packet payload.

If the numblkscod value is not 3:

1. Find the first frame with a strmtyp value of 0, a substreamid value of 0, and a convsync flag set to 1.

   This indicates the start of an access unit.

2. Use the frame as the first frame of the PES packet payload.
3. As frames are placed in the PES packet payload, keep track of the number of accumulated audio blocks.
4. Find the next frame with a strmtyp value of 0, a substreamid value of 0, and a convsync flag set to 1:

   - If the frame immediately preceding the next frame does not complete six blocks (or a multiple of six blocks), keep searching until completing six blocks or multiple of them.
   - If the frame immediately preceding the next frame completes six blocks (or a multiple of six blocks), the current access unit is completed.

5. If putting another access unit in the payload, restart the count of accumulated audio blocks, put the next frame in the PES packet payload, and continue with step 4.

   Alternatively, use the frame immediately preceding the frame that starts the next access unit to complete the current PES packet payload.

6. Use the frame that starts the next access unit as the first frame of the next PES packet payload.

6.2.3 Signaling a Dolby Digital Plus bitstream in an MPEG-2 transport stream

To signal the presence of Dolby Digital Plus in the transport stream, the PES packet header must have the stream_id field set to 0xBD (private_stream_1).

In addition, the PMT entry for a Dolby Digital Plus elementary stream must:

- Have the stream_type parameter set to 0x87 (Dolby Digital Plus)
- Include the E-AC-3_audio_descriptor() in the descriptor loop following the ES_info_length field

If multiple Dolby Digital Plus bitstreams are present in the transport stream, a separate PMT entry must be included for each stream.

Syntax of the Enhanced AC-3 descriptor

The syntax of the E-AC-3 descriptor is listed in the table.

Syntax	Word size in bits	Identifier	Value
E-AC-3_audio_descriptor()			
{			

Syntax	Word size in bits	Identifier	Value
descriptor_tag.....	8	uimsbf	0xCC
descriptor_length.....	8	uimsbf	
reserved.....	1	bslbf	1
bsid_flag.....	1	bslbf	
mainid_flag.....	1	bslbf	
asvc_flag.....	1	bslbf	
mixinfoexists_flag.....	1	bslbf	
substream1_flag.....	1	bslbf	
substream2_flag.....	1	bslbf	
substream3_flag.....	1	bslbf	
reserved.....	1	bslbf	1
full_service_flag.....	1	bslbf	
service_type.....	3	uimsbf	
number_of_channels.....	3	uimsbf	
language_flag.....	1	bslbf	
language_flag_2.....	1	bslbf	
reserved.....	1	uimsbf	0
if(bsid_flag==1){			
bsid.....	5	uimsbf	
else{			
zero_bits.....	5	uimsbf	00000
}			
if(mainid_flag==1){			
reserved.....	3	bslbf	111
priority.....	2	uimsbf	
mainid.....	3	uimsbf	
}			
if(asvc_flag==1){			
asvc.....	8	bslbf	
}			
if(substream1_flag==1){			
substream1.....	8	uimsbf	
}			
if(substream2_flag==1){			
substream2.....	8	uimsbf	
}			
if(substream3_flag==1){			
substream3.....	8	uimsbf	
}			
if(language_flag==1){			

Syntax	Word size in bits	Identifier	Value
language.....	3 X 8	uimsbf	
}			
if(language_flag_2==1){			
language_2.....	3 X 8	uimsbf	
}			
if(substream1_flag==1){			
substream1_lang.....	3 X 8	uimsbf	
}			
if(substream2_flag==1){			
substream2_lang.....	3 X 8	uimsbf	
}			
if(substream3_flag==1){			
substream3_lang.....	3 X 8	uimsbf	
}			
}			

Semantics for the Enhanced AC-3 descriptor

These are the semantics of the parameters within the E-AC-3 descriptor for a Dolby Digital Plus bitstream.

- The descriptor_tag must be set to is 0xCC.
- The descriptor_length field specifies the total number of bytes of the data portion of the descriptor following the byte defining the value of this field. The minimum length of the descriptor is three bytes, but it may be longer depending on the use of the subsequent descriptor flags.
- The bsid_flag must be set to 1 when the optional bsid field is present in the descriptor.
- The mainid_flag, asvc_flag, and mixinfoexists parameters must be set to 0.
- The substream1_flag must be set to 1 when the Dolby Digital Plus stream contains an additional associated audio service in independent substream 1. If an independent substream with a substreamid value of 1 is not present in the bitstream, the substream1_flag must be set to 0.
- The substream2_flag parameter must be set to 1 when the Dolby Digital Plus stream contains an additional associated audio service in independent substream 2. If an independent substream with a substreamid value of 2 is not present in the bitstream, the substream2_flag flag must be set to 0.
- The substream3_flag parameter must be set to 1 when the Dolby Digital Plus stream contains an additional associated audio service in independent substream 3. If an independent substream with a substreamid value of 3 is not present in the bitstream, this flag must be set to 0.
- The value of the one-bit full_service_flag indicates whether the audio service in independent substream 0 (and any dependent substreams associated with independent substream 0) of the Dolby Digital Plus stream is full-service audio, suitable for presentation, or whether this audio service is only a partial service that should be combined with another audio service before presentation. The value must be set according to these criteria:

- If the audio service is sufficiently complete for presentation without being combined with another audio service (for example, a visually impaired service containing all elements of the program, including music, effects, dialogue, and a narrative description of the visual program content), the full_service_flag must be set to a value of 1.
- If the service is not sufficiently complete must be combined with another audio service (for example, a visually impaired service containing only a narrative description of the visual program content, which must be combined with another audio service containing music, effects, and dialogue), the full_service_flag must be set to a value of 0.

- The service_type field indicates the type of audio service being conveyed in independent substream 0 (and any dependent substreams associated with independent substream 0) of the Dolby Digital Plus stream. The service_type field must be interpreted as listed in the table.

Table 4: Service type field

Field value	Description	Restrictions *	
		Full-service flag	Number of channels field
000	Complete Main (CM)	Must be set to 1	
001	Music and Effects (ME)	Must be set to 0	
010	Visually Impaired (VI)		
011	Hearing Impaired (HI)		
100	Dialogue (D)	Must be set to 0	
101	Commentary (C)		
110	Emergency (E)	Must be set to 1	Must be set to 000
111	Voiceover (VO)	Must be set to 0	Must be set to 000
111	Karaoke	Must be set to 1	Must be set to 010, 011, or 100

* The values of the service_type field are valid only if the conditions identified in the restrictions columns are satisfied.

- The number_of_channels field (three bits) indicates the number of channels present in independent substream 0 (and any dependent substreams associated with independent substream 0) of the Dolby Digital Plus stream. This field must be interpreted as listed in the number of channels field table:
  - The language_flag (one bit) indicates whether the language field (three bytes) is present in the descriptor. If the language field is present, then this bit must be set to 1. If the language field is not present, then this bit must be set to 0.
  - The language_flag_2 must be set to 0.
  - The bsid field (five bits) indicates the Dolby Digital Plus coding version. If the bsid field is included, the value of the field must be set to the same value as the bsid parameter in independent substream 0 of the Dolby Digital Plus stream.

Table 5: Number of channels field

Field value *	Description	Restrictions †	
		Full-service flag	Service type field
000	Mono		
001	1+1 mode		
010	Two channel		

Table 5: Number of channels field (continued)

Field value *	Description	Restrictions †	
		Full-service flag	Service type field
011	Two channel encoded in Dolby Surround		
100	Multichannel audio (greater than two channels and up to 5.1 channels)		
101	Multichannel audio (greater than 5.1 channels)	Must be set to 1	Must be set to 000
110	Reserved for future use		
111	Reserved for future use		

* For two-channel Dolby Digital Plus streams, the number_of_channels field should be set to 011 when the dsurmod parameter is set to 011 (encoded in Dolby Surround), and should be set to 010 if the dsurmod parameter is set to any other value, or is not present.

† The values of the number_of_channels field are valid only if the conditions identified in the restrictions column are satisfied.

- The substream1 field (eight bits) indicates the type of audio carried in independent substream 1 of the Dolby Digital Plus stream. The value assignments of each bit are indicated in the substream field bit value assignments table. If the Dolby Digital Plus bitstream contains an independent substream with a substreamid value of 1, the PMT must include the substream1 field.
- The substream2 field (eight bits) indicates the type of audio carried in independent substream 2 of the Dolby Digital Plus stream. The value assignments of each bit are indicated in the substream field bit value assignments table. If the Dolby Digital Plus bitstream contains an independent substream with a substreamid value of 2, the PMT must include the substream2 field.
- The substream3 field (eight bits) indicates the type of audio carried in independent substream 3 of the Dolby Digital Plus stream. The value assignments of each bit are indicated in the substream field bit value assignments table. If the Dolby Digital Plus bitstream contains an independent substream with a substreamid value of 3, the PMT must include the substream3 field.

Table 6: Substream field bit value assignments

Substream 1–3 bits	Description
b7 (most-significant bit)	Reserved (must be set to 1)
b6	Reserved (must be set to 0)
b5 to b3	Service type flags (see the *Service type flags* table)
b2 to b0	Number of channels flags (see the *Number of channels flags* table)

Table 7: Service type flags

Bit values			Description	Restrictions *
b5	b4	b3		Number of channels field
0	0	0	Reserved	
0	0	1	Music and Effects (ME)	
0	1	0	Visually Impaired (VI)	
0	1	1	Hearing Impaired (HI)	

Table 7: Service type flags (continued)

Bit values			Description	Restrictions *
b5	b4	b3		Number of channels field
1	0	0	Dialogue (D)	
1	0	1	Commentary (C)	
1	1	0	Reserved	
1	1	1	Voiceover (VO)	Must be set to 000

* The values of the service_type flags bit values are considered valid only if the conditions identified in the restrictions column are satisfied.

Table 8: Number of channels flags

Flags			Description
b2	b1	b0	
0	0	0	Mono
0	0	1	Reserved for future use
0	1	0	Two channel *
0	1	1	Two channel encoded in Dolby Surround
1	0	0	Multichannel audio (greater than two channels and up to 5.1 channels)
1	0	1	Reserved for future use
1	1	0	Reserved for future use
1	1	1	Reserved for future use

* For two-channel substreams, the number of channels field should be set to 011 when the dsurmod parameter is set to 011 (encoded in Dolby Surround), and should be set to 010 if dsurmod is set to any other value, or is not present.

- The substream1, substream2, and substream3 field must be set to values that are indicated in the tables in this section. The respective substream field may be ignored if other values are conveyed:
  - The language field (three bytes) defines the language of this audio service. The language field must contain a three-character code as specified by ISO 639-2. Each character is coded into eight bits according to ISO 8859-1 (ISO Latin-1) and inserted in order into the 24-bit field. The coding is identical to that used in the MPEG-2 ISO_639_language_code value in the ISO_639_language_descriptor specified in ISO/IEC 13818-1.
  - The additional_info field provides optional bytes for future use.

6.3 MPEG-2 transport stream segment transition requirements

At the transition point from one MPEG-2 transport stream segment to another, specific requirements must be followed to ensure that playback across the transition point continues seamlessly for both audio and video.

This is especially critical when the playback device chooses to switch to an alternative MPEG-2 transport stream segment (for example, to best match the available data rate of the IP connection).

6.3.1 Segmenting MPEG-2 transport streams

Perform transport stream segmentation based on the presentation time stamps (PTSs) in the transport stream. Ensure that the difference between the first (earliest) PTS in segment n and segment n + 1 is less than or equal to the segment duration indicated by the `#EXT-X-TARGETDURATION` attribute.

Do not split PES packet data between segments. Because A/V streams consist of interleaved PES packet data in most cases, reorganizing transport stream packet data may be required.[7]

Because random access is limited in HTTP Live Streaming to segment boundaries, a single Program Association Table (PAT) and Program Map Table (PMT) table is sufficient in the beginning of each segment. A PAT followed by a PMT is required for media decoding. Table repetition inside the segment is not required and may be omitted to save bandwidth.

For optimal bandwidth efficiency, we recommend the use of variable bit-rate transport stream segments over a constant bit rate. Convert constant bit-rate transport stream segments to variable bit rate by removing any null packets present in the transport stream segment.

6.3.2 Video and audio data offset

The duration of a Dolby Digital Plus access unit is always equal to 1,536 audio samples, or 32 ms at 48 kHz. The duration of a video frame varies, depending on the video frame rate. For example, a video frame rate of 25 fps equals a video frame duration of 40 ms, and a video frame rate of 29.97 fps equals a video frame duration of 33.367 ms. Consequently, video and audio PES packet boundaries are rarely (if ever) time aligned within the MPEG-2 transport stream. Therefore, at the end of an MPEG-2 transport stream segment, there will be an offset between the end of the last video and last audio PES packet of the segment. This is illustrated in the A/V presentation time stamp offset at MPEG-2 transport segment boundaries figure.

To ensure that playback across a segment transition is seamless, and to maintain A/V synchronization, the segmenter must meet these requirements when constructing an HTTP Live Streaming compliant MPEG-2 transport stream:

- Each segment contains only complete PES packets. Fragmentation of PES packets containing Dolby Digital Plus audio data across a segment boundary is not permitted.[8]
- To ensure that switching between audio and video streams encoded at different bit rates is seamless, all segments that correspond to the same presentation period of the multimedia presentation (segments containing alternative renditions of the same content, with each rendition encoded at a different bit rate) contain an identical number of video and audio access units.
- The first PTS of each audio stream in the segment is equal to or greater than the first video PTS of the segment.
- The time offset between the first video PTS and the first audio PTS of a segment is less than 2,880 PTS ticks.

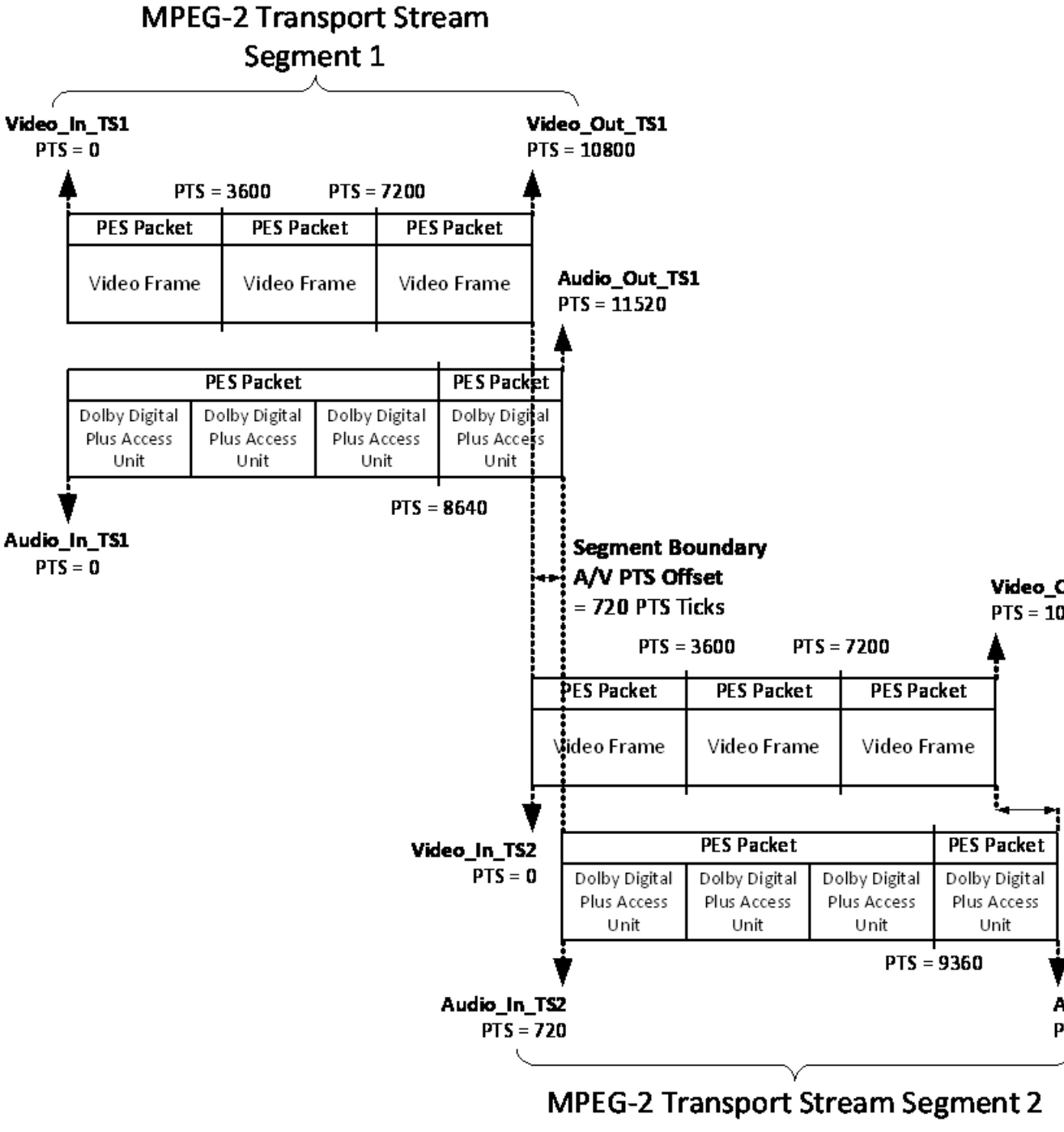
[7] In a variable bit-rate transport stream, this can be achieved relatively easily by moving the transport stream packets of the corresponding PES packet from segment n + 1 to the end of the segment n.

[8] This creates some dependency, as PES packetization is usually performed by the multiplexer. For cases in which more than one audio access unit is present in a PES, packet remultiplexing may be required in the segmenter.

- The time offset between the Audio_In time and Video_In time of a segment (the A/V PTS offset) is identical to the time offset between the Audio_Out time and Video_Out time of the previous segment.
- Segmenters may elect to reset the time base at the beginning of a segment. (The program clock reference [PCR] and PTS are not continuous across segment boundaries.) In this case, the discontinuity indicator in the first transport stream packet of a PID is designated as a PCR_PID, and the first packet of any audio elementary stream must be set to 1 (see ISO/IEC 13818-1).

The figure shows an example of two MPEG-2 transport stream segments containing 25 fps video and Dolby Digital Plus audio.

Figure 10: A/V presentation time stamp offset at MPEG-2 transport segment boundaries



For the purposes of illustration, the segments shown contain only a few frames of video and audio data. Segments used in a real-world HTTP Live Streaming application contain up to ten seconds of video and audio. This example assumes multiplexer and segmenter interaction to reset PTS values at segment boundaries.

In this example, the multiplexer places three Dolby Digital Plus access units (each with a duration equivalent to 2,880 PTS ticks) within a single PES packet, and places each video frame (with a duration equivalent to 3,600 PTS ticks) within its own PES packet. To ensure that the PTS offset between the end of the video stream and the end of audio stream is less than 2,880 ticks, the multiplexer places only a single Dolby Digital Plus access unit within a PES packet at the segment boundary, resulting in an A/V PTS offset of 720 PTS ticks.

At the start of the next segment, the PTS value of the first video PES packet is `0`, and the PTS value of the first Dolby Digital Plus PES packet (which again contains three Dolby Digital Plus access units) is `720`, ensuring that the offset at the end of the segment TS1 is maintained at the start of segment TS2.

Note: In real-world implementations, the multiplexer may choose not to reset the PTS to 0 at the start of each new segment, and instead use the `Video_Out` and `Audio_Out` PTS values from the previous segment as the PTS values of the first video and audio PES packets, respectively.

By maintaining the A/V PTS offset at each segment boundary, both the multiplexer and segmenter ensure that synchronization between the audio and video PTS is maintained, and that A/V synchronization during playback is also maintained. Additionally, by ensuring that only whole PES packets are present in the MPEG-2 transport stream, a player is able to switch between both audio and video streams at different data rates without any visible or audible interruption to playback.

6.4 Dolby Digital Plus packetization into an audio-only elementary stream

Dolby Digital Plus bitstreams can be packetized into HTTP Live Streaming compliant audio-only elementary streams by using pack audio with the `ID3` tag.

When multiplexing a Dolby Digital Plus audio elementary stream into HTTP Live Streaming compliant audio-only elementary stream segments, each audio elementary stream segment must signal the time stamp of its first sample with an `ID3 PRIV` tag at the beginning of the segment.

Additional requirements can be obtained from the HTTP Live Streaming page (see the *Resources* section).

The structure of an audio-only elementary stream segment is shown in the figure.

Figure 11: Audio-only elementary stream segments

The syntax of the `ID3 PRIV` tag is described in the table.

Table 9: Syntax of the ID3 PRIV tag

Syntax	Word size in bits	Identifier	Value
ID3_tag(){			
ID3V2_header()			
PRIV_frame()			
}			
ID3V2_header(){			
ID3V2_file_identifier	3x8	string	“ID3”
ID3V2_version	2x8	uimsbf	0x0400
ID3V2_flags	8	bslbf	0x00
ID3V2_size	4x8	bslbf	4x0b0xxxxxxx(0x0000003f)
}			
PRIV_frame(){			
frame_ID	4x8	string	"PRIV"
frame_size	4x8	uimsbf	4x0b0xxxxxxx(0x00000035)
frame_status_flags	8	bslbf	0x00
frame_format_flags	8	bslbf	0x00
Owner_identifier	45x8	string	"com.apple.streaming.transportStream Timestamp"
Private_data	8x8	bslbf	The upper 31 bits must be set to 0, the lower 33 bits is the PTS value
}			

6.5 Dolby Digital Plus packetization into a fragmented MP4 file

The latest HTTP Live Streaming specification recommends using fragmented MP4 files for delivery of a multimedia presentation that contains both audio and video content, with audio and video content packetized in separate fragmented MP4 files. For details about Dolby Digital Plus packetization into an MP4 file, refer to *Dolby Digital Plus and ISO base media file format*.

Related information

Dolby Digital Plus and ISO base media file format on page 57

6.6 HTTP Live Streaming playlist files

An HTTP Live Streaming master playlist file provides information for your product about the available content for a media presentation.

Specifically, HTTP Live Streaming master playlist files provide information about:

- The media segments that comprise the media presentation
- Available content variants that can be used in the content selection process (for example, content rendered for different bandwidths)
- Available content renditions (for example, different languages)

The master playlist allows your product to select from different versions of the media presentation. Your product can switch to a lower data rate (for example, when the available delivery bandwidth is reduced) or switch to a higher data rate to improve audio and video quality. Your product can also choose to play alternative content for a presentation, such as an alternative language version or an audio description for a visually impaired user.

HTTP Live Streaming playlist files are regular M3U playlists, extended by the addition of information specific to HTTP Live Streaming. This extended information is contained in lines that start with a #EXT prefix inside the playlist. A playlist contains uniform resource identifiers (URIs) that point to media files or to other playlists.

6.6.1 Codec type indication for Dolby Digital Plus

For media streams containing Dolby Digital Plus, the CODEC attribute of the EXT-X-STREAM-INF parameter is recommended for codec type indication.

If present, the CODEC attribute must include ec-3 for media presentations using Dolby Digital Plus, as shown in this example:

```
#EXT-X-STREAM-INF:BANDWIDTH=3464568,CODECS="avc1.640028,ec-3" example.m3u8
```

An object type indicator (OTI) value must not be appended to the ec-3 string.[9]

6.6.2 Codec type indication for Dolby Atmos content

For HTTP Live Streaming streams containing Dolby Digital Plus tracks with Dolby Atmos content, the CODEC attribute of the EXT-X-STREAM-INF parameter must include a value of ec-3 for indication of the audio codec. In addition, a CHANNELS attribute must be included in the #EXT-X-MEDIA parameter.

The value of the CHANNELS attribute is the number of decodable objects followed by a slash (/) and then JOC. JOC is the audio coding identifier for a Dolby Digital Plus bitstream with joint object coding data. The number of decodable objects is the decimal representation of the eight-bit bit field for complexity_index_type_a in the EC3SpecificBox of a Dolby Digital Plus audio track. The complexity_index_type_a parameter has a value range of 1 to 16.

For example, for a media presentation containing a Dolby Digital Plus bitstream with Dolby Atmos content, the CHANNELS attribute of the #EXT-X-MEDIA parameter must include the JOC identifier:

```
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="ATMOS",LANGUAGE="eng",
NAME="ATMOS Track",AUTOSELECT=YES,DEFAULT=YES,CHANNELS="16/JOC",
URI="../../DDP/ChID_voices_6ch_256kbps_ddp_sub.m3u8"
```

For backward compatibility, we recommend adding an equivalent 5.1 bitstream that is not Dolby Atmos to the same #EXT-X-MEDIA group. However, the same Dolby Atmos bitstream can also be signaled as a regular 5.1 one to ensure playback on all Dolby Digital Plus enabled systems.

6.6.3 Construction of the HTTP Live Streaming master playlist file

To construct an HTTP Live Streaming master playlist file for a media presentation with multiple audio codecs (such as Dolby Digital Plus and AAC), use the #EXT-X parameters.

[9] The object type indicator value is not defined for Dolby Digital Plus.

- For each audio track (in each codec), create an entry of `#EXT-X-MEDIA`:
  - Set the `TYPE` attribute to `AUDIO`.
  - Assign a user-defined value or name for the `GROUP-ID` attribute. A set of one or more `EXT-X-MEDIA` parameters with the same `GROUP-ID` value and the same `TYPE` value defines a group of renditions.
  - Set the `CHANNELS` attribute to a proper value to indicate the number of audio channels in the media content. For example, use `CHANNELS="2"` for stereo Dolby Digital Plus or stereo AAC, and use `CHANNELS="6"` for 5.1-channel Dolby Digital Plus or 5.1-channel Dolby Digital.
  - Specify the language code (for example, `"en"`, `"fr"`) for the `LANGUAGE` attribute.
  - Specify a name for each audio group with the `NAME` attribute.

    Whenever several audio tracks are encoded from the same source content (for example, the same language dub), the `NAME` attribute in the `EXT-X-MEDIA` parameter must be the same for all of them. The native Apple HTTP Live Streaming player automatically switches to the appropriate track from the available codec variants.
  - Set the `DEFAULT` attribute to `YES` if you want this rendition of the content to play back at the client side by default; otherwise, set it to `NO`.
  - Include other attributes as needed, such as `AUTOSELECT` and `URI`.
- For each content variant, create an entry of `EXT-X-STREAM-INF`:
  - Set the `CODEC` attribute to a comma-separated two-part value to indicate both the audio and video codec for playback.
  - Set the `AUDIO` attribute to a specific value that must match the value of the `GROUP-ID` attribute. This value indicates the set of audio renditions that should be used when playing back the media presentation.
  - Include other attributes as needed, such as `RESOLUTION`, `FRAME-RATE`, and so on, to describe other characteristics of the variant.

The ordering of the `EXT-X-STREAM-INF` entries in the master playlist determines implicit preference if the player considers two variant streams encoded with different codecs to be equally suitable for playback. A variant stream that appears earlier takes precedence in playback.

In order to enable playback of stereo Dolby Digital Plus on an iPhone or iPad with iOS 10, or on a Mac with macOS Sierra, you must list the stereo Dolby Digital Plus `EXT-X-STREAM-INF` entries first in the master playlist, and stereo AAC second. This ensures that Dolby Digital Plus plays back wherever it is supported, and automatically falls back to AAC on other devices.

6.6.4 HTTP Live Streaming playlist files with associated audio for receiver mixing

A single Dolby Digital Plus bitstream may include one or more associated audio programs in addition to the main audio program (program extensions).

For example, the associated audio program may be used to provide an audio description for the visually impaired or an audio presentation for the hearing impaired. The associated program is intended to be mixed with the main audio program after decoding. When providing content using these program extensions, we recommend providing the same content in all bandwidth variants.

If associated audio streaming is not always desired (for example, to provide different versions for different clients), we recommend that a rendition group that contains renditions with and

without associated audio streams be made available. In a playlist, alternative renditions of the same content are identified with the #EXT-X-MEDIA parameter. To properly signal Dolby Digital Plus renditions using the #EXT-X-MEDIA parameter:

- Provide a meaningful NAME attribute (for example, Movie for the rendition without an associated program or Movie with a spoken scene description for the visually impaired for a rendition with an associated program).

  Receivers may display the NAME attribute value for manual user selection and use the DDP-PROGRAM1, DDP-PROGRAM2, and DDP-PROGRAM3 attribute values to enable automatic selection of alternative content (for example, for visually impaired users).

- For a rendition that contains associated programs, indicate the associated programs in the <attribute-list> of the #EXT-X-MEDIA parameter:

  - Add the DDP-PROGRAM0 attribute to the <attribute-list> of the #EXT-X-MEDIA parameter. The DDP-PROGRAM0 attribute must indicate the eight-bit substream information with an unquoted hexadecimal integer number as specified in this table.

    Table 10: Dolby Digital Plus PROGRAM0 bits

Dolby Digital Plus PROGRAM0 bits	Description
b7 (most-significant bit)	Reserved (must be set to 1)
b6	Full_service_flag
b5 to b3	Service_type
b2 to b0	Number_of_channels

    The substream information for the DDP-PROGRAM0 attribute, listed in the table, is indicated in the same way as the comparable parameters of the E-AC-3 audio descriptor. (See the *Semantics for the enhanced AC-3 descriptor* section.)

  - If the Dolby Digital Plus stream contains an independent substream with a substreamid value of 1, add a DDP-PROGRAM1 attribute to the <attribute-list> of the #EXT-X-MEDIA parameter.
  - If the Dolby Digital Plus stream contains an independent substream with a substreamid value of 2, add a DDP-PROGRAM2 attribute to the <attribute-list> of the #EXT-X-MEDIA parameter.
  - If the Dolby Digital Plus stream contains an independent substream with a substreamid value of 3, DDP-PROGRAM3 attribute must be added to the <attribute-list> of the #EXT-X-MEDIA parameter.

  Note: We recommend also adding a DDP-PROGRAM0 attribute to the <attribute-list> of the #EXT-X-MEDIA parameter for the rendition that does not contain an associated program. The DDP-PROGRAM0 attribute describes the configuration of independent substream 0 (and, if present, any dependent substreams associated with independent substream 0) of the Dolby Digital Plus stream.

- The DDP-PROGRAM1, DDP-PROGRAM2, and DDP-PROGRAM3 attributes must indicate the eight-bit substream information with an unquoted hexadecimal integer number as is also used in the E-AC-3 audio descriptor. (See *Substream 1–3 field bit value assignments*.)

6.6.5 Examples of HTTP Live Streaming playlist files with Dolby Digital Plus

This topic contains examples of HTTP Live Streaming playlist files with Dolby Digital Plus.

The first example shows a master playlist listing two media playlists. The two media playlists contain Dolby Digital Plus bitstreams at different data rates. The first one contains Dolby Atmos content in Dolby Digital Plus bitstreams, signaled with CHANNELS attribute as 16/JOC; the second one contains legacy Dolby Digital Plus content (that is, not Dolby Atmos), signaled with the CHANNELS attribute as 6.

```
#EXTM3U
#EXT-X-VERSION:7
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="ATMOS",LANGUAGE="eng",NAME="English Audio
Track",AUTOSELECT=YES,DEFAULT=YES,\
CHANNELS="16/JOC",URI="../../ATMOS/ChID_voices_6ch_384kbps_ddp_joc_sub.m3u8"
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="DDP",LANGUAGE="eng",NAME="English Audio
Track",AUTOSELECT=YES,DEFAULT=YES,\
CHANNELS="6",URI="../../DDP/ChID_voices_6ch_256_ddp_sub.m3u8"

#EXT-X-STREAM-INF:BANDWIDTH=4214869,AVERAGE-BANDWIDTH=2865708,CODECS="avc1.4d4028,ec-3",\
RESOLUTION=1280x720,FRAME-RATE=25.000,AUDIO="ATMOS"
../../VIDEO/Living-Room_1280x720p_25fp_h264_sub.m3u8
#EXT-X-STREAM-INF:BANDWIDTH=4086869,AVERAGE-BANDWIDTH=2737708,CODECS="avc1.4d4028,ec-3",\
RESOLUTION=1280x720,FRAME-RATE=25.000,AUDIO="DDP"
../../VIDEO/Living-Room_1280x720p_25fp_h264_sub.m3u8

#EXT-X-I-FRAME-STREAM-INF:BANDWIDTH=572161,AVERAGE-
BANDWIDTH=177951,CODECS="avc1.4d4028",RESOLUTION=1280x720,\
URI="../../VIDEO/Living-Room_1280x720p_25fp_h264_iframe_index_sub.m3u8"
```

The second example contains eight video encodings, three audio encodings, and subtitles. The video encodings are different in data rate. The audio encodings include AAC, Dolby Digital Plus, and Dolby Digital Plus with Dolby Atmos content. For each content variant, a video encoding is referenced in association with an audio encoding. The master playlist file lists all combinations and permutations of video and audio encodings.

Note: A backslash (\) is used to indicate that the tag continues on the following line.

```
#EXTM3U
#EXT-X-VERSION:7
#EXT-X-INDEPENDENT-SEGMENTS

#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=7712564,BANDWIDTH=7774493,CODECS="avc1.64002a,mp4a.
40.2",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v8/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=6302077,BANDWIDTH=6319326,CODECS="avc1.64002a,mp4a.
40.2",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v7/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=4764431,BANDWIDTH=4787533,CODECS="avc1.64002a,mp4a.
40.2",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v6/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=3236616,BANDWIDTH=3242718,CODECS="avc1.640020,mp4a.
40.2",\
RESOLUTION=1280x720,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v5/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=2215351,BANDWIDTH=2266597,CODECS="avc1.640020,mp4a.
```

```
40.2",\
RESOLUTION=960x540,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v4/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=1290298,BANDWIDTH=1304252,CODECS="avc1.64001e,mp4a.
40.2",\
RESOLUTION=768x432,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v3/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=912621,BANDWIDTH=924620,CODECS="avc1.64001e,mp4a.
40.2",\
RESOLUTION=640x360,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v2/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=539420,BANDWIDTH=550248,CODECS="avc1.640015,mp4a.
40.2",\
RESOLUTION=480x270,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="AAC",SUBTITLES="sub1"
v1/prog_index.m3u8

#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=7744562,BANDWIDTH=7806491,CODECS="avc1.64002a,ec-3",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v8/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=6334075,BANDWIDTH=6351325,CODECS="avc1.64002a,ec-3",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v7/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=4796429,BANDWIDTH=4819532,CODECS="avc1.64002a,ec-3",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v6/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=3268615,BANDWIDTH=3274717,CODECS="avc1.640020,ec-3",\
RESOLUTION=1280x720,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v5/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=2247349,BANDWIDTH=2298596,CODECS="avc1.640020,ec-3",\
RESOLUTION=960x540,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v4/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=1322297,BANDWIDTH=1336250,CODECS="avc1.64001e,ec-3",\
RESOLUTION=768x432,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v3/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=944619,BANDWIDTH=956619,CODECS="avc1.64001e,ec-3",\
RESOLUTION=640x360,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v2/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=571419,BANDWIDTH=582247,CODECS="avc1.640015,ec-3",\
RESOLUTION=480x270,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="EC3",SUBTITLES="sub1"
v1/prog_index.m3u8

EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=7744562,BANDWIDTH=7806491,CODECS="avc1.64002a,ec-3",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-
CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
v8/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=6334075,BANDWIDTH=6351325,CODECS="avc1.64002a,ec-3",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-
CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
v7/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=4796429,BANDWIDTH=4819532,CODECS="avc1.64002a,ec-3",\
RESOLUTION=1920x1080,FRAME-RATE=59.940,CLOSED-
CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
v6/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=3268615,BANDWIDTH=3274717,CODECS="avc1.640020,ec-3",\
RESOLUTION=1280x720,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
```

```
v5/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=2247349,BANDWIDTH=2298596,CODECS="avc1.640020,ec-3",\
RESOLUTION=960x540,FRAME-RATE=59.940,CLOSED-CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
v4/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=1322297,BANDWIDTH=1336250,CODECS="avc1.64001e,ec-3",\
RESOLUTION=768x432,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
v3/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=944619,BANDWIDTH=956619,CODECS="avc1.64001e,ec-3",\
RESOLUTION=640x360,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
v2/prog_index.m3u8
#EXT-X-STREAM-INF:AVERAGE-BANDWIDTH=571419,BANDWIDTH=582247,CODECS="avc1.640015,ec-3",\
RESOLUTION=480x270,FRAME-RATE=29.970,CLOSED-CAPTIONS="cc1",AUDIO="ATMOS",SUBTITLES="sub1"
v1/prog_index.m3u8

#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=260759,BANDWIDTH=266759,CODECS="avc1.64002a",\
RESOLUTION=1920x1080,URI="v8/iframe_index.m3u8"
#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=212848,BANDWIDTH=217848,CODECS="avc1.64002a",\
RESOLUTION=1920x1080,URI="v7/iframe_index.m3u8"
#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=161178,BANDWIDTH=165275,CODECS="avc1.64002a",\
RESOLUTION=1920x1080,URI="v6/iframe_index.m3u8"
#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=128564,BANDWIDTH=136728,CODECS="avc1.640020",\
RESOLUTION=1280x720,URI="v5/iframe_index.m3u8"
#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=96846,BANDWIDTH=98483,CODECS="avc1.640020",\
RESOLUTION=960x540,URI="v4/iframe_index.m3u8"
#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=80061,BANDWIDTH=83389,CODECS="avc1.64001e",\
RESOLUTION=768x432,URI="v3/iframe_index.m3u8"
#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=63776,BANDWIDTH=64939,CODECS="avc1.64001e",\
RESOLUTION=640x360,URI="v2/iframe_index.m3u8"
#EXT-X-I-FRAME-STREAM-INF:AVERAGE-BANDWIDTH=39837,BANDWIDTH=40568,CODECS="avc1.640015",\
RESOLUTION=480x270,URI="v1/iframe_index.m3u8"

#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="AAC",LANGUAGE="eng",NAME="English",AUTOSELECT=YES,\
DEFAULT=YES,CHANNELS="2",URI="a1/prog_index.m3u8"
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="EC3",LANGUAGE="eng",NAME="English",AUTOSELECT=YES,\
DEFAULT=YES,CHANNELS="6",URI="a2/prog_index.m3u8"
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="ATMOS",LANGUAGE="eng",NAME="English",AUTOSELECT=YES,\
DEFAULT=YES,CHANNELS="16/JOC",URI="a3/prog_index.m3u8"

#EXT-X-MEDIA:TYPE=SUBTITLES,GROUP-ID="sub1",NAME="English",LANGUAGE="eng",\
DEFAULT=YES,AUTOSELECT=YES,FORCED=NO,URI="s1/eng/prog_index.m3u8"

#EXT-X-MEDIA:TYPE=CLOSED-CAPTIONS,GROUP-ID="cc1",NAME="English",LANGUAGE="eng",\
DEFAULT=YES,AUTOSELECT=YES,INSTREAM-ID="CC1"
```

The third example is a master playlist listing three media playlists. One playlist contains a Dolby Digital Plus audio bitstream, one a Dolby Digital bitstream, and the other an AAC bitstream.

```
#EXTM3U
#EXT-X-VERSION:7
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="DDP",LANGUAGE="eng",NAME="English Audio
Track",AUTOSELECT=YES,\
DEFAULT=YES,CHANNELS="6",URI="../../DDP/ChID_voices_6ch_256kbps_ddp_sub.m3u8"
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="DD",LANGUAGE="eng",NAME="English Audio
Track",AUTOSELECT=YES,\
DEFAULT=YES,CHANNELS="6",URI="../../DD/ChID_voices_6ch_640kbps_dd_sub.m3u8"
```

```
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="AAC",LANGUAGE="eng",NAME="English Audio
Track",AUTOSELECT=YES,\
DEFAULT=YES,CHANNELS="2",URI="../../AAC/ChID_voices_2ch_64kbps_aac_sub.m3u8"

#EXT-X-STREAM-INF:BANDWIDTH=4389571,AVERAGE-BANDWIDTH=3294620,CODECS="avc1.4d4028,ec-3",\
RESOLUTION=1280x720,FRAME-RATE=25.000,AUDIO="DDP"
../../VIDEO/Living-Room_1280x720p_25fp_h264_sub.m3u8
#EXT-X-STREAM-INF:BANDWIDTH=4005571,AVERAGE-BANDWIDTH=2910620,CODECS="avc1.4d4028,ac-3",\
RESOLUTION=1280x720,FRAME-RATE=25.000,AUDIO="DD"
../../VIDEO/Living-Room_1280x720p_25fp_h264_sub.m3u8
#EXT-X-STREAM-INF:BANDWIDTH=3813571,AVERAGE-BANDWIDTH=2718620,CODECS="avc1.4d4028,mp4a.
40.2",\
RESOLUTION=1280x720,FRAME-RATE=25.000,AUDIO="AAC"
../../VIDEO/Living-Room_1280x720p_25fp_h264_sub.m3u8

#EXT-X-I-FRAME-STREAM-INF:BANDWIDTH=572161,AVERAGE-BANDWIDTH=177951,CODECS="avc1.4d4028",\
RESOLUTION=1280x720,URI="../../VIDEO/Living-
Room_1280x720p_25fp_h264_iframe_index_sub.m3u8"
```

The fourth example is a master playlist listing two media playlists. One playlist contains an English Dolby Digital Plus audio bitstream and the other a French Dolby Digital Plus audio bitstream.

```
#EXTM3U
#EXT-X-VERSION:7
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="AUDIO",LANGUAGE="eng",NAME="Audio Track DDP
eng",AUTOSELECT=YES,\
DEFAULT=YES,CHANNELS="6",URI="../../DDP/ChID_voices_eng_6ch_256kbps_ddp_sub.m3u8"
#EXT-X-MEDIA:TYPE=AUDIO,GROUP-ID="AUDIO",LANGUAGE="fra",NAME="Audio Track DDP
fra",AUTOSELECT=YES,\
DEFAULT=NO,CHANNELS="6",URI="../../DDP/ChID_voices_fra_6ch_256kbps_ddp_sub.m3u8"

#EXT-X-STREAM-INF:BANDWIDTH=4086869,AVERAGE-BANDWIDTH=2737708,CODECS="avc1.4d4028,ec-3",\
RESOLUTION=1280x720,FRAME-RATE=25.000,AUDIO="AUDIO"
../../VIDEO/Living-Room_1280x720p_25fp_h264_sub.m3u8

#EXT-X-I-FRAME-STREAM-INF:BANDWIDTH=572161,AVERAGE-BANDWIDTH=177951,CODECS="avc1.4d4028",\
RESOLUTION=1280x720,URI="../../VIDEO/Living-
Room_1280x720p_25fp_h264_iframe_index_sub.m3u8"
```

7 Dolby Digital Plus and ISO base media file format

This section describes the process of analyzing and preparing a Dolby Digital Plus bitstream for online delivery within an ISO base media file.

As Dolby Digital Plus features a high level of flexibility and potentially complex bitstream configurations, this section defines constraints on Dolby Digital Plus that minimize impact on decoding devices while ensuring that a wide range of applications can be supported. To simplify the design of ISO base media file format compliant demultiplexers, this section presents a specific method for grouping Dolby Digital Plus frames together for storage within an ISO base media file track sample.

7.1 Contents of the MP4 sample

Building an MP4 sample from a Dolby Digital Plus bitstream must meet certain conditions.

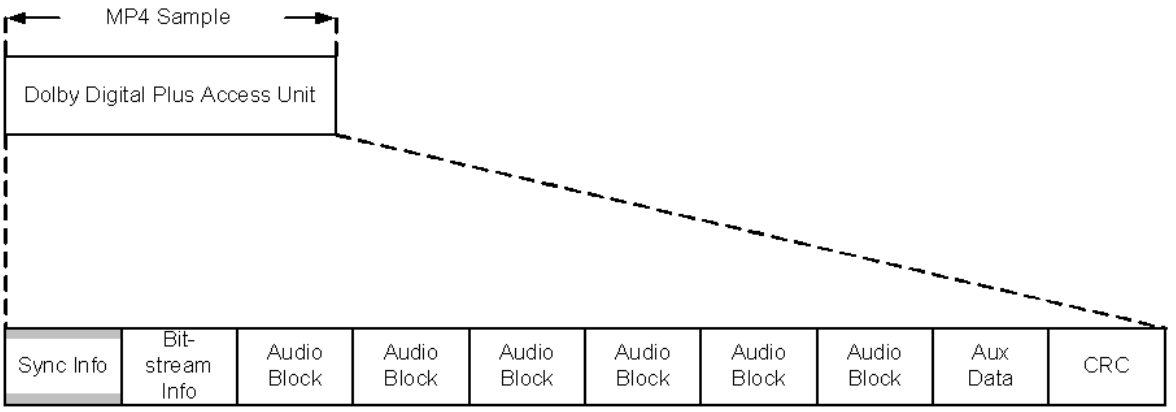
When multiplexing a Dolby Digital Plus bitstream into an MP4 sample, the product must:

- Ensure that each MP4 sample contains one access unit. One access unit consists of all of the parts of the bitstream required by the Dolby Digital Plus decoder to produce 1,536 samples of decoded audio for each audio channel present in the bitstream.
- Ensure that each MP4 sample contains only one complete Dolby Digital Plus access unit.
- Ensure that a Dolby Digital Plus access unit does not span multiple MP4 samples.

These constraints ensure the correct operation of a downstream Dolby Digital Plus decoding device, particularly when this device is capable of converting the Dolby Digital Plus stream to Dolby Digital. This conversion requires the correct set of six blocks of audio data to produce a Dolby Digital frame.

How Dolby Digital Plus data is structured within an MP4 sample depends on the configuration of the Dolby Digital Plus bitstream. This figure shows the construction of an MP4 sample that contains a single Dolby Digital Plus access unit consisting of six audio blocks.

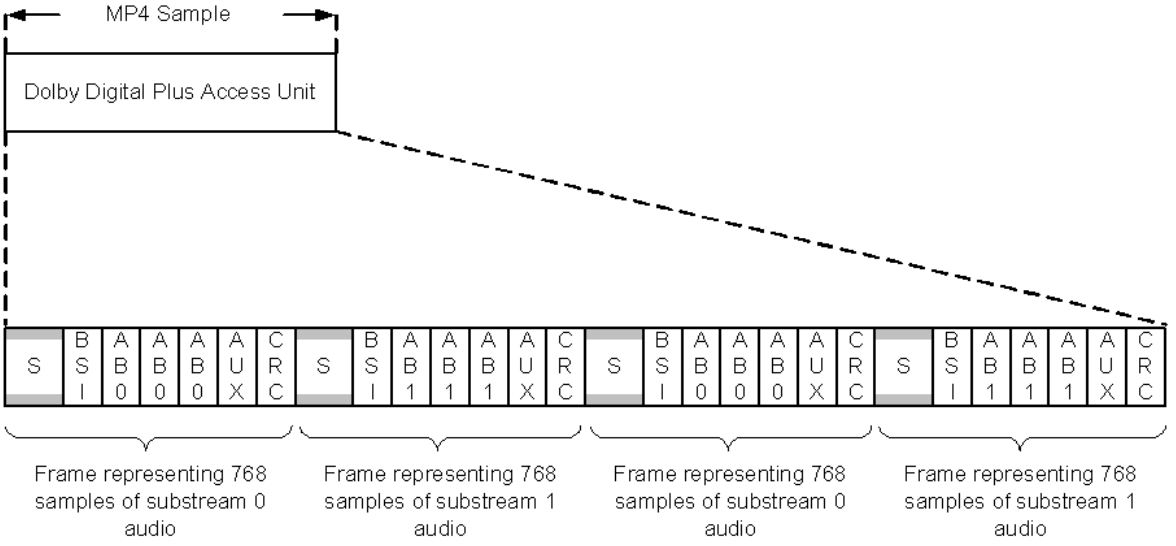
Figure 12: MP4 sample with a single substream with six blocks per frame



The six audio blocks represent 1,536 samples of audio from a single substream (substream 0).

This figure shows an MP4 sample that contains a single access unit of Dolby Digital Plus audio consisting of four frames.

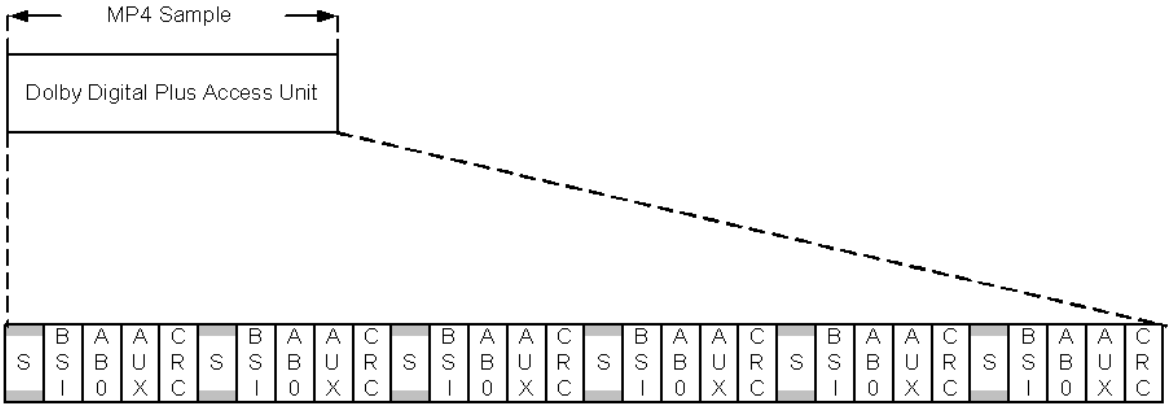
Figure 13: MP4 sample with two substreams with three blocks per frame



Each frame contains three audio blocks (denoted AB0 for substream 0 and AB1 for substream 1), each representing 256 samples of PCM audio from all channels in each substream.

This figure shows an MP4 sample that contains a single Dolby Digital Plus access unit consisting of six frames. Each frame contains one audio block, each representing 256 samples of PCM audio from every channel in the substream.

Figure 14: MP4 sample with a single substream with one block per frame



7.2 Parsing a Dolby Digital Plus bitstream

A product must use several bitstream information fields in the Dolby Digital Plus bitstream to identify access units and to properly signal the Dolby Digital Plus bitstream information in the MP4 file.

To locate the start of an access unit, the product checks the numblkscod, strmtyp, and substreamid values. The multiplexing processing logic depends on the number of blocks per frame (indicated by the numblkscod value) in the Dolby Digital Plus bitstream.

If the numblkscod value is 3:

1. Find the first frame with a strmtyp value of 0 and a substreamid value of 0. This indicates the start of an access unit.
2. Use the frame as the first frame of the MP4 sample.
3. Search for the next frame with a strmtyp value of 0 and substreamid value of 0. This indicates the start of the next access unit.
4. Use the frame immediately preceding the frame that starts the next access unit to complete the MP4 sample.
5. Use the frame that starts the next access unit as the first frame of the next MP4 sample.

If the numblkscod value is not 3:

1. Find the first frame with a strmtyp value of 0, a substreamid value of 0, and a convsync flag set to 1.

   This indicates the start of an access unit.
2. Use the frame as the first frame of the MP4 sample.
3. As frames are placed in the MP4 sample, keep track of the number of accumulated audio blocks.
4. Find the next frame with a strmtyp value of 0, a substreamid value of 0, and a convsync flag set to 1.

   If the frame immediately preceding the next frame completes six blocks (or a multiple of six blocks), the current sample is completed. Otherwise, keep searching until six blocks have been accumulated.

7.3 Signaling Dolby Digital Plus bitstreams in an ISO base media file

The basic structures defined within ISO/IEC 14496-12 to identify audio tracks are used with specific extensions (defined in Annex F of ETSI TS 102 366) to provide detailed information on the characteristics of a Dolby Digital Plus stream.

The information is included in the stbl box of the moov box of the ISO base media file. The locations and hierarchy of the boxes that must be included in stbl to identify a Dolby Digital Plus stream within an ISO base media file are listed in the table.

Table 11: Sample table box hierarchy for Dolby Digital Plus audio tracks

Nesting Level				
4	5	6	7	Reference
stbl				ISO/IEC 14496-12
	stsd			
		EC3SampleEntry (header type set to ec-3)		Annex F of ETSI TS 102 366
			EC3SpecificBox (header type set to dec3)	Annex F of ETSI TS 102 366 and *Deriving the contents of the EC3Specific box* section of this information set
	stts			ISO/IEC 14496-12
	stsc			
	stsz			
	stz2			
	stco			
	co64			

In the table, the value of the nesting level provided for each box is based on the structure of the complete ISO base media file, beginning with a nesting value of 0 for the ftyp and moov boxes. The table also includes a reference to where each box is defined.

For an unencrypted ISO base media file, the EC3SampleEntry box header type value is ec-3. The value of the EC3SpecificBox header type is dec3.

7.4 Signaling Dolby Digital Plus bitstreams with Dolby Atmos content in an ISO base media format file

In addition to the definitions in Annex F of ETSI TS 102 366, other information must be added to the EC3SpecificBox in a backward-compatible way to describe the characteristics of the Dolby Atmos content carried by a Dolby Digital Plus stream.

The extensions specific to audio objects are highlighted in the table. These extensions must be used to signal the presence of the Dolby Atmos content within an ISO base media file.

Syntax	Word size in bits	Identifier
EC3SpecificBox ()		
{		
BoxHeader.Size.....	32	uimsbf
BoxHeader.Type.....	32	uimsbf
data_rate.....	13	uimsbf
num_ind_sub.....	3	uimsbf
for (i = 0; i < num_ind_sub + 1; i++)		
{		
fscod.....	2	uimsbf
bsid.....	5	uimsbf

Syntax	Word size in bits	Identifier
reserved.....	1	bslbf
asvc.....	1	bslbf
bsmod.....	3	uimsbf
acmod.....	3	uimsbf
lfeon.....	1	bslbf
reserved.....	3	uimsbf
num_dep_sub.....	4	uimsbf
if (num_dep_sub > 0)		
{		
chan_loc.....	9	uimsbf
}		
else		
{		
reserved.....	1	bslbf
}		
}		
flag_ec3_extension_type_reserved.....	7	bslbf
flag_ec3_extension_type_a.....	1	bslbf
if (flag_ec3_extension_type_a == 1)		
{		
complexity_index_type_a.....	8	umisbf
}		
additional_data[].....	8	uimsbf
}		

These extensions are Dolby Atmos related.

flag_ec3_extension_type_a

The one-bit field must be set to 1 to indicate that Dolby Atmos content is carried in the first independent substream (I0) of a 5.1 Dolby Digital Plus stream.

flag_ec3_extension_type_reserved

This seven-bit field must be reserved for future application.

For content creation referring to this version of the specification, the reserved bits must be set to 0. For playback referring to this version of the specification, the reserved bits shall be ignored.

complexity_index_type_a

This eight-bit field indicates the decoding complexity of a Dolby Digital Plus bitstream carrying Dolby Atmos content. If the flag_ec3_extension_type_a bit in the addbsi has a value of 0, the complexity_index_type_a will not be available in the E-AC-3 descriptor. If the flag_ec3_extension_type_a bit in the addbsi has a value of 1, the complexity_index_type_a in

the E-AC-3 descriptor must take the value of the complexity_index_type_a field found in the addbsi of the Dolby Digital Plus bitstream.

A larger value indicates higher complexity.

additional_data[ ]

These optional bytes must be reserved for future use.

Other fields must conform to Annex F of ETSI TS 102 366.

7.5 Deriving the contents of the EC3SpecificBox

This topic provides additional details on the information specific to Dolby Digital Plus and how information is derived from the Dolby Digital Plus bitstream parameters.

The data inside the EC3SpecificBox can be calculated by using information in the Dolby Digital Plus bitstream parameters, including:

- frmsiz
- fscod
- numblkscod
- substreamid
- strmtyp
- bsmod
- chanmap
- addbsi

data_rate

The data_rate parameter indicates the data rate (in kbps) of the entire Dolby Digital Plus bitstream. The value is the sum of the data rates of all the substreams in the Dolby Digital Plus bitstream. When a bitstream uses variable data-rate encoding, data_rate indicates the maximum data rate of the bitstream.

The data rate of each substream is calculated using this equation:

$$\text{data\_rate\_sub} = \frac{(\text{frmsiz}+1)*\text{fs}}{\text{numblks}*16}$$

In this equation:

- frmsiz is the value of the frmsiz parameter in the Dolby Digital Plus frame.
- fs is the sampling frequency of the Dolby Digital Plus bitstream (in kHz). (The fs value is derived from the fscod parameter in the Dolby Digital Plus frame.)
- numblks is the number of audio blocks per frame. (The numblks value is derived from the numblkscod parameter in the Dolby Digital Plus frame.)

num_ind_sub

The num_ind_sub parameter indicates the number of independent substreams present in the Dolby Digital Plus bitstream. The value of num_ind_sub is equal to the value of the substreamid parameter found in the last independent substream of the bitstream. During bitstream parsing,

this will be the frame with a strmtyp value of 0 that precedes the frame with both a strmtyp value of 0 and a substreamid value of 0 (indicating that this frame belongs to the first independent substream of the bitstream).

bsmod

The bsmod parameter is optional in a Dolby Digital Plus bitstream. Detecting it requires additional parsing.

If the Dolby Digital Plus bitstream contains only one independent substream, inclusion of the value of bsmod in the EC3SpecificBox is optional. If parsing yields more than one independent substream, the value of bsmod must be included in the EC3SpecificBox for each substream to ensure that the system parsing the DASH-compliant file can quickly identify the audio services present in the bitstream.

This field itself is not optional. When there is no bsmod, the value must be set to zero.

num_dep_sub

The num_dep_sub parameter indicates the number of dependent substreams associated with an independent substream. Its value is equal to the value of the substreamid parameter found in the frame with a strmtyp value of 1 (that is, in the dependent substream) immediately preceding a frame with a strmtyp value of 0 (that is, in the independent substream).

chan_loc

The chan_loc field indicates channel locations (beyond the standard 5.1 channels) that are carried by dependent substreams associated with an independent substream. The contents of the chan_loc field are determined by parsing the chanmap bit field in every dependent substream associated with a particular independent substream, and setting the corresponding channel locations in the chan_loc field to a value of 1.

Because this field is used by the system only to indicate the unique channel locations present in the bitstream, it is not necessary to reflect replacement channels in this field. Therefore, duplicate channel locations in the chanmap field indicate replacement channels and can be ignored.

flag_ec3_extension_type_a

The flag_ec3_extension_type_a parameter indicates whether Dolby Atmos content is present in a bitstream. Its value is equal to the value of the flag_ec3_extension_type_a bit found in the addbsi.

complexity_index_type_a

The complexity_index_type_a parameter indicates the decoding complexity of a Dolby Digital Plus bitstream carrying Dolby Atmos content. If the flag_ec3_extension_type_a bit in the addbsi has a value of 0, the complexity_index_type_a will not be available in the EC3SpecificBox. If the flag_ec3_extension_type_a bit in the addbsi has a value of 1, the complexity_index_type_a in the EC3SpecificBox shall take the value of the complexity_index_type_a field found in the addbsi of the bitstream.

Glossary

A/V
Audio/video.

AAC
Advanced Audio Coding. A perceptual audio coding system that is described by ISO/IEC 14496-3.

access unit
All of the frames required by a Dolby Digital Plus decoder to produce 1,536 samples of decoded audio for each channel present in the bitstream (including all substreams). The first frame of an access unit has a substream type of 0 and a substream ID value of 0. If the access unit consists of frames that contain one, two, or three blocks of audio, the first frame of the access unit has the convsync parameter set to 1.

AVC
Advanced Video Coding. See H.264.

bed object
An object with positional metadata that does not change over time and is described by a predefined speaker position.

block
A portion of a frame.

BSI
Bitstream information. Information included in an encoded audio bitstream that describes the audio (metadata) or provides instructions to a decoder (or other device in the audio reproduction chain) on how to process the audio. Dolby Digital Plus and Dolby Digital BSI are detailed in ETSI TS 102 366 and ATSC A/52.

CRC
Cyclic redundancy check.

DASH
Dynamic Adaptive Streaming over HTTP. An adaptive bit-rate streaming protocol that enables high-quality streaming of media content over the Internet delivered from HTTP.

DISC
Dolby Interoperability Support Center.

dynamic object
An object with positional metadata that may vary over time and is described by three coordinates (x, y, z).

elementary stream
A bitstream that is the output of an audio or video encoder and contains only one type of data, such as audio or video.

E-AC-3
Enhanced AC-3.

Dolby Digital Plus, also known as Enhanced AC-3 or E-AC-3, is a digital audio compression coding system for transport and storage of multichannel digital audio specified in Annex E of ATSC A/52 and Annex E of ETSI TS 102 366.

The file extension for a Dolby Digital Plus file is .ec3.

frame
In audio, a series of PCM samples or encoded audio data representing the same time interval for all channels in the configuration. Metadata pertaining to the frame can be carried within the frame or separately, depending on context.

frame set
Six consecutive blocks of Dolby Digital Plus audio data from a single substream. A frame set always represents 1,536 samples of audio data from a single substream.

GUID
Globally unique identifier. A unique reference number used as an identifier in software.

HDMI
High-Definition Multimedia Interface. A high-speed, high-capacity format for transferring digital information and the specific hardware interface for the format.

HLS
HTTP Live Streaming. An adaptive streaming protocol for delivery of media content developed by Apple.

ISO
International Organization for Standardization.

joint object coding
An algorithm used to efficiently code object-based audio content.

master playlist
A playlist where all of the URI lines in the playlist point to media playlists. A playlist that contains URIs that point to alternative content for a presentation, such as alternative language versions of the content.

media assets
A collection of files that contains a multimedia presentation formatted for adaptive streaming. Generally, a media asset consists of multiplexed and fragmented media and one or more files that describe how to play back the media (for example, playlist or manifest files).

media presentation
A collection of files that contains media content prepared for adaptive streaming. The presentation includes media files that contain the content and files that describe how to access and play the content.

MP4 sample
A single ISO base media file track sample, as defined in section 3.1.10 of *ISO/IEC 14496-12*.

MPD
Media Presentation Description. A manifest used in MPEG Dynamic Adaptive Streaming over HTTP (MPEG-DASH) to describe the available streaming content, its various alternatives, URL addresses, and other characteristics, as well as segments that contain the actual multimedia bitstreams in the form of chunks, in single or multiple files.

MPEG
Moving Picture Experts Group. An ISO/IEC working group that develops video and audio encoding standards. Also the name of a family of digital video and audio coding standards.

MPEG-4
An MPEG standard (ISO/IEC 14496) for a group of audio and video coding formats and related technologies.

object
An audio signal plus associated object audio metadata.

object audio metadata
Information used for rendering an audio object. Comprises metadata such as positional metadata, content metadata, or metadata for personalization. Each object must at least have associated positional metadata containing specific information for the renderer. The specification of positional metadata is different for each object type.

object audio renderer
Renders object-based audio to a specific speaker layout. The input is composed of objects, and the outputs are speaker feeds.

PAT
Program Association Table. Program Association Table of an MPEG-2 transport stream.

PCR
Program clock reference. A periodically transmitted value of 42 bits that provides a sample of the system time clock in the encoder and which is used to properly demultiplex packets and to ensure that audio and video are synchronized.

PES
Packetized elementary stream. An elementary stream that is split into small chunks (packets) for transmitting and combining multiple streams within a transport stream. Each PES is identified by a unique packet identifier (PID).

PID
Packet identifier. A unique code that identifies a packetized elementary stream (PES) within a transport stream. The PID is contained in the transport stream packet header and is listed in the service information (SI) for a transport stream.

playlist
An extended .m3u8 file that contains one or more uniform resource identifiers (URIs). A URI can point to another playlist or to a media file.

PMT
Program Map Table. A table within an MPEG-2 transport stream that defines the set of elementary streams associated with a specific program.

presentation
References to AC-4 substreams to be decoded and presented simultaneously.

presentation configuration
Set of metadata to describe how a presentation must be decoded.

PTS
Presentation time stamp. The presentation time stamp is contained in the packetized elementary stream (PES) packet header that indicates when an access unit should be decoded and presented for output. The PTS is used in combination with other time stamp parameters to synchronize audio and video.

substream
A decodable unit that represents a specific channel configuration (mono, stereo, or 5.1) and contains audio data and corresponding metadata.

substream ID
A metadata field in a Dolby Digital Plus bitstream for numbering a substream. The substream ID, in combination with the substream type, identifies a substream within a Dolby Digital Plus stream, as defined in Annex E of ETSI TS 102 366.

substream type
A metadata field in a Dolby Digital Plus bitstream that describes a substream. There are different types of substreams that make up a Dolby Digital Plus bitstream. As defined in Annex E of ETSI TS 102 366, independent substreams (type 0) may be decoded independently of any other substreams that might exist in the bitstream. Dependent substreams (type 1) must be decoded in conjunction with the independent substream with which it is associated.

time slice
A collection of Dolby Digital Plus frames that represents the audio data from the same point in time across multiple substreams. A time slice may represent one, two, three, or six blocks of audio data, depending on the number of blocks used per frame.

transport stream
As defined in ISO/IEC 13818-1, a packetized bitstream that is used to transmit audio and video information. A transport stream is made up of multiplexed program elementary streams.

transport stream segment
A single .ts file that is part of an HTTP Live Streaming (HLS) transport stream.

URI
Uniform Resource Identifier. A group of characters identifying a resource on a network (typically, the Internet).

variant playlist
A playlist that contains Uniform Resource Identifiers (URIs) that point to alternative content for a presentation, such as alternative language versions of the content. A variant playlist lists URIs for each variant presentation so that a playback client can switch between playback of the streams dynamically based on parameters such as language, bit rate, and channel configuration.