Dolby Vision

Streams within the MPEG-DASH format

Version 2.0
13 November 2018

Copyright

For information, contact:

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.

Contents

1 Introduction to Dolby Vision streams within the MPEG-DASH format

This document specifies the required data formatting and signaling between a server and its clients to enable Dolby Vision to be transported within MPEG Dynamic Adaptive Streaming over HTTP (MPEG-DASH) in conjunction with the ISO base media file format.

1.1 MPEG-DASH

MPEG-DASH is an adaptive bit-rate streaming technology that partitions a media presentation into segments and delivers them from a server to a client via HTTP.

The content on the server consists of two parts: a Media Presentation Description (MPD), and a sequence of small HTTP-based file segments, each of which can be downloaded independently. The MPD is a manifest of the available content, its various alternatives, their URL addresses, and other characteristics (including timing information and media attributes 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. Each segment contains a short interval of playback time (typically ten seconds). Segments may be played back sequentially to ensure that a program that is many hours in duration can be streamed. MPEG-DASH enables reuse of commonly available HTTP servers/caches to deliver media content to end users. MPEG-DASH supports commonly available container and media formats/codecs.

1.2 Dolby Vision stream multiplexing in MPEG-DASH

For transmission using MPEG-DASH, the base-layer and reference-picture-unit substreams are combined into a single Dolby Vision stream. The Dolby Vision stream can be multiplexed into a fragmented MP4 or an MPEG transport stream.

For detailed information about the process of combining the base layer and reference-picture-unit substreams into one single Dolby Vision stream, refer to the *Dolby Vision VES multiplexing specification*.

1.3 Standards and Dolby documents

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

Standards

- ISO/IEC 23009-2:2014, *Information Technology—Dynamic Adaptive Streaming over HTTP (DASH)—Part 1: Media Presentation Description and Segment Formats*, available from http://www.iso.org.
- ISO/IEC 14496-12:2015, *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 14496-15:2014, *Information Technology—Coding of Audio-Visual Objects, Part 15: Carriage of Network Abstraction Layer (NAL) Unit Structured Video in ISO Base Media File Format*, available from http://www.iso.org
- ISO/IEC 14496-10:2009, *Information Technology—Coding of Audio-Visual Objects, Part 10: Audio*, available from http://www.iso.org.
- ISO/IEC 23008-2:2013, *Information Technology—High Efficiency Coding and Media Delivery in Heterogeneous Environments, Part 2: High Efficiency Video Coding*, 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.

Dolby documents

- *Dolby Vision Video Elementary Stream Multiplexing Specification* (dovi_proenc_ves_muxing_spec.pdf), located in the documentation package of the *Dolby Vision Professional Encoder Software Integration Development Kit*.
- *Dolby Vision Streams Within the ISO Base Media File Format*, available from https://www.dolby.com/us/en/technologies/dolby-vision/dolby-vision-for-creative-professionals.html#5.
- *Dolby Vision Profiles and Levels*, available from https://www.dolby.com/us/en/technologies/dolby-vision/dolby-vision-for-creative-professionals.html#5.

1.4 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 Dolby Vision encryption requirements

Dolby Vision is agnostic if the encryption follows ISO/IEC 23001-7.

To facilitate stream reformatting before decryption, it is required that at least the length field and the `nal_unit_type` field (the first byte after the length field) of each unit is left unencrypted, and recommended that the entire slice headers remain unencrypted. In addition, it should be noted that:

- For other structured video sample description stream formats (for example, avc3, hvc1, and hev1), only video slice data should be encrypted. The size and type headers of a video slice should be unencrypted. For the other types, no encryption is needed.
- There are multiple units per sample, requiring multiple pieces of clear and encrypted data per sample.
- units that do not contain video slice data need not be encrypted, and should not be encrypted if they contain information that must be accessed prior to decryption, such as caption information contained in SEI units.

Although some of the statements about encryption are recommendations in ISO/IEC 23001-7, some will make the following items as requirements moving forward:

- units that do not contain any audio or video data must remain entirely unencrypted. This means that the Dolby Vision reference picture unit must remain entirely unencrypted.
- units that contain audio/video data must have the slice header unencrypted.
- In the MP4 box structure, the fourCC string (for example, dvh1 for profile 5) must be replaced with encv.

3 Dolby Vision profiles and levels

The Dolby Vision format provides many features, not all of which apply to every application. Subsets of Dolby Vision features are stipulated in profiles and levels. Develop your application to support one or more of these profiles and levels as needed.

3.1 Dolby Vision profiles

A Dolby Vision profile is a subset of Dolby Vision feature configurations predefined by Dolby.

This table only lists profiles that are supported in online streaming formats. There are more profiles applicable to other use cases. For details, refer to *Dolby Vision profiles and levels*.

Profile ID	BL codec	Dolby Vision fourCC string
5	10-bit HEVC	dvhe or dvh1
8	10-bit HEVC	dvhe or dvh1

In this table,

- **BL codec**, `10-bit HEVC`, indicates H.265 main10 profile.
- **Dolby Vision FourCC string** is the codec fourCC string used within the `CODEC` tag.

  Dolby Vision codecs use strings begin with d. As defined in *Dolby Vision Streams Within the ISO Base Media File Format*, Dolby Vision-specific configuration boxes may be used with standard codecs strings for certain profiles. Codecs other than `HEVC` may be supported in the future, for which additional Dolby Vision bitstream profile IDs will be added.

3.2 Dolby Vision levels

A Dolby Vision level specifies the maximum pixel rate, decoded bitstream video width, and bit rate supported by a product within a given bitstream profile.

Typically, there is a limit on the maximum number of pixels a product can process per second within a given bitstream profile; the levels defined here generally correspond to the product processing capability. Although not listed, noninteger frame rates are supported.

Table 1: Dolby Vision levels

Level ID	Maximum pixel rate (pps)	Maximum decoded bitstream video width (pixels)	Example decoded bitstream resolution @ frame rate (fps)	Maximum bit rates	
				Main tier (Mbps)	High tier (Mbps)
01	22,118,400	1280	1280 × 720 @ 24	20	50
02	27,648,000	1280	1280 × 720 @ 30	20	50
03	49,766,400	1920	1920 × 1080 @ 24	20	70
04	62,208,000	2560	1920 × 1080 @ 30	20	70
05	124,416,000	3840	1920 × 1080 @ 60	20	70

Table 1: Dolby Vision levels (continued)

Level ID	Maximum pixel rate (pps)	Maximum decoded bitstream video width (pixels)	Example decoded bitstream resolution @ frame rate (fps)	Maximum bit rates	
				Main tier (Mbps)	High tier (Mbps)
06	199,065,600	3840	3840 × 2160 @ 24	25	130
07	248,832,000	3840	3840 × 2160 @ 30 Also supports: 1920 × 1080 @ 120 *	25	130
08	398,131,200	3840	3840 × 2160 @ 48	40	130
09	497,664,000	3840	3840 × 2160 @ 60	40	130
10	995,328,000	3840	3840 × 2160 @ 120	60	240
11	995,328,000	7680	• 7680 × 4320 @ 30 • 3840 × 2160 @ 120	60	240
12	1,990,656,000	7680	7680 × 4320 @ 60	120	480
13	3,981,312,000	7680	7680 × 4320 @ 120	240	800

* This frame rate can be used for Dolby Vision bitstreams packetized in a single program compliant to MPEG-2 transport streams.

The columns in this table include:

- **Maximum pixel rate (PPS)**: This column lists imposed limits on arithmetic combinations of decoded bitstream resolution and frame rate (decoded bitstream resolution multiplied by frame rate: horizontal pixels × vertical pixels × frame rate). The maximum pixels per second is a constant for a given level. The decoded bitstream resolution is inversely proportional to the frame rate, meaning that the decoded bitstream resolution can be reduced for obtaining higher frame rate, and vice versa. Note that the decoded bitstream resolution here is for baseband video, irrespective of the particular video compression codec that is used.
- **Example decoded bitstream resolution @ frame rate (fps)**: Baseband picture horizontal and vertical pixels followed by frame rate.
- **Maximum bit rates**: This column indicates the maximum bit rate of the encoded bitstream.
- **High tier**: Note that for Dolby Vision bitstream profile 7, Blu-ray Disc Association specifications allow a maximum high tier bit rate of 100 Mbps for each level. Similarly, there may be other Dolby Vision enabled systems that limit or require different maximum bit rates. Additionally, high tier may be required for some applications that use temporal subscale layers.

4 Signaling Dolby Vision in a Media Presentation Description file

Information about the Dolby Vision streams encapsulated in MPEG-DASH format can be signaled within a Media Presentation Description (MPD) file.

4.1 Media Presentation Description

A MPD file is a hierarchical XML document that provides information for an MPEG-DASH client about the available content for a media presentation.

The media presentation contains the encoded audio and video streams that are segmented into chunks for online delivery. The MPD is a manifest of the available content. The MPEG-DASH client uses the information in the MPD for constructing the HTTP URLs that then allow it to access segments containing the actual audio and video content.

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 that contain the actual audio and video content and controlling playback.

4.2 Media Presentation Description with Dolby Vision

In MPEG-DASH, the main way of signaling the type of video to be streamed is by using the mimeType attribute in conjunction with the codecs attribute. Both attributes are registered in the MPEG-DASH MPD file.

4.2.1 Adaptation sets

A media presentation typically consists of multiple media content components, for example, a Dolby Vision video component and a Dolby Digital Plus audio component. Each media content component is represented by an `AdaptationSet` element in the MPD file.

In addition to the Dolby Vision adaptation set, Dolby recommends including a separate adaptation set that refers to an SDR version of the same content as the Dolby Vision stream. This alternate adaptation set ensures that the media presentation can be decoded and played back by a device that does not support Dolby Vision.

The frame rate of the Dolby Vision stream is indicated by the `frameRate` attribute in the `AdaptationSet` element.

4.2.2 Representations

An adaptation set consists of a group of interchangeable versions of one media content component. Each version is described with a `Representations` element in the MPD file.

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 Vision streams in one adaptation set.

The resolution of the Dolby Vision stream is indicated by the `width` and `height` attributes in the `Representations` element.

4.2.3 mimeType and codecs attribute values

Use the `mimeType` and codecs attributes to signal the codec, profile, and level information of the referenced Dolby Vision stream.

The `mimeType` and codecs attributes can be placed within either the `AdaptationSet` or `Representations` element.

The `mimeType` attribute describes the encapsulation format used to store the Dolby Vision streams present in the adaptation set. For adaptation sets that conform to ISO/IEC 14496-12, the `mimeType` attribute must be set to `video/mp4` (for ISO base media files that contain both audio and video tracks).

The codecs attribute indicates the codec in which the Dolby Vision stream is encoded. For a Dolby Vision stream, the video codec value is composed in this format:

```
[Dolby_Vision_fourCC].[Dovi_Profile_ID].[Dovi_Level_ID]
```

In this string:

- The `[Dolby_Vision_fourCC]` is a four-character value as listed in the *Sample description entry code* table that indicates the codec type. This value can be obtained from Dolby Vision sample entries, as described in the Dolby Vision Streams Within the ISO Base Media File Format specification.
- The `[Dovi_Profile_ID]` is a two-digit value representing the Dolby Vision profile ID as listed in the *Dolby Vision profiles* section. This value can be obtained from Dolby Vision decoder configuration box, as described in the Dolby Vision Streams Within the ISO Base Media File Format specification.
- The `[Dovi_Level_ID]` is a two-digit value representing the Dolby Vision level ID as listed in the *Dolby Vision levels* section. This value can be obtained from Dolby Vision decoder configuration box, as described in the Dolby Vision Streams Within the ISO Base Media File Format specification.

For a Dolby Vision stream that is not backward compatible, only a Dolby Vision fourCC code is referenced in the MPD; for a Dolby Vision stream that is backward compatible, both a standard High-Efficiency Video Coding (HEVC) fourCC code and a Dolby Vision fourCC code must be referenced in the MPD.

Table 2: Sample description entry code

Codec	Codec variant	SDR fourCC code	Dolby Vision fourCC code
HEVC	Parameter sets (VPS, PPS, or SPS) are stored either in the sample entries or as part of the samples, or in both.	hev1	dvhe
	Parameter sets (VPS, PPS, or SPS) are stored in the sample entries only.	hvc1	dvh1

For example:

```
mimeType="video/mp4" codecs="dvhe.05.07"
```

In this example, the highlighted string indicates that the video element is signaled as a single-layer, non-backward-compatible Dolby Vision stream encoded as 10-bit HEVC video with a maximum resolution of 3840 × 2160 at 30 fps. This corresponds to Dolby Vision profile ID 5 and level ID 7.

For a backward-compatible Dolby Vision stream, besides referencing the Dolby Vision codec, the MPD must also contain a reference to the backward compatible codec in a separate adaptation set. The CODECS attribute uses the codec values described in ISO/IEC 14496-15.

```
[BL_compatible_codec]
```

In the string, the [BL_compatible_codec] takes a backward-compatible codec value format that begins with a fourCC code and may optionally be followed by profile and level information; for details, refer to RFC 6381.

Related information

Dolby Vision profiles on page 7
Dolby Vision levels on page 7

4.3 Media Presentation Description file examples

Examples of MPEG-DASH MPD files that contain Dolby Vision video elements, as well as Dolby Digital Plus audio media elements, are provided in this section.

4.3.1 Profile 8 Dolby Vision stream

This MPD example describes a media presentation that consists of a Dolby Vision profile 8.1 video component with a Dolby Digital Plus audio component. The essence of the Dolby Vision track is a profile 8.1 Dolby Vision stream encoded as 10-bit HEVC video with a resolution of 3840 × 2160 at 30 fps, where the bitstream is backward-compatible with HDR10 decoders.

```
<?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="PT0H1M2.550S"
        minBufferTime="PT1.2S"
```

```
        profiles="urn:mpeg:dash:profile:isoff-on-demand:2011">
        <BaseURL>./</BaseURL>
        <Period>
          <!-- Video -->
          <AdaptationSet mimeType="video/mp4" codecs="hvc1.2.4.L153.b0"
                subsegmentAlignment="true" subsegmentStartsWithSAP="1">
             <Representation id="1" bandwidth="17677192"
              width="3840" height="2160" frameRate="30">
                <BaseURL>video_DV_p81.mp4</BaseURL>
                <SegmentBase indexRange="812-1095">
                  <Initialization range="0-811"/>
                </SegmentBase>
             </Representation>
          </AdaptationSet>
          <!-- Video -->
          <AdaptationSet mimeType="video/mp4" codecs="dvh1.08.07"
                subsegmentAlignment="true" subsegmentStartsWithSAP="1">
             <Representation id="1" bandwidth="17677192"
              width="3840" height="2160" frameRate="30">
                <BaseURL>video_DV_p81.mp4</BaseURL>
                <SegmentBase indexRange="812-1095">
                  <Initialization range="0-811"/>
                </SegmentBase>
             </Representation>
          </AdaptationSet>
          <!-- Audio -->
          <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="und"
           subsegmentAlignment="true" subsegmentStartsWithSAP="1">
             <Representation id="2" bandwidth="192000">
               <AudioChannelConfiguration
             schemeIdUri="tag:dolby.com,2014:dash:audio_channel_configuration:2011"
             value="F801"/>
                <BaseURL>audio.mp4</BaseURL>
                <SegmentBase indexRange="652-875">
                  <Initialization range="0-651"/>
                </SegmentBase>
             </Representation>
          </AdaptationSet>
        </Period>
</MPD>
```

In this MPD example, the HEVC fourCC code for the base layer is hvc1, which determines the Dolby Vision fourCC code must be set to dvh1. In case the HEVC fourCC code for the base layer is hev1, the Dolby Vision fourCC code must be set to dvhe, as indicated in the *mimeType and codecs attribute values* section; consequently the video codecs attribute would be `hev1.2.4.L153.b0,dvhe.08.07`.

Related information

mimeType and codecs attribute values on page 10

4.3.2 Profile 5 Dolby Vision stream

This MPD example describes a media presentation that consists of a Dolby Vision profile 5 video component with a Dolby Digital Plus audio component. The essence of the Dolby Vision track is a profile 5 Dolby Vision stream encoded as 10-bit HEVC video with a resolution of 3840 × 2160 at 30 fps, where the base layer is not backward compatible.

```
<?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="PT0H1M2.550S"
        minBufferTime="PT1.2S"
        profiles="urn:mpeg:dash:profile:isoff-on-demand:2011">
        <BaseURL>./</BaseURL>
        <Period>
          <!-- Video -->
          <AdaptationSet mimeType="video/mp4" codecs="dvhe.05.07"
           subsegmentAlignment="true" subsegmentStartsWithSAP="1">
            <Representation id="1" bandwidth="17677192" width="3840"
             height="2160" frameRate="30">
               <BaseURL>h265_single_track_NBC.mp4</BaseURL>
               <SegmentBase indexRange="812-1095">
                  <Initialization range="0-811"/>
               </SegmentBase>
            </Representation>
          </AdaptationSet>
          <!-- Audio -->
          <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="und"
subsegmentAlignment="true"
        subsegmentStartsWithSAP="1">
             <Representation id="2" bandwidth="192000">
               <AudioChannelConfiguration
                schemeIdUri="tag:dolby.com,2014:dash:audio_channel_configuration:2011"
                value="F801"/>
               <BaseURL>audio.mp4</BaseURL>
               <SegmentBase indexRange="652-875">
                  <Initialization range="0-651"/>
               </SegmentBase>
            </Representation>
          </AdaptationSet>
        </Period>
</MPD>
```

In this MPD example, the base layer is non-backward-compatible, and the Dolby Vision fourCC code is dvhe. In case the Dolby Vision fourCC code is dvh1, the video codecs attribute must be set to dvh1.05.07.

Related information

mimeType and codecs attribute values on page 10

Glossary

AVC
: Advanced Video Coding. See H.264.

HEVC
: High-Efficiency Video Coding. See H.265.

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-DASH
: MPEG 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.