Article

HLS Authoring Specification for Apple Devices

Describes the requirements for live and VOD audio-video content delivery using HTTP Live Streaming (HLS) to Apple devices.

On This Page
Overview
Topics
See Also

Overview

For a deeper discussion of the features available in HLS, please refer to Apple's streaming resource page at http://developer.apple.com/streaming/, which contains pointers to the overview document, the HLS specification, technical notes, FAQs, presentations, and examples.

About HLS Authoring

The HLS specification is a published RFC [RFC8216]. However, HLS continues to evolve, so there is an updated draft specification - draft-pantos-hls-rfc8216bis [HLS2]. This document always uses the most recent version of the draft standard.

> **Note**
>
> HLS was originally specified in draft-pantos-http-live-streaming [HLS1]. That document was superseeded by RFC 8216.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

Find additional information for HTTP Live Streaming Specifications here: HLS Authoring Specification for Apple Devices Appendices.

General Authoring Requirements

A stream that matches these requirements should be compatible with iOS, tvOS, or macOS.

Video

1. **Video encoding requirements**

1.1. All video MUST be encoded using H.264/AVC or HEVC/H.265.

1.2. The container format for H.264 video MUST be fragmented MP4 (fMP4) files or MPEG transport streams.

1.3. Profile and Level for H.264 MUST be less than or equal to High Profile, Level 4.2.

1.4. For H.264 you SHOULD use High Profile in preference to Main or Baseline Profile.

1.5. The container format for HEVC video MUST be fMP4.

1.6. Profile, Level, and Tier for HEVC MUST be less than or equal to Main10 Profile, Level 5.0, High Tier.

1.7. High Dynamic Range (HDR) HEVC video MUST be HDR10 or DolbyVision.

1.8. For HDR10 video, the SEI NAL units (i.e. static metadata) SHOULD be in the HEVC Configuration Box ('hvcC') and not in the individual sample data.

1.9. Profile and Level for Dolby Vision MUST be Profile 5 (single layer 10-bit HEVC) and less than or equal to Level 7.

1.10. You SHOULD use video formats in which the parameter sets are stored in the sample descriptions, rather than the samples. (i.e., Use 'avc1', 'hvc1', or 'dvh1' rather than 'avc3', 'hev1', or 'dvhe'.)

1.11. For backward compatibility, content SHOULD NOT use a higher level than required by the content resolution and frame rate.

1.12. For backward compatibility, some video content SHOULD be encoded with H.264.

1.13. Key frames (IDRs) SHOULD be present every two seconds.

1.14. All interlaced source content MUST be de-interlaced.

1.15. You SHOULD de-interlace 30i content to 60p instead of 30p.

1.16. Live/linear video from NTSC or ATSC source SHOULD be 60 or 59.94 fps.

1.17. Live/linear video from PAL source SHOULD be 50 fps.

1.18. VOD content SHOULD use a natural frame rate for the content. Any of the typical frame rates: 23.976, 24, 25, 29.97, 30, 50, 59.94, and 60 fps are supported.

1.19. Frame rates above 60 fps SHALL NOT be used.

1.20. For HDR content, frame rates less than or equal to 30 fps SHOULD be provided.

1.21. Streams SHOULD use a single color space: one of either Rec. 601, Rec. 709, DCI-P3, or Rec. 2020.

1.22. For VOD, the color space SHOULD be the original color space of the material.

1.23. If multiple video streams are provided (H.264, HEVC, HDR), then each stream SHOULD provide all anticipated bandwidths. Clients SHOULD NOT be required to switch codecs.

1.24. For backward compatibility, SDR streams MUST be provided. (See also item 6.16)

1.25. There are many possible choices of bit rates for variants. The following tables provide one possible set of bit rate variants. (See On bit rates for variants for additional considerations.):

Table 2 Video average bit rate (kb/s) table 1

16:9 aspect ratio	H.264/AVC	Frame rate
416 x 234	145	≤ 30 fps
640 x 360	365	≤ 30 fps
768 x 432	730	≤ 30 fps
768 x 432	1100	≤ 30 fps
960 x 540	2000	same as source
1280 x 720	3000	same as source
1280 x 720	4500	same as source
1920 x 1080	6000	same as source
1920 x 1080	7800	same as source

Table 3 Video average bit rate (kb/s) table 2

16:9 aspect ratio	HEVC/H.265 30 fps	HDR (HEVC) 30 fps	Frame rate
640 x 360	145	160	≤ 30 fps
768 x 432	300	360	≤ 30 fps
960 x 540	600	730	≤ 30 fps
960 x 540	900	1090	≤ 30 fps
960 x 540	1600	1930	same as source
1280 x 720	2400	2900	same as source
1280 x 720	3400	3850	same as source
1920 x 1080	4500	5400	same as source
1920 x 1080	5800	7000	same as source
2560 x 1440	8100	9700	same as source
3840 x 2160	11600	13900	same as source
3840 x 2160	16800	20000	same as source

> **Note**
>
> The above bit rates are initial encoding targets for typical content delivered via HLS. We recommend you evaluate them against your specific content and encoding workflow then adjust accordingly.
>
> 30i source content is considered to have a source frame rate of 60 fps. 24fps HEVC content should use bit rates reduced by about 20% from the values above.

1.26. For VOD content, the average segment bit rate MUST be within 10% of the AVERAGE-BANDWIDTH attribute. (See Declared versus measured values of bandwidths.)

1.27. For VOD content, the measured peak bit rate MUST be within 10% of the BANDWIDTH attribute.

1.28. For live/linear content, the average segment bit rate over a long (~1 hour) period of time MUST be less than 110% of the AVERAGE-BANDWIDTH attribute.

1.29. For live/linear content, the measured peak bit rate MUST be less than 125% of the BANDWIDTH attribute.

1.30. For VOD content the peak bit rate SHOULD be no more than 200% of the average bit rate.

1.31. Different variants MAY have different frame rates.

1.32. The default video variant(s) SHOULD be the 2000 kb/s (average bit rate) variant. (Defaults are the first variant listed in the master playlist within a group of variants having compatible audio.)

1.33. All video variants SHOULD have identical aspect ratios.

Audio

2. **Audio encoding requirements**

2.1. Audio data SHOULD be provided as an elementary audio stream or in fMP4.

2.2. Stereo audio formats are:

- AAC-LC
- HE-AAC v1
- HE-AAC v2
- Multichannel formats that only carry stereo

2.3. Stereo audio in some AAC format MUST be provided.

2.4. You SHOULD NOT use HE-AAC if your audio bit rate is above 64 kb/s.

2.5. Supported multichannel audio formats are:

- Dolby Digital (AC-3)
- Dolby Digital Plus (E-AC-3)
- Dolby Digital Plus with Dolby Atmos

2.6. If Dolby Digital Plus is provided and your streams are delivered to devices that do not support Dolby Digital Plus then Dolby Digital MUST be provided also. See Audio Compatibility.

2.7. Stereo audio MAY be provided at multiple bit rates

2.8. Multichannel audio MUST be provided at a single bit rate. (For example, don't provide Dolby Digital 5.1 at two bit rates, but you may provide Dolby Digital 5.1 and Dolby Digital Plus 5.1.)

2.9. Overall bit rate recommendations are as follows:

Audio channels	Format	Total (kb/s)
2.0 (stereo)	AAC	32 to 160
2.0 (stereo)	Dolby Digital Plus	96 to 160
5.1 (surround)	Dolby Digital	384
5.1 (surround)	Dolby Digital Plus	192
7.1 (surround)	Dolby Digital Plus	384
Nominally 16	Dolby Digital Plus with Dolby Atmos	384 to 768

2.10. You MAY change channel layout within a stream.

2.11. Channel layout changes MUST be in conjunction with an EXT-X-DISCONTINUITY tag.

2.12. If you provide Descriptive Video Service (DVS), also called Audio Description, it MUST be marked with the attribute CHARACTERISTICS=`"public.accessibility.describes-video"`.

2.13. If you provide DVS, the AUTOSELECT attribute MUST have a value of "YES".

2.14. The name of DVS audio SHOULD indicate that the stream is DVS content.

2.15. If you provide alternate audio or DVS, you MUST provide it for the entire content duration even if it only really exists for a portion of the main content.

2.16. Alternate audio or DVS MAY reuse the audio segments of regular content when the alternate content is not available.

2.17. A single media segment MUST NOT contain multiple audio streams.

2.18. Loudness management/Dynamic Range Control SHOULD be present unless all content (including ads) has been encoded at the same audio levels.

2.19. Dolby Digital and Dolby Digital Plus loudness SHOULD be specified by the dialnorm field (ATSC A/52:2012).

2.20. AAC dialog loudness SHOULD be specified by either AAC `prog_ref_level` (ISO 14496-3 sub-clause 4.5.2.7), as specified by SCTE 193-1 section 7.4.1; OR by the `loudnessInfo()` payload as specified by ISO 23003-4, in which case `samplePeakLevel` or `truePeakLevel` MUST be present, `measurementSystem` MUST be 2, and `methodDefinition` MUST be 1 or 2.

2.21. HE-AAC in fMP4 files MUST use explicit signaling of SBR data.

Ads and Pre-/Mid-/Post-Rolls

3. **Ad requirements** (see also Media Playlists section)

3.1. Your ads and other inserted media SHOULD be included in the media playlists.

3.2. Inserted media SHOULD be encoded with the same codecs as the other content in the Media Playlist.

3.3. Inserted media SHOULD be encoded with the same aspect ratio as the other content in the Media Playlist.

3.4. Inserted media SHOULD be at or under the specified bandwidths for that variant.

Accessibility

4. **Accessibility requirements**

4.1. Captions SHOULD be provided with your streams to make content accessible to the deaf or hard of hearing.

4.2. Supported caption formats are:

- CEA-608 closed captions
- CEA-708 closed captions
- WebVTT subtitles
- IMSC1 subtitles (text profile only)

4.3. Closed captions (if any) MUST be included in the video media segments.

4.4. The presence of closed captions MUST be declared via an EXT-X-MEDIA tag that contains a LANGUAGE attribute.

4.5. If a subtitles track is intended to provide accessibility for the deaf and hard of hearing it MUST be marked with the attribute CHARACTERISTICS=`"public.accessibility.describes-music-and-sound"`. (Subtitles with this attribute value are treated the same as closed captions.)

4.6. If a subtitles track is intended to provide accessibility for the deaf and hard of hearing the AUTOSELECT attribute MUST have a value of "YES".

4.7. The LANGUAGE attribute MUST be included in the EXT-X-MEDIA tag for a subtitles track.

4.8. For information on including Audio Descriptions with your streams, see items 2.12 through 2.16.

Subtitles

5. **Subtitle requirements**

5.1. Subtitles MAY be provided.

5.2. Subtitles MUST be WebVTT (according to the HLS specification) or IMSC1 in fMP4.

5.3. WebVTT subtitles MUST be in text files, with an X-TIMESTAMP-MAP according to the HLS specification.

5.4. IMSC1 subtitles MUST be in fMP4 files.

5.5. The subtitle playlist MUST exist for the entirety of the main content, even if the actual subtitles only exist for a portion of the content.

5.6. For live/linear content, target durations for subtitle playlists MUST be identical to other media.

5.7. For VOD content, target durations of subtitle playlists MAY be longer than the other media.

5.8. If the content has forced subtitles and regular subtitles in a given language, the regular subtitles track in that language MUST contain both the forced subtitles and the regular subtitles for that language.

5.9. If your videos contain text burnt into the video and you have access to a version without the burn-in, you SHOULD use Forced Subtitles instead. (This allows you to easily translate into multiple languages. An example of when you might use forced subtitles is a science fiction film, where alien languages are translated into English.)

5.10. The kind of subtitles SHOULD be specified in the CODECS attribute of the associated EXT-X-STREAM-INF tags.
You SHOULD use "stpp.ttml.im1t" to identify IMSC1 subtitles. You MAY use "wvtt" to identify WebVTT subtitles.

5.11. Forced subtitles SHOULD always have `AUTOSELECT=YES`.

Trick Play

6. **Trick play requirements**

6.1. I-frame playlists (EXT-X-I-FRAME-STREAM-INF) MUST be provided to support scrubbing and scanning UI.

6.2. You SHOULD have one frame per second "dense" I-frame renditions. These are dedicated renditions that only contain I-frames.

6.3. Alternatively, you MAY use the I-frames from your normal content, but trick play performance is improved with a higher density of I-frames.

6.4. If you provide multiple bit rates at the same spatial resolution for your regular video then you SHOULD create the I-frame playlist for that resolution from the same source used for the lowest bit rate in that group.

6.5. The bit rate of I-frame playlists SHOULD be the bit rate of a normal playlist of the same resolution times the fps of the I-frame playlist divided by eight. (See I-frame bit rates versus normal bit rates.)

6.6. You SHOULD provide multiple I-frame Media Playlists with different bit rates.

6.7. As with normal video, there are many possible choices of bit rates for dense I-frame variants. The following tables provide one possible set of variants. (See On bit rates for variants.):

Table 5 Trick play average video bit rate (kb/s) table 1

16:9 aspect ratio	H.264/AVC
640 x 360	45
768 x 432	90
960 x 540	250
1280 x 720	375
1920 x 1080	580

16:9 aspect ratio	HEVC/H.265	HDR (HEVC)
768 x 432	40	55
960 x 540	75	94
960 x 540	200	238
1280 x 720	300	360
1920 x 1080	525	650

> **Note**
>
> These bit rates are based on the assumption of a 1 fps dense I-frame track. The values should be modified based on the actual frame rate of the I-frame rendition. (See item 6.5.)
>
> The above tables are average bit rate. Earlier versions of this specification used peak bit rates.

6.8. I-frame playlists MUST contain the EXT-X-I-FRAMES-ONLY tag.

6.9. The peak segment bit rate MUST be calculated according to the HLS specification.

6.10. If using fMP4, I-frame segments MUST include the 'moof' header associated with the I-frame