

I-frame playlist with HLS CMAF

Sanchayan Maity



Agenda



- HTTP Live Streaming (HLS)
- Playlist examples
- Understanding MP4 boxes
- Implementation

- Developed by Apple and released in 2009
- Standardised in [RFC 8216](#)
- HTTP-based adaptive bit-rate streaming communications protocol
- HTTP traffic, unlike UDP-based protocols such as RTP
- Delivered over widely available HTTP-based content delivery networks

Media playlist



```
#EXTM3U
#EXT-X-VERSION:7
#EXT-X-INDEPENDENT-SEGMENTS
#EXT-X-TARGETDURATION:10
#EXT-X-PLAYLIST-TYPE:VOD
#EXT-X-MAP:URI="init00000.mp4"
#EXTINF:10,
segment00000.m4s
#EXTINF:10,
segment00001.m4s
#EXTINF:10,
segment00002.m4s
#EXTINF:10,
segment00003.m4s
#EXT-X-ENDLIST
```

I-frame playlist



- Encoded video frames whose encoding doesn't depend on any other frame.
- I-frame Playlists can be used for
 - trick play, such as fast forward
 - rapid reverse
 - scrubbing

I-frame playlist



```
...  
#EXT-X-I-FRAMES-ONLY  
#EXT-X-MAP:URI="init00000.mp4"  
#EXT-X-BYTERANGE:2110@1332  
#EXTINF:10,  
segment00000.m4s  
#EXT-X-BYTERANGE:1785@1332  
#EXTINF:10,  
segment00001.m4s  
#EXT-X-BYTERANGE:1616@1332  
#EXTINF:10,  
segment00002.m4s  
#EXT-X-BYTERANGE:1610@1332  
#EXTINF:10,  
segment00003.m4s  
#EXT-X-ENDLIST
```

- Formally MPEG-4 Part 14
- Multimedia container format
- Can store audio, video, subtitles, images
- Supported codecs?
 - H265, H264, AV1
 - AAC

MP4 boxes



```
[ftyp] size=8+16
  major_brand = mp42
  minor_version = 1
  compatible_brand = mp42
  compatible_brand = avc1
[moov] size=8+37058
  [mvhd] size=12+96
    timescale = 600
    duration = 36057
    duration(ms) = 60095
  [trak] size=8+9497
    [tkhd] size=12+80, flags=1
      enabled = 1
      id = 1
      duration = 36057
      width = 0.000000
      height = 0.000000
    [edts] size=8+28
      [elst] size=12+16
```


MP4 boxes (ii)



```
entry_count = 1
entry/segment duration = 36057
entry/media time = 0
entry/media rate = 1
[mdia] size=8+9361
  [mdhd] size=12+20
    timescale = 22050
    duration = 1326080
    duration(ms) = 60139
    language = eng
  [hdlr] size=12+46
    handler_type = soun
    handler_name = Apple Sound Media Handler
  [minf] size=8+9263
    [smhd] size=12+4
      balance = 0
    [dinf] size=8+28
      [dref] size=12+16
        [url ] size=12+0, flags=1
```

```
    location = [local to file]
[stbl] size=8+9203
[stsd] size=12+91
    entry_count = 1
[mp4a] size=8+79
    data_reference_index = 1
    channel_count = 2
    sample_size = 16
    sample_rate = 22050
[esds] size=12+39
    [ESDescriptor] size=5+34
        es_id = 0
        stream_priority = 0
    [DecoderConfig] size=5+20
        stream_type = 5
        object_type = 64
        up_stream = 0
        buffer_size = 6144
        max_bitrate = 0
```

MP4 boxes (iv)



```
        avg_bitrate = 64000
        DecoderSpecificInfo = 13 90
        [Descriptor:06] size=5+1
...
[trak] size=8+10236
    [tkhd] size=12+80, flags=1
        enabled = 1
        id = 2
        duration = 36057
        width = 640.000000
        height = 360.000000
    [edts] size=8+28
        [elst] size=12+16
            entry_count = 1
            entry/segment duration = 36057
            entry/media time = 0
            entry/media rate = 1
    [mdia] size=8+10100
        [mdhd] size=12+20
```

MP4 boxes (v)



```
timescale = 600
duration = 36057
duration(ms) = 60095
language = eng
[hdlr] size=12+46
  handler_type = vide
  handler_name = Apple Video Media Handler
[minf] size=8+10002
  [vmhd] size=12+8, flags=1
    graphics_mode = 0
    op_color = 0000,0000,0000
  [dinf] size=8+28
    [dref] size=12+16
      [url ] size=12+0, flags=1
        location = [local to file]
  [stbl] size=8+9938
    [std ] size=12+134
      entry_count = 1
    [avc1] size=8+122
```

MP4 boxes (vi)



```
data_reference_index = 1
width = 640
height = 360
compressor =
[avcC] size=8+36
  Configuration Version = 1
  Profile = Baseline
  Profile Compatibility = e0
  Level = 30
  NALU Length Size = 4
  Sequence Parameter = [27 42 e0 1e a9 18 14 05 ff 2e 00 d4 18 04 1a db 0a d7 bd f0 10]
  Picture Parameter = [28 de 09 c8]
...
[trak] size=8+10741
...
[trak] size=8+6345
[free] size=8+0
[free] size=8+0
[mdat] size=8+5473758
```

Fragmented MP4



Fragmented structure (`moof` + `mdat` pairs) allows for,

- Low-latency streaming
- HTTP-based adaptive streaming (DASH, HLS)
- Random access at fragment boundaries
- Efficient live encoding and delivery

- Top level boxes (also called `atoms`)
 - `ftype` (file type box)
 - `moov` (Movie box)
- Boxes within `moov`
 - `mvhd` (Movie header box)
 - `trak` (Track box)
 - `tkhd` (Track header)
 - `mdia` (Media box)
- Fragment boxes
 - `moof` (Movie fragment box)
 - `mfhd` (Movie fragment header)
 - `traf` (Track fragment box)
 - `tfhd` (Track fragment header)
 - `tfdt` (Track fragment decode time)
 - `trun` (Track run box)

Init segment



```
[ftyp] size=8+20
  major_brand = cmf2
  minor_version = 0
  compatible_brand = iso6
  compatible_brand = cmfc
  compatible_brand = cfsd
[moov] size=8+732
  [mvhd] size=12+108, version=1
    timescale = 3000
    duration = 0
    duration(ms) = 0
  [trak] size=8+564
    [tkhd] size=12+92, version=1, flags=7
      enabled = 1
      id = 1
      duration = 0
      width = 320.000000
      height = 240.000000
    [mdia] size=8+452
```


Init segment (ii)



```
[mdhd] size=12+32, version=1
    timescale = 3000
    duration = 0
    duration(ms) = 0
    language = und
[hdlr] size=12+33
    handler_type = vide
    handler_name = VideoHandler
[minf] size=8+355
    [vmhd] size=12+8, flags=1
        graphics_mode = 0
        op_color = 0000,0000,0000
    [dinf] size=8+28
        [dref] size=12+16
            [url ] size=12+0, flags=1
                location = [local to file]
    [stbl] size=8+291
        [std] size=12+195
            entry_count = 1
```

Init segment (iii)



```
[avc1] size=8+183
  data_reference_index = 1
  width = 320
  height = 240
  compressor =
  [avcC] size=8+42
    Configuration Version = 1
    Profile = Baseline
    Profile Compatibility = c0
    Level = 20
    NALU Length Size = 4
    Sequence Parameter = [67 42 c0 14 da 05 07 ec 05 a8 30 30 35 28 00 00 03 00 08 00 00
03 01 e4 78 a1 55]
    Picture Parameter = [68 ce 3c 80]
.....
[mvex] size=8+32
  [trex] size=12+20
    track id = 1
    default sample description index = 1
```

Init segment (iv)



```
default sample duration = 0  
default sample size = 0  
default sample flags = 0
```

fMP4 segment



```
[styp] size=8+20
[moof] size=8+1288
  [mfhd] size=12+4
    sequence number = 1
  [traf] size=8+1264
    [tfhd] size=12+12, flags=20028
      track ID = 1
      default sample duration = 100
      default sample flags = 1010000
    [tfdt] size=12+4
      base media decode time = 0
    [trun] size=12+1212, flags=205
      sample count = 300
      data offset = 1436 <--- Offset from the start of moof to data
      first sample flags = 2000000
  [mdat] size=8+192171
```

```
[trun] size=12+1212, flags=205
  sample count = 300 <--- Note the sample count
  data offset = 1436
  first sample flags = 2000000
  entries:
    ( 0) sample_size = 2100 <-- I-frame (not always!)
    ( 1) sample_size = 529
    ( 2) sample_size = 662
  ....
    (298) sample_size = 705
    (299) sample_size = 742
```

- `cmafmux`
 - `enable-keyframe-meta` property for adding `meta` on buffers pushed downstream
 - `chunk-mode` property for chunking on key frames/I-frame
- `iframe-playlist-location` property on `hlscmafsink`
- **Uses existing media segments or single media file referenced by media playlist.**

```
videotestsrc pattern=ball num-buffers=2048 ! \  
video/x-raw,format=I420,framerate=30/1 ! videorate ! \  
x264enc bitrate=2000 speed-preset=ultrafast tune=zerolatency key-int-max=60 ! \  
h264parse config-interval=-1 ! queue ! \  
hlscmafsink target-duration=10 iframe-playlist-location=iframe-playlist.m3u8
```

I-frame playlist (A recap)



```
...  
#EXT-X-I-FRAMES-ONLY  
#EXT-X-MAP:URI="init00000.mp4"  
#EXT-X-BYTERANGE:2110@1332  
#EXTINF:10,  
segment00000.m4s  
#EXT-X-BYTERANGE:1785@1332  
#EXTINF:10,  
segment00001.m4s  
#EXT-X-BYTERANGE:1616@1332  
#EXTINF:10,  
segment00002.m4s  
#EXT-X-BYTERANGE:1610@1332  
#EXTINF:10,  
segment00003.m4s  
#EXT-X-ENDLIST
```


I-frame playlist (multiple key frames)



```
#EXT-X-I- FRAMES- ONLY
#EXT-X-MAP:URI="init000000.mp4"
#EXT-X-BYTERANGE:2108@372
#EXTINF:2,
segment000000.m4s
#EXT-X-BYTERANGE:1616@39466
#EXTINF:2,
segment000000.m4s
#EXT-X-BYTERANGE:1741@79426
#EXTINF:2,
segment000000.m4s
#EXT-X-BYTERANGE:1712@120252
#EXTINF:2,
segment000000.m4s
#EXT-X-BYTERANGE:1747@160068
#EXTINF:2,
segment000000.m4s
```

```
[styp] size=8+20
[moof] size=8+328
  [mfhd] size=12+4
    sequence number = 1
[traf] size=8+304
  [tfhd] size=12+12, flags=20028
    track ID = 1
    default sample duration = 100
    default sample flags = 1010000
  [tfdt] size=12+4
    base media decode time = 0
  [trun] size=12+252, flags=205
    sample count = 60    <--- Note the sample count
    data offset = 476
    first sample flags = 2000000
[mdat] size=8+38750
[moof] size=8+328
  [mfhd] size=12+4
    sequence number = 2
```

fMP4 boxes (ii)



```
[traf] size=8+304
  [tfhd] size=12+12, flags=20028
    track ID = 1
    default sample duration = 100
    default sample flags = 1010000
  [tfdt] size=12+4
    base media decode time = 6000
  [trun] size=12+252, flags=205
    sample count = 60
    data offset = 476
    first sample flags = 2000000
[mdat] size=8+39616
[moof] size=8+328
  [mfhd] size=12+4
    sequence number = 3
[traf] size=8+304
  [tfhd] size=12+12, flags=20028
    track ID = 1
    default sample duration = 100
```

fMP4 boxes (iii)



```
    default sample flags = 1010000
[tfdt] size=12+4
    base media decode time = 12000
[trun] size=12+252, flags=205
    sample count = 60
    data offset = 476
    first sample flags = 2000000
[mdat] size=8+40482
[moof] size=8+328
    [mfhd] size=12+4
        sequence number = 4
    [traf] size=8+304
        [tfhd] size=12+12, flags=20028
            track ID = 1
            default sample duration = 100
            default sample flags = 1010000
        [tfdt] size=12+4
            base media decode time = 18000
        [trun] size=12+252, flags=205
```

fMP4 boxes (iv)



```
    sample count = 60
    data offset = 476
    first sample flags = 2000000
[mdat] size=8+39472
[moof] size=8+328
    [mfhd] size=12+4
        sequence number = 5
[traf] size=8+304
    [tfhd] size=12+12, flags=20028
        track ID = 1
        default sample duration = 100
        default sample flags = 1010000
    [tfdt] size=12+4
        base media decode time = 24000
[trun] size=12+252, flags=205
    sample count = 60
    data offset = 476
    first sample flags = 2000000
[mdat] size=8+37936
```

fMP4 boxes - Summary



- From initialization segment, `timescale` = 3000
- `tfhd` -> `default sample duration` = 100
- `mfhd` -> sequence numbers = 1, 2, 3, 4, 5
- `tfdt` -> `base media decode time` = 0, 6000, 12000, 18000, 24000
- `trun` -> sample count = 60

Each fragment

- Contains 60 video frames, each with a duration of 100 time units
- Total fragment duration = 6000 time units
- $6000 \text{ time units} \div 3000 \text{ timescale} = 2 \text{ seconds per fragment}$
- $100 \text{ time units} / 3000 \text{ timescale} = 33.33\text{ms} \sim 30 \text{ fps}$

```
Fragment 1: 0      → 6000  = 0.0s → 2.0s
Fragment 2: 6000   → 12000 = 2.0s → 4.0s
Fragment 3: 12000  → 18000 = 4.0s → 6.0s
Fragment 4: 18000  → 24000 = 6.0s → 8.0s
Fragment 5: 24000  → 30000 = 8.0s → 10.0s
```

- Needs a Multi-variant/master playlist.

```
#EXT-X-STREAM-INF:BANDWIDTH=2048000,CODECS="avc1.42e00a"  
playlist.m3u8
```

```
#EXT-X-I-FRAME-STREAM-INF:BANDWIDTH=2048000,CODECS="avc1.42e00a",URI="iframe-playlist.m3u8"
```

As per Section 3.3 of RFC 8216,

An fMP4 Segment in a Playlist containing the EXT-X-I-FRAMES-ONLY tag (Section 4.3.3.6) MAY omit the portion of the Media Data Box following the intra-coded frame (I-frame) sample data.

How to test



- Muxer test for checking output with chunk on key frame.
- For HLS CMAF, for every I-frame segment
 - Presence of `moof` and `mdat`
 - `trun` data offset
 - `moof` size + `moof` header == `trun` data offset
 - `moof` size + `moof` header + size of first `trun` entry (I-frame) == Input size

Questions



- Reach out on
 - Email:
 - contact@centricular.com
 - sanchayan@centricular.com