Source   **Telecom ParisTech**
Status   **For consideration at the 108th MPEG Meeting**
Title   **Relaxing Movie Fragment Sequence Number**
Author   Cyril Concolato, Jean Le Feuvre

1   Introduction

ISO/IEC 14496-12 defines how to store media streams in files, including for long running presentations using the concept of "movie fragment". Movie fragments have been designed for progressive linear storage and playback, however new scenarios are problematic with the current standard. This contribution presents such scenarios and proposes a solution to overcome the problem.

2   Scenarios and Problems

2.1   Reminder

The current ISO/IEC 14496-12 standard specifies in the 'mfhd' box (Movie Fragment Header) :

The movie fragment header contains a sequence number, as a safety check. The sequence number usually starts at 1 and must increase for each movie fragment in the file, in the order in which they occur. This allows readers to verify integrity of the sequence; it is an error to construct a file where the fragments are out of sequence.

NOTE There is no requirement that the sequence numbers be consecutive, only that the value in a given movie fragment be greater than in any preceding movie fragment.

2.2   Movie Fragments and Media Source Extensions

The W3C Media Source Extensions (MSE) specification [1] defines a mechanism to push media segments in the decoding buffer of an HTML 5 video element for playback with an HTML5 web page. MSE enables pushing/concatenating those media segments with different modes:
   - media segments may come from a single source or from multiple sources, as one way to achieve ad-insertion for instance. The same media parser (called SourceBuffer in MSE terminology) is used if the Initialisation Segment is the same between both sources. However, the media segments having been generated from different encoders, there is no guarantee that the sequence_number field respects the growing rule.
   - media segments may be concatenated in "sequence" irrespective of their original timestamp. For instance, one may construct a bitstream taking the segments generated from an initial file in reverse order. Obviously, in this case, the rule is not respected.

2.3 Movie Fragments, DASH and editing

Movie fragments are heavily used in MPEG-DASH. In case a DASH presentation has been generated and a new track needs to be added later on under the same Initialization Segment, adding new fragments require renumbering all existing fragments or provisioning for a maximum number of fragments and leaving empty ranges of sequence numbers.

In a similar way, when using separate tracks for storage of multi-layered videos, the packager has to be very cautious about the number of generated segments (one per layer) to respect the sequence number rule. When converting an RTP layered multicast or a multi-PID layered MPEG-2 TS, this means that the processing of each layer cannot be made independently (read "in different threads") efficiently, since each stream packager has to known what the other packager is doing. While this can be pre-configured when generating time-aligned segments, it becomes quite complex when non time-aligned segments are used (for example, smaller GOP size on the base than on the enhancement). The benefit of the sequence number constraint for validation is not even obvious, since readers usually read *tfdt* information to validate the timing of the segments.

There are also cases where it can be useful to assemble multiplexed media segments from already authored single-track sources and a common initialization segment. The constraint imposes a rewrite of the movie sequence number in case the packager wants to change media interleaving. Here again, the validation of the segment will most likely be done through *tfdt*.
Finally, insertions of movie fragments containing non-frequent data (subtitles, captions, meta-data tracks) may be done after packaging of AV data done by the encoders, and would therefore require rewriting of sequence numbers.

2.4 Proposal

In practice, players (such as GPAC, VLC, QuickTime?…) already ignore the 'sequence_number' field and only throw warnings. However, it may be that some embedded players check it and complain in case the rule is not respected. We propose therefore to define the 'iso8' brand, in which we require that players ignore the 'sequence_number' field.

3 Conclusion

We recommend MPEG to adopt the proposed modification and to include it in the next amendment to ISOBMFF.

[1] https://dvcs.w3.org/hg/html-media/raw-file/tip/media-source/media-source.html
[2] https://dvcs.w3.org/hg/html-media/raw-file/tip/media-source/isobmff-byte-stream-format.html