INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG2016/M39258
October 2016, Chengdu, CN

Source **Telecom ParisTech**
Status **For consideration during MPEG #116**
Title **On Initial TFDT value**
Author Cyril Concolato, Jean Le Feuvre

1 Introduction

The tfdt box provides "an integer equal to the sum of the decode durations of all earlier samples in the media ". It is unclear if a file whose first tfdt indicates a non-zero value is conformant or not. It seems not according to CMAF discussion. This contribution reports the result of playback experiments with such file.

2 Experiment and results

We did the following experiment:

- Segment a 10s mp4 file into 1s segment files.
- Create a new mp4 file by concatenating the initialization segment with last media segment i.e. with media data from time 9 to 10, leading to a file with a first tdft whose baseMediaDecodeTime is 9.
- Play the mp4 in different players

Results are given in the table below:

Player	Behavior
GPAC MP4Client	10 seconds timeline. Plays immediately for 1s the last second.
VLC 2.2.1	
Google Chrome 53	
Firefox Nightly 49	
Microsoft Windows Media Player 12 Edge	Shows the full timeline, plays from time 0 showing the first frame until it is reached and then plays normally.
QuickTime 7.7.5	Black screen for 1s.

Almost all players support this file.

Note, in the case of Media Source Extension-based playback, two behaviors are possible depending on the selected mode "sequence" or "segments". In segments mode, the content is placed on the timeline at the time indicated in the tfdt. If no other data is appended, the playback does not start unless a seek is made to a buffered time, i.e. after tfdt. In "sequence" mode, the tfdt is ignored and the decode time of the first sample in the sequence is adjusted to the user specified

"timestampOffset" value (sync between tracks has to be handled explicitly) and other samples follow according to sample durations.

3 Discussion

First, we believe that the file (as discussed above) should be considered conformant. This is not clear enough in the specification.

Second, we believe that the playback interoperability should be improved.

Two options are possible for playback. This is simple for a single track file:

- wait for the time indicated in the tdft to start playing (MSE-segment mode)
- start playback immediately from the time indicated in the tfdt (MSE-sequence mode)

This can be generalized if the file contains multiple tracks. Two options are:

- wait for the minimum time indicated in the first tdft in all tracks to start playing
- start playback immediately from the same minimum time

However, we do not want to change the playback behavior if previous samples were processed if the moov had samples. In this case, as already specified, the tfdt documents or extends the duration of the previous sample. However, in the case reported here, there is no previous sample.

We suggest recommending a standard playback behavior. If not acceptable, we suggest allowing authors to indicate the desired playback behavior to be overridden by the surrounding environment (such as MSE mode, or use of presentation time offset in DASH, if any)

To indicate the author preference, without having to modify the existing data (moov, ftyp …), we suggest inserting a top-level box after the 'moov' prior to the fragment 'moof', defined as follows:

```
aligned(8) class AutoSeekFragmentsBox extends Box('sfra') {}
```

with the following semantics:
When present, it indicates that the author prefers that the playback of this file behaves as if the player was ordered to seek at the minimum movie (aka presentation) time of all corresponding tfdt times of all tracks.

4 Conclusion

We recommend adopting the proposed modifications (conformance and playback behavior) in an amendment to ISO/IEC 14496-12.