

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

**ISO/IEC JTC1/SC29/WG11 MPEG2012/M26900  
October 2012, Shanghai, China**

**Source**     **Telecom ParisTech**  
**Status**    **For consideration at the 102<sup>nd</sup> MPEG meeting**  
**Title**     **On Fragmentation of long lasting samples**  
**Authors**  **Jean Le Feuvre, Cyril Concolato**

## **1 Introduction**

ISOBMF provides tools for signaling carousel-like repetition of meta data, using sample dependency flags as indicated in section 8.6.4.1 of ISO/IEC 14496-12. This feature instructs a client that it can ignore a sample with such flags set if it had already processed a sync sample or a sample with `sample_depends_on=2`. This typically allows a segmenter to duplicate a long lasting sample (meta-data or text, or even a video frame in a VFR stream) over several fragments or segments, depending on the durations of the sample/segment/fragment.

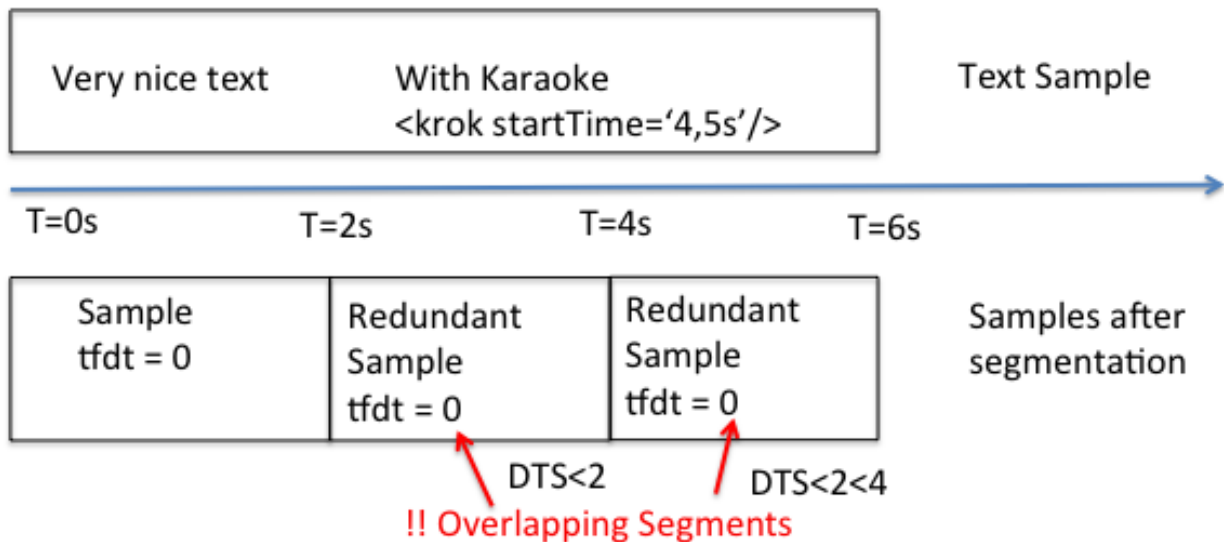
While this is fine for many meta-data or text formats, there are situations in which this replication does not guarantee correct media reconstruction, for example when the sample contains timing information relative to its CTS (karaoke in 3GPP timed text, potential timing info in TTML or WebVTT...).

One approach to solve this problem is to edit the sample, split it into as many samples as desired, adjust the internal timing whenever needed and rewrite the samples. This has however a few drawbacks:

- it is not guaranteed that such an edition is always possible, depending on the underlying meta-data format
- it breaks generic (media-agnostic) fragmentation of ISOBMF files, which is not desirable.

## **2 Proposed Solution**

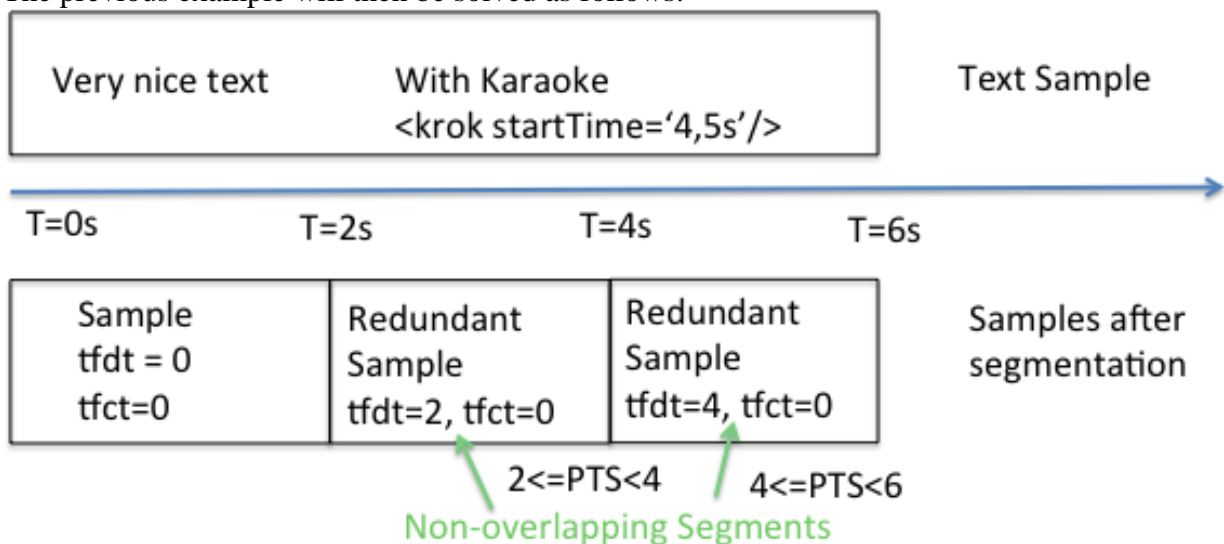
We propose to add signaling of the timestamp of the original sample whenever the sample is cloned/duplicated in movie fragments, so that a player can adjust any composition timing of the sample data if needed. One way to do so is to shift the `TFDT.baseMediaDecodeTime` value. However this would require a reword of the definition of non-overlapping segments in DASH, as a DTS signaled in TFDT of sample N could be less than the CTS+duration of the sample N-1 (N is a clone of N-1), and even less than the DTS of sample N (N and N-1 are clones of sample N-2 or less), as shown in the following figure.



We therefore suggest the addition of a new box in the `traf` header, the Track Fragment Original Composition Time, which gives the original composition timestamp (in media timescale) of the first sample in the `traf`, and which can be ignored if the first sample is not flagged as a ‘repeated’ sample or if the first sample is discarded because of its sample dependency flags.

```
aligned(8) class TrackFragmentOriginalCompositionTimeBox extends FullBox('tfct',
version, 0) {
  if (version==1) {
    unsigned int(64) originalMediaCompositionTime;
  } else { // version==0
    unsigned int(32) originalMediaCompositionTime;
  }
}
```

The previous example will then be solved as follows:



### 3 Conclusion

In this contribution, we have proposed a box, which allows exact reconstruction of the original media timeline in case of sample carouseling, which can be used for any media type supporting the carousel mechanism in ISO/BMF. This box allows proper carriage of 3GPP timed text in

ISOBMF segments, or segmenting of WebVTT files without having to rewrite the WebVTT sample. We suggest adopting this signaling in ISOBMF specification.