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

**ISO/IEC JTC1/SC29/WG11 MPEG2013/M31448**
**October 2013, Geneva (CH)**

**Source** **Telecom ParisTech**
**Status** **For consideration at the 106nd MPEG meeting**
**Title** **Follow-up on Uniform Signaling for Timeline Alignment**
**Author** Jean Le Feuvre

# 0 Foreword

This contribution is the logical continuation of long online, offline or face-2-face discussions on the topic of "hybrid delivery" aver the last year, and the author lost count of all participants in the debates- let them be thanked. The author would like to address special thanks to the experts that patiently reviewed this contribution, especially Mr David Singer and Ali C. Begen.

# 1 Introduction

In Incheon meeting an exploration on Uniform Signaling for Timeline Alignment was started, in order to investigate the required tools at the systems level to allow a media presentation packaged and delivered in one format over one network type to be "augmented" by another media presentation, possibly using different packaging and/or delivery means. More specifically, the exploration welcomes feedback on the topics of:

- timeline alignment of media packaged in different containers and delivered over different networks
- Discovery of the add-on media enhancing an existing presentation
- Tools allowing event signaling related to the add-on or enabling pre-fetch of add-on media

MPEG has a key role to play in this area, and should propose guidelines and technologies in a well-identified manner in order to ensure interoperability in the connected media CE market. Other standard bodies such as HbbTV (ETSI) are currently looking for solutions to recommend in their products for what they usually refer to as "Hybrid Delivery", and plan to release their new technologies before mid 2014.

In this regard, this contribution reviews existing MPEG and IETF technologies that can be used to achieve the goal listed in the exploration activity for all the topics previously stated.

# 2 Technical Reviews

In this document, we only consider the following packaging formats:

- MPEG-2 Transport Stream
- ISO Base Media File Format

- MPEG-DASH
- IETF RTP+RTCP

This is however not exhaustive and can be extended for new standards such as MMT.

## *2.1 Timeline Alignment*

We investigate how timeline is described and can be aligned for each of the above standards. timing information are usually found:

**NPT (Normal Play Time), as defined in section 3.6 RFC2326**: a timestamp indicating the stream absolute position relative to the beginning of the presentation, in a lossless way. Matching NPTs is a straight process of comparing the timestamps.

**NTP (Network Time Protocol):** absolute clock tick that can be associated with a media frame. Matching NTPs of several streams is trickier since there is no guarantee on the NTP clock precision of the different senders.

**TC (Time Code):** another form of NTP with more details on the source, typically taking into account frame rates and drop-frames video formats. It is designed to provide frame-accurate description of a media. An overview of time codes is given in section 5 of RFC5484 In the following, we use $TC_S$ (resp. $TC_{L)}$ to indicate 24 bits (resp. 64-bits) timecodes, as defined in section 6.2 of RFC5484.

**Question:**

- Do we want to also consider Precision Time Protocol as a potential tool to describe timing ?

It should be noticed that TimeCodes and NPTs can be matched together, while NTP cannot be aligned with TCs or NPTs without external mean (such as RTSP giving the correspondence between NPT and RTP timestamp, which in turn is matched to NTP through RTCP)

| Standard | NPT | NTP | $TC_S$ / $TC_L$ |
|---|---|---|---|
| ISOBMFF | Yes (MediaTime) | No | No (Could be added in a meta-data track) |
| MPEG-2 TS<br>TEMI WD (w13661) | Yes | Yes | Yes |
| MPEG-DASH | Yes<br>(Media Presentation Timeline of active period) | No | No (Could be added in a meta-data track) |
| RTP+RTCP | No | Yes | Yes with RFC5484 |
| RTP+RTCP+RTSP | Yes | Yes | Yes with RFC5484 |

We therefore recommend the following definitions:

"In the context of timeline alignment of multiple media, the Unified Time Line (UTL) in each container is defined as follows:

- For ISOBMFF, UTL of a media is the track timeline, e.g. media timeline after applying potential edit lists, which are required to take into accounts time shifts introduced by some coding patterns.
- For MPEG-DASH, UTL of a media is the Media Presentation Time of this media (ie AdaptationSet), computed as

  **period@start + CTS(media) – period@presentationTimeOffset**.

  If a period is explicitly designed as the add-on material though an fragment URI, the Media Presentation Time is computed as

  **CTS(media) – period@presentationTimeOffset**.

  - **NOTE:** we should be careful that period start may not precisely match a media frame since it is not a fractional number.

- For MPEG-2 TS, UTL of a media is carried by a TEMI stream, as defined in 13818-1 Annex T.
- For RTP+RTCP, UTL of a media is described according to RFC5484, in which case SDP configuration shall be present. However, in cases where pure NTP synchronization only is required, UTL of a media shall be derived from NTP timestamps.
- For RTP+RTCP+RTSP, UTL of a media is the NPT as indicated in the RTSP PLAY responses, or may b described according to RFC5484 if signaled in the SDP. If both NPT and RFC5484 are used in the RTSP+RTP session, they shall be identical (same time origin, same speed).

Note: For media services requiring frame-accurate decoding or presentation, it is strongly recommended to use UTL that do not rely on NTP, as different senders may have different NTP precisions."

From this list we can derive the following matching table describing whether it is possible to achieve frame-accurate alignment of media timelines of main presentation and add-on media, based on the container/transport format.

| Standard | ISOBMFF | TS+TEMI | MPEG-DASH | RTP+RTCP | RTP+RTCP+RTSP |
|---|---|---|---|---|---|
| ISOBMFF | YES | YES | YES | NO (needs TCs) | Yes with NPT |
| TS+TEMI | | YES | YES | YES | YES |
| MPEG-DASH | | | YES | NO (needs TCs) | Yes with NPT |
| RTP+RTCP | | | | YES | YES |
| RTP+RTCP+RTSP | | | | | YES |

## *2.2 Add-on Location*

In order to locate the add-on from the main media, some extensions might be needed in the existing standards.

### 2.2.1 Add-on Location in MPEG-2 TS

The TEMI stream (13818-1 Annex T under specification) gives the possibility to associate one or several add-ons to an existing program in a transport stream, by giving their mime types and URI.

### 2.2.2 Add-on Location in ISOBMFF

There is currently no specific tool in ISOBMFF for external media location that could be added to the file. One could use a meta info, but the box is more dedicated to storage of files and XML data in ISOBMFF, whereas the need here is related to hyperlinking. We recommend standardizing a meta-data track to list potential enhancements to the media presentation in the file. One possible syntax for the sample could be a collection of boxes, including an Add-on Declaration Box:

Add-on Declaration Box

Box Type: 'addo'
Container: Add-on Meta Data sample
Mandatory: No
Quantity: Zero or One

The add-on declaration box is used to declare optionnal presentations that could be used as enhancements for the presentation contained in this file.

**Syntax**

```
aligned(8) class AddOnDeclarationBox extends FullBox('addo', version = 0, 0) {

     unsigned int(32) entry_count;
     int i;
     for (i=0; i < entry_count; i++) {
          string mimeType;
          string URI;
     }
}
```

**Semantics**

mimeType: optional string containing the mimeType of the add-on, optionally with codec parameters.
URI: string containing the identification or location of the add-on.

There are also cases where track identification is required at the file format level, for example if the file contains an add-on coded in a scalable codec. The primary media has a reference to the add-on, but the add-on file needs to point back to the scalable base. Obviously, pointing back with the same URL is not very convenient, since moving the base would require rewriting the add-on. Furthermore, the base media presentation may be moved to any place!
We suggest using an external track alias box, which could be at the moov level:

```
aligned(8) class ExternalTrackDeclarationBox extends FullBox('trax', version = 0,
0) {
 unsigned int(32) originalTrackID;
 if (originalTrackID==0) {
     string mimeType;
     unsigned int(8) lang[3] ;
 }
 TrackReferenceBox tref; //optionnal
}
```

The `originalTrackID` indicates the ID of the track referred to in the main media presentation file. If the `originalTrackID` is unknown or the external track definition does not refer to an ISOBMFF file, `mimeType` shall be set to identify the media stream in the main media presentation; if the main media is localized, it is recommended to also indicate its language in `lang`.

The `tref` box indicates the track dependencies between the referred track and the tracks in this file. This allows describing coding dependencies in the add-on file, even when the base layer track is not in the file.

**Question:**

- do we want to signal something like DASH "role" to hint the client whether the add-on can be used or not ? Some experts believe this should be done.
- do we want the possibility to signal alternate add-ons (e.g. same content but different formats) and complementary add-ons (different contents)? Some experts believe this should be done.

### 2.2.3 Add-on Location in MPEG-DASH

MPEG-DASH defines all the tools needed to gather media from several sources in the MPD. Therefore, if MPEG-DASH is used as the entry point presentation, it may links to add-on content through AdaptationSets or Representations. In the context of this standard, all add-ons to the DASH presentation are DASH media (media segments) and DASH descriptions (MPD).

There can be cases where DASH is used as an add-on to an existing presentation and where one AdaptationSet is a scalable enhancement of an existing stream in that presentation. In this case, it is suggested to use the predefined URLs as dependencyId of the lowest layer in the AdaptationSet:

- dash://video
- dash://audio

**Question:**

- in cases where several audio or video streams are present in the base media presentation, shouldn't we add role and languages ?

### 2.2.4 Add-on Location in SDP

It could be useful for pure multicast distribution to have potential add-on indicated in the SDP. This should be done using "u=" line in SDP, which defines a URI as follows:

```
"The URI should be a pointer to additional information about the    session.
This field is OPTIONAL, but if it is present it MUST be    specified before
the first media field.  No more than one URI field    is allowed per session
description.
"
```

One limitation is that only one URI can be associated with the source.

We can also recommend using something like
x-addon=mimeType URL

for each potential add-on for the session.

**Question:**

- do we want dynamic add-ons that may vary over time ? If so, should this be signaled through SAP ? Through new extension RTCP reports?

## *2.3 Various signaling*

This section investigates what tools are needed in the main content to carry events related to the add-on. More specifically, we investigate:

- Announcement of upcoming add-on for prefetching
- Splicing
- Forcing reload of add-on descriptions, typically for MPEG-DASH MPD reload.

### 2.3.1 MPEG-2 TS

The TEMI stream (13818-1 Annex T under specification) gives the possibility to signal one or several upcoming add-ons, signal splicing in order to optimize resource fetching and force reloading of the target description (DASH MPD).

### 2.3.2 ISOBMFF

The ISOBMFF container being static, all associated resources are known in advances. There should therefore be no need of extra signaling for dynamic events, these should be part of the add-on technology used (e.g., emsg box in DASH).

### 2.3.3 MPEG-DASH

When MPEG-DASH is the base presentation, all enhancements to the presentation are also in MPG-DASH formats, and all signaling such as splicing, prefetch information and reloading are already available in the DASH standard.

### 2.3.4 RTP/RTSP

The current design only allows for static URLs for add-on, there is no need for prefetching and splicing indications. There could be a need to car a “force reload” event for the MPD, maybe this should be done via a header extension?

# 3 Conclusion

We suggest starting a WD for USTA gathering the different contributions on the topic, welcome more contributions and maybe start liaising with other standard bodies on this activity.