



OPEN IPTV FORUM e.V.

Open IPTV Forum e.V. Secretariat, 650 Route des Lucioles, 06921 Sophia Antipolis Cedex, France
Phone : + 33 492 94 43 83 Fax : +33 492 38 52 90 Email: contact@oipf.tv

18th June, 2010

Liaison letter

To: 3GPP SA4

3GPP Liaison inputs, 3GPP Liaison@etsi.org

Kari Järvinen, Chair 3GPP SA4, kari.ju.jarvinen@nokia.com

cc. Paolino Usai, Secretary 3GPP SA4, paolo.usai@etsi.org

cc. Stephen Hayes, Chair 3GPP TSG SA, stephen.hayes@ericsson.com

cc. Maurice Pope, Secretary 3GPP TSG SA, maurice.pope@etsi.org

Open IPTV Forum contacts:

Paul Szucs, Solution Working Group chair, paul.szucs@eu.sony.com

Nilo Mitra, Architecture Working Group chair, nilo.mitra@ericsson.com

Catherine Hammond, Liaison Officer, liaisons@oipf.tv

Dear 3GPP SA4 Colleagues,

OIPF Solution WG would like to thank 3GPP SA4 for their liaison letter out of their Vancouver meeting at the end of April.

In the meantime we have provisionally finalised our work on the OIPF HAS specification, also taking into account 3GPP SA4's advice, and would like to provide you with some details of the OIPF Release 2 Solution specification for HTTP Adaptive Streaming (HAS), for which the technical review in OIPF has just been launched.

We would like to highlight some aspects of the OIPF HAS specification and would appreciate your review and comment at the SA4 meeting on 21-24 June:

1. Usage with MP4 file format systems layer

- For every Representation, a [TS26234] Initialisation Segment is made available.
 - For all non-partial Representations, a reference to the Initialisation Segment is present in a <InitialisationSegmentURL> element in the <Representation>-element.
 - For all partial Representations a reference to a single common Initialisation Segments are present in an <InitialisationSegmentURL> in the <SegmentInfoDefault> element.
 - An Initialisation Segment is delivered with MIME type "video/mp4".
 - Initialisation Segments are be formatted as specified in [TS26234], section 12.4.2.2. For every media stream of the (set of partial) Representation(s), the *moov*-box in the Initialisation Segments contains a *trak*-box describing the samples of the media streams in compliance with [ISOFF].
- Every Representation consists of Media Segments that are formatted as specified in [TS26234], section

Disclaimer

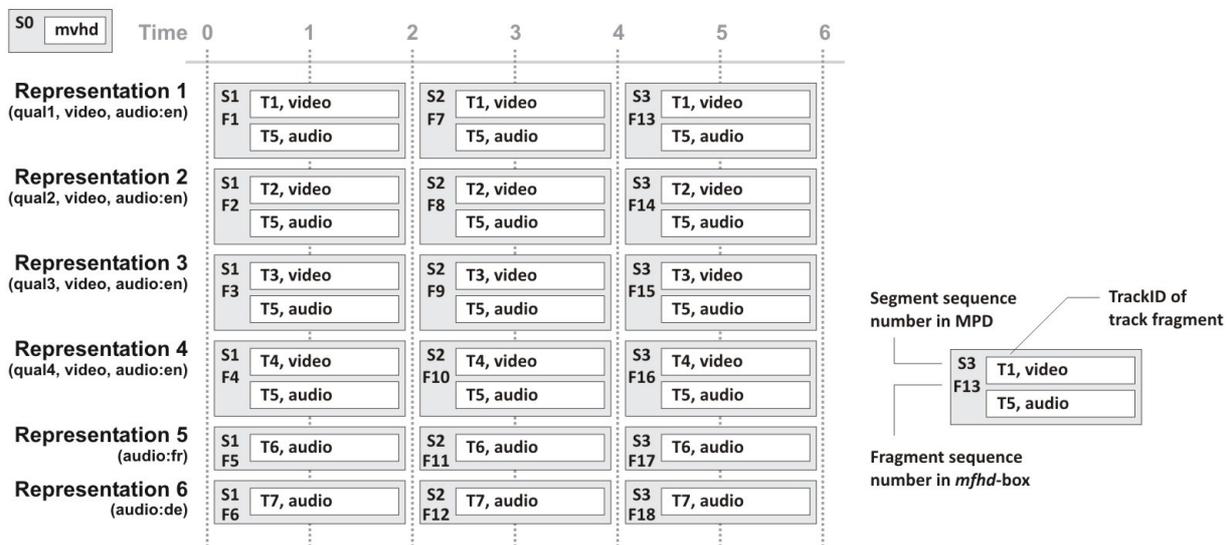
The Open IPTV Forum e.V. accepts no liability whatsoever for any use of this document.

12.4.2.3.

- A Media Segment is delivered with MIME type “video/vnd.3gpp.segment” as specified in [TS26244]
- To allow a terminal to seek to any Segment with a certain index and start playback with perfect audio/video synchronization, every *traf*-box of a track that contains audio should contain a [TS26244] *tfad*-box. The contents of the box SHALL be such that if the terminal starts the playback of the audio samples of a Segment as specified in the box, then the audio and video of the Segment are played in perfect sync.
- The Initialisation Segment and the Media Segments are formatted such that a file that consists of the Initialisation Segment and an arbitrary selection of Media Segments of the (set of partial) Representation(s), stored in order of the *sequence_number* in their *mfhd*-box, is an [ISOFF] compliant file. (Note that this statement assumes that [ISOFF] allows for ‘gaps’ in the *sequence_numbers* of consecutive *moof*-boxes; i.e. the difference in the *sequence_number* of consecutive *moof*-boxes may be larger than one).

Via Partial Representations, services are allowed to offer the various elementary streams of a presentation as separate downloads/streams. In this case it is required that there is one single Initialisation Segment describing the samples in all Media Segments of all Partial Representations and that the concatenation of the Initialisation Segment and the Media Segments is an [ISOFF] compliant file. This section illustrates how such requirement can be met by working out the following example in combination with the MP4 system layer.

In this example a service offers a video in four bitrates and audio in three languages (English, French and German), where two languages (French and German) are offered for separate retrieval as separate partial Representations. The figure depicts a potential allocation of movie and track fragments to segments and representations for the first few segments of this example.



Partial Representation Example

Each Segment has a sequence number in the MPD and contains a single movie fragment with a certain sequence number in the *mfhd*-box. Segments of Representations 1 through 4 contain both audio (English) and video track fragments. In this example the service has chosen to put each video track on its own TrackID, such that each track can have its own set of configuration parameters in the *mvhd*-box.

If a terminal selects to retrieve German audio and starts playing the video at a quality 3 while increasing the quality with each consecutive segment, then it will retrieve the sequence of Segments as depicted here:.



Partial Representation Retrieval

When stored as depicted (Initialisation Segment first, Media Segments in sequence of the Fragment sequence number) this is a valid [ISOFF] file. Note that to play this file, the terminal will need to switch to those video tracks that were selected for download.

Note that the MPD could also include four additional non-partial Representations, that reference the same Media Segments as Representations 1 through 4 in this example, and the same (or a different) Initialisation Segment. In this way the same service (and the same HTTP caches!) can be used for terminals that do not support partial representations.

2. Component element definition

The Media Presentation Description (MPD) SHALL be as specified in [TS26234] section 12.2, with the following extensions and additional requirements:

- The MPD is an XML file that validates against the following schema. Note that this XML schema imports the schema specified in [TS26234]. This means that an MPD that does not use any of the OIPF specific extensions will validate against both the schema defined in [TS26234] as well as the following MPD schema:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns="urn:oipf:adaptive_streaming"
  targetNamespace="urn:oipf:adaptive_streaming"
  xmlns:pss="urn:3GPP:ns:PSS:AdaptiveHTTPStreamingMPD:2009">

  <xs:import namespace="urn:3GPP:ns:PSS:AdaptiveHTTPStreamingMPD:2009" schemaLocation="3GPP-MPD-006.xsd" />

  <xs:element name="Component" type="ComponentType"/>
  <xs:element name="Components" type="ComponentsType"/>
  <xs:complexType name="ComponentsType">
    <xs:sequence>
      <xs:element minOccurs="1" maxOccurs="unbounded" ref="Component"/>
    </xs:sequence>
  </xs:complexType>
  <xs:complexType name="ComponentType">
    <xs:attribute name="id" type="xs:string" use="optional"/>
    <xs:attribute name="type" type="xs:string" use="required"/>
    <xs:attribute name="lang" type="xs:string" use="optional"/>
    <xs:attribute name="angle" type="xs:string" use="optional"/>
    <xs:attribute name="channels" type="xs:unsignedInt" use="optional"/>
    <xs:attribute name="impaired" type="xs:boolean" use="optional"/>
    <xs:attribute name="adMix" type="xs:boolean" use="optional"/>
  </xs:complexType>
</xs:schema>
```

- A <pss:Representation> element may carry the @group-attribute set to a non-zero value. In this case the attribute indicates that the <Representation>-element is not necessarily a complete Representation, but consists of one or more individual components (video, audio, subtitles, etc.) which may be

downloaded and provided to the terminal in addition to content being downloaded from other <Representation> elements. In this case the <Representation> element contains one or more <Component> elements, as specified below, for each elementary stream contained in the <Representation>. Note that it is the responsibility of the application on the terminal to select the desired components and to initialize the terminal accordingly. Item 3 below contains an informative description of how this can be done. The value of the @group-attribute will be the same for Representations that contain at least one same Component. Two Representations of the same type but with completely different Components (e.g. audio at two different languages) will have different values.

Element/Attribute	Description	Optionality
Component	This element contains a description of a component.	
@id	the value of this identifier is a container specific identifier for the elementary stream (e.g. the PID for MPEG2-TS or the trackID for MP4).	O
@type	specify the component type. It shall be one of Video, Audio or Subtitles	M
@lang	specify language code according to RFC 5646 [r3] for audio and subtitles stream.	O
@angle	specify camera angle for video stream.	O
@channels	specify audio channels for audio stream.	O
@impaired	specify if the stream is audio description for visually impaired or subtitles for hearing impaired	O
@adMix	in the case where the type attribute is Audio and the impaired attribute is true, specify if this audio stream must be mixed with the main audio stream	O

Component Element and Attributes

3. Component stream management

The <Representation> element defined in [TS26234] corresponds to a particular version of the full content item with all its elements (video, audio, subtitles, etc). The different representations listed in the MPD correspond also to full, alternate versions that differ in a particular aspect (bitrate, language, etc). This means that the terminal will at every moment download and present segments of only one representation. While this provides a quite simple and straightforward model it has an important lack of flexibility in the following sense: if there are many alternatives for a particular component (e.g. audio in different languages) and there are also a number of different bitrate alternatives, all combinations shall be available at the server.

For instance, if a service provides 3 audio languages and the video in 4 bitrate levels, then it will need to provide 12 different representations; however, there will be groups of 3 representations which share exactly the same bulky video (they only differ in audio). This causes an important waste of storage space in the server. Even if the server can be optimized with respect to this (e.g. to build the segments in real time from the elementary streams stored separately in its disks), this cannot be done in the HTTP caches, which in general are not aware of the protocol and the similarities between the content pointed to by different URLs.

In order to solve this problem, OIPF includes the concept of Partial Representations in the MPD. Presence of the @partial attribute in the <Representation> element indicates that the representation does not include all elements available for presentation at the terminal, but only a subset of them (e.g. "audio in Spanish"). An OIPF terminal will be able to identify the representations that it requires and will download their segments independently, composing the final combination at the terminal side. In case of the example service above, the server may serve 4 representations with 4 different bitrate versions of a movie with English audio, and separately it can serve 2 additional representations, each with just the audio in a different language. This way, all the 12 combinations mentioned above will be possible (all bitrates at all languages) but with roughly a third of the required storage in the server and the HTTP caches.

As the multiplicity of representation happens in two dimensions (bitrate versions and alternate content), a new attribute 'SwitchGroup' has been added to the <Representation> element. The terminal uses this attribute to identify representations which contain effectively the same alternative version of the content, only at different bitrates. Representations that have different values in this element have different content.

This component-aware scenario relates to the process for selecting and displaying the desired set of components. This process also takes place for content that is delivered through other mechanisms, and is not specific for content that is delivered via the HTTP adaptive protocol described in this document. In the context of OIPF (for example using the DAE “Extensions to video/broadcast for playback of selected components”), this process operates based on the information from the MPEG2-TS or MP4 metadata. This information is also available in the context of HTTP Adaptive Streaming in the Initialisation Segment. This leads to the following informative process for component selection in case of content that is delivered using this specification:

1. Retrieve the MPD. If the MPD includes both, decide if you want to play the partial or non-partial Representations. It is RECOMMENDED to use the partial representations.
2. In case of a non-partial Representation:
 - a) Based on meta-data in the MPD (typically the @bandwidth-attribute), select an initial Representation.
 - b) If present, retrieve the Initialisation Segment of the Representation.
 - c) Retrieve Media Segments of the chosen Representation.
 - d) Find the elementary streams in the downloaded Initialisation Segment / Media Segments. Typically select one video and one audio stream. If there are options, select from those.
 - e) Setup the “player” to play the selected elementary streams. Play them.
 - f) While playing, allow the user to select from other/additional elementary streams in the Initialisation Segment / Media Segments.
 - g) If the terminal wants to switch to a different bitrate, select an alternate non-partial Representation and continue from step 2b.
3. In case of a partial Representation:
 - a) Based on the metadata in the MPD (typically the @bandwidth-attribute and the <Component> element) select the initial Representations.
 - b) If present, retrieve the Initialisation Segment of the Period.
 - c) Retrieve Media Segments of the chosen Representations.
 - d) Based on the @id’s of the <Components> elements, or using information from the Initialisation Segment, setup the “player” to play the selected elementary streams. Play them.
 - e) While playing, allow the user to select from other/additional Components/elementary streams. If other/additional streams are selected, continue from step 3c.
 - f) If the terminal wants to switch to a different bitrate select an alternate partial Representation in the same switchGroup and continue from step 3c.

Note that the Initialisation Segment will always contain the full description of all component alternatives, so it will be guaranteed that there are no identifiers conflicts between them (e.g. two languages with the same MPEG-TS PID or MP4 trackID). The parsing of this Initialisation Segment and the corresponding settings of the chipset to select the appropriate components will be responsibility of the application (the media player).

4. Content protection

For the signalling of content protection method in the MPD (HAS, section 3.3), we have constrained the @schemeIDUri attribute to use DVB CA System Identifiers, since we apply this resource throughout the OIPF Solution, and this resource is managed by DVB Services srl.



OPEN IPTV FORUM e.V.

As a further issue regarding the standards landscape, we would like to inform 3GPP SA4 that OIPF intends to respond to the MPEG Call for Proposals on “HTTP Streaming”, whereby we will highlight the collaboration of our two organisations in producing a OIPF HAS specification harmonised with 3GPP Release 9.

We look forward to your feedback at your earliest convenience, and to a continued fruitful collaboration.

Actions

To SA4: OIPF requests SA4 to review and comment on the approaches adopted in the OIPF HAS specification detailed above.

Yours sincerely,

Paul Szucs
Nilo Mitra
Catherine Hammond