

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

**ISO/IEC JTC1/SC29/WG11 MPEG2014/M33210  
March 2014, Valencia, Spain**

**Source**    **Telecom ParisTech, Canon Research Centre France**  
**Status**    **For consideration at the 108<sup>th</sup> MPEG Meeting**  
**Title**      **Usage of DASH SRD for HEVC Tiling**  
**Author**    Cyril Concolato, Jean Le Feuvre, Franck Denoual, Frédéric Mazé, Eric Nassor

## **1 Introduction**

During its 107<sup>th</sup> meeting, MPEG issued two documents: the Working Draft for the Amendment 2 of MPEG-DASH, which defines how to provide spatial information relating the representations given in an MPD, for instance when a video is tiled and the tiles are offered in the MPD; and the Working Draft for the Amendment 1 of the carriage of NAL unit structured video in the ISO Base Media File Format, which defines how to store HEVC Tile tracks in the ISOBMFF. This contribution reviews the scenarios and requirements for the streaming with DASH of tiled videos (including HEVC). It then presents how to achieve such scenarios, given the requirements, based on these two documents, and finally proposes some clarifications to the DASH specification.

## **2 Scenarios and requirements**

### **2.1 Scenarios**

In this document we will consider 3 scenarios, taken from the Spatial Relationship Descriptions (SRD) CE document, as follows:

- A. In the first scenario, a 4k video is made available for DASH streaming. The author, knowing that some users will not have the necessary bandwidth to stream the video at this resolution, also offers alternate resolutions for this video (HD, SD ...), as usual in DASH, but also offers alternate videos, at the initial resolution, representing only spatial parts of the video (2x2 tiles), coded as independent videos. The 5 videos may be coded using AVC, HEVC or else.
- B. In a second scenario, the same presentation as in the first scenario is delivered but the full video is encoded using HEVC and the tiles are encoded as HEVC tiles of the full video, with decoding dependencies.
- C. In a third scenario, alternate representations are additionally provided for each tile and for the full video, using HEVC scalability.

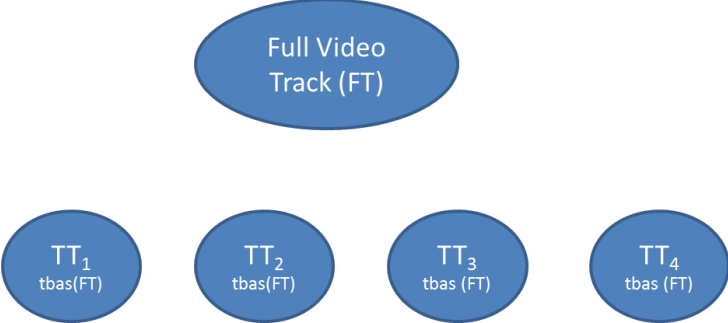
### **2.2 Requirements**

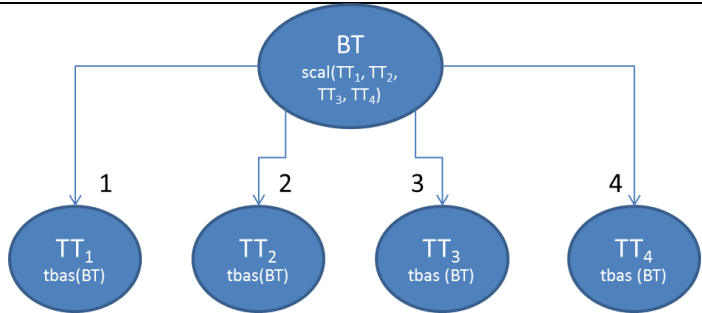
To realize the above scenarios, we believe the following requirements are important:

1. We have to signal to a DASH client that representations are spatially related to each other. For this purpose, we will use DASH SRD.

2. We want to preserve the DASH basic feature, i.e. the possibility to seamlessly switch between different qualities of the full frame video or between the different qualities of the tile videos, when they are available.
3. We want to be able to download a tile (independent video, HEVC tile, or HEVC scalable layer) without downloading the full frame video or other non-dependent tiles.
4. We want to allow a graceful degradation of the streaming for non-SRD-aware clients.

### **3 Streaming video tiles using MPEG-DASH**

Use Case	DASH MPD	Remarks
<p>Scenario A: Independent videos representing tiles of another video</p> 	<pre> &lt;AdaptationSet&gt;   &lt;Representation id="FT" ...&gt;     &lt;SupplementalProperty       schemeIdUri="urn:mpeg:dash:srd:2014"       value="1 0 0 W H"/&gt;     &lt;Role       schemeIdUri="urn:mpeg:dash:role:2011"       value="main"/&gt;   &lt;/Representation&gt; &lt;/AdaptationSet&gt; &lt;AdaptationSet&gt;   &lt;Representation id="TT1" ...&gt;     &lt;SupplementalProperty       schemeIdUri="urn:mpeg:dash:srd:2014"       value="1 0 0 W/2 H/2"/&gt;     &lt;Role       schemeIdUri="urn:mpeg:dash:role:2011"       value="alternate"/&gt;   &lt;/Representation&gt; &lt;/AdaptationSet&gt; ... &lt;AdaptationSet&gt;   &lt;Representation id="TT4" ...&gt;     &lt;SupplementalProperty       schemeIdUri="urn:mpeg:dash:srd:2014"       value="1 W/2 H/2 W/2 H/2"/&gt;     &lt;Role       schemeIdUri="urn:mpeg:dash:role:2011"       value="alternate"/&gt;   &lt;/Representation&gt; &lt;/AdaptationSet&gt; </pre>	<p>In this example, each video (full-frame and tiles) is represented by an Adaptation Set, possibly with multiple representations if the video is offered in different qualities, resolutions, or bitrates.</p> <p>Each representation is associated with a SupplementalProperty descriptor containing SRD, because they can be understood or played even if the SRD are ignored. These descriptors could be replaced by EssentialDescriptors for the tile videos if the author wishes to restrict the selection for non SRD aware clients.</p> <p>The Role descriptors indicate which Representation is the Full Video ('main' role) and which are the Tiled Representations (using 'alternate', or 'supplementary') to indicate the fallback behavior for non SRD-aware clients.</p> <p>Note that SRD aware clients will be able to understand that the tile videos are indeed tiles and that they may present them together.</p> <p>Note also that the @dependencyId attribute is not used because the video are independently decodable.</p>
<p>Scenario B: Tiling using HEVC Tile tracks (without scalability)</p>		<p>In this scenario, the full frame video and the tiles are encoded using HEVC and HEVC tiles. The videos are stored in the ISOBMFF using an HEVC track containing extractors to several HEVC tile tracks. Each (base or tile) track is exposed as a representation in DASH. This enables the streaming of a tile</p>



```

<AdaptationSet>
  <Representation id="BT"
    dependencyId="TT1 TT2 TT3 TT4" ...>
    <SupplementalProperty
      schemeIdUri="urn:mpeg:dash:srd:2014"
      value="1 0 0 W H"/>
    <Role
      schemeIdUri="urn:mpeg:dash:role:2011"
      value="main"/>
    </Representation>
</AdaptationSet>
<AdaptationSet>
  <Representation id="TT1" ...>
    <SupplementalProperty
      schemeIdUri="urn:mpeg:dash:srd:2014"
      value="1 0 0 W/2 H/2"/>
    <Role
      schemeIdUri="urn:mpeg:dash:role:2011"
      value="alternate"/>
    </Representation>
</AdaptationSet>
...
<AdaptationSet>
  <Representation id="TT4" ...>
    <SupplementalProperty
      schemeIdUri="urn:mpeg:dash:srd:2014"
      value="1 W/2 H/2 W/2 H/2"/>
    <Role
      schemeIdUri="urn:mpeg:dash:role:2011"
      value="alternate"/>
    </Representation>
</AdaptationSet>
<AdaptationSet>
  <Representation id="BT" ...>
    <!-- no SRD if no VCL -->
    <Role
      schemeIdUri="urn:mpeg:dash:role:2011"
  
```

without downloading all other tiles.

There are several options to write an MPD for this scenario, depending on how the videos are encoded, and on what the author wants to express or highlight.

### Option 1

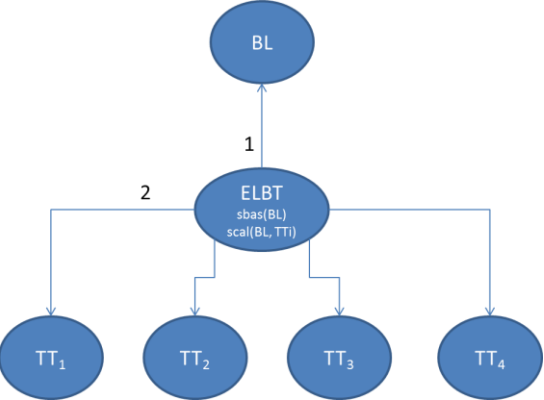
In this option, the author places @dependencyId on the representation corresponding to the base track, which has an SRD indicating that this is the full frame video. The @dependencyId enables the correct playback of the full video for both SRD-aware and non-SRD-aware clients.

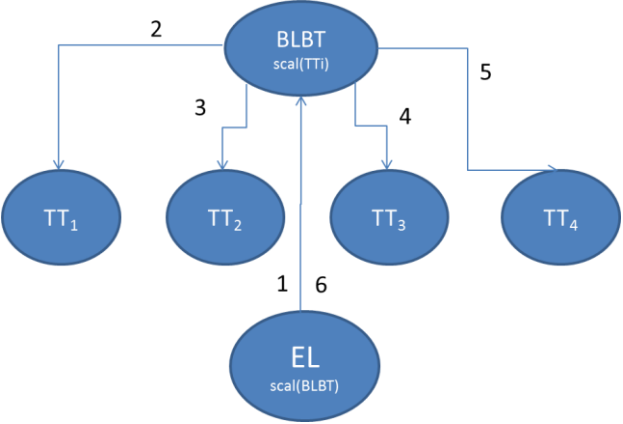
This option may be used only if tile tracks are indeed independently decodable (i.e. if the tile tracks and base tracks have a common initialization segment, if the parameter sets are not in-band in the base track, and if SEI messages are duplicated in the base track and in the tile tracks; or if Parameter Sets and SEI messages are duplicated and in-band in all tracks). For all other cases, such MPD would be non-conformant, and option 2 below may be used.

### Option 2

In this option, the author places @dependencyId on the representations

	<pre>         value="alternate"/&gt;       &lt;/Representation&gt;     &lt;/AdaptationSet&gt;   &lt;AdaptationSet&gt;     &lt;Representation id="TT1"       dependencyId="BT" ...&gt;       &lt;SupplementalProperty         schemeIdUri="urn:mpeg:dash:srd:2014"         value="0 0 W/2 H/2"/&gt;       &lt;/Representation&gt;     &lt;/AdaptationSet&gt;   ...   &lt;AdaptationSet&gt;     &lt;Representation id="TT4"       dependencyId="BT" ...&gt;       &lt;SupplementalProperty         schemeIdUri="urn:mpeg:dash:srd:2014"         value="W/2 H/2 W/2 H/2"/&gt;       &lt;/Representation&gt;     &lt;/AdaptationSet&gt;    &lt;AdaptationSet&gt;     &lt;Representation id="All"       dependencyId="TT1 TT2 TT3 TT4" ...&gt;       &lt;SupplementalProperty         schemeIdUri="urn:mpeg:dash:srd:2014"         value="0 0 W H"/&gt;       &lt;Role         schemeIdUri="urn:mpeg:dash:role:2011"         value="main"/&gt;       &lt;/Representation&gt;     &lt;/AdaptationSet&gt; </pre>	<p>corresponding to the tile tracks, if the tile tracks cannot be decoded without the base track.</p> <p>It should be noted that in this case, the @dependencyId are not used as usual, because they are not placed on the representation corresponding to the track containing the extractors.</p> <p>Note also that if the base track does not contain any VCL NAL units, only extractors, the base track on its own is not playable and must be ‘disabled’ (and the ‘main’ role cannot be placed anymore).</p> <p>It is currently not possible to disable a representation in an MPD, to indicate that it is only there for dependency purposes. To solve this problem, we suggest using the following options:</p> <ul style="list-style-type: none"> <li>- setting the @width and @height to 0 on the representation. Normally, no DASH client should select such representation (unless it considers that it’s an authoring mistake)</li> <li>- using an EssentialDescriptor, with an SRD indicating w=h=0.</li> <li>- using an EssentialDescriptor with another schemeIdUri (e.g. urn:dash:disabled)</li> <li>- defining a new role value.</li> </ul> <p>Finally, it should be noted that if the main role needs to be placed, for fall back of non-SRD aware clients, an additional representation (‘All’ in the example) may be added, with @dependencyId set to the Tile</p>
--	---	--

		<p>representations. This representation is a ‘virtual’ representation, needed only to signal the dependencies and should in fact contain no segment. However, the DASH standard does not have that option. We suggest therefore that authors create a single ‘dummy’/degenerate segment containing only ‘styp’ and no data and to use that as media segment.</p>
<p>Scenario C: Scalability with tiling on the enhancement layer only</p>  <pre> graph TD     BL((BL)) -- 1 --&gt; ELBT((ELBT sbas(BL) scal(BL, TTi)))     ELBT -- 2 --&gt; TT1((TT1))     ELBT -- 2 --&gt; TT2((TT2))     ELBT -- 2 --&gt; TT3((TT3))     ELBT -- 2 --&gt; TT4((TT4))   </pre>	<pre> &lt;AdaptationSet&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="main"/&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 0 0 W H"/&gt;   &lt;Representation id="BL" ...&gt; &lt;/Representation&gt;   &lt;Representation id="EnhancedAll"     dependencyId="TT1 TT2 TT3 TT4" ...&gt; &lt;/Representation&gt; &lt;/AdaptationSet&gt; &lt;AdaptationSet&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="alternate"/&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 0 0 0 0"/&gt;   &lt;Representation id="ELBT"     dependencyId="BL" ...&gt; &lt;/Representation&gt; &lt;/AdaptationSet&gt; &lt;AdaptationSet&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 0 0 W/2 H/2"/&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="alternate"/&gt;   &lt;Representation id="TT1" ...&gt;     dependencyId="ELBT" ...&gt; &lt;/Representation&gt;   </pre>	<p>For the following scenarios, scalability is introduced. Two scalable layers are used.</p> <p>In this case, the base layer is not tiled while the enhancement layer is tiled.</p> <p>The MPD for this case can be written with Option 1 or Option 2 as described above, depending on whether tile tracks can be played independently of the base track. The MPD proposed here is written according to Option 2 defined above. It groups the base and enhanced ‘virtual’ representations in one Adaptation Set with one SRD. Representations indicating that tiles can be played are placed in other Adaptation Sets. And the ‘disabled’ representation is placed in separate Adaptation Set.</p> <p>Note that, in this example, each tile track in the enhancement layer has a dependency on the base track in the enhancement layer. When selecting the “EnhancedAll” Representation, there is a problem with the concatenation of segments.</p> <p>In this example, TT1 (TT2, TT3 and TT4) all depend on ELBT, which both</p>

	<pre> &lt;AdaptationSet&gt; ... &lt;AdaptationSet&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 W/2 H/2 W/2 H/2"/&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="alernate"/&gt;   &lt;Representation id="TT4" ...&gt;     <b>dependencyId="ELBT" ...</b>   &lt;/Representation&gt; &lt;/AdaptationSet&gt; </pre>	<p>depend on BL. If the concatenation rule is followed strictly, the DASH client will pass the following segments to the media engine:  BL+ELBT+TT1+BL+ELBT+TT2 ...  We see that BL and ELBT are duplicated in this concatenation. We propose to clarify the concatenation rule with the following text:</p> <p>“If following recursively the dependencies indicated in the @dependencyId of Representations, the DASH client encounters twice the same Representation, the corresponding media segments shall only be concatenated once, when first encountered.”</p>
<p>Scenario C: Scalability with tiling on the base layer only</p> 	<pre> &lt;AdaptationSet&gt;   &lt;Representation id="BLBT" ...&gt;     &lt;!-- no SRD if no VCL --&gt;     &lt;Role       schemeIdUri=" urn:mpeg:dash:role:2011"       value="alternate"/&gt;     &lt;/Representation&gt;   &lt;/AdaptationSet&gt;   &lt;AdaptationSet&gt;     &lt;Role       schemeIdUri=" urn:mpeg:dash:role:2011"       value="main"/&gt;     &lt;SupplementalProperty       schemeIdUri=" urn:mpeg:dash:srd:2014"       value="1 0 0 W H"/&gt;     &lt;Representation id="All"       <b>dependencyId="TT1 TT2 TT3 TT4" ...</b>     &lt;/Representation&gt;     &lt;Representation id="EL"       <b>dependencyId="BLBT" ...</b>     &lt;/Representation&gt;   &lt;/AdaptationSet&gt;   &lt;AdaptationSet&gt;     &lt;SupplementalProperty       schemeIdUri=" urn:mpeg:dash:srd:2014" </pre>	<p>In this case, only the enhancement layer is tiled. The storage in the ISOBMFF uses recursive extractors as defined in the new ISOBMFF amendment. The enhancement layer track uses extractors on the base track of the base layer, which in turns extracts from the tile tracks.</p> <p>The MPD for this case can be written with Option 1 or Option 2 as described above, depending on whether tile tracks can be played independently of the base track. The MPD presented in the left column uses Option 2. It uses the ‘virtual’ and ‘disabled’ representations.</p>

	<pre> value="1 0 0 W/2 H/2"/&gt; &lt;Role   schemeIdUri=" urn:mpeg:dash:role:2011"   value="alernate"/&gt; &lt;Representation id="TT1"   <b>dependencyId="BLBT"...</b>&gt; &lt;/Representation&gt; &lt;AdaptationSet&gt; ... &lt;AdaptationSet&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 W/2 H/2 W/2 H/2"/&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="alernate"/&gt;   &lt;Representation id="TT4"     <b>dependencyId="BLBT"...</b>&gt;   &lt;/Representation&gt; &lt;/AdaptationSet&gt; </pre>	
<p>Scenario C: Scalability in the base track of the base layer and in the enhancement layer</p> 	<pre> &lt;AdaptationSet&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="supplementary"/&gt;   &lt;EssentialProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 0 0 0 0"/&gt;   &lt;Representation id="BLBT" w=h=0...&gt;   &lt;/Representation&gt;   &lt;Representation id="ELBT" w=h=0...&gt;   &lt;/Representation&gt; &lt;/AdaptationSet&gt; &lt;AdaptationSet&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="main"/&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 0 0 W H"/&gt;   &lt;Representation id="BaseAll"     <b>dependencyId="BLTT1 ... BLTT4" ...</b>&gt;   &lt;/Representation&gt;   &lt;Representation id="EnhancedAll"     <b>dependencyId="ELTT1 ... ELTT4" ...</b>&gt;   &lt;/Representation&gt; </pre>	<p>In this scenario, we have two scalable layers, each layer being tiled. We apply Option 2.</p> <p>We have one Adaptation Set containing the Base and Enhancement ‘disabled’ representations. We have then an Adaptation Set with two ‘virtual’ representations for the full video in both scalable qualities and then one Adaptation Set per tile, each containing two representations (for the two scalable layers).</p>



	<pre> &lt;/AdaptationSet&gt; &lt;AdaptationSet&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="alternate"/&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 0 0 W/2 H/2"/&gt;   &lt;Representation id="BLTT1"     <b>dependencyId="BLBT"...</b>   &lt;/Representation&gt;   &lt;Representation id="ELTT1"     <b>dependencyId="BLTT1 ELBT"...</b>   &lt;/Representation&gt; &lt;/AdaptationSet&gt; ... &lt;AdaptationSet&gt;   &lt;Role     schemeIdUri=" urn:mpeg:dash:role:2011"     value="alternate"/&gt;   &lt;SupplementalProperty     schemeIdUri=" urn:mpeg:dash:srd:2014"     value="1 W/2 H/2 W/2 H/2"/&gt;   &lt;Representation id="BLTT4"     <b>dependencyId="BLBT"...</b>   &lt;/Representation&gt;   &lt;Representation id="ELTT4"     <b>dependencyId="BLTT4 ELBT"...</b>   &lt;/Representation&gt; &lt;/AdaptationSet&gt; </pre>	
--	---	--

## 4 Conclusion

In this contribution, we presented how to author MPD enabling the streaming of tiled videos using MPEG-DASH and its SRD amendment, for different scenarios (independently coded videos, HEVC tiled tracks and HEVC scalable and tiled tracks) with the ability to download only the necessary data and with fall back behavior for non SRD aware clients. To enable these scenarios, two clarifications are necessary to be added to the DASH standard:

- clarification on how to disable a representation when it is not meaningful to be presented on its own but is a complementary representation on which another representation depends;
- clarification on how to concatenate media segments when scalable and tile tracks are used

We recommend MPEG to adopt the first clarification in the Amendment 1 to DASH, to issue a corrigendum with the second one and to update MPEG-DASH Part 3 with the above examples and guidelines.