Source	**Telecom ParisTech, Canon Research Centre France**
Status	**For consideration at the 107th MPEG Meeting**
Title	**Multi-layer HEVC and Tile Tracks**
Author	Jean Le Feuvre, Cyril Concolato, Franck Denoual, Frédéric Mazé, Eric Nassor

1 Introduction

The WD for multi-layer HEVC and tiling (w13960) introduces the notion of Tile Tracks to store HEVC tiles in different tracks; the WD also defines storage of multi-layer HEVC in the ISOBMFF file format. This contribution reviews the usage of extractors in multi-layer configurations, and proposes to relax some existing rules on extractors.

NOTE: *the contribution does not address how dependencies between tiles across layers are expressed in the file format; this exploration is the subject of contribution m32286.*
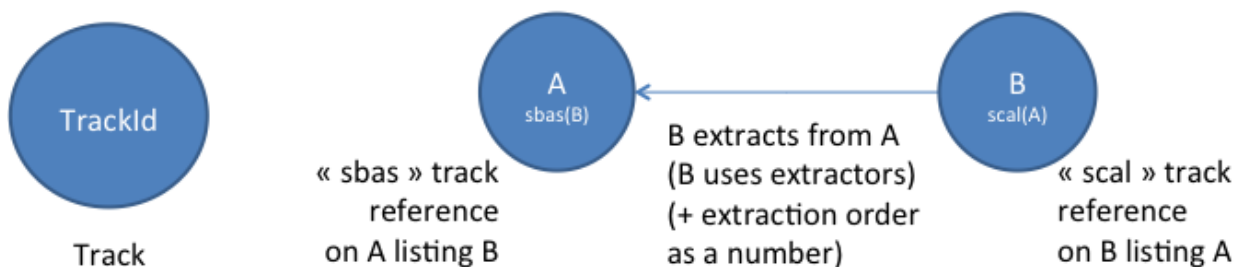
2 Multi-layer and tiling

Several configurations are possible for tiling and multi-layer, as follows:
- both base and enhancement layer not tiled;
- base layer tiled and enhancement layer not tiled;
- base layer not tiled and enhancement layer tiled;
- both base and enhancement layer tiled.

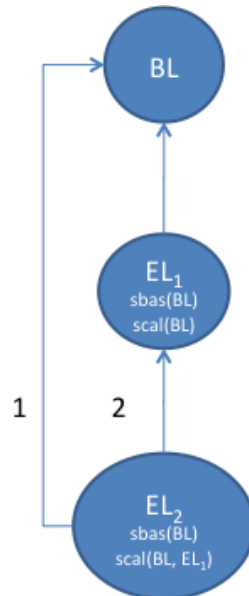In this contribution, the following conventions will be used:

This contribution addresses only one enhancement layer; it is easy for the reader of this contribution to extend the configurations to include more layers.

2.1 Reminders

The regular use case without any tiled layer is as seen with SVC/MVC. The track structure will look as follows:


Each sample in the enhancement layer will have an extractor to the base layer NALUs, followed by the NALUs of this layer. Each additional layer in a dedicated track will result in an additional extractor in the sample.

The single layer (HEVC) tiled use case as currently defined in the WD and refined in contribution m32284 is as follows:
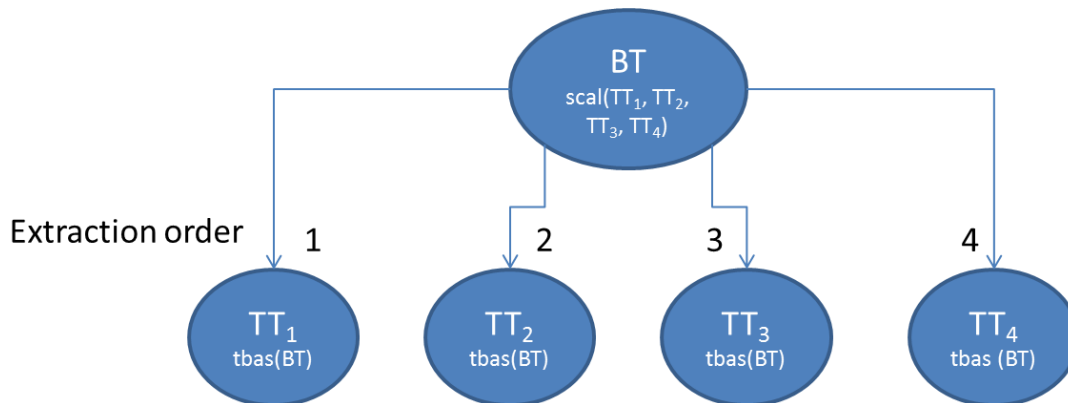

2.2 Non-tiled base and tiled enhancement

The use case of non-tiled base and tiled enhancement is similar to above single layer (HEVC) tiled use case, with the addition of an extractor at the beginning of each sample in the enhancement layer, as the entire base sample will be extracted.
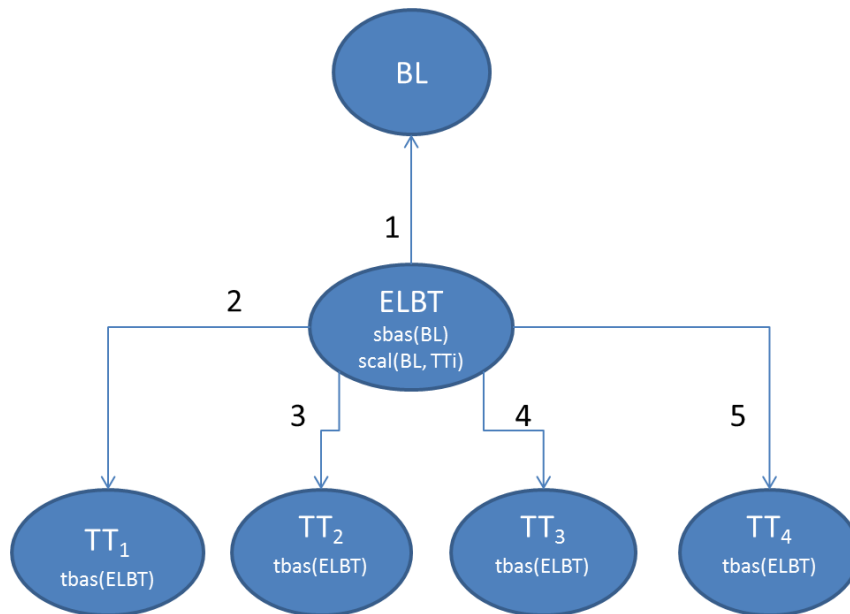
2.3 Tiled Base and non-tiled enhancement

In this use case, the base layer is tiled, with one track per tile and a "base" track containing extractors to each tile track and common info (SEI messages, parameter sets). The enhancement layer is not tiled. The track layout will look as follows:
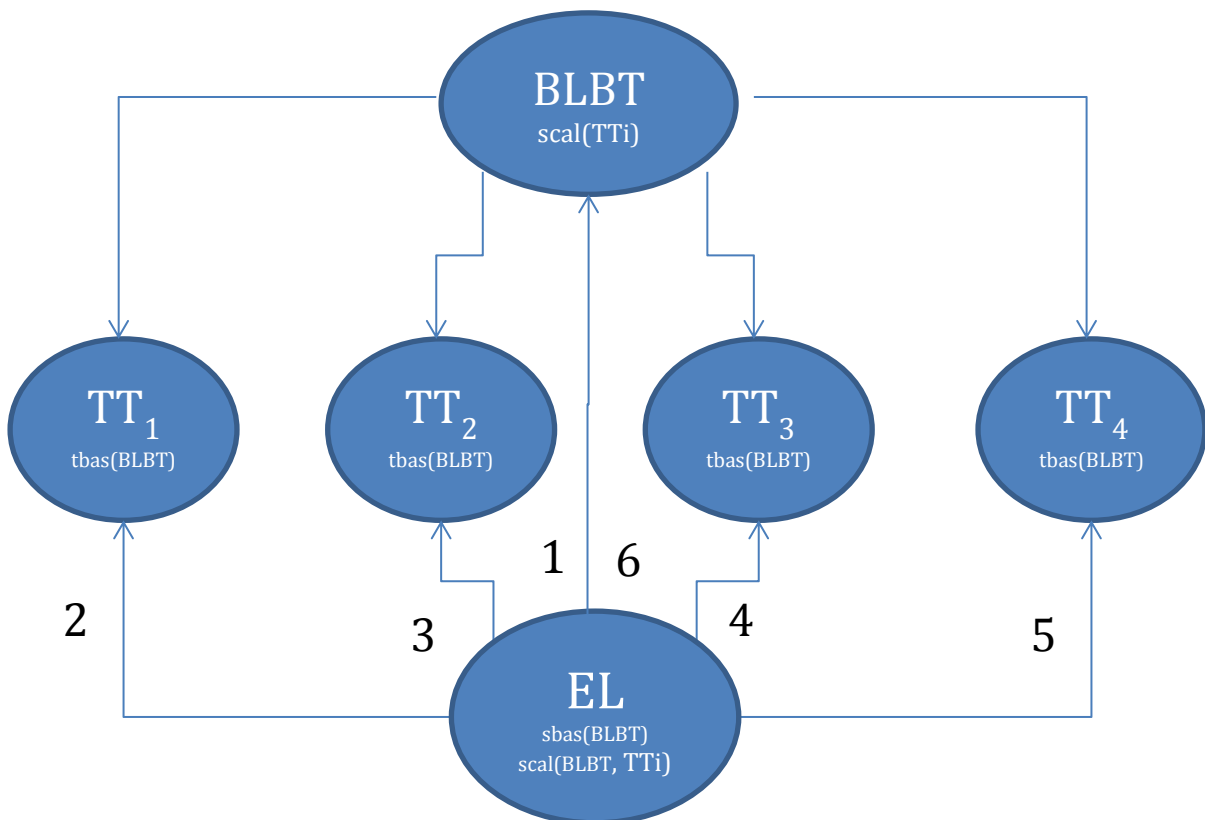
The sample at the EL layer will be composed of:

- Extractors to common data from base layer (SEI & co, extractor '1')
- Extractors to each tile track of the base layer ('2'-'5' extractors)
- (Optionally, extractors to any SEI suffix message in the base, '6' in the diagram)

- NAL units from the enhancement layer

Note that we do not describe the extractors present in the base layer, but there are indeed present. This complex structure and duplicate of extractor information is driven by the extractor rule stating that *an extractor must not reference, directly or indirectly, another extractor.*

2.4 Tiled Base and Tiled enhancement

In this example, each layer is tiled, with each tile in its own track and a base track for common information of the layer. The track structure looks as follows:
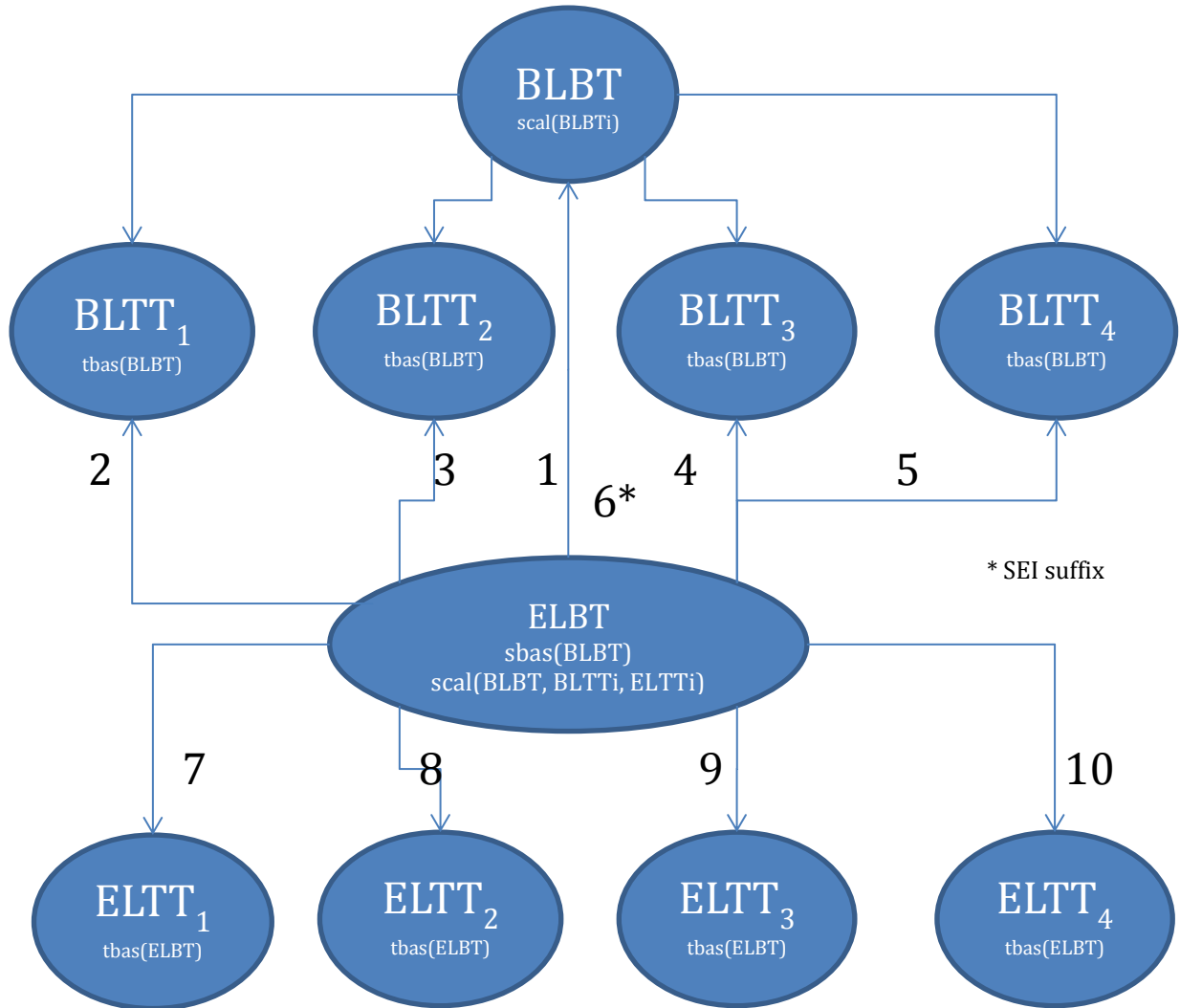

This rather complex structure (but conformant to the current spec) costs us a very large number of extractors in the ELBT sample: each set of NALUs are extracted directly since *an extractor must not reference, directly or indirectly, another extractor.*
For a 2x2 tiling at 25 Hz, the cost for extractors at enhancement layer is: 10 extractors * 12 Bytes * 25 * 8 = 24kbps.

This is moreover bad design since the same extractor pattern for tiles at Layer(i) will be repeated in all upper layers. The overhead will increase with each upper layer.

3   Proposed modification

We suggest relaxing the constraint on recursive extractors. For the above example, the result would be as follows:
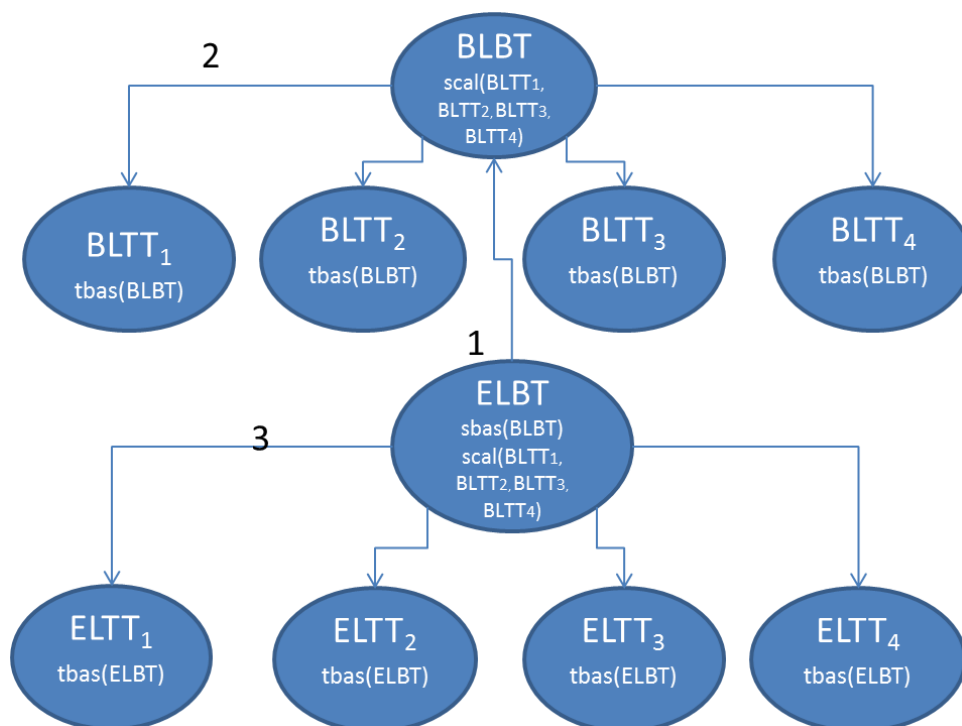
Instead of referencing each tile from the base layer explicitly, the enhancement layer only has one base layer extractor per sample.

Compared to previous example, for a 2x2 tiling at 25 Hz, the cost for extractors at enhancement layer is: 5 extractors * 12 Bytes * 25 * 8 = 14.4kbps (saving 40%).

Adding a new layer will not increase the encapsulation overhead, since only one extractor to the lower layer will be needed.

We suggest the following modification:

Replace in Annex A.3.1

"Those bytes must not contain extractors; an extractor must not reference, directly or indirectly, another extractor."

With

"Unless explicitly allowed, those bytes must not contain extractors (e.g. an extractor does not usually reference, directly or indirectly, another extractor)."

And add to the semantics of `data_length` (usually the number of bytes to copy):
"if data_length is equal to the size of the referred sample as indicated by the *stsz* or *trun* table, extractors present in the referred sample shall be recursively solved. There shall not be any cycle in the extractor paths."

For backward compatibility, we could allow this only for hvc2/shv2/… track types.

4 Conclusion

We suggest adding the proposed modifications to the WD on multi-layer HEVC storage "Enhanced support of HEVC and MVC+D".