

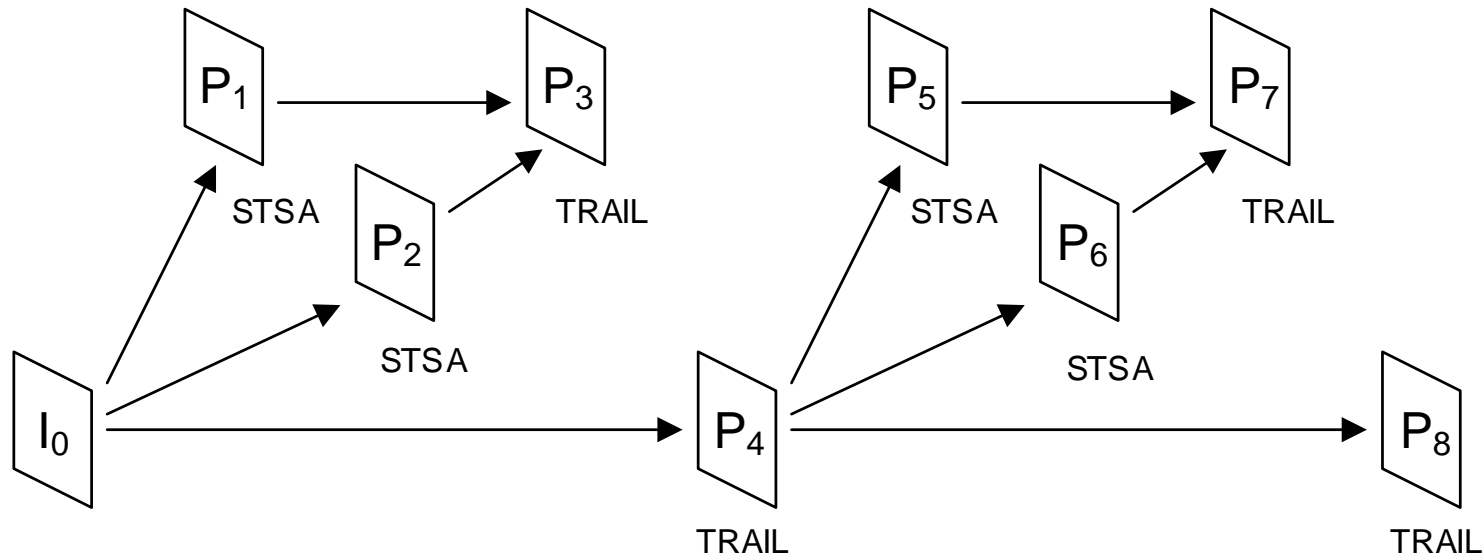
JVET-P0359



AHG17: An STSA_NUT restriction on PPS and APS availability

Rickard Sjöberg, Martin Pettersson. Mitra Damghanian
Ericsson

Background: The STSA picture



- The STSA_NUT in VVC indicates pictures to which temporal sub-layer up-switching is possible
- For example, assume you are decoding temporal sub-layer n and lower
 - If there is an STSA picture belonging to temporal sub-layer $(n+1)$ you can decode the STSA picture and pictures following the STSA picture belonging to sub-layers $(n+1)$ and lower

Background: Activation process



At the previous meeting, the parameter set activation processes were replaced by requirements that parameter sets shall be available before they are referred:

7.4.3.4 Picture parameter set RBSP semantics

A PPS RBSP shall be available to the decoding process prior to it being referred, included in at least one access unit with TemporalId less than or equal to the TemporalId of the PPS NAL unit or provided through external means, and the PPS NAL unit containing the PPS RBSP shall have nuh_layer_id equal to the nuh_layer_id of the coded slice NAL unit that refers it.

All PPS NAL units with a particular value of pps_pic_parameter_set_id within an access unit shall have the same content.

7.4.3.5 Adaptation parameter set RBSP semantics

Each APS RBSP shall be available to the decoding process prior to it being referred, included in at least one access unit with TemporalId less than or equal to the TemporalId of the coded slice NAL unit that refers it or provided through external means.

Let aspLayerId be the nuh_layer_id of an APS NAL unit. If the layer with nuh_layer_id equal to aspLayerId is an independent layer (i.e., vps_independent_layer_flag[GeneralLayerIdx[aspLayerId]] is equal to 1), the APS NAL unit containing the APS RBSP shall have nuh_layer_id equal to the nuh_layer_id of a coded slice NAL unit that refers it. Otherwise, the APS NAL unit containing the APS RBSP shall have nuh_layer_id either equal to the nuh_layer_id of a coded slice NAL unit that refers it, or equal to the nuh_layer_id of a direct dependent layer of the layer containing a coded slice NAL unit that refers it.

All APS NAL units with a particular value of adaptation_parameter_set_id and a particular value of aps_params_type within an access unit shall have the same content

Background: Sub-bitstream extraction process



The sub-bitstream extraction process only requires sub-bitstreams consisting of all NAL units equal to or lower than `tIdTarget` to be conforming:

10 Sub-bitstream extraction process

Inputs to this process are a bitstream, a target `nuh_layer_id` list `lIdTargetList`, and a target highest TemporalId value `tIdTarget`. Output of this process is a sub-bitstream.

It is a requirement of bitstream conformance for the input bitstream that any output sub-bitstream that is the output of the process specified in this clause with the bitstream, `lIdTargetList` containing of any particular `nuh_layer_id` value in the range of `vps_layer_id[0]` to `vps_layer_id[vps_max_layers_minus1]`, inclusive, when `vps_all_independent_layers_flag` is equal to 1, or `lIdTargetList` containing all `nuh_layer_id` value in the range of `vps_layer_id[0]` to `vps_layer_id[vps_max_layers_minus1]`, inclusive, when `vps_all_independent_layers_flag` is equal to 0, and `tIdTarget` equal to any value in the range of 0 to 6, inclusive, as inputs, and that satisfies the following conditions shall be a conforming bitstream:

- *The output sub-bitstream contains at least one VCL NAL unit with `nuh_layer_id` equal to each of the `nuh_layer_id` values in the `lIdTargetList`.*
- *The output sub-bitstream contains at least one VCL NAL unit with TemporalId equal to `tIdTarget`.*
NOTE – A conforming bitstream contains one or more coded slice NAL units with TemporalId equal to 0.

The output sub-bitstream is derived as follows:

- *Remove all NAL units with TemporalId greater than `tIdTarget`.*
- *Remove all NAL units with `nuh_layer_id` not included in `lIdTargetList`.*

Observations on the current VVC draft



- Any VVC bitstream consisting of multiple temporal layers must be constructed such that any bitstream created by removal of a complete temporal layer and all temporal layers above it would be conforming.
- For APS and PPS, there are requirements that every APS and PPS that is referred by a NAL unit shall be available
 - This means the APS and PPS must precede the NAL unit in decoding order
 - Also, the APS and PPS must belong to the same or lower temporal layer (otherwise an output from the sub-bitstream extraction process would be non-conforming)
- We believe that the current situation is as follows:
 - If a decoder is not decoding all temporal layers in the bitstream and a temporal layer up-switch at an STSA picture may occur in the future, it must process (store or decode) all APSes and PPSes of the full bitstream, including APSes and PPSes of higher temporal ID than is currently decoded
 - This means that a network node that prunes the bitstream must forward these higher temporal layer APSes and PPSes to the decoder
 - Note that any layer access unit must include a picture, which means that a decoder receiving a bitstream from a network node must be fed with pictures of higher temporal layer than it is decoding.

An example



- Assume an encoder generating a bitstream consisting of the following NAL units:

AU number	NAL unit number	NAL unit type	TemporalId	PPS id	APS id	PPS id referred	APS id referred
0	0	DPS	0	-	-		
0	1	VPS	0	-	-		
0	2	SPS	0	-	-		
0	3	PPS	0	0	-		
0	4	APS	0	-	0		
0	5	IDR_N_LP	0	-	-	0	0
1	6	PPS	1	1	-	-	-
1	7	APS	1	-	1	-	-
1	8	STSA	1	-	-	1	1
2	9	TRAIL	0	-	-	0	0
3	10	STSA	1	-	-	1	1

- We propose to enable a network node to start forwarding NAL units of a higher temporal layer when it encounters an STSA access unit without having to forward any NAL units of that higher temporal layer that precedes the STSA access unit in decoding order
- In the example, a network node that starts forwarding temporal sub-layer 0 only, and starts to forward temporal sub-layer 1 NAL units at NAL unit number 10, should not need to forward NAL units 6, 7 and 8, i.e. we propose that the stream above should not be allowed

Proposal



We propose to add the following restriction text in the semantics of the `nal_unit_type` syntax element:

- When the current access unit has a `nal_unit_type` equal to `STSA_NUT`, the following applies:
 - The current access unit shall not refer to a PPS or an APS that precedes the first NAL unit of the current access unit in decoding order and has `TemporalId` equal to the `TemporalId` of the current access unit
 - For any following access unit that has `TemporalId` equal to that of the current access unit and follows the current access unit in decoding order, the following access unit shall refer to a PPS or an APS that has `TemporalId` equal to that of the current access unit that precedes the first NAL unit of the current access unit in decoding order.

