|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  10th Meeting: Stockholm, SE, 11-20 July 2012 | Document: JCTVC-J0063\_r4 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Syntax for NAL Packet Priority** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Eun-Seok Ryu  Yan Ye  Yuwen He  Yong He 9710 Scranton Rd, Suite 250  San Diego, CA 92121, USA | Tel: Email: | +1-858-210-4808 [eun.ryu@interdigital.com](mailto:eun.ryu@interdigital.com)  [yan.ye@interdigital.co,m](mailto:yan.ye@interdigital.co,m)  [yuwen.he@interdigital.com](mailto:yuwen.he@interdigital.com)  [yong.he@interdigital.com](mailto:yong.he@interdigital.com) |
| *Source:* | InterDigital Communications, LLC | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

Pictures in the same temporal level in hierarchical-B structure can have different influence on error propagation and on decoded video quality. Currently in HEVC working draft 7, NAL unit header does not indicate packet priority within the same temporal layer. This contribution proposes two syntax options to indicate such priority of a NAL unit.

# Problem Statement

In video compression and transmission, packet prioritization is of utmost importance for the role it plays in UEP (Unequal Error Protection), packet dropping for bandwidth adaptation, as well as QP (Quantization Parameter) control for enhanced video quality, to name a few. As such, packet priority is essential for optimal QoS handling in applications such as video streaming. However, the current HEVC draft specification [1] does not provide sufficient information for prioritization. For example, priority ID is not specified. And although temporal ID (*temp\_id*) in the NAL header is provided, video packets with the same temporal ID could have different priorities.

## Uniform Prioritization

Fig. 1 denotes the current uniform prioritization in hierarchical-B structure [2] with four dyadic stages in temporal domain. Although pictures in lower temporal levels have higher priority, in the current HEVC working draft, however, there are no additional methods to assign different priorities to pictures at the same temporal level.



Fig. . The current uniform prioritization in hierarchical-B structure with four temporal levels.

Fig. 1 shows the random access setting in the common test condition [4]. In Fig. 1, POC 2 and POC 6 have the same priority. However, uniform prioritization at the same temporal level presents a problem because the importance of two POCs could be different according to both the RPS (Reference Picture Set) and the size of the reference picture list. The following sections explain why these two pictures have different importance.

## The Features of Picture Referencing

To compare the importance of two POCs located in the same temporal level, we define *Position A* as the pictures with POC equal to 2 + *N* × GOP (Group of Picture), and *Position B* as the pictures with POC equal to 6 + *N* × GOP, where GOP is 8 and *N* represents the number of GOP(s). Then, POC 2, 10, 18, and 26 belong to *Position A*, and POC 6, 14, 22, and 30 belong to *Position B*.

In the RA setting, *Position A* and *Position B* appear in the reference picture lists (L0 and L1) with different frequencies during each intra period. shows the reference picture lists for each POC and the total number of times *Position A* and *Position B* were referenced, using intra period equal to 32 as an example. In the table, POCs in *Position A (blue)* were referenced 12 times and POCs in *Position B (red)* 16 times during the intra period (= 32 POCs). Thus, compared to *Position A,* *Position B* is referenced more, indicating the pictures in *Position B* are more likely to cause severe error propagation if they are dropped during transmission.

Table 1. The amount of appearance in reference picture list (RA setting, GOP 8, IntraPeriod 32).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Picture** | **QP Offset** | **Reference**  **Buffer size**  **(L0 and L1)** | **Temporal**  **ID** | **Reference Picture Set**  **(RPS)** | | | | **Reference Picture Lists** | | | | | | | |
| **L0** | | | | **L1** | | | |
| 0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 8 | 1 | 4 | 0 | -8 | -10 | -12 | -16 | 0 |  |  |  | 0 |  |  |  |
| 4 | 2 | 2 | 0 | -4 | -6 | 4 |  | 0 | 8 |  |  | 8 | 0 |  |  |
| 2 | 3 | 2 | 0 | -2 | -4 | 2 | 6 | 0 | 4 |  |  | 4 | 8 |  |  |
| 1 | 4 | 2 | 0 | -1 | 1 | 3 | 7 | 0 | **2** |  |  | **2** | 4 |  |  |
| 3 | 4 | 2 | 0 | -1 | -3 | 1 | 5 | **2** | 0 |  |  | 4 | 8 |  |  |
| 6 | 3 | 2 | 0 | -2 | -4 | -6 | 2 | 4 | **2** |  |  | 8 | 4 |  |  |
| 5 | 4 | 2 | 0 | -1 | -5 | 1 | 3 | 4 | 0 |  |  | **6** | 8 |  |  |
| 7 | 4 | 2 | 0 | -1 | -3 | -7 | 1 | **6** | 4 |  |  | 8 | **6** |  |  |
| 16 | 1 | 4 | 0 | -8 | -10 | -12 | -16 | 8 | **6** | 4 | 0 | 8 | **6** | 4 | 0 |
| 12 | 2 | 2 | 0 | -4 | -6 | 4 |  | 8 | **6** |  |  | 16 | 8 |  |  |
| 10 | 3 | 2 | 0 | -2 | -4 | 2 | 6 | 8 | **6** |  |  | 12 | 16 |  |  |
| 9 | 4 | 2 | 0 | -1 | 1 | 3 | 7 | 8 | **10** |  |  | **10** | 12 |  |  |
| 11 | 4 | 2 | 0 | -1 | -3 | 1 | 5 | **10** | 8 |  |  | 12 | 16 |  |  |
| 14 | 3 | 2 | 0 | -2 | -4 | -6 | 2 | 12 | **10** |  |  | 16 | 12 |  |  |
| 13 | 4 | 2 | 0 | -1 | -5 | 1 | 3 | 12 | 8 |  |  | **14** | 16 |  |  |
| 15 | 4 | 2 | 0 | -1 | -3 | -7 | 1 | **14** | 12 |  |  | 16 | **14** |  |  |
| 24 | 1 | 4 | 0 | -8 | -10 | -12 | -16 | 16 | **14** | 12 | 8 | 16 | **14** | 12 | 8 |
| 20 | 2 | 2 | 0 | -4 | -6 | 4 |  | 16 | **14** |  |  | 24 | 16 |  |  |
| 18 | 3 | 2 | 0 | -2 | -4 | 2 | 6 | 16 | **14** |  |  | 20 | 24 |  |  |
| 17 | 4 | 2 | 0 | -1 | 1 | 3 | 7 | 16 | **18** |  |  | **18** | 20 |  |  |
| 19 | 4 | 2 | 0 | -1 | -3 | 1 | 5 | **18** | 16 |  |  | 20 | 24 |  |  |
| 22 | 3 | 2 | 0 | -2 | -4 | -6 | 2 | 20 | **18** |  |  | 24 | 20 |  |  |
| 21 | 4 | 2 | 0 | -1 | -5 | 1 | 3 | 20 | 16 |  |  | **22** | 24 |  |  |
| 23 | 4 | 2 | 0 | -1 | -3 | -7 | 1 | **22** | 20 |  |  | 24 | **22** |  |  |
| 32 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 28 | 2 | 2 | 0 | -4 | -6 | 4 |  | 24 | **22** |  |  | 32 | 24 |  |  |
| 26 | 3 | 2 | 0 | -2 | -4 | 2 | 6 | 24 | **22** |  |  | 28 | 32 |  |  |
| 25 | 4 | 2 | 0 | -1 | 1 | 3 | 7 | 24 | **26** |  |  | **26** | 28 |  |  |
| 27 | 4 | 2 | 0 | -1 | -3 | 1 | 5 | **26** | 24 |  |  | 28 | 32 |  |  |
| 30 | 3 | 2 | 0 | -2 | -4 | -6 | 2 | 28 | **26** |  |  | 32 | 28 |  |  |
| 29 | 4 | 2 | 0 | -1 | -5 | 1 | 3 | 28 | 24 |  |  | **30** | 32 |  |  |
| 31 | 4 | 2 | 0 | -1 | -3 | -7 | 1 | **30** | 28 |  |  | 32 | **30** |  |  |
| The number of appearances in reference picture list (L0 and L1)  \*counted once if the ref. POC. number is in both L0 and L1 | | | | | | | | **Position A** | | | | **Position B** | | | |
| **12** | | | | **16** | | | |

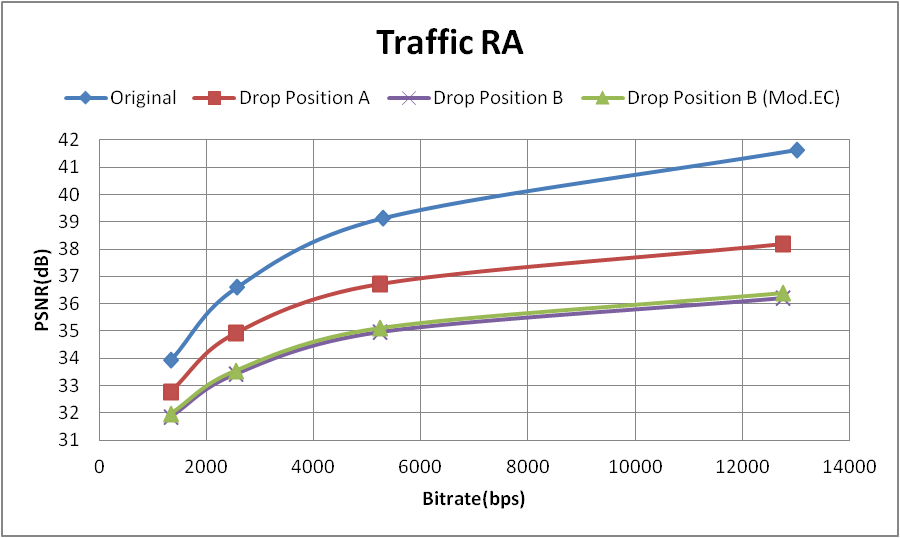
## The effect of Error Propagation

In order to quantify video quality degradation, we created packet dropping tests with encoded bitstreams; i.e., we dropped a picture per every IntraPeriod in *Position A* for 1st test and a picture per every IntraPeriod in *Position B* for 2nd test. Specifically, the Position A picture or the Position B picture in the second GOP of an intra period is dropped; that is, for the first intra period in the sequence, this corresponds to dropping POC 10 (Position A) vs. POC 14 (Position B), respectively. To measure the difference of objective video quality in two tests, the HM 6.1 video decoder decodes the damaged bitstream using frame copy. In HM 6.1, if a packet is lost, the decoder copies from the previous picture. This means for position A picture loss, a higher quality picture (POC 8) is copied compared to position B picture loss (where POC 12 is copied). To provide a more fair comparison, we also modified the HM6.1 decoder such that POC 16 instead of POC 12 is copied in the case of Position B loss. The test results for this case are provided below as “modified EC”.

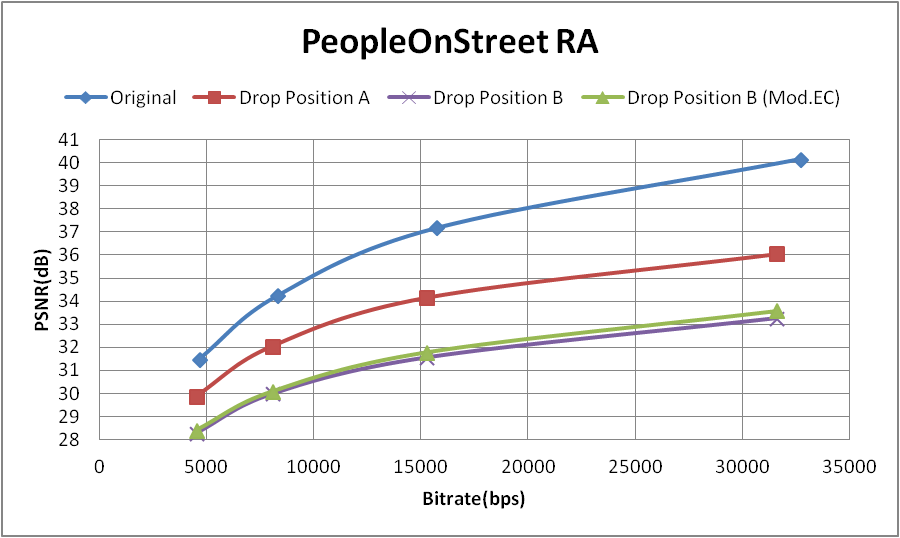
The testing was conducted on Windows 64-bit operating system, and using HM 6.1 decoder [3] (as well as a modified version as explained above) and three test sequences (Traffic, PeopleOnStreet, and ParkScene) from HEVC common test conditions [4]. The PSNR degradation caused by dropping pictures in *Position A* was less than the PSNR degradation caused by dropping pictures in *Position B*. The difference in error propagation effect is up to 2.8dB in luma PSNR (PeopleOnStreet, QP22). Figure 2 shows the corresponding rate-distortion curves, showing that the picture in *Position B* (purple-colored curve) is more important than the picture in *Position A* (red-colored curve). The curves in Figure 2 also show that the different impact on error propagation due to Position A loss and Position B loss largely holds regardless of the different error concealment strategies used.

Table 2. Comparison of *Position A* drop with *Position B* drop (RA setting).

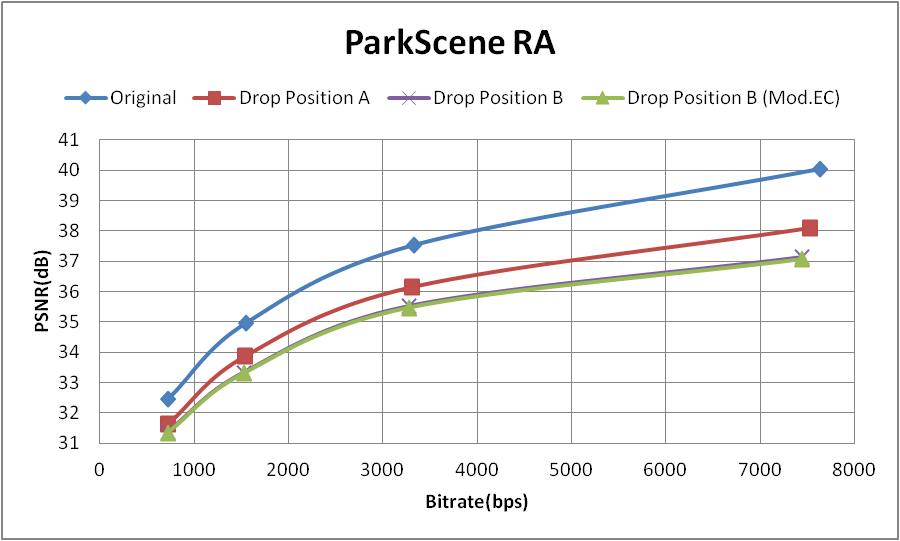
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | **Drop Position B** | | | | | | | | **Drop Position A** | | | | | | | | | **Drop Position B (mod. EC)** | | | | |
| Sequence | QP | | Kbps | | Y psnr | | U psnr | | V psnr | | Kbps | | | Y psnr | | U psnr | | V psnr | | kbps | | Y psnr | U psnr | V psnr |
| Traffic | 22 | | 12752 | | 36.19 | | 40.66 | | 42.78 | | 12751 | | | 38.18 | | 40.94 | | 43.39 | | 12752 | | 36.37 | 40.68 | 42.83 |
| 27 | | 5231 | | 34.96 | | 39.27 | | 41.42 | | 5231 | | | 36.72 | | 39.36 | | 41.69 | | 5231 | | 35.10 | 39.28 | 41.47 |
| 32 | | 2532 | | 33.42 | | 38.01 | | 40.18 | | 2532 | | | 34.91 | | 38.06 | | 40.30 | | 2532 | | 33.54 | 38.03 | 40.20 |
| 37 | | 1321 | | 31.86 | | 37.16 | | 39.34 | | 1321 | | | 32.77 | | 37.19 | | 39.44 | | 1321 | | 31.96 | 37.18 | 39.37 |
| PeopleOnStreet | 22 | | 31596 | | 33.25 | | 43.64 | | 43.72 | | 31614 | | | 36.04 | | 44.04 | | 44.01 | | 31596 | | 33.57 | 43.65 | 43.68 |
| 27 | | 15247 | | 31.55 | | 42.09 | | 42.59 | | 15257 | | | 34.15 | | 42.43 | | 42.88 | | 15247 | | 31.78 | 42.14 | 42.59 |
| 32 | | 8050 | | 29.97 | | 40.79 | | 41.54 | | 8054 | | | 32.04 | | 40.98 | | 41.68 | | 8050 | | 30.08 | 40.83 | 41.57 |
| 37 | | 4515 | | 28.27 | | 39.71 | | 40.60 | | 4517 | | | 29.88 | | 39.88 | | 40.74 | | 4515 | | 28.40 | 39.75 | 40.63 |
| ParkScene | 22 | | 7441 | | 37.13 | | 41.46 | | 43.01 | | 7523 | | | 38.10 | | 41.75 | | 43.25 | | 7441 | | 37.06 | 41.45 | 43.00 |
| 27 | | 3274 | | 35.52 | | 39.97 | | 41.19 | | 3304 | | | 36.15 | | 40.05 | | 41.21 | | 3274 | | 35.46 | 39.97 | 41.18 |
| 32 | | 1527 | | 33.33 | | 38.49 | | 39.73 | | 1539 | | | 33.88 | | 38.59 | | 39.82 | | 1527 | | 33.30 | 38.50 | 39.74 |
| 37 | | 718 | | 31.36 | | 37.44 | | 38.95 | | 723 | | | 31.64 | | 37.50 | | 39.00 | | 718 | | 31.33 | 37.44 | 38.95 |



(a)



(b)



(c)

Fig. . The RD curves by *Position A* drop and *Position B* drop ((a),(b),(c): Drop 1 picture per IntraPeriod)

# Proposed Syntax for NAL Packet Priority

This section proposes three syntax options to indicate the priority of a NAL unit.

* Option 1: This method changes *nal\_ref\_flag* to *nal\_priority\_flag*, and uses it to indicate the packet priority in the same temporal level. This method adds some priority indication using 1-bit flag (*nal\_priority\_flag*) without major modification of current NAL header syntax. For example, in Fig. 1, if picture with POC 6 is more important than picture with POC 2 in the same temporal level, this method could set nal\_priority\_flag to 1 (higher importance) to POC 6 and set *nal\_priority\_flag* to 0 (lower importance) to POC 2. The location of *nal\_priority\_flag* could be moved to another location in NAL header. As an example, it could be located next to *temporal\_id*.

Table 3. Option 1: (7.3.1) NAL unit syntax.

|  |  |
| --- | --- |
| nal\_unit( NumBytesInNALunit ) { | **Descriptor** |
| **forbidden\_zero\_bit** | f(1) |
| ~~nal\_ref\_flag~~ | ~~u(1)~~ |
| nal\_priority\_flag // *indicates NAL priority in same temporal level* | u(1) |
| **nal\_unit\_type** | u(6) |
| **temporal\_id** | u(3) |
| **reserved\_one\_5bits** | u(5) |
| …… |  |
| } |  |

* Option 2: Provide the priority information in AU (Access Unit) delimiter by modifying the current syntax (Table 4).

Table 4. Option 2: (7.3.2.8) Access unit delimiter RBSP syntax

|  |  |
| --- | --- |
| access\_unit\_delimiter\_rbsp( ) { | **Descriptor** |
| **pic\_type** | u(3) |
| priority\_id | u(4) |
| rbsp\_trailing\_bits( ) |  |
| } |  |

# References

1. B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, T. Wiegand, High efficiency video coding (HEVC) text specification draft 7, Document no. JCTVC-I1103, May 2012.
2. H. Schwarz, D. Marpe, T. Wiegand, Analysis of hierarchical B pictures and MCTF, IEEE International Conference on Multimedia and Expo, pp. 1929-1932, 2006.
3. HEVC Reference Software ver. 6.1 (HM 6.1), url: <http://hevc.kw.bbc.co.uk/trac/browser/tags/HM-6.1?rev=2385>
4. F. Bossen, Common test conditions and software reference configurations, Document no. JCTVC-H1100, Feb. 2012.

# Patent rights declaration(s)

**InterDigital Communications may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**

# Appendix: DIS text

Option 1:

|  |
| --- |
| **Section 7.4.1:**  **~~nal\_ref\_flag~~** ~~equal to 1 specifies that the content of the NAL unit contains a sequence parameter set, a picture parameter set, an adaptation parameter set or a slice of a reference picture.~~  ~~nal\_ref\_flag equal to 0 for a NAL unit containing a slice indicates that the slice is part of a non-reference picture.~~  ~~nal\_ref\_flag shall be equal to 1 for video parameter set, sequence parameter set, picture parameter set or adaptation parameter set NAL units. When nal\_ref\_flag is equal to 0 for one VCL NAL unit of a particular picture, it shall be equal to 0 for all VCL NAL units of the particular picture.~~  ~~nal\_ref\_flag shall be equal to 1 for NAL units with nal\_unit\_type equal to 4, 5, 6, 7, or 8.~~  ~~nal\_ref\_flag shall be equal to 0 for all NAL units having nal\_unit\_type equal to 29, 30, or 31.~~  **nal\_priority\_flag** equal to 1 specifies that the NAL unit has a higher priority than the other NAL units in the same temporal level.  nal\_priority\_flag equal to 0 specifies that the NAL unit has a normal priority in the same temporal level.  NOTE 2 - When the pictures have different temporal levels (temporal\_id), the pictures in lower temporal level always have higher priorities than the pictures in higher temporal level. However, the pictures in same temporal level also need to be distinguished with each other. The nal\_priority\_flag indicates whether the picture is more important than the other pictures in the same temporal level. |
| **3.72 non-reference picture**: A *picture* ~~coded with nal\_ref\_flag equal to 0. A~~ *~~non-reference picture~~* that is not used for *inter prediction* of any other *pictures*.  **3.98 reference picture**: A *picture* ~~with nal\_ref\_flag equal to 1. A~~ *~~reference picture~~* ~~contains samples~~ that may be used for *inter prediction* in the *decoding process* of subsequent *pictures* in *decoding order*.  **section 7.4.1.1**  2. The resulting sequence of bytes is then prefixed as follows.  – The sequence of bytes is prefixed with the first byte of the NAL unit containing the syntax elements forbidden\_zero\_bit, ~~nal\_ref\_flag~~ nal\_priority\_flag, and nal\_unit\_type, where nal\_unit\_type indicates the type of RBSP data structure the NAL unit contains.  **section 7.4.1.2.4**  – ~~nal\_ref\_flag~~ nal\_priority\_flag differs in value ~~with one of the nal\_ref\_flag values being equal to 0~~.  **section 8.3.5**  – ~~If nal\_ref\_flag of the current picture is equal to 0, the current picture is marked as "unused for reference".~~  ~~– Otherwise, the current picture is marked as "used for short-term reference".~~  **section D.2.18**  **~~sop\_desc\_nal\_ref\_flag~~sop\_desc\_nal\_priority\_flag[** i ] specifies the ~~nal\_ref\_flag~~ nal\_prioirty\_flag value of the i-th picture in decoding order within the SOP.  Let ~~nalRefFlag[ j ]~~nalPriFlag[ j ], tId[ j ], stRpsIdx[ j ], and picOrderCntVal[ j ] be the values of ~~nal\_ref\_flag~~ nal\_prioity\_flag, temporal\_id, st\_rps\_idx, and PicOrderCntVal that are in effect for the j-th picture in decoding order starting from j equal to 0 indicating the first picture of the SOP where this SEI message resides. Let currSeqParamSet be the previous seq\_parameter\_set\_rbsp with seq\_parameter\_set\_id equal to sps\_id in decoding order. long\_term\_ref\_pics\_present\_flag shall be equal to 0 in currSeqParamSet.  It is a requirement of bitstream conformance that when the SOP description SEI message, the following constraints shall apply for each picture i from picture 0 to picture num\_pics\_in\_sop\_minus1 when picOrderCntExp[ i ] is equal to picOrderCntVal[ j ], where j is greater than 0 and picOrderCntVal[ j − 1 ] is less than or equal to picOrderCntExp[ num\_pics\_in\_sop\_minus1 ]:  – ~~nalRefFlag~~nalPriFlag[ j ] shall be equal to sop\_desc\_nal\_~~ref~~priority\_flag[ i ] |

Option 2:

|  |
| --- |
| **Section 7.4.2.8:**  **priority\_id** specifies a priority identifier for the following NAL unit(s) until the next access unit delimiter is present. The priority identifier indicates the priority of NAL unit(s) in the same temporal level. |