ISO/IEC JTC 1/SC 29/WG 03 N0536

**ISO/IEC JTC 1/SC 29/WG 03  
MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** Technologies under Consideration for NAL File Format

**Status:** Approved

**Date of document:** 2022-04-29

**Source:** ISO/IEC JTC 1/SC 29/WG 03

**No. of pages:** 9 (with cover page)

**Email of Convenor:** young.L @ samsung . com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg3>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 03 MPEG SYSTEMS**

**ISO/IEC JTC 1/SC 29/WG 03 N0536**

**April 2022, Virtual**

|  |  |
| --- | --- |
| **Title** | **Technologies under Consideration for NAL File Format** |
| **Source** | **WG 03, MPEG Systems** |
| **Status** | **Approved** |
| **Serial Number** | **21429** |

[1 Signalling of non-reference layer (from m56045) 1](#_Toc102172658)

[2 APS Roll Recovery (from m54403) 2](#_Toc102173096)

[3 14496-15 (NAL video file formats) errata items (from m55192) 3](#_Toc102173097)

[4 EDRAP and Random access (from m56766) 4](#_Toc102173098)

[5 Generic Codecs Parameter (from m59046) 6](#_Toc102173120)

# Signalling of non-reference layer (from m56045)

We propose signaling of a sample group for non-reference samples. Non-reference samples are never used as a reference samples. A layer non reference flag is signaled in the sample group. It is asserted that knowing information about non-reference pictures sample group to indicate samples which are not used as a reference is useful. For example, this information can be used to drop samples and not decode them if the playback is lagging on a resource-constrained system. Additionally, this information can be used by media aware network element to drop samples if there is bandwidth starvation.

The proposed addition is shown below compared to WG03N0035 [1].

* + 1. **Non reference sample group** 
       1. **Definition**

Group Types: 'nref'  
Container: Sample Group Description Box ('sgpd')  
Mandatory: No  
Quantity: Zero or one

This sample group is used to mark non reference samples. A non reference sample is never used as a reference sample for any other sample. An accompanying instance of the SampleGroupDescriptionBox with the same grouping type shall be present. The grouping\_type\_parameter is not defined for the SampleToGroupBox with grouping type 'nref'.

* + - 1. **Syntax**

class NonReferenceEntry() extends VisualSampleGroupEntry ('nref')  
{  
 bit(7) reserved = 0;  
 unsigned int(1) layer\_non\_ref\_only\_flag;  
}

layer\_non\_ref\_only\_flag equal to 1 specifies that for this sample a current picture is never used as a reference picture for any picture within its own layer and may or may not be used a reference picture for inter-layer prediction for pictures of a different layer. layer\_non\_ref\_only\_flag equal to 0 specifies that the current sample is never used as a reference sample.



# APS Roll Recovery (from [m54403](http://wg11.sc29.org/doc_end_user/current_document.php?id=75250&id_meeting=183))

11.8.X APS Roll Recovery

11.8.X.1 Definition

The 'apsr' sample group indicates that a VVC sync sample from a VVC track requires additional gathering of prefix and suffix APS NAL units from preceding samples and rewriting these as APS prefix NAL units to be a self-contained sync sample. This avoids having to duplicate APS information at each sync sample in the track.

A sync sample not belonging to an 'apsr' sample group does not require any additional processing to gather the dependent APS (i.e., all APS required are in the sample entry or in the sample). The 'apsr' sample group shall only be present in a VVC track or a VVC base track with no dependencies to a VVC non-VCL track; it shall not be present in VVC subpicture tracks or VVC non-VCL tracks.

The aps\_roll\_count shall be such that all samples described by the roll operation are available in the track, track fragment or ISOBMFF segment (as indicated by the roll\_type) being processed.

A sample associated to an 'apsr' sample group description entry shall be a sync sample or a sample with SAP type 3 or 4 (potentially associated to a 'roll' sample group description entry).

A sample associated to an 'apsr' sample group description entry with aps\_roll\_count not equal to 0 shall be considered as a SAP type 4.

11.8.X.2 Syntax

aligned(8) class APSRollRecoveryEntry () extends VisualSampleGroupEntry('apsr')  
{  
 unsigned int(2) roll\_type;  
 unsigned int(2) roll\_mode;  
 unsigned int(4) reserved=0;  
 if (roll\_type == 0) {  
 unsigned int(16) aps\_roll\_count;  
 }

X.3 Semantics

roll\_type indicates the pre-roll distance for APS NAL units when producing a sync sample with sample number N belonging to this group. The following values are defined:

0: APS NAL units are gathered starting from the sample located aps\_roll\_count samples before the sample belonging to the group

1: APS NAL units are gathered starting from the first sample of the track or track fragment

2: APS NAL units are gathered starting from the first sample of the associated ISOBMFF segment

3: reserved

roll\_mode indicates which samples in the identified roll sample window should be analyzed for APS NAL unit gathering. The following values are defined:

0: the required APS NAL units may be present in any sample

1: the required APS NAL units are only present in the first sample

2: the required APS NAL units are only present in samples that are either sync samples or samples marked as 'rap '

3: the required APS NAL units are only present in samples of the current track fragment

aps\_roll\_count indicates the number of samples to rewind for APS gathering; a value of 0 indicates that the associated sample contains all APS NAL units for its processing.

# 14496-15 (NAL video file formats) errata items (from [m55192](https://wg11.sc29.org/doc_end_user/current_document.php?id=76249&id_meeting=184))

[Editor’s note] The following changes have been captured in the current TuC document in lack of a Defect Report on 14496-15 issued at MPEG #132. This paragraph is thus meant to be moved in a more appropriate document at the next meeting.

1. Search and replace "RBSP payload" (1 instance) with "NAL unit payload", because in the context where the phase is used, the start code emulation bytes that are part of "NAL unit payload" but not part of "RBSP payload" should be considered.
2. Search and replace "byte stream payload" (2 instances) with "NAL unit payload", for similar reason as above.
3. Change, in clause 8.4.1.1.1, the following:

"When the sample entry name is 'hvc1', the default and mandatory value of array\_completeness is 1 for arrays of all types of parameter sets, and 0 for all other arrays. When the sample entry name is 'hev1', the default value of array\_completeness is 0 for all arrays."

to

"When the sample entry name is 'hvc1', the ~~default and mandatory~~ value of array\_completeness ~~is~~ shall be equal to 1 for the arrays of all types of parameter sets~~, and 0 for all other arrays~~. When the sample entry name is 'hev1', the ~~default~~ value of array\_completeness shall be equal to 1 for the arrays of all types of parameter sets ~~is 0 for all arrays~~."

1. Change, in clause 9.5.3.1.1, the following:

"When the sample entry name is 'lhv1', the default and mandatory value of array\_completeness is 1 for arrays of all types of parameter sets, and 0 for all other arrays. When the sample entry name is 'lhe1', the default value of array\_completeness is 0 for all arrays."

to

"When the sample entry name is 'lhv1', the ~~default and mandatory~~ value of array\_completeness ~~is~~ shall be equal to 1 for the arrays of all types of parameter sets~~, and 0 for all other arrays~~. When the sample entry name is 'lhe1', the ~~default~~ value of array\_completeness shall be equal to 1 for the arrays of all types of parameter sets ~~is 0 for all arrays~~."

1. Change, in clause 8.3.3.1.3, the following:

array\_completeness when equal to 1 indicates that all NAL units of the given type are in the following array and none are in the stream; when equal to 0 indicates that additional NAL units of the indicated type may be in the stream; the default and permitted values are constrained by the sample entry name.

to the following:

array\_completeness when equal to 1 indicates that all NAL units of the given type are in the following array and none are in the stream; when equal to 0 indicates that additional NAL units of the indicated type may be in the stream; the ~~default and~~ permitted values are constrained by the sample entry name.

1. Change, in clause 8.3.3.1.1, the following:

The level indication general\_level\_idc shall indicate a level of capability equal to or greater than the highest level indicated for the highest tier in all the parameter sets.

to the following:

The level indication general\_level\_idc shall indicate a level of capability equal to or greater than the highest level ~~indicated for the highest tier~~ in all the parameter sets.

Because the highest level of the highest tier could be lower than the highest level of the lowest tier, while level determines spatial resolution etc., which is of vital importance for determining the required decoding capability.

# EDRAP and Random access (from m56766)

The ‘edrp’ sample group from ISO/IEC 14496-12 TuC **Error! Reference source not found.** proposes to signal additional random access point in bitstream that could outperform DRAP in coding efficiency thank to the possibility to refer to previous EDRAP for inter prediction as represented in the figure below:

Logo

Description automatically generated

*Figure 3: Inter prediction reference relationship among the RAP pictures in the EDRAP case.*

In current version of the Part-12 TuC, no information is provided in case an EDRAP depends on APS NAL units present in prior samples.

For example, the figure below is an example of such VVC bitstream.

A picture containing logo

Description automatically generated

When Random Accessing from the EDRAP6 sample, the ‘edrp’ sample group indicates that only IDR0 and EDRAP2 are needed for reference to decode the bitstream (yellow or grey arrows). The ‘edrp’ sample group is silent about possible references to APS NAL units in previous samples (red arrows). For example, EDRAP6 may reference APSs that are not part of the sample used as references in the EDRAP as per section 11.3.4 of ISO/IEC 14496-15 (recalled hereafter for convenience):

*When the sample entry name is 'vvc1'and the track does not have a track reference of type 'vvcN', the following applies:*

* *If the sample is a sync sample, all APSs needed for decoding that sample shall be included either in the sample entry or in the sample itself.*
* ***Otherwise (the sample is not a sync sample), all APSs needed for decoding the sample shall be included either in the sample entry or in any of the samples since the previous sync sample to the sample itself, inclusive.***

A screenshot of a video game

Description automatically generated

From the discussion at MPEG #134 (<http://mpegx.int-evry.fr/software/MPEG/Systems/FileFormat/NALuFF/-/issues/133>), the following observation has been made:

All non-VCL NAL units needed to decode the EDRAP should be referenced from the EDRAP sample group.

# Generic Codecs Parameter (from m59046)

## EVC Codecs Parameter

DASH and other applications require defined values for the Codecs parameter specified in IETF RFC 6381 for ISO BMFF Media tracks. The 'codecs' parameter string for the EVC codec is defined as follows:

<sample entry 4CC>.<key1><value1>.<key2><value2>.….<keyN><valueN>

Keys are defined as 4CCs. An set of keys and the associated value pairs are defined in **Table *2***. Additional keys may be specified as 4CCs. Preferably, keys are aligned with ISO/IEC 23091-2. [Ed. (MH/KB): If the keys derived from ISO/IEC 23091-2 are kept in this document, a normative reference to ISO/IEC 23091-2 has to be added.]

If a specific key is not provided, then the value takes the default value specified in the table, or the value is unknown if no default is specified.

NOTE: The parameters from 'vbit' onwards in **Table *2*** are not EVC specific and are applicable to any video codec. It is foreseen that these values are moved to ISO/IEC 14496-12 in future versions and may then be applicable to other video codecs as well. [Ed. (MH): FI\_106-207 resolution: Keys from vbit onwards moved to ISO/IEC 14496-12.]

**Table 2 - Definition of Set of Keys and Values defined for EVC**

|  |  |  |  |
| --- | --- | --- | --- |
| **Key** | **Key Definition** | **Value** | **Default Value** |
| 'vprf' | Defines the video profile | profile\_idc | 1 |
| 'vlev' | Defines the video level | level\_idc | 51 |
| 'vtoh' | Defines the toolset | toolset\_idc\_h in hex decimal | x1FFFFF |
| 'vtol' | Defines the toolset | toolset\_idc\_l in hex decimal | x000000 |
| 'vbit' | Defines the video bit depth for luma and chroma | Value is a 2 digit decimal . The first digit is the luma bit depth minus 8. The second digit is the chroma bit depth minus 8 | unknown |
| 'vcss' | Defines the chroma subsampling | The subsampling scheme is expressed as a three part ratio J:a:b (e.g. 4:2:2) that describes the number of luminance and chrominance samples in a conceptual region that is J pixels wide, and 2 pixels high. The parts are (in their respective order):   * J: horizontal sampling reference (width of the conceptual region). Usually, 4. * a: number of chrominance samples (Cr, Cb) in the first row of J pixels. * b: number of changes of chrominance samples (Cr, Cb) between first and second row of J pixels.   Then the value is as follows. The first digit is J, the second digit is a and the third digit is b. | 420 |
| 'vcpr' | Defines colour primaries (ColourPrimaries) as defined in ISO/IEC 23091-2. | Defines colour primaries (ColourPrimaries) values as defined in ISO/IEC23091-2. Value is a 2 digit decimal with a possible preceding 0. | 01 |
| 'vtrc' | Defines transfer characteristics (TransferCharacteristics) as defined in ISO/IEC23091-2. | Defines transfer characteristics (TransferCharacteristics) values as defined in ISO/IEC23091-2. Value is a 2 digit decimal with a possible preceding 0. | 01 |
| 'vmac' | Defines matrix coefficients (MatrixCoefficients) as defined in ISO/IEC23091-2. | Defines matrix coefficients (MatrixCoefficients) values as defined in ISO/IEC23091-2. Value is a 2 digit decimal with a possible preceding 0. | 01 |
| 'vfrf' | Defines VideoFullRangeFlag as defined in ISO/IEC23091-2. | Defines VideoFullRangeFlag as defined in ISO/IEC23091-2. Value is a 1 digit decimal restricted to values 1 or 0. | 0 |
| 'vfpq' | Defines video frame packing type (VideoFramePackingType) as defined in ISO/IEC23091-2 together with the QuincunxSamplingFlag. | Defines video frame packing type (VideoFramePackingType) values as defined in ISO/IEC23091-2. Value is a 2 digit decimal with the first one being the QuincunxSamplingFlag and the second digit the VideoFramePackingType. | If not present, then no frame packing is used. |
| 'vpci' | Defines Packed content interpretation type (PackedContentInterpretationType) as defined in ISO/IEC23091-2. | Packed content interpretation type (PackedContentInterpretationType) as defined in ISO/IEC23091-2. Value is a 1 digit. | If not present, then no packed content is used. |
| 'vsar' | Defines Sample aspect ratio indicator (SampleAspectRatio) as defined in ISO/IEC23091-2. | Defines Sample aspect ratio indicator (SampleAspectRatio) values as defined in ISO/IEC23091-2. Value is a 2 digit decimal with a possible preceding 0. | 01 |

For example, codecs="evc1.vprf3.vlev51.vto1FFFFF.vtoh000000.vbit20.vcss420.‌vcpr09.vtrc16.vmac09.vsar01" represents EVC Main Profile, level 5.1, with 4:2:0 chroma subsampling co-located with (0, 0) luma sample, a restricted tool set, ITU-R BT.2100 color primaries, ITU-R BT.2100 PQ transfer characteristics, ITU-R BT.2100 YCbCr color matrix and sample aspect ratio 1:1. [Ed. (KB): If the keys derived from ISO/IEC 23091-2 are kept in this document, these ITU-Rs should be added to the Bibiography.]

All keys in **Table *2*** must be recognized, if the 'evc1' sample entry is recognized.

If a key is not recognized, the key value pair is ignored