Video Services Forum (VSF)
Technical Recommendation
TR-01:2018

Transport of JPEG 2000 Broadcast Profile video in MPEG-2 TS over IP

June 5, 2018
TR-01:2018-06-05

http://www.videoservicesforum.org

Executive Summary

JPEG 2000 is used in broadcast transmission applications for cost-effective, high quality real-time transport of television video signals over IP networks. This VSF Technical Recommendation (TR), in addition to defining profiles for streaming of JPEG 2000 Broadcast Profile video, adds information for the interoperable transport of audio and ancillary data in an MPEG-2 Transport Stream. This TR also defines an optional Forward Error **C**orrection scheme.

The primary objective of this TR is to enable interoperability between products from different equipment manufacturers.

The VSF has published a previous version of this TR known as *VSF TR-01:2013, Transport of JPEG 2000 Broadcast Profile video in MPEG-2 TS over IP*. This document, *VSF TR-01:2018*, adds the following capabilities to the original version:

1. **Ultra-low Latency**: Enables Ultra-Low Latency (ULL) encoding of professional video using horizontal, independent JPEG 2000 stripes. End-to-end latencies of less than one video frame are possible.

2. **Higher video resolutions, frame rates and bit depths**: Beyond supporting up to 120fps frame rates and 12 bit depths it also provides additional implementation options for 4K and higher resolutions by adding "block mode".

3. **Broader Color Space and Mastering Display Metadata:** Expands color space to include color primaries, transfer characteristics and matrix coefficients, in accordance with the values defined in Rec. ITU-T H.273. Adds mastering display metadata as described in SMPTE ST 2086:2018

4. **Interoperability**: Addresses issues with two incompatible definitions of the JPEG 2000 elementary stream header; one in Annex M of Rec. ITU-T T.800 | ISO/IEC 15444-1, and the other in Annex S of Rec. ITU-T H.222.0 | ISO/IEC 13818-1.

   The issue is resolved by establishing Annex S of Rec. ITU-T H.222.0 | ISO/IEC 13818-1 as the single authoritative source for the header used in this document.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the Recommendation set forth in this document, and to provide supporting documentation.

Table of Contents

1. Introduction

In June of 2012, the Video Services Forum, Inc. (VSF) created an Activity Group titled the "JPEG 2000 Broadcast Profile Technical Recommendation Group". The output of that group was published as *VSF TR-01:2013, Transport of JPEG 2000 Broadcast Profile video in MPEG-2 TS over IP*.

A follow-on Activity Group was formed in October 2014 to, a) develop an interoperable method for the provision of low-latency JPEG 2000 transport, b) to add specifications for the compression and transmission of Ultra-High Definition formats, and c) to address interoperability issues with the original Technical Recommendation. This document is the result of the work of that group.

1.1 Contributors

Contributors to this document include:

- John Dale, Activity Group Chairman, Media Links, Inc.
- Antonin Descampe, intoPIX SA
- Brad Gilmer, Gilmer & Associates, Inc.
- Inge Hillested, Nevion
- Rakesh Jalali, Evertz
- Andrew Krupiczka, ESPN
- Jean-Baptiste Lorent, intoPIX SA
- Chuck Meyer, Grass Valley, a Belden Brand
- Feng Wang, Imagine

1.2 About the Video Services Forum

The Video Services Forum, Inc. (www.videoservicesforum.org) is an international association dedicated to video transport technologies, interoperability, quality metrics and education. The VSF is composed of service providers, users and manufacturers. The organization's activities include:

- providing forums to identify issues involving the development, engineering, installation, testing and maintenance of audio and video services;
- exchanging non-proprietary information to promote the development of video transport service technology and to foster resolution of issues common to the video services industry;
- identification of video services applications and educational services utilizing video transport services;
- promoting interoperability and encouraging technical standards for national and international standards bodies.

The VSF is an association incorporated under the Not For Profit Corporation Law of the State of New York. Membership is open to businesses, public sector organizations and individuals worldwide.

For more information on the Video Services Forum, contact:

Bob Ruhl
Operations Manager
Video Services Forum
+1 929-279-1995, bob.ruhl1@verizon.net.

2. Conformance Notation

Normative text is text that describes elements of the design that are indispensable or that contain the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative", or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3. Normative References

[1] AES: AES3:2009, “Digital input-output interfacing — Serial transmission format for two-channel linearly-represented digital audio data”

[2] ANSI/SCTE 127 2007 “Carriage of Vertical Blanking Interval (VBI) Data in North American Digital Television Bitstreams”

[3] ETSI EN 301 775 “Digital Video Broadcasting (DVB); Specification for the carriage of Vertical Blanking Information (VBI) data in DVB bitstreams”

[4] Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017: "Information Technology - Generic Coding of moving pictures and associated audio information: Systems"

[5] Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2018 AMD1: " Ultra-Low Latency and 4k and higher resolution support for transport of JPEG 2000 video "

[6] Rec. ITU-T H.273 (2016) | ISO/IEC 23001-8:2016 : "Coding-independent code points for video signal type identification"

[7] Rec. ITU-T T.800 (2015) | ISO/IEC 15444-1:2016 : "Information technology - JPEG 2000 image coding system: Core coding system"

[8] SMPTE ST 299-2: 2010 “Extension of the 24-Bit Digital Audio Format to 32 Channels for 3Gb/s Bit – Serial Interfaces”

[9] SMPTE ST 302M-2007: "Television - Mapping of AES3 Data into MPEG-2 Transport Stream".

[10] SMPTE ST 337:2015 “Television - Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface".

[11] SMPTE ST 338:2016 “Format for Non-PCM Audio and Data in an AES3 – Data Types”.

[12] SMPTE ST 2022-1:2007 “Forward Error Correction for Real-Time Video/Audio Transport Over IP Networks”

[13] SMPTE ST 2022-2:2007 “Unidirectional Transport of Constant Bit Rate MPEG-2 Transport Streams on IP Networks”

[14] SMPTE ST 2038:2008 “Carriage of Ancillary Data Packets in MPEG-2 Transport Stream”

[15] SMPTE ST 2063:2012 “Stereoscopic 3D Full Resolution Contribution Link Based on MPEG-2 TS”

[16] SMPTE ST 2086:2018 “Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images”

4. Acronyms

3G	3 Gbit/s (Serial Digital Interface)
AES	Audio Engineering Society
ES	Elementary Stream
ETSI	European Telecommunications Standards Institute
FEC	Forward Error Correction
HD	High Definition
HDR	High Dynamic Range

IEC	International Electrotechnical Commission
IP	Internet Protocol
ISO	International Organization for Standardization
ITU	International Telecommunication Union
J2K	JPEG 2000
JPEG	Joint Photographic Experts Group
MPEG	Moving Picture Experts Group
PES	Packetized Elementary Stream
PID	Packet Identifier
PTS	Presentation Time Stamp
RTP	Real-time Transport Protocol
S3D	Stereoscopic 3D
SD	Standard Definition
SDI	Serial Digital Interface
SMPTE	Society of Motion Picture Television Engineers
TR	Video Services Forum Technical Recommendation[1]
TS	Transport Stream
ULL	Ultra-Low Latency
UDP	User Datagram Protocol

5. Definitions

Codestream	Compressed image data representation that includes all necessary data to allow (lossless or lossy) reconstruction of the sample values of a digital image
Device	Hardware or software application that can include multiple Senders and Receivers
Interoperability	An end user or service provider can transport a signal using devices from different manufacturers that state compliance with this Technical Recommendation with the expectation that they will successfully achieve their business objective
J2K block	A JPEG 2000 codestream or codestreams corresponding to a rectangular portion of the active video frame, divided horizontally and vertically in the spatial domain
J2K block mode	An optional mode dividing each video frame into rectangular blocks, each of which is encoded independently as a J2K block
J2K stripe	A JPEG 2000 codestream or codestreams corresponding to a rectangular portion of the full raster of the video frame or, if the optional J2K block mode is being used, then the JPEG 2000 codestream or codestreams corresponding to the full raster of the portion of the video frame covered by a J2K block.

[1] Note that the term Technical Recommendation is also used by other organizations such as ETSI.

J2K stripe mode — Optional mode dividing video frame or J2K block into a succession of horizontal stripes, and formatted as a J2K Video Access Unit. Each J2K stripe is encoded as an independent JPEG 2000 codestream, or if the video is interleaved, as two independent JPEG 2000 codestreams.

J2K video elementary stream – Video elementary stream consisting of a succession of J2K Video Access Units

J2K Video Access Unit – The JPEG 2000 codestream or codestreams comprising a decodable and randomly accessible image or block, preceded by the *J2K_elsm_header* that contains all the necessary information to decode the J2K Video Access Unit

Receiver — Element within a device that terminates one RTP stream from the Network

Sender — Element within a device that originates one RTP stream into the Network

6. System Overview (Informative)

An end-user or service provider of broadcast transmission services may utilize devices that implement this Technical Recommendation (TR) for the unidirectional transport of various television signals over IP. Examples of these signals include SDI, RTP Streams, native server or camera outputs and associated audio.

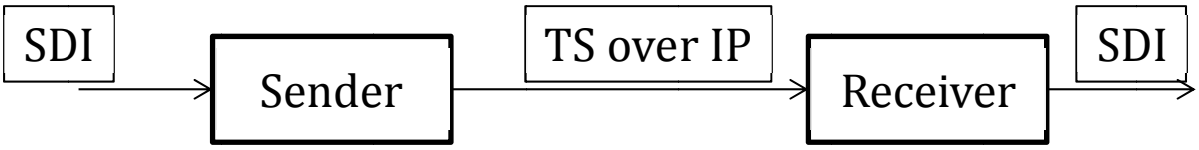

Figure1: Example system for transmission of SDI signal over IP

As shown in figure 1 above, a Sender with an SDI video input interface extracts the active video, audio and ancillary data components from the SDI signal. The active video is sent to a JPEG 2000 compression engine. The JPEG 2000 code stream is multiplexed into an MPEG-2 Transport Stream (TS) together with its associated audio and ancillary data. (The system supports transparent pass-through of linear PCM and non-PCM audio, and transparent pass-through of ancillary data.) The Sender encapsulates the TS in an RTP stream and transmits this stream using Internet Protocol to a receiving device ("Receiver"). The Receiver de-encapsulates the RTP/IP stream, de-multiplexes the TS, decodes the JPEG 2000 codestream, and places the video together with associated audio and ancillary data onto the SDI output.

This TR specifies the syntax and semantics of the signal between the Sender and the Receiver, and in so doing places constraints on the behavior of the Sender; it also specifies some minimum requirements for the Receiver. These requirements of the Sender and Receiver are needed for interoperability.

TR-01:2018 includes a mechanism to signal the use of a number of new features. To signal the use of new features, the 16 bits of the *profile_and_level* field from TR-01:2013 are split into two parts:

- The most significant bit in the *profile_and_level* field is used as the *extended_capability_flag*. If it is set to '1', it indicates that the stream complies with TR-01:2018.

  If it is set to '0', this most significant bit indicates that the Sender is transmitting a legacy TR-01:2013 stream.

- The 15 least significant bits correspond to the 15 least significant bits of the 16-bit RSiz parameter included in the JPEG 2000 codestreams that will follow.

  Senders populate these bits as specified in Table A.10 of Rec. ITU-T T.800 (2015) | ISO/IEC 15444-1:2016.

If the *extended_capability_flag* is set to '1':

- the J2K video stream uses an extended color specification (through three bytes that define the chromaticity parameters, as described below).

- one or more of the following capabilities may be enabled: stripes (through the J2K stripe

mode), blocks (through the J2K block mode), or inclusion of mastering display metadata. (The Sender indicates the exact list of enabled capabilities by use of subsequent flags in the *J2K_video_descriptor*, as described in Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1.)

7. Signaling a VSF TR-01 Stream

The amendment on transport of JPEG 2000 on MPEG-2 TS (Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1) defines a J2K _video _descriptor containing a *profile_and_level* field. The most significant bit of the *profile_and_level* field shall be used as an *extended_ capability_flag:*

- Senders shall set this *extended_capability_flag to* '1' to indicate that the stream uses an extended color specification based on Rec. ITU-T H.273 (2016) | ISO/IEC 23001-8:2016, that J2K stripes or J2K blocks may be enabled, and that Mastering Display Metadata may be included, as described in this document, TR-01:2018.

- Senders shall set the *extended_capability_flag* to '0', to indicate that the emitted stream complies with TR-01:2013.

The remaining 15 least significant bits of the *profile_and_level* field correspond to the 15 least significant bits of the 16-bit RSiz parameter included in the JPEG 2000 codestreams that will follow.

Senders shall populate these 15 least significant bits as specified in Table A.10 of Rec. ITU-T T.800 (2015) | ISO/IEC 15444-1:2016.

Note: By setting the *extended_capability_field* to '1', the 16 bits of the legacy *profile_and_level* field will have a value outside of the range permitted by TR-01:2013, enabling rapid detection of compatibility between the incoming stream and the Receiver's capability. Implementers are cautioned that there is no guarantee of how a legacy TR-01:2013 Receiver will behave if it encounters a TR-01:2018 stream wherein the extended capability flag is set to '1'.

8. Backward and Forward Compatibility Between TR-01:2013 and TR-01:2018

Backwards compatibility:

- TR-01:2018 Receiver processing a TR-01:2013 stream shall properly interpret the TR-01:2013 *profile_and_level* field and shall generate an identical result at its output as a TR-01:2013 Receiver.
- TR-01:2013 streams generated by a TR-01:2018 Sender shall be exactly the same as streams generated by TR-01:2013 Senders.

Forward compatibility:

A TR-01:2013 Receiver that encounters a TR-01:2018 stream should gracefully exit after parsing the *profile_and_level* field and seeing a value out of the range 0x0101-0x04ff (the legal range specified in TR-01:2013).

9. Profile, Format, Frame Rate & Min/Max Bit Rate

Table 1 below contains a listing of defined profiles and corresponding video formats, frame rates and min/max bit rates. A short name is provided for each profile. All formats listed in italic are formats that were originally defined in VSF TR-01:2013.

Note: horizontal pixel values and Rsiz values are specified as part of establishing specific interoperability points.

Senders implementing a particular format and frame rate shown in Table 1 below shall ensure that the codestreams emitted from their devices are within the Min/Max Bit Rate range described.

Profile	Short Name	Format/Frame Rate	Min/Max Bit Rate (Mbps)
1	*SD*	*576i/25* *480i/29.97*	25-200 25-200
2	*HD*	*720p/50* *720p/59.94* *1080i/25* *1080i/29.97* *1080p/23.98** *1080p/24** *1080p/25**	75-200 75-200 75-200 75-200 75-200 75-200 75-200
3	*3G*	*1080p/50* *1080p/59.94* 1080p/100 1080p/119.88 1080p/120	100-400 100-400 200-800 200-800 200-800
4	*S3D-HD*	*720p/50* *720p/59.94* *1080i/25* *1080i/29.97* *1080p/23.98** *1080p/24** *1080p/25**	75-200 each channel, L+R 150-400 75-200 each channel, L+R 150-400 75-200 each channel, L+R 150-400 75-200 each channel, L+R 150-400 75-200 each channel, L+R 150-400 75-200 each channel, L+R 150-400 75-200 each channel, L+R 150-400
5	*S3D-3G*	*1080p/50* *1080p/59.94*	100-400 each, L+R 200-800 100-400 each, L+R 200-800
6	UHD4	2160p/23.94* 2160p/24* 2160p/25* 2160p/29.97* 2160p/30* 2160p/50 2160p/59.94 2160p/100 2160p/119.88 2160p/120	200-800 200-800 200-800 200-800 200-800 400 – 1000 for 1G Interface, 400-1600 for 10G Interface 400 – 1000 for 1G Interface, 400-1600 for 10G Interface 800 – 1000 for 1G Interface, 800-3200 for 10G Interface 800 – 1000 for 1G Interface, 800-3200 for 10G Interface 800 – 1000 for 1G Interface, 800-3200 for 10G Interface

7	UHD8	4320p/23.94*	800 – 1000 for 1G Interfaces, 800-3200 for 10G Interface
		4320p/24*	800 – 1000 for 1G Interfaces, 800-3200 for 10G Interface
		4320p/25*	800 – 1000 for 1G Interfaces, 800-3200 for 10G Interface
		4320p/29.97*	800 – 1000 for 1G Interfaces, 800-3200 for 10G Interface
		4320p/30*	800 – 1000 for 1G Interfaces, 800-3200 for 10G Interface
		4320p/50	1400 - 6400
		4320p/59.94	1400 - 6400
		4320p/100	2400 - 10000
		4320p/119.88	2400 - 10000
		4320p/120	2400 - 10000

Table 1. Video Format, Frame Rate and Bit Rate

Table 1 notes:
Video format is given as active lines, scanning (interlaced or progressive) and frame rate:

- 59.94‖ is equivalent to ―60/1.001
- 29.97‖ is equivalent to ―30/1.001
- 119.88‖ is equivalent to ―120/1.001
- 23.98‖ is equivalent to ―24/1.001

* Optional film frame rates

10. Essence Service Components

This section details of all of the service components including JPEG 2000 video, audio and metadata including specific restrictions for each service component in order to improve inter-operability.

10.1 JPEG 2000 Video

10.1.1 MPEG2 TS and JPEG 2000 Codestream (Informative)

This is an informative summary of the organization of a TR-01:2018 video stream within a MPEG2 TS packet stream, according to Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1. It may be useful to refer to figure 2 below.

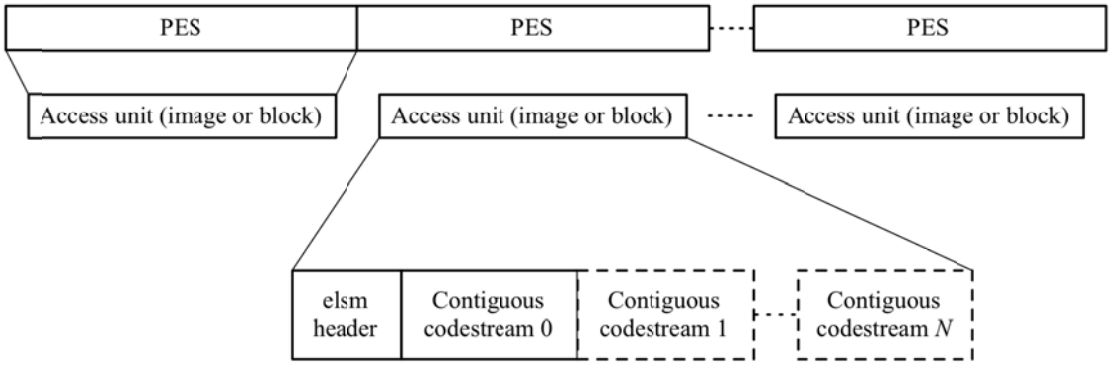

Figure 2: Structure and order of J2K Video Access Units

To transport JPEG 2000 streams on MPEG2 TS, the following signaling elements and headers are required:

- A *J2K_video_descriptor* that includes the description of a *J2K video elementary stream*. This descriptor is included for each *J2K video elementary stream* component in the PMT (Program Map Table) with *stream_type* equal to 0x21, as described in Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1.

  Note: at this layer in MPEG2TS, there are descriptions of other components such as audio and ancillary data.

- A *J2K video elementary stream* consisting of a succession of J2K Video Access Units, each of them embedded in a Packetized Elementary Stream (PES). Each such access unit includes a J2K Elementary Stream Header (*J2K_elsm_header*) followed by one or more J2K Codestreams (containing, for example, Image, Block, Stripes, 3D L, or 3D R) as illustrated in figure 2 above. The number of J2K Codestreams '*N*' is given in Table 2 below.

Content	J2K stripe mode	*N*
Progressive	Disabled	1
Interlaced	Disabled	2
Progressive	Enabled	*strp_max_idx* + 1
Interlaced	Enabled	2*(*strp_max_idx*+1)

Table 2: "*N*", Number of contiguous J2K codestreams in a single J2K Video Access Unit

Note: The *strp_max_idx* is a field present in the *J2K_video_descriptor* and in each *J2K_elsm_header* (when J2K stripe mode is enabled). *strp_max_idx* corresponds to the total number of stripes in the frame or block, minus one.

10.1.2 JPEG 2000 Codestream Restrictions

Senders shall comply with the "*Broadcast Contribution Single Tile Profile*" as specified in Rec. ITU-T T.800 | ISO/IEC 15444-1.

Each single frame, field, block or stripe shall be encoded as a single tile.

Senders shall not use the "*Broadcast Contribution Multi-Tile Profile*" or the "*Broadcast Contribution Multi-Tile Reversible Profile*".

Receivers shall be able to decode compressed JPEG 2000 codestreams if those codestreams comply with the "Broadcast Contribution Single Tile Profile" as specified in Rec. ITU-T T.800 | ISO/IEC 15444-1.

Senders shall set the values of SIZ marker segment in the JPEG 2000 codestream according to table 3A, table 3B, and table 3C. Note: table entries in italics represent values taken from VSF TR-01:2013.

Color components shall be transported as one codestream (Note: this is the case where Csiz=3).

For YCbCr representations, Senders shall use the following order of components in the JPEG 2000 codestream:

- Y (component index shall be set to 0)
- Cb (component index shall be set to 1)
- Cr (component index shall be set to 2)

As shown in table 3A below, Senders shall set the $XRsiz^i$ parameters according to the sub-sampling that is being used, depending upon the number of components.

Allowed Configurations	$XRsiz^i$ Parameter Value
4:2:2 Subsampling	($XRsiz^0$=1, $XRsiz^1$=2, $XRsiz^2$=2)
4:2:2:4 Subsampling	($XRsiz^0$=1, $XRsiz^1$=2, $XRsiz^2$=2, $XRsiz^3$=1)
4:4:4 Subsampling	($XRsiz^0$=1, $XRsiz^1$=1, $XRsiz^2$=1)
4:4:4:4 Subsampling	($XRsiz^0$=1, $XRsiz^1$=1, $XRsiz^2$=1, $XRsiz^3$=1)

Table 3A. $XRsiz^i$ Parameter For Component Sub-sampling

For the allowed configurations described in Table 3A, $YRsiz^i$ shall be set to '1' for all values of *i*.

All components shall have the same bit depth.

Bit depth, as specified by the $Ssiz^i$ value shall be set to either of the values in table 3B below.

Allowed Configurations	$Ssiz^i$ Parameter Value
10-bit	($Ssiz^i$=9, i=0,1,2,3)
12-bit	($Ssiz^i$=11, i=0,1,2,3)

Table 3B. $Ssiz^i$ Parameter For Component Bit Depth

If there is an Alpha channel, it shall either be transported in its own codestream, or the Alpha channel shall be transported in the same codestream as the color components.

Csiz shall be set according to Table 3C below.

Allowed Configurations	Csiz Parameter Value
Three color components	*Csiz = 3*
Alpha Channel only	Csiz = 1
Three Color Components and Alpha Channel	Csiz = 4

Table 3C. Csiz Parameter For Number of Color Components

Senders shall use a single *codeblock size* value for the image.

Senders and Receivers shall support a codeblock size of 32x32 (xcb=ycb=5)

Note: During the drafting of this document (2017), a survey was conducted regarding actual deployments of TR-01:2013. No implementer had chosen a codeblock size other than 32 x 32.

Senders shall observe the following constraints regarding marker segments:

- The TLM marker segments shall be used as described in Rec ITU-T T.800 | ISO IEC 15444-1

- The COC marker segments shall not be used

Note: Because of other constraints, the COC marker segments do not add any additional relevant information than what is already provided by the COD marker segments.

- The QCC marker segment(s) may be used (Note: utilizing different quantization parameters for the different components can improve rate allocation.)

- The PLM, PLT, SOP and EPH markers and marker segments shall not be present

Senders supporting stereoscopic operation shall comply with SMPTE ST 2063, and shall ensure that the sets of JPEG 2000 encoding parameters used for the stereoscopic image pairs in Profiles "S3D-HD" and " S3D-3G" are strictly identical.

When *J2K stripe mode* is enabled, the number of decomposition levels N_L shall be constrained by the stripe height h_s as follows:

- $h_s <= 256 : 1 <= N_L <= 3$
- $h_s <= 1024: 1 <= N_L <= 4$
- $h_s <= 2048: 1 <= N_L <= 5$
- $h_s <= 4096: 1 <= N_L <= 6$
- $h_s <= 8192: 1 <= N_L <= 7$

10.1.3 JPEG 2000 Video descriptor

Senders shall construct the *J2K video_descriptor* as specified in Clause 2.6.80 of Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1.

10.1.4 JPEG 2000 Block Mode Coding

Note: Block mode allows implementers to split a video image into multiple video blocks by dividing the video horizontally and vertically in the spatial domain. The main use case for this block mode is to encode different quadrants or blocks of a video image separately in UHD applications and transport them as a single J2K Transport stream. As an example, a 4Kp/59.94 video can be divided into four 1080p/59.94 video blocks.

For Profiles 1 through 5 shown in table 1 above, senders shall not use block mode coding.

Block mode coding is optional. If being used, it shall be enabled in the *J2K_video_descriptor*.

For Profiles 6 and 7 in table 1 above, senders shall select from one of the permitted values for block size shown in table 4 below.

Profile	Short Name	Block Size
1	*SD*	N/A
2	*HD*	N/A

3	*3G*	N/A
4	*S3D-HD*	N/A
5	*S3D-3G*	N/A
6	UHD4	UHD4 2x2, UHD4 2x1 Note: block size for 2x2 is1920x1080 and block size for 2x1 is 1920x2160
7	UHD8	UHD8 4x4, UHD8 4x1 Note: block size for 4x4 is 1920x1080 and block size for 4x1 is 1920x4320

Table 4 – Block sizes for UHD4 and UHD8

Senders shall encode each stream of video blocks into its own JPEG 2000 packetized elementary stream (PES) and each of these Packetized Elementary Streams shall be assigned a unique Program Identifier (PID) within a given Transport Stream.

All video blocks from a given stream shall be placed into the same MPEG2 Transport Stream.

There shall be no requirement for unique PIDs across different Transport Streams.

The JPEG 2000 block coding shall be implemented as specified in Section S.3 of Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1: "Ultra-Low Latency and 4k and higher resolution support for transport of JPEG 2000 video".

When block mode is enabled, Senders shall set the *extended_capability_flag* and the *block_flag* in the *J2K_video_descriptor* to '1', and they shall ensure the following fields are present and appropriately set in the descriptor: *full_horizontal_size, full_vertical_size, blk_width, blk_height, max_blk_idx_h, max_blk_idx_v, blk_idx_h, blk_idx_v*.

When block mode is disabled, Senders shall set the *block_flag* in the *J2K_video_descriptor* to '0', and they shall not include any of the block-related fields in the descriptor.

10.1.5 J2K Stripe Mode Coding

Note: TR01:2018 supports use cases where Ultra Low Latency (Ultra Low Latency means end to end transport latency of under one video frame) is required. This is accomplished by means of a striping mechanism. To use this mechanism, a Sender shall divide each picture (each single

frame, field or block) horizontally into an integer number of stripes as shown in table 5 below. Depending on the complexity of the video stream content to be encoded, when using the striping mechanism described in this document, implementers may find that an increase in configured target transport bit-rate is required (as compared to not using striping) in order to ensure sufficient and consistent video quality, particularly at stripe boundaries.

Stripe Mode shall be optional.

If Stripe Mode Coding is used, it shall be enabled in the *J2K_video_descriptor*.

If J2K Block Mode is enabled, the value of the *strp_height* field located in the *J2K_video_descriptor* and in each Elementary Stream header shall be the same across all J2K Blocks.

When using Stripe Mode, a Sender shall divide each picture (each single frame, field or block) horizontally into an integer number of stripes as shown in table 5 below.

Profile	Short Name	Video Format	Number of Stripes	*strp_max_idx* Value
1	*SD*	All	Entire Frame	Not Allowed
2	*HD*	*720p/50*	2-6	*1-5*
		720p/59.94	2-6	*1-5*
		1080i/25	2-4	*1-3*
		1080i/29.97	2-4	*1-3*
		*1080p/23.98**	4-9	*3-8*
		*1080p/24**	4-9	*3-8*
		*1080p/25**	4-9	*3-8*
3	*3G*	*1080p/50*	4-9	*3-8*
		1080p/59.94	4-9	*3-8*
		1080p/100	4-9	*3-8*
		1080p/119.88	4-9	*3-8*
		1080p/120	4-9	*3-8*
4	*S3D-HD*	All	Entire Frame	Not Allowed
5	*S3D-3G*	All	Entire Frame	Not Allowed
6	UHD4	2160p/23.94*	2-16	1-15
		2160p/24*	2-16	1-15
		2160p/25*	2-16	1-15
		2160p/29.97*	2-16	1-15
		2160p/30*	2-16	1-15
		2160p/50	2-16	1-15
		2160p/59.94	2-16	1-15
		2160p/100	2-16	1-15
		2160p/119.88	2-16	1-15
		2160p/120	2-16	1-15
7	UHD8	4320p/23.94*	2-32	1-31
		4320p/24*	2-32	1-31
		4320p/25*	2-32	1-31
		4320p/29.97*	2-32	1-31
		4320p/30*	2-32	1-31
		4320p/50	2-32	1-31
		4320p/59.94	2-32	1-31
		4320p/100	2-32	1-31
		4320p/119.88	2-32	1-31
		4320p/120	2-32	1-31

Table 5. ULL Striping

Table 5 Notes:
Video format is given as active lines, scanning (interlaced or progressive) and frame rate:

- 59.94‖ is equivalent to ―60/1.001
- 29.97‖ is equivalent to ―30/1.001
- 119.88‖ is equivalent to ―120/1.001
- 23.98‖ is equivalent to ―24/1.001

JPEG 2000 stripe coding shall be implemented as specified in Section S.4 of Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1: "Ultra-Low Latency and 4k and higher resolution support for transport of JPEG 2000 video".

Note: For video formats such as 720p, 1080p and 2160p, there are certain stripe configurations that give equal sized stripes, and other stripe configurations in which the last (bottom) stripe size is different from the others. This document allows the implementer flexibility in the choice of stripe configurations to allow for various system latencies.

When JPEG 2000 stripe mode is enabled, Senders shall set the *extended_capability_flag* and *stripe_flag* in the *J2K_video_descriptor* to '1', and they shall ensure the following fields are present and appropriately set in the descriptor:

- *strp_max_idx*
- *strp_height*

In the *J2K_elsm_header*, Senders shall ensure that the Stripe Coding box (*j2k_strp*) is present, and they shall ensure the Time Coding box (*j2k_tcod*) is not present.

Senders shall transmit stripes in raster order, from top to bottom.

In the Stripe Coding box, Senders shall ensure that the parameters *strp_max_idx* and *frame_vertical_size* are set appropriately, as defined in Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1.

Stripes shall never have a height lower than 120, even for the last one.

When JPEG 2000 stripe mode is disabled, Senders shall set the *stripe_flag* in the *J2K_video_descriptor* to '0', and they shall not include any of the stripe-related fields in the descriptor.

Senders shall ensure that the Stripe Coding box (*j2k_strp*) is not present in the *J2K_elsm_header*, and they shall ensure that the Time Coding box (*j2k_tcod*) is present.

10.1.6 JPEG 2000 Elementary Stream

Senders shall construct JPEG 2000 video elementary streams and encapsulate those streams in an MPEG-2 Transport Stream according to Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1.

Senders shall set the following parameters as indicated:

- *stream_id* = private_stream_1
- *stream_type* = '0x21'

Senders shall include the *J2K_video_descriptor* in the Program Map Table (PMT) for the JPEG 2000 service component.

Senders shall ensure that the J2K Video Access Unit contains the correct number of contiguous JPEG 2000 codestreams, as given in table 2.

Note: This number depends on the kind of content transported (progressive or interlaced) and the use of J2K stripe mode (see above).

10.1.6.1 JPEG 2000 Elementary Stream Header

Senders shall construct the *J2K_elsm_header* as specified in table S.1 of Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1, including the order of the elementary stream header boxes.

Note: Because some syntax elements of the *J2K_elsm_header* depend on the values of flags in the video descriptor, a Receiver is advised to parse the *J2K_video_descriptor* (located in the PMT of the Transport Stream) in order to determine the value of *extended_capability_flag*, *interlaced_video* and *stripe_flag* prior to parsing the *J2K_elsm_header*.

10.1.6.2 Field Coding

For interlaced video signals, Senders shall set the *interlaced_video* flag in the *J2K_video_descriptor* to '1'.

Senders shall ensure that the Field Coding box is present in the *J2K_elsm_header*.

Senders shall transmit fields in temporal order.

Senders shall ensure that the codestream corresponding to the field with the top-most line is located first in the *J2K Video Access Unit*.

In the Field Coding box, Senders shall ensure that the parameters below are set to the values shown:

- Fic = 2
- Fio = 1

If J2K stripe mode is enabled, Senders shall ensure that for a given stripe, the second field is immediately following the first one.

Example: if there are two stripes, the following order shall be used to build the J2K Video Access Unit: "[stripe_0_field_0] [stripe_1_field_0]… [stripe_N-1_field_0] [stripe_0_field_1] [stripe_1_field_1]… [stripe_N-1_field_1].

For progressive video signals, Senders shall set the interlaced_video flag in the *J2K_video_descriptor* to '0', and they shall ensure that the Field Coding box is not present in the *J2K_elsm_header*.

10.1.6.3 Maximum Bitrate and Codestream size

In accordance with Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1, the *interlaced_video* flag and the *stripe_flag* in the *J2K_video_descriptor* shall be used to determine

the values for the *brat_auf1* and *brat_auf2* fields.

Senders shall set the *brat_auf1* and *brat_auf2* according to the values indicated in table 6 below.

Content	stripe mode	brat_auf1	brat_auf2
Interlaced	Enabled	All zero	All zero
Interlaced	Disabled	Size of 1st field codestream	Size of 2nd field codestream
Progressive	Enabled	All zero	Field not present
Progressive	Disabled	Size of frame codestream	Field not present

Table 6: *brat_auf1* and *brat_auf2* field values

Note: In the case where stripe mode is enabled, a Receiver can use the values indicated in the Maximum Bitrate field (brat_max_br) to determine the size of the JPEG 2000 codestream(s). Alternatively, a receiver can use the tile-part length as indicated in the TLM marker segment, or the value of the Psot parameter in the SOT marker segment in the JPEG 2000 main header (note that the latter may be zero).

Note: In the case where stripe mode is disabled and AUF fields are required in the *J2K_elsm_header*, implementers of this Technical Recommendation are advised that, since the exact codestream sizes for both fields of an interlaced frame (AUF_1 and AUF_2) needs to be known prior to transmission of a J2K Video Access Unit, an additional field period of buffering latency can be expected on the encoder side in addition to the actual encoding latency.

10.1.6.4 Frame rate

The frame rate of the JPEG 2000 video ES is signaled using the NUM and DEN parameters of the Frame Rate box located in the *J2K_elsm_header*, as well as in the *J2K_video_descriptor*.

Senders shall signal the frame rate of the JPEG 2000 video ES using the values shown in table 7 below.

Frame rate	Interlaced (I) or Progressive (P)	Numerator NUM	Denominator DEN
24/1.001	P	24 000	1001
24	P	24	1
25	I	25	1
30/1.001	I	30 000	1001
50	P	50	1
60/1.001	P	60 000	1001
100	P	100	1
120/1001	P	120 000	1001

Table 7 – Signaling of Supported Frame Rates

10.1.7 Color specification

Depending on the value of the *extended_capbility_flag*, the color space shall be indicated using one of the two methods described below.

Method 1. When the *extended_capability_flag* is set to '0' (i.e. for TR01:2013 stream), Senders

shall signal the Broadcast Color Specification Code in the *J2K_video_descriptor* and in the *Broadcast_Color_Box* in the *J2K_elsm_header* (*bcol_colcr* field) using the values shown in table 8 below.

Method 2. When the *extended_capability_flag* is set to '1' (i.e. for a TR-01:2018 stream), the color information shall be specified using four fields (three 8-bit fields and one 1-bit field), namely *color_primaries*, *transfer_characteristics*, *matrix_coefficients*, and *video_full_range_flag*, as described in Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1. These fields shall be coded according to the semantics with the same name defined in Rec. ITU-T H.273 | ISO/IEC 23001-8.

		Method 1	**Method 2**
Profile	**Description**	**Value**	
		extended_capability_flag = 0	*extended_capability_flag* = 1
1	*SD*	*0x02*	See table 9 below referencing (ITU-T H.273 | ISO/IEC 23001-8
2	*HD*	0x03	
3	*3G*	0x03	
4	*S3D-HD*	0x03	
5	*S3D-3G*	0x03	
6	UHD4	N/A	
7	UHD8	N/A	

Table 8. Signaling of supported Color Space specifications

Table 8 Notes:
- The value "0x02" indicates the use of Rec. ITU-R BT 601-7
- The value "0x03" indicates the use of Rec. ITU-R BT 709-6

10.1.7.1 Example Color Space Table (Informative)

Color space	Color primaries code	Transfer characteristics code	Matrix coefficients code	Video full range flag
Rec. ITU-R BT.601-7 625	5	6	5	0
Rec. ITU-R BT.601-7 525	6	6	6	0
Rec. ITU-R BT.709-6	1	1	1	0
Rec. ITU-R BT.2020-2	9	14 (10bit)	9 (non-constant luminance)	0
Rec. ITU-R BT.2020-2	9	15 (12bit)	9 (non-constant luminance)	0
Rec. ITU-R BT.2020-2	9	14 (10bit)	10 (constant luminance)	0
Rec. ITU-R BT.2020-2	9	15 (12bit)	10 (constant luminance)	0
Rec. ITU-R BT.2100-1	9	16	9 (Y'CbCr)	0
Rec. ITU-R BT.2100-1	9	16	14 (ICtCp)	0

Table 9 (Informative) – Selected examples of extended color space specification

Table 9 above provides example signaling code values for Table 8, when the *extended_capability_flag* is set to '1'. This illustrates extended color specification VSF TR-01:2018 compliant signaling. These code values are taken from Rec. ITU-T H.273 | ISO/IEC 23001-8. VSF TR-01:2013 did not make use of this extension.

10.1.8 Mastering Display Metadata

Optionally, Senders have the opportunity to specify the Mastering Display Metadata in the *J2K_video_descriptor*. To do so, Senders shall set both the *extended_capability_flag* and the *mdm_flag* to '1', as described in Rec. ITU-T H.222.0 (2017) | ISO/IEC 13818-1:2017 AMD1.

Mastering Display Metadata shall then be specified using the following fields:

- *X_c0, Y_c0, X_c1, Y_c1, X_c2, Y_c2, X_wp, Y_wp, L_max* and *L_min*, defined in SMPTE ST2086:2014 "Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images"

- *MaxFALL* and *MaxCLL*, defined in ANSI/CTA 861-G:2016 "A DTV Profile for Uncompressed High Speed Digital Interfaces"

If these fields have unknown values at the time the stream is generated, they shall not be included in the descriptor and the *mdm_flag* shall be set to '0'.

10.1.9 JPEG 2000 Still Pictures

Senders shall ensure that the JPEG 2000 video ES does not contain any JPEG 2000 still pictures.

Senders shall set the still_mode field in the *J2K_video_descriptor* to '0'.

10.1.10 Stereoscopic 3D

Senders shall format Stereoscopic 3D image pairs for transport in compliance with SMPTE ST 2063.

10.2 Audio

Senders shall ensure that audio signals shall be sampled at a rate of 48 kHz, and that the sample clock shall be synchronous to the video pixel clock.

Senders shall ensure that audio signals are formatted in accordance with AES3: AES3:2009.

AES3 signals shall consist of PCM audio samples or of non-PCM compressed audio signals.

10.2.1 Audio Transport (PCM and Non-PCM signals)

SMPTE ST 302 shall be used for the transport of all audio signals.

Senders and Receivers that are compliant with TR-01:2013 shall support Row 1 of table 10 below.

In addition, Senders and Receivers that are compliant with TR-01:2018 shall support Profiles one through five in Table 10 below.

Receivers under this Technical Recommendation shall be capable of simultaneously receiving and processing the number of AES3 channel pairs shown in Table 10 below.

Profile	Short Name	Bit Depth and AES Channel Pairs
1	*SD*	Bit depth is 20 or 24 bits Channel Pairs 1 - 4 1 - 4 channel pair per PID
2	*HD*	Bit depth is 20 or 24 bits Channel Pairs 1 - 8 1 - 4 channel pair PID
3	*3G*	Bit depth is 20 or 24 bits Channel Pairs 1 - 16 1 - 4 channel pair PID
4	*S3D-HD*	Bit depth is 20 or 24 bits Channel Pairs 1 - 8 1 - 4 channel pair per PID

5	*S3D-3G*	Bit depth is 20 or 24 bits Channel Pairs 1 - 16 1 - 4 channel pair per PID
6	UHD4	Bit depth is 24 bits Channel Pairs 1 - 16 1 - 4 channel pair per PID
7	UHD8	Bit depth is 24 bits Channel Pairs 1 - 16 1 - 4 channel pair per PID

Table 10: Supported PCM Audio Profiles

Receivers should incorporate a selection mechanism that allows the user to choose which audio signals (channel pair) to process, from amongst those sent.

Senders should assign ascending MPEG-2 transport stream PID values to ST 302 audio elementary streams such that the first audio stream has the lowest PID and the last audio stream has the highest PID.

Unless overridden by the user, the Receiver should receive the first audio stream from the lowest value audio PID and last audio stream from the highest PID.

Senders shall mark each audio stream with MPEG-2 Presentation Time Stamps (PTS) corresponding to video frames in the source video as required by SMTPE ST 302.

Upon receipt and processing of multiple SMPTE ST 302 audio streams, Receivers shall synchronize the audio streams such that, upon presentation, audio samples from J2K Video Access Units containing the same PTS value shall be emitted synchronously (phase aligned) with each other.

10.2.2 AES3 as Transport of Linear PCM (Uncompressed) Audio (Informative)

Note: The number of 48 kHz audio samples corresponding to a given frame of video is not an integer in some frame rates. Therefore the size of the PES packets for carrying said audio samples will vary.

Senders and Receivers shall support 48 kHz audio at all supported video frame rates, without altering the number of audio samples per frame.

10.2.3 AES3 as Transport of Non-PCM Data

Senders and Receivers shall support the transparent transport of non-PCM compressed audio signals as audio elementary streams in compliance with AES3 SMPTE ST 302 and SMPTE ST 337.

Senders and Receivers may support other non-PCM digital signals as defined in SMPTE ST 338.

Note: SMPTE ST 337 documents a common use-case where compressed audio bitstreams are packaged into an AES3 signal for transport. Senders and Receivers need to take care to ensure that non-PCM audio signals identified in SMPTE 337 and ST 338 are transported in such a

manner as to not disrupt their contents.

Sample-Rate-Conversion operations shall not be performed on non-PCM audio signals.

Note: Implementers are cautioned that SMPTE ST 337 and 338 are updated frequently.

10.2.4 Maintenance of A/V Sync

Receivers shall use the PTS values in both audio and video PES to extract each signals time base in order to maintain A/V synchronization within +/- 2 ms.

Note: When synchronizing the output video to a local (GenLock) time base, A/V synchronization error of +- ½ video frame time can be expected. Implementers are referred to Rec. ITU-R BT.1359-1 "Relative Timing of Sound and Vision for Broadcasting", as well as CEA-CEB20 "A/V Synchronization Processing Recommended Practice" for additional guidance on this topic. For the specific case of Dolby E, implementers are referred to SMPTE RDD 19.

10.3 Ancillary Data

Most SDI signals (both SD and HD) include Horizontal Ancillary (HANC) and Vertical Ancillary (VANC) data packets formatted in accordance with SMPTE ST 291-1.

Subject to the exceptions and limits noted below, Senders and Receivers shall support the transport of HANC and VANC data using the method specified in SMPTE ST 2038.

Note: Implementers might want to transport some data commonly found in HANC and VANC such as teletext or timecode using alternative Standardized transport methods available in MPEG-2 TS or J2K video PES. Senders may make use of these alternative methods, although support for SMPTE ST 2038 is required. Receivers are not required to support any transport method other than SMPTE ST 2038.

If an Implementer chooses to provide an alternative transport method for ancillary data, Senders shall only use one method at a time.

Senders shall not emit ancillary data using both SMPTE ST 2038 and through the alternative method simultaneously.

Note: Implementers may choose to provide controls that permit the user to choose between ST 2038 and an alternative transport method.

Note: VANC and HANC packets can appear on any line, and the method specified for transport herein allows for them to be placed back into the line and space from whence they came.

10.3.1 HANC and VANC data which are excluded from transport

Although embedded audio is formatted as HANC data, Senders shall use the method identified in section 9.2 for the transport of all audio signals.

Senders shall not use the methods in this section for audio.

The Embedded Audio Control Packet defined in SMPTE ST 299-1 should not be transmitted by Senders, and should be ignored by Receivers.

Receivers shall generate a locally correct Embedded Audio Control Packet based on their specific configuration.

EDH, CRC, and Line Number information, while present in the ancillary data spaces, is not formatted as ANC packets under SMPTE ST 291-1 and therefore Senders shall not send this data.

10.3.2 Limits on the total amount of ANC data to be transported

Note: If all of the available ancillary data spaces are packed full of well-formed packets, a very significant amount of transport bandwidth might be required to transport all of the data. In practice, while the use of ANC data to carry information is expected to increase during the lifetime of this TR, this document places practical limits on the amount of data to be transported in order to foster interoperability in balance with reasonable implementation.

For each of the profiles defined, table 11 below indicates the MINIMUM number of 10-bit words of ANC data, carried using SMPTE ST 2038, which shall be supported by compliant Senders and Receivers.

Senders shall create streams that meet these minimum bit rates, and Receivers shall be able to output ANC data streams that meet these bit rates.

Senders and Receivers that are compliant with this TR may support the transport of higher amounts of ANC data than that specified in table 11 below.

Profiles	Description	Number of 10-bit words/sec
1-7	*SD, HD, 3G, S3D-HD, S3D-3G, UHD4, UHD8*	
	Number of 10-bit words of ANC Data to be transported per second	104800
	Worst-Case PES bit rate for ANC Data (bits/field)	26106
	ES Buffer Size (Bn) (bits)	4 * 26106
	Transport Maximum bit rate (Rmax) (bits/second)	2,500,000

Table 11 – Amount of ANC Data Transported Using ST 2038 to be Supported by the Different Profiles

Note: In table 11, 104800 10-bit words per second is equivalent to 8 ANC data packets per frame, each having 7 header words and 255 User Data Words, at 50 frames per second. Since the size of ANC data packets is variable, more than 8 packets can, of course, be transported.

Note: Users should be aware that ST 2038 AND data is encapsulated in TS packets (188 bytes). The resulting TS bit rate can be substantially higher than the underlying ANC data rate. For example, a 2-byte payload of CEA 608 data becomes 188 bytes in the TS layer (ignoring the header overhead).

The ANC data transmitted by Senders shall be compliant with the T-STD model as specified in ISO/IEC 13818-1, using an elementary stream buffer size (Bn) as articulated in table 11. The transport buffer TBn for the ANC service is specified to be 512 bytes in ISO/IEC 13818-1.

For ST 2038 ANC data, the transport buffer shall be drained (Rxn) at 1.2 times [Rmax] as per table 11.

Note: the PES bit rate calculation in table 11 assumes 20 lines of ANC data per field (frame in progressive formats), and 10 separate ANC data packets per each line for the purpose of the ST 2038 overhead estimation (14 bytes PES header, plus 4 bytes per each packet header, 10 packets per line). A worst-case field rate of 60 fields per second is used for this calculation.

Note: The Transport Maximum bit rate calculation assumes the worst-case PES rate, and adds TS header overhead, plus an average 92 bytes stuffing per PES for PES alignment as indicated in ST 2038. The calculated value of 2449560 is rounded up to 2500000 for the purposes of this TR.

10.3.3 Prioritization of certain ANC signals (Informative)

Senders might want to provide a mechanism for filtering the ancillary data types that should be sent. In case of over-subscription, Senders might also wish to provide prioritization for the ANC data types being sent. In no particular order, examples of important ANC data types include:

- Closed Captions (SMPTE ST 334-1)
- Time Code (SMPTE ST 12-2)
- AFD/Bar data (SMPTE ST 2016-3)
- Audio metadata (SMPTE ST 2020-2, SMPTE ST 2020-3)
- ANSI/SCTE 104 Messages (SMPTE ST 2010)
- DVB/SCTE VBI data (SMPTE ST 2031, OP-47, SMPTE RDD 8)

10.4 Vertical Blanking Interval Data for Standard-Definition Signals

Note: While both High-Definition and Standard-Definition systems routinely carry associated information as ANC data packets, Standard-Definition (SD) digital signals may carry important digital data representations of analogue waveforms. These waveforms are bandwidth-limited to permit a compatible analog conversion, and are located within the Vertical Blanking Interval. Transport of these VBI signals is expected as part of the use case for transporting SD signals and is described in this section.

Certain VBI signals have well-known transport mappings which are required under this standard; in addition a pass-through mechanism is defined for transport of arbitrary waveform VBI data.

10.4.1 VBI signals with specific transport mappings

Senders shall support the transport of VBI data using the method specified in ETSI EN 301 775 and ANSI/SCTE 127 subject to the exceptions and limits noted below.

Receivers shall recreate VBI data in accordance with the ETSI standard as well.

Senders and Receivers may transport VBI data using other methods specified in standards for MPEG-2 TS or J2K video PES transport.

Implementers may provide controls to permit the user to choose whether an alternate transport method is used or not.

Senders shall support the transport of VBI data formatted in accordance with CEA-608B (often found on line 21 of field 1 and field 2 in 525-line systems) using the method described in ETSI EN 301 775.

Receivers shall output VBI data which has been transported in accordance with the method described in the ETSI standard.

Senders shall support the transport of VBI data formatted as teletext (including subtitles in teletext) using the method described in ETSI EN 301 775.

Receivers shall output VBI data formatted as teletext (including subtitles in teletext) which has been transported in accordance with the method described in the ETSI standard.

Senders shall support the transport of Wide Screen Signaling (WSS) data in VBI data using the method described in ETSI EN 301 775.

Receivers shall output WSS data that has been transported in accordance with the method described in the ETSI standard.

10.4.2 Arbitrary sample-based transport mapping

Note: In addition to the well-known signals described above, there are additional waveforms defined in ETSI EN 301 755 and in extensions defined in SCTE 127, as well as proprietary or unknown waveforms in the VBI area that might need to be transported. Transporting the entire VBI region using the method below could be onerous in terms of bit rate. However, transport service providers in coordination with their customers might choose to select some lines of the VBI to transport using the sample-based method below.

Senders shall support the user selectable transport of information represented in VBI data utilizing the "Data field for monochrome 4:2:2 samples" method described in ETSI EN 301 775 section 4.9.

Receivers shall output the information in accordance with the method described in the ETSI standard.

10.4.3 Limits on the transport of VBI signals

The following limits are applied to the transport of VBI signals.

Senders shall be required to format for transport, and Receivers shall be required to output, up to and including:

- two lines per field of CEA-608 data

- 12 lines per field of Teletext (including subtitles in teletext)
- 1 line per field of Wide-Screen Signaling (WSS)
- 2 lines per field of Monochrome 4:2:2 samples

11. Interoperability Points

This document establishes the interoperability points shown in table 12, part 1 and table 12, part 2 below. Senders and Receivers that claim compliance with an interoperability point shall comply with all of the technical parameters specified in table 12, part 1 and table 12, part 2 below for that interoperability point (table row number).

Note: Implementers are cautioned that the values in the table below are only to establish interoperability between vendors. Please refer to Section 9, "*Profile, Format, Frame Rate and Minimum/Maximum Bit Rate*" for actual rate ranges permitted by this TR.

Note: Additional interoperability points may be established by other documents including future VSF Technical Recommendations.

Interop Point	Profile/Short Name	extended_capability_flag	Format & Frame Rate*	Bit Rate **	Capability Rsiz parameter***	Bit Depth	Color Sampling	Color Space
1	SD	0	480i/29	50Mbps	0x0101	10bit	4:2:2	BT 601-6
2	SD	0	576i/25	50Mbps	0x0101	10bit	4:2:2	BT 601-6
3	HD	0	720px1280/59	75Mbps	0x0102	10bit	4:2:2	BT-709
4	HD	0	720px1280/50	75Mbps	0x0102	10bit	4:2:2	BT-709
5	HD	0	1080ix1920/29	75Mbps	0x0102	10bit	4:2:2	BT-709
6	HD	0	1080ix1920/25	75Mbps	0x0102	10bit	4:2:2	BT-709
7	3G	0	1080px1920/59	200Mbps	0x0104	10bit	4:2:2	BT-709
8	3G	0	1080px1920/50	200Mbps	0x0104	10bit	4:2:2	BT-709
9	S3D-HD	0	720px1280/59	75Mbps Each (L,R)	0x0102	10bit	4:2:2	BT-709
10	S3D-HD	0	720px1280/50	75Mbps Each (L,R)	0x0102	10bit	4:2:2	BT-709
11	S3D-HD	0	1080ix1920/29	75Mbps Each (L,R)	0x0102	10bit	4:2:2	BT-709
12	S3D-HD	0	1080ix1920/25	75Mbps Each (L,R)	0x0102	10bit	4:2:2	BT-709
13	S3D-3G	0	1080px1920/59	200Mbps Each (L,R)	0x0104	10bit	4:2:2	BT-709
14	S3D-3G	0	1080px1920/50	200Mbps Each (L,R)	0x0104	10bit	4:2:2	BT-709
15	HD	1	720px1280/59	125Mbps	0x0102	10bit	4:2:2	BT-709
16	HD	1	720px1280/50	125Mbps	0x0102	10bit	4:2:2	BT-709
17	HD	1	1080ix1920/29	125Mbps	0x0102	10bit	4:2:2	BT-709
18	HD	1	1080ix1920/25	125Mbps	0x0102	10bit	4:2:2	BT-709
19	3G	1	1080px1920/59	200Mbps	0x0104	10bit	4:2:2	BT-709
20	3G	1	1080px1920/50	200Mbps	0x0104	10bit	4:2:2	BT-709
21	UHD4	1	2160px3840/59	800Mbps	0x0106	10bit	4:2:2	BT-709
22	UHD4	1	2160px3840/50	800Mbps	0x0106	10bit	4:2:2	BT-709
23	UHD4	1	2160px3840/59	800Mbps	0x0106	10bit	4:2:2	BT-709
24	UHD4	1	2160px3840/50	800Mbps	0x0106	10bit	4:2:2	BT-709
25	UHD4	1	2160px3840/59	800Mbps	0x0106	10bit	4:2:2	BT-709
26	UHD4	1	2160px3840/50	800Mbps	0x0106	10bit	4:2:2	BT-709
27	UHD4	1	2160px3840/59	800Mbps	0x0106	10bit	4:2:2	BT-709
28	UHD4	1	2160px3840/50	800Mbps	0x0106	10bit	4:2:2	BT-709

Table 12, Part 1: Interoperability Points

Interop Point	Striping	Blocking	Code Block Tx	Code Block Rx	Audio	Audio	Audio ****	B-B Target Latency	Carry over from TR-01 2013
1	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
2	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
3	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
4	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
5	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
6	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
7	None	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID		Just audio change
8	None	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID		Just audio change
9	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
10	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
11	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
12	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
13	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
14	None	None	32 x 32	32 x 32	20- bit	1 ChPr	1 PID		Yes
15	6	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	6-8ms	ULL Application
16	6	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	7-9ms	ULL Application
17	5	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	20-27ms	ULL Application
18	5	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	24-32ms	ULL Application
19	9	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	6-8ms	ULL Application
20	9	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	7-9ms	ULL Application
21	None	4 x 1080 Sq. Div.	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	3-4 frames	UHD Application
22	None	4 x 1080 Sq. Div.	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	3-4 frames	UHD Application
23	None	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	3-4 frames	UHD Application
24	None	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	3-4 frames	UHD Application
25	9	4 x 1080 Sq. Div.	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	6-8ms	ULL UHD Application
26	9	4 x 1080 Sq. Div.	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	7-9ms	ULL UHD Application
27	9	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	6-8ms	ULL UHD Application
28	9	None	32 x 32	32 x 32	24- bit	1 ChPr	1 PID	7-9ms	ULL UHD Application

Table 12 Part 2: Interoperability Points

Table 12 notes:

* Video format is given as active lines, scanning (interlaced or progressive) and frame rate (- 59.94 is equivalent to 60/1.001, while 29.97 is equivalent to 30/1.001)

** Bit rate to use in interop compatibility test (not as a recommended use case). Refer to TR-01:2018 for actual rate range

*** Rsiz values are calculated based on the "format & frame rate", "bitrate", and "color sampling" columns.

**** Audio to be analyzed on a single channel pair, however multi-channel support up to maximum in TR01:2018 could be included

Video format is given as active lines, scanning (interlaced or progressive) and frame rate (59.94 is equavlent to

12. IP Encapsulation and Forward Error Correction

Senders shall map the MPEG-2 TS for transport over an IP network in accordance with SMPTE ST 2022-2.

If FEC is implemented, Senders shall construct the FEC stream in accordance with SMPTE ST 2022-1.

Receivers shall be able to accept IP streams that are compliant with SMPTE ST 2022-2.

If FEC is implemented, Receivers shall be able to accept and process FEC streams constructed in accordance with SMPTE ST 2022-1.

Although SMPTE 2022-2 allows for 1, 4 and 7 TS packets per IP datagram, in order to be compliant with this TR, Senders and Receivers shall support 7 TS packets per IP datagram.

Senders and Receivers may support 1 and 4 TS packets per IP datagram, if desired.

13. Bibliography (Informative)

- CEA-CEB20 “A/V Synchronization Processing Recommended Practice”
- CEA-608-B-2000 “Line 21 data services”
- Rec. ITU-R BT 601-7 “Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios”
- Rec. ITU-R BT 709-6 “Parameter values for the HDTV standards for production and international programme exchange”
- Rec. ITU-R BT.1359-1 “Relative Timing of Sound and Vision for Broadcasting”
- Rec. ITU-R BT.2020-2 “Parameter values for ultra-high definition television systems for production and international programme exchange”
- Rec. ITU-R BT.2100-1 “Image parameter values for high dynamic range television for use in production and international programme exchange”
- SMPTE RDD 8:2008 “Storage and Distribution of Teletext Subtitles and VBI Data for High-Definition Television”
- SMPTE RDD 19:2011 “Guidelines on the Use of Dolby® E with Video Signals at Frame Rates Greater than 30 Hz”
- SMPTE ST 12-2:2008 “Television – Transmission of Time Code in the Ancillary Data Space”
- SMPTE ST 259:2008 “Television – SDTV Digital Signal/Data – Serial Digital Interface”
- SMPTE ST 291-1:2011 “Ancillary Data Packet and Space Formatting”
- SMPTE RP 291-2:2013 “Ancillary Data Space use – 4:2:2 SDTV and HDTV component systems and 4:2:2 2048 x 1080 Production Image formats”
- SMPTE ST 292-1:2011 “1.5 Gb/s Signal/Data Serial Interface”
- SMPTE ST 334-1:2007 “Vertical Ancillary Data Mapping of Caption Data and Other Related Data”
- SMPTE ST 340:2008 “Format for Non-PCM Audio and Data in AES3 – ATSC A/52B Digital Audio compression Standard for AC-3 Data Types”
- SMPTE ST 372:2011 “Dual Link 1/5 Gb/s Digital Interface for 1920 x 1080 and 2048 x 1080 Picture Formats”
- SMPTE ST 424:2006 “Television – 3 Gb/s Signal/Data Serial Interface”
- SMPTE ST 425:2011 “Source Image Format and Ancillary Data Mapping for the 3 Gb/s Interface”
- SMPTE ST 2010:2008 Vertical Ancillary Data Mapping of ANSI/SCTE 104 Messages
- SMPTE ST 2016-3:2009 “Vertical Ancillary Data Mapping of Active Format Description and Bar Data”
- SMPTE ST 2020-1:2008 “Format of Audio Metadata and Description of the Asychronous Serial Bitstream Transport"
- SMPTE ST 2020-2:2008 “Vertical Ancillary Data Mapping of Audio Metadata - Method A”
- SMPTE ST 2020-3-2008 “Vertical Ancillary Data Mapping of Audio Metadata - Method B”
- SMPTE ST 2031-2:2007 “Carriage of DVB/SCTE VBI Data in VANC”
- SMPTE ST 2041-2:2010 “Format for Non-PCM Audio in AES3 - MPEG-2 AAC and HE AAC Audio in ADTS”SMPTE ST 2041-3:2010 “Format for Non-PCM Audio and Data in AES3 - MPEG-4 AAC and HE AAC Compressed Digital Audio in ADTS and LATM/LOAS Wrappers”
- ANSI/SCTE 127 2007 “Carriage of Vertical Blanking Interval (VBI) Data in North American

Digital Television Bitstreams”
– UltraHD Forum: Phase A Guidelines, August 25, 2017, Revision 1.4*

**Bibliography Note: The latest revision of Guidelines from UltraHD Forum in Table 11, lists JPEG 2000 video as one of possible contribution level video formats, which could be used in both HD and UHD broadcast production.*