Video Services Forum (VSF)
Technical Recommendation TR-08

Transport of JPEG XS
Video in ST 2110-22

August 9, 2021
VSF TR-08:2021

INTELLECTUAL PROPERTY RIGHTS

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.

THIS RECOMMENDATION IS BEING OFFERED WITHOUT ANY WARRANTY WHATSOEVER, AND IN PARTICULAR, ANY WARRANTY OF NONINFRINGEMENT IS EXPRESSLY DISCLAIMED. ANY USE OF THIS RECOMMENDATION SHALL BE MADE ENTIRELY AT THE IMPLEMENTER'S OWN RISK, AND NEITHER THE FORUM, NOR ANY OF ITS MEMBERS OR SUBMITTERS, SHALL HAVE ANY LIABILITY WHATSOEVER TO ANY MPLEMENTER OR THIRD PARTY FOR ANY DAMAGES OF ANY NATURE WHATSOEVER, DIRECTLY OR INDIRECTLY, ARISING FROM THE USE OF THIS RECOMMENDATION.

LIMITATION OF LIABILITY

VSF SHALL NOT BE LIABLE FOR ANY AND ALL DAMAGES, DIRECT OR INDIRECT, ARISING FROM OR RELATING TO ANY USE OF THE CONTENTS CONTAINED HEREIN, INCLUDING WITHOUT LIMITATION ANY AND ALL INDIRECT, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES (INCLUDING DAMAGES FOR LOSS OF BUSINESS, LOSS OF PROFITS, LITIGATION, OR THE LIKE), WHETHER BASED UPON BREACH OF CONTRACT, BREACH OF WARRANTY, TORT (INCLUDING NEGLIGENCE), PRODUCT LIABILITY OR OTHERWISE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THE FOREGOING NEGATION OF DAMAGES IS A FUNDAMENTAL ELEMENT OF THE USE OF THE CONTENTS HEREOF, AND THESE CONTENTS WOULD NOT BE PUBLISHED BY VSF WITHOUT SUCH LIMITATIONS.

Executive Summary

This VSF Technical Recommendation (TR), in addition to defining profiles for streaming of JPEG XS video over SMPTE ST 2110-22, adds information for the interoperable transport of audio and ancillary data over other relevant SMPTE ST 2110 standards. The JPEG XS compression method is used in low latency transmission applications for cost-effective, high quality, real-time transport of television video signals over IP networks. The term "XS" is meant to convey the "extra small", "extra speed" nature of the compression method.

Recipients of this document are invited to submit technical comments. The VSF also requests that recipients notify us 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

This document describes an encapsulation of JPEG XS codestream(s) into SMPTE ST 2110-22 for transmission over Internet Protocol-based networks. It also addresses potential system interoperability issues with other media components within the ST 2110 system.

In February of 2020, the VSF created the JPEG XS Activity Group. This group was formed to develop an interoperable method for the provision of low-latency JPEG XS coded video for WAN & LAN transport. The group was also charged with defining interoperability points for the compression and transmission of High Definition and Ultra-High Definition formats.

1.1 Contributors

Contributors to this document include:

- John Dale, Activity Group Chairman, Media Links, Inc.
- Brad Gilmer, Gilmer & Associates, Inc.
- Andrew Krupiczka, Disney
- Jean-Baptiste Lorent, intoPIX SA
- Tim Bruylants, intoPIX SA
- Danny Pierini, Matrox
- John Mailhot, Imagine Communications
- John Schilberg, Utah Scientific

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

AES: AES3:2009 (r2019): “Digital input-output interfacing — Serial transmission format for two-channel linearly-represented digital audio data”

AMWA BCP-002-01: Natural Grouping of NMOS Resources v1.0.0

AMWA BCP-003-01: Secure Communication in NMOS Systems v1.0.0

AMWA BCP-003-02: Authorization in NMOS Systems v1.0.0

AMWA BCP-003-03: Certificate provisioning in NMOS Systems v1.0.0

AMWA IS-04 NMOS Discovery and Registration Specification (Stable) v1.3.1

AMWA IS-05 NMOS Device Connection Management Specification (Stable) v1.1.1

ANSI/CTA 861-H:2021 “A DTV Profile for Uncompressed High-Speed Digital Interfaces”

IEEE 1588-2008: “IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems” (PTP)

IETF draft-ietf-payload-rtp-jpegxs: “RTP Payload Format for ISO/IEC 21122 (JPEG XS)” or the IETF Standards-Track RFC which replaces it

IETF RFC 7273 “RTP Clock Source Signalling”

IETF RFC 8866 “SDP: Session Description Protocol”

ISO/IEC 21122-1:20191: “Information technology — JPEG XS low-latency lightweight image coding system — Part 1: Core coding system”

ISO/IEC 21122-2:20191: “Information technology — JPEG XS low-latency lightweight image coding system — Part 2: Profiles and buffer models”

ISO/IEC 21122-3:2019[1]: “Information technology — JPEG XS low-latency lightweight image coding system — Part 3: Transport and container formats”

JT-NM TR-1001_2020v1.1 System Environment and Device Behaviors for ST 2110 Media Nodes in Engineered Networks –Networks, Registration, and Connection Management

Rec. ITU-R BT.2020-2: “Parameter values for ultra-high definition television systems for production and international programme exchange”

Rec. ITU-R BT.2100-2: “Image parameter values for high dynamic range television for use in production and international programme exchange”

Rec. ITU-T H.273 (2016) | ISO/IEC 23091-2:2019: "Information Technology - Coding-independent code points - Part 2: Video"

SMPTE ST 2022-7:2019: “Seamless Protection Switching of RTP Datagrams”

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

SMPTE ST 2110-10:2017: “Professional Media over Managed IP Networks: System Timing and Definitions”

SMPTE ST 2110-21:2017: “Professional Media over Managed IP Networks: Traffic Shaping and Delivery Timing for Uncompressed Active Video”

SMPTE ST 2110-22:2019: “Professional Media over Managed IP Networks: Constant Bit-Rate Compressed Video”

SMPTE ST 2110-30:2017: “Professional Media over Managed IP Networks: PCM Digital Audio”

SMPTE ST 2110-31:2018: “Professional Media over Managed IP Networks: AES3 Transparent Transport”

SMPTE ST 2110-40:2018: “Professional Media over Managed IP Networks: SMPTE ST 291-1 Ancillary Data”

Note that Level 1K-1 and 4:2:0 color sampling will be supported by ISO/IEC 21122 2nd editions which are in progress with expected publication January-March 2022

Note: Joint ITU and ISO/IEC documents refer to exactly the same standard text, and may share the same title, however in some cases they do not.

4. Acronyms

ACL	Audio Conformance Level
AES	Audio Engineering Society
ANC	Ancillary Data
API	Application Programming Interface
Bpp	Bits per pixel
ETSI	European Telecommunications Standards Institute
FEC	Forward Error Correction
FHD	Full High Definition
HD	High Definition
HDR	High Dynamic Range
HLG	Hybrid-Log Gamma
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPMX	Internet Protocol Media Experience
ISO	International Organization for Standardization
ITU	International Telecommunication Union
JPEG	Joint Photographic Experts Group
JPEG XS	Joint Photographic Experts Group 21122 Coding Standard
LAN	Local Area Network
PCM	Pulse Code Modulation
PQ	Perceptual Quantize
RGB	Red Green Blue
RTP	Real-Time Protocol
SDP	Session Description Protocol
SDR	Standard Dynamic Range
SDI	Serial Digital Interface
SMPTE	Society of Motion Picture and Television Engineers
TR[2]	Video Services Forum Technical Recommendation[1]
UHD	Ultra High Definition
UHD1	Ultra high resolution with a resolution of 3840 × 2160, which is found in ITU-R BT

---

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

	2020
UHD2	Ultra high resolution with a resolution of 7680 × 4320, which is found in ITU-R BT 2020
VSF	Video Services Forum
VS	Video Support Super Box
YCbCr	Luminance Component, Blue-Difference and Red-Difference Chroma Components
WAN	Wide Area Network

5. Definitions

Device	"Device" as defined in SMPTE ST 2110-10.
Media Node	"Media Node" as defined in JT-NM TR-1001
Receiver	"Receiver" as defined in SMPTE ST 2110-10.
Sender	"Sender" as defined in SMPTE ST 2110-10.
JPEG XS RTP Spec	IETF RTP Payload Format for ISO/IEC 21122 (JPEG XS) specification.

6. System Overview/Use Cases (Informative)

An end-user or service provider of broadcast transmission services can utilize devices that implement this Technical Recommendation (TR) for the unidirectional transport of real time television signals over an IP network within a facility or between facilities. The signals may include video, audio, and SMPTE 291 formatted Ancillary Data packets.

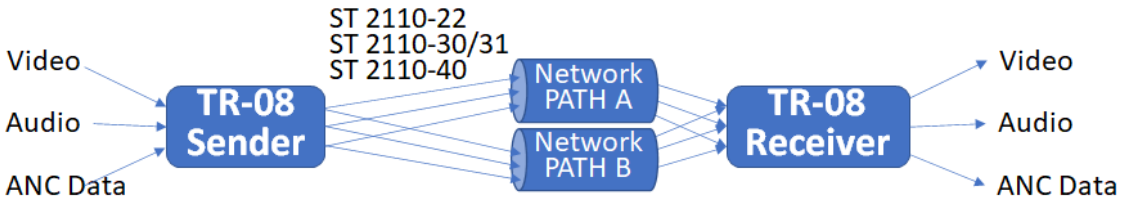

Figure 1: Example system for transmission of real time television signal over IP

As shown in figure 1 above, the Sender has video, audio, and Ancillary Data inputs. These inputs are presumed to be time-aligned at their presentation to the Sender. The video is compressed using a JPEG XS compression algorithm. The JPEG XS codestream is converted into a Real-Time Protocol (RTP) stream in accordance with the JPEG XS RTP Spec and meeting the requirements of SMPTE ST 2110-22. Audio signals are converted into SMPTE ST 2110-30 or SMPTE ST 2110-31 streams, and ANC Data is packaged into SMPTE ST 2110-40 streams.

The system defined in this TR supports linear PCM audio using SMPTE ST 2110-30, and non-PCM audio (and non-audio signals represented in AES3 wrappers) using SMPTE ST 2110-31.

The Sender transmits the RTP streams over the network in accordance with SMPTE ST 2110-10, including optional support of the SMPTE ST 2022-7 redundancy model and separate network paths. The Receiver de-encapsulates the RTP/IP streams, recovering the Audio, ANC Data, and the JPEG XS codestream. The JPEG XS codestream is further decoded to video essence in accordance with provisions of this recommendation. The reconstructed video, audio, and ANC Data signals are time-aligned using the RTP timestamps (correcting for different RTP Clock rates) before

presentation.

The target end to end transmission latency, (less network and transit delay, and not including buffering for delay path differential), for the real time transmission of all essence components including audio and ancillary data is approximately one tenth of a video frame.

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

Receivers under this TR-08 must make the relevant video signal metadata (colorimetry, transfer characteristic, frame rate, sampling structure, image dimensions, etc) available to downstream devices using the technical standards appropriate to the internal or external interfaces provided.

A number of different "Capability Sets", each with several "Conformance Levels", are specified in section 10 of this recommendation. These capability sets include interoperation of Senders and Receivers within a facility under a common clock domain, and also interoperation of Senders and Receivers in different facilities, perhaps separated by some distance, and operating on potentially different clock domains.

7. Organization and Signaling of a VSF TR-08 Package

Streams compliant with this Technical Recommendation shall be organized into a "TR-08 Package" consisting of one SMPTE ST 2110-22 video stream(s), zero or more SMPTE ST 2110-30, zero or more ST 2110-31 audio stream(s), and zero or more SMPTE ST 2110-40 ANC Data streams.

Media Nodes should comply with the requirements and behaviors specified in Section 11 of JT-NM TR-1001_2020v1.1 System Environment and Device Behaviors for ST 2110 Media Nodes in Engineered Networks – Networks, Registration, and Connection Management except where overridden by provisions of this Technical Recommendation.

Media Nodes shall register their Senders and Receivers using the Registration API defined in the AMWA NMOS IS-04 Discovery and Registration Specification.

Receivers shall be capable of connecting to Senders' streams when directed as per the AMWA NMOS IS-05 Device Connection Management Specification.

Media Nodes shall indicate appropriate grouping of Senders and Receivers as per AMWA BCP-002-01 Natural Grouping Of NMOS Resources Specification. (Example: a video and associated audio stream from a field camera.)

Optionally Media Nodes may implement security on the NMOS APIs. If a media note implements security, the implementation shall comply with AMWA NMOS BCP-003 specifications.

Video signal shall be compressed using a JPEG XS codec that conforms to ISO/IEC 21122-1 as described in Section 8.1 of this TR.

The JPEG XS codestream shall be encapsulated into RTP using the payload format described in the JPEG XS RTP specification.

The RTP Stream shall meet all of the requirements of SMPTE ST 2110-22, and by extension, shall meet all of the requirements of SMPTE ST 2110-10.

Details of the required elements of the IS-05 "transport-tile" (an RFC 4566 SDP object) shall be as described in the JPEG XS RTP specification and shall meet the requirements of SMPTE ST 2110-10.

PCM Audio signals shall be organized into SMPTE ST 2110-30 streams or SMPTE ST 2110-31 streams, depending on application requirements.

Non-PCM audio signals and any AES3-encapsulated non-audio signals shall be encapsulated into SMPTE ST 2110-31 streams.

Ancillary Data signals shall be encapsulated into SMPTE ST 2110-40 streams.

Each essence component shall be synchronous to the clock source declared in the *ts-refclk* element of the Sender's Session Description Protocol (SDP) object in compliance with RFC 7273 RTP Clock Source Signaling.

8. Essence Service Components

This section establishes specific restrictions for JPEG XS video, audio and metadata in order to improve interoperability between Senders and Receivers from different implementers.

8.1 JPEG XS Video

8.1.1 JPEG XS Codestream Restrictions

JPEG XS Codestreams emitted by Senders that are compliant with this TR shall adopt the "High444.12" profile as specified in ISO/IEC 21122-2, with the following restrictions and changes:

- **Number of components, chroma sampling format and alpha channel**: If there is an alpha channel, it shall be transported in its own codestream.

  Codestream parameters shall be set according to one of the allowed configurations shown in Table 1. below.

Allowed Configurations	**N_c value (= number of components)**	**$s_x[c]$ and $s_y[c]$ values (= horizontal and vertical sampling factors)**	**Ppih**	**Cpih**
3 color components 4:2:0 YCbCr sampling	$N_c = 3$	$s_x[0] = s_y[0] = 1$ $s_x[1] = s_x[2] = 2$ $s_y[1] = s_y[2] = 2$	0x0000	0
3 color components 4:2:2 YCbCr sampling	$N_c = 3$	$s_x[0] = 1$ $s_x[1] = s_x[2] = 2$ $s_y[0] = s_y[1] = s_y[2] = 1$	0x4a40	0

3 color components 4:4:4 YCbCr sampling	$N_c = 3$	$s_x[0] = s_x[1] = s_x[2] = 1$ $s_y[0] = s_y[1] = s_y[2] = 1$	0x4a40	0 (No RCT)
3 color components 4:4:4 RGB sampling	$N_c = 3$	$s_x[0] = s_x[1] = s_x[2] = 1$ $s_y[0] = s_y[1] = s_y[2] = 1$	0x4a40	1 (RCT)
alpha channel only 4:0:0 sampling	$N_c = 1$	$s_x[0] = s_y[0] = 1$	0x4a40	0

Table 1. Number of components and chroma sampling

- **Color transformation**:
  If the three components of the image use the YCbCR digital representation, Cpih shall be set to 0 (no color transformation).

  Senders shall use the following order of components in the JPEG XS codestream, and the component index shall be set as follows: Y (component index shall be set to 0), Cb (component index shall be set to 1), Cr (component index shall be set to 2).

  In the case of RGB color components with 4:4:4 sampling, Cpih shall be set to 1. This indicates that the three color components of the image can be reversibly transformed to YCbCr digital representation by JPEG XS.

  Senders shall use the following order of components in the JPEG XS codestream, and the component index shall be set as follows: R (component index shall be set to 0), G (component index shall be set to 1), B (component index shall be set to 2).

- **Input bit depth**: B[c] shall be set to 8, 10, or 12 for all values of c (i.e. all components) based on the Conformance Level and selected Capability Set shown in Table 5 in Section 10.

- **Number of horizontal wavelet transformations**: $N_{L,x}$ shall be set to 5

- **Number of vertical wavelet transformations**: $N_{L,y}$ shall be set to 2

- **Quantizer type**: Qpih shall be set to 1 (uniform quantizer)

- **Level**: shall be set to either 1k-1, 2k-1, 4k-2, or 8k-2, depending on the targeted capability set based on the Conformance Level and selected Capability Set shown in Table 5 in Section 10. Implementers shall select the lowest level possible given the width, height and frame rate of the image.

  Note: The selected level defines constraints on the maximum dimensions and framerate of the images in the uncompressed domain, and as such, sets a lower bound on the throughput in the decoded domain.

- **Sublevel**: shall be set to Sublev3bpp or Sublev6bpp when the bpp is less than or equal to 3bpp, and shall be set to Sublev6bpp when the bpp exceeds 3bpp. Implementations compliant with this TR shall not set Sublevel to any other values.

Compression bit rate shall be further constrained to a maximum of 4 bpp, even in the case where the Sublevel is set to Sublev6bpp (see capability sets defined in Section 10).

Note: The selected sublevel defines constraints on the maximum number of bits per pixel for an encoded image and as such, sets a lower bound on the throughput in the encoded domain that a conforming decoder implementation supports.

8.1.2 JPEG XS Stream

The JPEG XS Stream shall be mapped into IP packets in accordance with the JPEG XS RTP specification with the following constraints:

The *pacKetization mode (K)* bit shall be set to "0" in the RTP Payload Header. This sets the codestream packetization mode to 'codestream'.

The number of bytes of *Payload Data* in a packet shall be a multiple of 8 bytes. If the JPEG XS data falls short of the 8-byte multiple, then bytes of 0x00 shall be appended to the JPEG XS data in the last packet to achieve the 8-byte multiple.

8.1.2.1 Field coding and frame rate

Field coding and frame rate shall be set by the *frat* field, whose semantics are defined in ISO/IEC 21122-3.

Senders shall set the *frat* field in the JPEG XS Stream as follows:

For interlaced signals, *top-field-first* mode shall be used. Fields shall be transmitted in temporal order of appearance in all cases. The *bottom-field-first* mode shall not be used.

For interlaced signals, the *Interlace_Mode* of the *frat* field shall be set to '1' (*top-field-first*). For progressive signals, the *Interlaced_Mode* of the *frat* field shall be set to '0' (*progressive*).

The *frat* field shall always be set to signal the frame rate and shall not indicate an unknown frame rate.

In addition, the frame rate can be signaled as described in the JPEG XS RTP specification as part of the SDP. If there is conflict between the SDP and the JPEG XS codestream, the values in the JPEG XS codestream shall prevail.

Sender implementations compliant with this document shall ensure that the information signaled via the JPEG XS RTP specification shall always match the information contained in the SDP object.

8.1.2.2 Color Specification & Dynamic Range

Color information shall be specified in the JPEG XS Stream using the Color Specification descriptor as defined in ISO/IEC 21122-3 and in the JPEG XS RTP specification.

Color information shall be signaled in the SDP object as described in the JPEG XS RTP specification. If there is conflict between the information in the SDP object and the JPEG XS codestream, or if the SDP indicates a colorimetry value of "UNSPECIFIED", the values in the

JPEG XS codestream shall prevail.

Sender implementations compliant with this document shall ensure that the information signaled via the JPEG XS RTP specification shall always match the information contained in the SDP object.

Note: Table 2. below summarizes the signaling code values a sender might employ for commonly used color spaces.

Color space	Color primaries code	Transfer characteristics code	Matrix coefficients code	Video full range flag	Notes
Rec. ITU-R BT.709-6	1	1	1	0	BT 709 SDR
Rec. ITU-R BT.2100-2	9	16	9 (Y'CbCr)	0	PQ HDR with BT 2020
Rec. ITU-R BT.2100-2	9	18	9 (Y'CbCr)	0	HLG with BT 2020

Table 2. (Informative) – Selected examples of color space specification

8.1.2.3 Mastering Display Metadata

Mastering Display Metadata is defined in SMPTE ST 2108-1 and ST 2108-2. This information may be carried as specified in the JPEG XS RTP specification in the Mastering Display Metadata box within the JPEG XS Video Support Super Box (VS), or it may be transported in a ST 2110-40 stream.

In the case where the Mastering Display Metadata is present at a decoder in multiple locations, the information in the Mastering Display Metadata box shall take priority.

If a Sender does not have access to the Mastering Display Metadata at the time the stream is generated, then the Mastering Display Metadata box shall not be included in the VS box.

Mastering Display Metadata shall 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*, as defined in SMPTE ST 2086:2018 "Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images"

- *MaxFALL* and *MaxCLL*, as defined in ANSI/CTA 861-H:2021 "A DTV Profile for Uncompressed High-Speed Digital Interfaces"

If the Mastering Display Metadata is unknown or not included in the video input when the stream is generated, the Mastering Display Metadata shall not be included in the in the JPEG XS video descriptor, and the *mdm_flag* shall be set to '0'.

8.2 Audio Transport (PCM and Non-PCM signals)

Audio signals shall be sampled at a rate of 48 kHz, using a sampling clock which is synchronous to the signaled ts-refclk clock source in the SDP.

PCM Audio signals shall be organized into SMPTE ST 2110-30 streams or SMPTE ST-2110-31 streams or a combination thereof. Any non-PCM audio signals, and/or AES3-formatted non-audio signals, shall be packaged into SMPTE ST 2110-31 streams.

Senders and Receivers that are compliant with this TR shall support the Audio Conformance Level specified in Section 10 of this document for the specified use case, referring to the requirements enumerated in Table 3 below.

Audio Conformance Level	Bit Depth, Sampling, PIDs and AES Channel Pairs per PID
A	24 bits, 48 kHz sampling Up to four (4) audio streams each stream may be SMPTE 2110-30 (Level A) or 2110-31 (Level A) Up to 16 total channels (or channel-equivalents)
B	24 bits, 48 kHz sampling Up to eight (8) audio streams each stream may be SMPTE 2110-30 (Level A) or 2110-31 (Level A) Up to 32 total channels (or channel-equivalents)

Table 3. Audio Capability Sets

Receivers should incorporate a selection mechanism that allows the user to choose the mapping of audio channels from JPEG XS codestreams to subsequent devices or processes. If Receivers implement a re-mapping capability, then AMWA NMOS IS-08 shall be used for controlling this mapping.

8.3 Ancillary Data

Most SDI signals include Horizontal Ancillary (HANC) and Vertical Ancillary (VANC) data packets formatted in accordance with SMPTE ST 291-1. Senders compliant with this recommendation shall transport the ANC data packets (subject to the restrictions below) using SMPTE ST 2110-40. Receivers shall recover the ANC data and present it to the downstream application.

Since the JPEG XS RTP specification includes a box structure for video specific metadata as a prefix to the JPEG XS codestream, it may be possible that a contradiction could occur between this ancillary metadata, metadata included within the SDP parameters and a potential SMPTE 2110-40 stream. Implementers shall follow best practices as indicated in this document to avoid duplication of metadata. If duplication is unavoidable, then metadata shall be consistent in each location with

no difference. In regards to video metadata specifically, if a receiver detects a conflict, then the JPEG XS RTP specification box metadata shall prevail.

Note: This document does not preclude applications from sending additional information through other stream formats in parallel with the streams described herein; in particular the VSF IPMX activity is developing some parallel stream formats for certain application metadata.

8.3.1 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 8.2 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 shall be ignored by Receivers if present in the ST 2110-40 stream. Receivers shall generate a locally correct Embedded Audio Control Packet based on their specific configuration if they are creating an SDI output.

EDH, CRC, and Line Number information, while present in the ancillary data spaces of SDI, are not formatted as ANC packets under SMPTE ST 291-1 and therefore shall not appear in SMPTE ST 2110-40 streams.

9. IP Encapsulation, Forward Error Correction, and Receiver Timing

Senders and Receivers compliant with this recommendation shall be able to create and process, respectively, IP streams that are compliant with SMPTE ST 2110-22.

If SMPTE ST 2022-5 FEC is implemented in a Sender, that Sender shall construct the FEC stream in accordance with SMPTE ST 2022-5, and signal that FEC is being used, as specified in SMPTE ST 2110-10.

If SMPTE ST 2022-5 FEC is implemented in a Receiver, that Receiver shall be able to process FEC streams constructed in accordance with SMPTE ST 2022-5 and shall make use of the signaling specified in 2110-10.

If SMPTE ST 2022-7 redundant streams are implemented in a Sender, then that Sender shall construct the redundant streams as described in SMPTE ST 2022-7 and shall signal them as specified in SMPTE ST 2110-10.

If SMPTE ST 2022-7 redundant streams are implemented in a Receiver, then that Receiver shall process the redundant streams which have been constructed in accordance with SMPTE ST 2022-7, and shall make use of the signaling specified in ST 2110-10.

Note: other FEC approaches may be used and if so, will be signaled in accordance with their defining documents; in particular the VSF IPMX activity FEC approach may be used.

Receivers shall implement one or more of the Timing Recovery Modes defined in Table 4 below.

Timing Recovery Mode	Definition (normative)
SYNC	Receivers shall be able to consume streams where the Sender's *ts-refclk* is equivalent to the Receiver's time source. Video Senders and Receivers shall conform to type W in SMPTE ST 2110-21. Note: Senders that implement the 2110TPNL traffic shape meet the requirements of 2110TPW.
ASYNC	Receivers shall be able to consume streams where the Sender's ts-refclk differs from the Receiver's time source. Video Senders shall conform to type W in SMPTE ST 2110-21, while video Receivers shall be of type A in SMPTE ST 2110-21.
IPMX	Video Senders shall conform to type W in SMPTE ST 2110-21, while Receivers shall implement the timing model under development in the VSF IPMX working group.

Table 4. Enumerated Names for the Timing Recovery Modes

10. Capability Sets and Interoperability Points

Senders and Receivers which claim conformance with this recommendation shall support one or more configuration(s). A configuration is defined by the intersection of a Conformance Level and a Capability Set, as specified in Table 5 below.

Senders and Receivers supporting a specific configuration in Table 5 below shall support all interoperability points listed for that configuration in Appendix B of this document.

Note: Appendix B contains a table of capability sets indexed by conformance level. The table provides reasonable coverage of common formats and features used in professional video applications at the time of the publication of this recommendation.

Conformance Level		Capability Set A Intra-facility Use	Capability Set B Interfacility	Capability Set C Intra-campus with Multimedia extensions	Capability Set D Intra-campus with Multimedia extensions
ALL	Timing	SYNC	ASYNC	ASYNC and IPMX	IPMX
FHD	Video	JPEG XS YCbCr 4:2:2 only bit depth = 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 1920 image height <= 1080	JPEG XS YCbCr 4:2:2 only bit depth = 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 1920 image height <= 1080	JPEG XS YCbCr 4:2:2, YCbCr 4:4:4, RGB 4:4:4 bit depth <= 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 2048 image height <= 1200	JPEG XS YCbCr 4:2:0 YCbCr 4:2:2, YCbCr 4:4:4, RGB 4:4:4 bit depth <= 12 maximum rate = 4bpp frame rate <= 60Hz image width <= 2048

					image height <= 1200
	Audio	ACL-A	ACL-A	ACL-A	ACL-A
UHD1 (4K)	Video	JPEG XS YCbCr 4:2:2 only bit depth = 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 3840 image height <= 2160	JPEG XS YCbCr 4:2:2 only bit depth = 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 3840 image height <= 2160	JPEG XS YCbCr 4:2:2, YCbCr 4:4:4, RGB 4:4:4 bit depth <= 10 maximum rate 4bpp frame rate <= 60Hz image width <= 4096 image height <= 2160	JPEG XS YCbCr 4:2:0[4] YCbCr 4:2:2, YCbCr 4:4:4, RGB 4:4:4 bit depth <= 12 maximum rate = 4bpp frame rate <= 60Hz image width <= 4096 image height <= 2160
	Audio	ACL-B	ACL-B	ACL-A	ACL-A
UHD2 (8K)	Video	JPEG XS YCbCr 4:2:2 only bit depth = 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 7680 image height <= 4320	JPEG XS YCbCr 4:2:2 only bit depth = 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 7680 image height <= 4320	JPEG XS YCbCr 4:2:2, YCbCr 4:4:4, RGB 4:4:4 bit depth <= 10 maximum rate = 4bpp frame rate <= 60Hz image width <= 8192 image height <= 4320	JPEG XS YCbCr 4:2:0[4] YCbCr 4:2:2, YCbCr 4:4:4, RGB 4:4:4 bit depth <= 12 maximum rate = 4bpp frame rate <= 60Hz image width <= 8192 image height <= 4320
	Audio	ACL-B	ACL-B	ACL-B	ACL-B

Table 5. Capability Sets

Appendix A (Informative) Session Description and NMOS Examples

The following Session Description Protocol (SDP) object describes a 1080i@29.97Hz JPEG XS stream, sent in duplicate using 2022-7. Note the media subtype "jxsv" and the bits-per-frame indication b=AS:4396 (4396 kilobits per frame on average, or about 2 bpp after overheads). The optional profile, level, and sublevel terms are also shown.

```
v=0
o=- 108902 53 IN IP4 10.0.81.53
s=jpeg-xs-test-session
i=jpeg-xs Testing 238.10.0.29:20000
t=0 0
a=recvonly
a=group:DUP AMBER BLUE
m=video 20000 RTP/AVP 96
c=IN IP4 238.10.0.29/32
b=AS:4396
a=source-filter: incl IN IP4 238.10.0.29 10.0.81.53
a=rtpmap:96 jxsv/90000
a=fmtp:96 packetmode=0; profile=High422.12; level=2k-1;
sublevel=Sublev3bpp; sampling=YCbCr-4:2:2; width=1920; height=1080;
interlace; exactframerate=30000/1001; depth=10; TCS=SDR;
colorimetry=BT709; PM=2110GPM; SSN=ST2110-22:2019; TP=2110TPNL
a=ssrc:12345 cname:nmos@nmos.tv
a=ts-refclk:ptp=IEEE1588-2008:74-83-EF-FF-FF-71-F8-AD:127
a=mediaclk:direct=0
a=mid:AMBER
m=video 20000 RTP/AVP 96
c=IN IP4 238.100.0.29/32
b=AS:4396
a=source-filter: incl IN IP4 238.100.0.29 10.0.81.54
a=rtpmap:96 jxsv/90000
a=fmtp:96 packetmode=0; profile=High422.12; level=2k-1;
sublevel=Sublev3bpp; sampling=YCbCr-4:2:2; width=1920; height=1080;
interlace; exactframerate=30000/1001; depth=10; TCS=SDR;
colorimetry=BT709; PM=2110GPM; SSN=ST2110-22:2019; TP=2110TPNL
a=ssrc:12345 cname:nmos@nmos.tv
a=ts-refclk:ptp=IEEE1588-2008:74-83-EF-FF-FF-71-F8-AD:127
a=mediaclk:direct=0
a=mid:BLUE
```

In registering senders and receivers using AMWA IS-04, the JPEG XS capable receiver must indicate so using the same subtype in its "caps" parameter (part of the receiver object):

```
"caps": { "media_types": [ "video/raw", "video/jxsv" ] }
```

In the NMOS Flow object, again the media_type must indicate video/jxsv. Additional parameters may be added to the NMOS Flow schema for the average rate per frame in a future NMOS version.

Appendix B (Informative) Capability Sets and Interoperability Points

Interoperability Points Capability Set A

Inter-op Points	Cap-abil-ity Set	Timing	Con-form-ance Level	Format & Frame Rate*	Sampling Points	Pixels per Second	Max Coding Efficiency Mbps	Max Coding Efficiency Bpp	Min Coding Efficiency Mbps	Min Coding Efficiency Bpp	Bit Depth	Color Sampling	Color Space	Audio Conform-ance Level	JPEG XS Profile			Reference Uncom-pressed Mbps Video, Mbps
															Profile	Level	Sublevel**	
1	A	SYNC	FHD	720px1280/59	921,600	55,240,759	83	1.5	221	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	1K-1	Sublev6bpp	1105
2	A	SYNC	FHD	720px1280/50	921,600	46,080,000	69	1.5	184	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	1K-1	Sublev6bpp	922
3	A	SYNC	FHD	1080ix1920/29	2,073,600	62,145,854	93	1.5	249	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	1K-1	Sublev6bpp	1243
4	A	SYNC	FHD	1080ix1920/25	2,073,600	51,840,000	78	1.5	207	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	1K-1	Sublev6bpp	1037
5a	A	SYNC	FHD	1080px1920/59	2,073,600	124,291,708	186	1.5	497	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	2486
5b	A	SYNC	FHD	1080px1920/59	2,073,600	124,291,708	186	1.5	497	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	A	High 444.12	2K-1	Sublev6bpp	2486
5c	A	SYNC	FHD	1080px1920/59	2,073,600	124,291,708	186	1.5	497	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	A	High 444.12	2K-1	Sublev6bpp	2486
6a	A	SYNC	FHD	1080px1920/50	2,073,600	103,680,000	156	1.5	415	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	2074
6b	A	SYNC	FHD	1080px1920/50	2,073,600	103,680,000	156	1.5	415	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	A	High 444.12	2K-1	Sublev6bpp	2074
6c	A	SYNC	FHD	1080px1920/50	2,073,600	103,680,000	156	1.5	415	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	A	High 444.12	2K-1	Sublev6bpp	2074
7a	A	SYNC	UHD1	2160px3840/59	8,294,400	497,166,833	746	1.5	1989	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	B	High 444.12	4K-2	Sublev6bpp	9,943
7b	A	SYNC	UHD1	2160px3840/59	8,294,400	497,166,833	746	1.5	1989	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	4K-2	Sublev6bpp	9,943
7c	A	SYNC	UHD1	2160px3840/59	8,294,400	497,166,833	746	1.5	1989	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	4K-2	Sublev6bpp	9,943
8a	A	SYNC	UHD1	2160px3840/50	8,294,400	414,720,000	622	1.5	1659	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	B	High 444.12	4K-2	Sublev6bpp	8,294
8b	A	SYNC	UHD1	2160px3840/50	8,294,400	414,720,000	622	1.5	1659	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	4K-2	Sublev6bpp	8,294

8c	A	SYNC	UHD1	2160px3840/50	8,294,400	414,720,000	622	1.5	1659	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	4K-2	Sublev6bpp	8,294
9a	A	SYNC	UHD2	4320px7680/59	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	ACL-B	High 444.12	8K-2	Sublev6bpp	39,773
9b	A	SYNC	UHD2	4320px7680/59	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	ACL-B	High 444.12	8K-2	Sublev6bpp	39,773
9c	A	SYNC	UHD2	4320px7680/59	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	ACL-B	High 444.12	8K-2	Sublev6bpp	39,773
10a	A	SYNC	UHD2	4320px7680/50	33,177,600	1,658,880,000	2488	1.5	6636	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	ACL-B	High 444.12	8K-2	Sublev6bpp	33,178
10b	A	SYNC	UHD2	4320px7680/50	33,177,600	1,658,880,000	2488	1.5	6636	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	ACL-B	High 444.12	8K-2	Sublev6bpp	33,178
10c	A	SYNC	UHD2	4320px7680/50	33,177,600	1,658,880,000	2488	1.5	6636	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	ACL-B	High 444.12	8K-2	Sublev6bpp	33,178

* "Video format is given as active lines, scanning (interlaced or progressive) and frame rate (59.9 is equivalent to 60/1.001, while 29.97 is equivalent to 30/1.001)

** Sublevel: shall be set to Sublev3bpp or Sublev6bpp when the bpp is less than or equal to 3bpp, and shall be set to Sublev6bpp when the bpp exceeds 3bpp. Implementations compliant with this TR shall not set Sublevel to any other values.

Interoperability Points Capability Set B

Inter-op Points	Cap-abil-ity Set	Timing	Con-form-ance Level	Format & Frame Rate*	Sampling Points	Pixels per Second	Max Coding Efficiency Mbps	Max Coding Efficiency Bpp	Min Coding Efficiency Mbps	Min Coding Efficiency Bpp	Bit Depth	Color Sampling	Color Space	Audio Conform-ance Level	JPEG XS Profile			Reference Uncom-pressed Mbps Video, Mbps
															Profile	Level	Sublevel**	
1	B	ASYNC	FHD	720px1280/59	921,600	55,240,759	83	1.5	221	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	-A	High 444.12	1K-1	Sublev6bpp	1105
2	B	ASYNC	FHD	720px1280/50	921,600	46,080,000	69	1.5	184	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	1K-1	Sublev6bpp	922
3	B	ASYNC	FHD	1080ix1920/29	2,073,600	62,145,854	93	1.5	249	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	1K-1	Sublev6bpp	1243
4	B	ASYNC	FHD	1080ix1920/25	2,073,600	51,840,000	78	1.5	207	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	1K-1	Sublev6bpp	1037
5a	B	ASYNC	FHD	1080px1920/59	2,073,600	124,291,708	186	1.5	497	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	2486
5b	B	ASYNC	FHD	1080px1920/59	2,073,600	124,291,708	186	1.5	497	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	A	High 444.12	2K-1	Sublev6bpp	2486
5c	B	ASYNC	FHD	1080px1920/59	2,073,600	124,291,708	186	1.5	497	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	A	High 444.12	2K-1	Sublev6bpp	2486
6a	B	ASYNC	FHD	1080px1920/50	2,073,600	103,680,000	156	1.5	415	4.00	10bit	4:2:2	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	2074

6b	B	ASYNC	FHD	1080px1920/50	2,073,600	103,680,000	156	1.5	415	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	A	High 444.12	2K-1	Sublev6bpp	2074
6c	B	ASYNC	FHD	1080px1920/50	2,073,600	103,680,000	156	1.5	415	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	A	High 444.12	2K-1	Sublev6bpp	2074
7a	B	ASYNC	UHD1	2160px3840/59	8,294,400	497,166,833	746	1.5	1989	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	B	High 444.12	4K-2	Sublev6bpp	9,943
7b	B	ASYNC	UHD1	2160px3840/59	8,294,400	497,166,833	746	1.5	1989	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	4K-2	Sublev6bpp	9,943
7c	B	ASYNC	UHD1	2160px3840/59	8,294,400	497,166,833	746	1.5	1989	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	4K-2	Sublev6bpp	9,943
8a	B	ASYNC	UHD1	2160px3840/50	8,294,400	414,720,000	622	1.5	1659	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	B	High 444.12	4K-2	Sublev6bpp	8,294
8b	B	ASYNC	UHD1	2160px3840/50	8,294,400	414,720,000	622	1.5	1659	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	4K-2	Sublev6bpp	8,294
8c	B	ASYNC	UHD1	2160px3840/50	8,294,400	414,720,000	622	1.5	1659	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	4K-2	Sublev6bpp	8,294
9a	B	ASYNC	UHD2	4320px7680/59	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	B	High 444.12	8K-2	Sublev6bpp	39,773
9b	B	ASYNC	UHD2	4320px7680/59	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	8K-2	Sublev6bpp	39,773
9c	B	ASYNC	UHD2	4320px7680/59	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	8K-2	Sublev6bpp	39,773
10a	B	ASYNC	UHD2	4320px7680/50	33,177,600	1,658,880,000	2488	1.5	6636	4.00	10bit	4:2:2	Rec. ITU-R BT.2020-2	B	High 444.12	8K-2	Sublev6bpp	33,178
10b	B	ASYNC	UHD2	4320px7680/50	33,177,600	1,658,880,000	2488	1.5	6636	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	8K-2	Sublev6bpp	33,178
10c	B	ASYNC	UHD2	4320px7680/50	33,177,600	1,658,880,000	2488	1.5	6636	4.00	10bit	4:2:2	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	8K-2	Sublev6bpp	33,178

* "Video format is given as active lines, scanning (interlaced or progressive) and frame rate (59.9 is equivalent to 60/1.001, while 29.97 is equivalent to 30/1.001)

** Sublevel: shall be set to Sublev3bpp or Sublev6bpp when the bpp is less than or equal to 3bpp, and shall be set to Sublev6bpp when the bpp exceeds 3bpp. Implementations compliant with this TR shall not set Sublevel to any other values.

Interoperability Points Capability Set C

Inter-op Points	Cap-ability Set	Timing	Con-form-ance Level	Format & Frame Rate*	Sampling Points	Pixels per Second	Max Coding Efficiency Mbps	Max Coding Efficiency Bpp	Min Coding Efficiency Mbps	Min Coding Efficiency Bpp	Bit Depth	Color Sampling ***	Color Space	Audio Con-form-ance Level	JPEG XS Profile			Reference Uncom-pressed Mbps Video, Mbps
															Profile	Level	Sublevel**	
Capability set C includes all items 1-10c in Capability Set B, Below are only the additional interop points																		
1a	C	ASYNC	FHD	1080px1920/59*	2,073,600	124,291,708	186	1.5	497	4.00	8bit	RGB 4:4:4	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	2983
1b	C	ASYNC	FHD	1080px1920/50	2,073,600	103,680,000	156	1.5	415	4.00	8bit	RGB 4:4:4	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	2488
1c	C	ASYNC	FHD	1080px1920/60	2,073,600	124,416,000	187	1.5	498	4.00	10bit	RGB 4:4:4	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	3732
1d	C	ASYNC	FHD	1080px1920/59*	2,073,600	124,291,708	186	1.5	497	4.00	8bit	YCbCr 4:4:4	Rec. ITU-R BT.709-6	A	High 444.12	2K-1	Sublev6bpp	2983
2a	C	ASYNC	FHD	1200px1920/60	2,304,000	138,101,898	207	1.5	552	4.00	8bit	Full Range RGB 4:4:4	Rec. ITU-R BT.709-6	A	High 444.12	4K-1	Sublev6bpp	3318
2b	C	ASYNC	FHD	1200px1920/50	2,304,000	115,200,000	173	1.5	461	4.00	8bit	Full Range RGB 4:4:4	Rec. ITU-R BT.709-6	A	High 444.12	4K-1	Sublev6bpp	2765
3a	C	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1991	4.00	8bit	RGB 4:4:4	Rec. ITU-R BT.2020-2	B	High 444.12	4K-2	Sublev6bpp	11,932
3b	C	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1991	4.00	10bit	YCbCr 4:4:4	Rec. ITU-R BT.2100-2	B	High 444.12	4K-2	Sublev6bpp	14,915
3c	C	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1991	4.00	10bit	YCbCr 4:4:4	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	4K-2	Sublev6bpp	14,915
3d	C	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1991	4.00	10bit	YCbCr 4:4:4	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	4K-2	Sublev6bpp	14,915
3e	C	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1991	4.00	8bit	RGB 4:4:4	Rec. ITU-R BT.2020-2	B	High 444.12	4K-2	Sublev6bpp	11,932

4a	C	ASYNC	UHD 2	4320px7680/59*	33,177,600	1,988,667,333	2986	1.5	7963	4.00	10bit	YCbCr 4:4:4	Rec. ITU-R BT.2100-2	B	High 444.12	8K-2	Sublev6bpp	59,660
4b	C	ASYNC	UHD 2	4320px7680/59*	33,177,600	1,988,667,333	2986	1.5	7963	4.00	10bit	YCbCr 4:4:4	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	8K-2	Sublev6bpp	59,660
4c	C	ASYNC	UHD 2	4320px7680/59*	33,177,600	1,988,667,333	2986	1.5	7963	4.00	10bit	YCbCr 4:4:4	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	8K-2	Sublev6bpp	59,660

* "Video format is given as active lines, scanning (interlaced or progressive) and frame rate (59.9 is equivalent to 60/1.001, while 29.97 is equivalent to 30/1.001)

** Sublevel: shall be set to Sublev3bpp or Sublev6bpp when the bpp is less than or equal to 3bpp, and shall be set to Sublev6bpp when the bpp exceeds 3bpp. Implementations compliant with this TR shall not set Sublevel to any other values.

*** Full Range is indicated, on 2a and 2b, where not indicated all others are limited range (RGB or YCbCr)

Interoperability Points Capability Set D

Inter-op Points	Cap-ability Set	Timing	Con-form-ance Level	Format & Frame Rate*	Sampling Points	Pixels per Frame	Max Coding Efficiency Mbps	Max Coding Efficiency Bpp	Min Coding Efficiency Mbps	Min Coding Efficiency Bpp	Bit Depth	Color Sampling ****	Color Space	Audio Con-form-ance Level	JPEG XS Profile			Reference Uncom-pressed Mbps Video, Mbps
															Profile***	Level	Sublevel**	
Capability set D includes all items 1-10c in Capability Set B, all items in Capability Set C 1a-4c. Below are only the additional interop points																		
1a	D	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1989	4.00	8bit	YCbCr 4:2:0	Rec. ITU-R BT. 2020-2	B	Custom	4K-2	Sublev6bpp	5,966
1b	D	ASYNC	UHD 1	2160px3840/50	8,294,400	414,720,000	622	1.5	1659	4.00	8bit	YCbCr 4:2:0	Rec. ITU-R BT. 2020-2	B	Custom	4K-2	Sublev6bpp	4,977
1c	D	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1989	4.00	10bit	YCbCr 4:2:0	Rec. ITU-R BT. 2020-2	B	Custom	4K-2	Sublev6bpp	7,458
1d	D	ASYNC	UHD 1	2160px3840/59*	8,294,400	497,166,833	746	1.5	1989	4.00	12bit	YCbCr 4:2:0	Rec. ITU-R BT. 2020-2	B	High 444.12	4K-2	Sublev6bpp	8,949
2a	C	ASYNC	UHD 2	4320px7680/59*	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	YCbCr 4:2:0	Rec. ITU-R BT.2100-2	B	High 444.12	8K-2	Sublev6bpp	29,830
2b	C	ASYNC	UHD 2	4320px7680/59*	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	YCbCr 4:2:0	Rec. ITU-R BT.2100-2 (PQ)	B	High 444.12	8K-2	Sublev6bpp	29,830
2c	C	ASYNC	UHD 2	4320px7680/59*	33,177,600	1,988,667,333	2983	1.5	7955	4.00	10bit	YCbCr 4:2:0	Rec. ITU-R BT.2100-2 (HLG)	B	High 444.12	8K-2	Sublev6bpp	29,830

* "Video format is given as active lines, scanning (interlaced or progressive) and frame rate (59.9 is equivalent to 60/1.001, while 29.97 is equivalent to 30/1.001)

** Sublevel: shall be set to Sublev3bpp or Sublev6bpp when the bpp is less than or equal to 3bpp, and shall be set to Sublev6bpp when the bpp exceeds 3bpp. Implementations compliant with this TR shall not set Sublevel to any other values.

*** The Custom profile used in the table is identical to the High 444.12 profile, but allowing for 4:2:0 sampling. Because it is not an official JPEG XS profile, it is signaled with Ppih set to 0x0000.

**** Full Range is indicated, where all others are limited range (RGB or YCbCr).