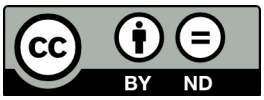
Video Services Forum (VSF)
Technical Recommendation TR-05

Essential Formats and Descriptions for Interoperability of SMPTE ST 2110-20 Video Signals

June 23, 2018



INTELLECTUAL PROPERTY RIGHTS

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THE FOREGOING NEGATION OF DAMAGES IS A FUNDAMENTAL ELEMENT OF THE USE OF THE CONTENTS HEREOF, AND THESE CONTENTS WOULD NOT BE PUBLISHED BY VSF WITHOUT SUCH LIMITATIONS.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 Technical Recommendation provides a list of essential formats for uncompressed video essence transport using SMPTE ST 2110-20:2017. Each format is defined using a set of SDP parameters that can be used to precisely specify an IP packet stream in a way that will facilitate interoperability between devices.

Product suppliers and users are encouraged to use the colloquial format names and Format Groups defined in this document in order to simplify product documentation and requirement descriptions. An example of how the colloquial format names and Format Group names could be used is given in Section 6.3 of this document.

Recipients of this document are requested to submit 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

SMPTE ST 2110-20 specifies a standard protocol for real-time, RTP-based transport of uncompressed active video essence over IP networks. A wide variety of video formats can be accommodated by that standard, permitting thousands of different permutations to potentially be created for various uses.

The VSF has two goals that it seeks to achieve by publishing this document. First, it seeks to establish a set of logical groups of video formats that have similar processing requirements for hardware and software systems. Second, it wants to produce a collection of precise and complete definitions of those commonly used formats that are an essential part of most modern television production workflows.

This document is not meant to restrict the development and use of formats that are not listed herein.

1.1. Contributors

The following individuals participated in the Video Services Forum SVIP working group which developed this technical recommendation.

Al Kovalick (Media Systems Consulting)	Dotan Levi (Mellanox)	Merrick Ackermans (MVA Broadcast Consulting)
Albert Faust (Lawo)	Eric Fankhauser (Evertz)	Michael Bany (Fox)
Andrew Krupiczka (ESPN)	Franz Baumann (IRT)	Michael Walter (Axiom Media Systems)
Arne Boenninghoff (Riedel)	Glenn Reitmeier (NBC Universal)	Paul Briscoe (Televisionary Consulting)
Bob Ruhl (VSF)	Jed Deame (Nextera Video)	Peter Brightwell (BBC)
Brad Gilmer (Gilmer and Associates)	John Mailhot (Imagine)	Robert Welch (Arista)
Carl Ostrom (System Resource)	John Williamson (Cisco)	Thomas Bause Mason (SMPTE)
Chuck Meyer (Grass Valley)	Kenneth K. Jonassen (Bridge Technologies)	Thomas Edwards (FOX)
Danny Pierini (Matrox)	Leigh Whitcomb (Imagine)	Wes Simpson (Telecom Product Consulting)
David Horton (Ericsson)	Luann Linnebur (Nevion)	

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 or this document, please call +1 929-279-1995 or e-mail opsmgr@videoservicesforum.org.

2. Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains 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 and 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. References (normative)

SMPTE ST 2110-20:2017 Professional Media Over Managed IP Networks: Uncompressed Active Video

RFC 4566 Session Description Protocol, Internet Engineering Task Force

4. Definitions

Format Group	A collection of video essence formats that are closely related in function and have similar bit rates. When a Format Group name is used in a product specification or a requirement list all the members of the group are supported.
Integer	**INT(x)** shall return the largest integer not greater than **x**.
SDP	Media Type parameters within Session Description Protocol, as defined in RFC 4566

5. Common Video Formats

This document contains a list of video signal format descriptions that are commonly used in professional media production. The list is intended to promote interoperability among products from multiple suppliers, by precisely specifying all of the SDP Media Type parameters that are associated with each video format. As industry practices evolve, this document may be updated to include new, widely-adopted formats.

This list in no way is meant to restrict the development and use of other formats not defined herein by any party at any time. For image formats carried in SDI as per ST 259, ST 292-1 or ST 425-1 that are not defined in this document, SMPTE ST 2022-8 encapsulation may also be used

5.1 Intended Use

Product suppliers and users are encouraged to use the formats and groupings defined in this specification in order to simplify product documentation and requirement descriptions. For example, if a product specification indicates support for the 1080p Format Group, then all of the

formats listed in that group in this specification shall be supported by that product. If a product only supports a set of individual formats from this document, then only those specific formats shall be listed in that product's documentation.

Each citation of this specification should include the year of the document revision. Thus, a reference should be in the form "1080i Format Group (VSF TR-05:2018)."

5.2 Format Groupings

Each video format defined in this document is a member of a Format Group. The groupings are made in consideration of the performance requirements for processing the signals in real time. When a Format Group name is used in a product specification or a requirement list all the members of the group shall be fully implemented. The Format Groups are NOT meant to be cumulative, so, for example, a product is perfectly free to implement the formats in the UHD-1 SDR Format Group without implementing the 1080i Format Group.

Note that additional Format Groups may be added in future revision of this specification. However, current Format Groups will not be changed in future revisions of this document.

6. Video Formats

Each format listed in this document shall be implemented exactly as defined, conforming strictly to the SDP Media Type parameter lists as provided herein. All SDP syntax and usage shall comply with RFC 4566.

Each video format listed below is provided with a colloquial name. It is intended that format names will be used in documents that comply with this recommendation.

6.1 High Definition Category

The High Definition Category consists of three groups. Products designated as supporting the High Definition Category shall support all three of the listed Format Groups.

Alternatively, products may support any combination of the three Format Groups independently by listing only those Format Groups that are supported.

6.1.1 720p Format Group

6.1.1.1 Format 720p50

```
width=1280; height=720; exactframerate=50; sampling=YCbCr-
4:2:2; depth=10; TCS=SDR; colorimetry=BT709; PM=2110GPM;
SSN=ST2110-20:2017;
```

6.1.1.2 Format 720p59.94

```
width=1280; height=720; exactframerate=60000/1001;
sampling=YCbCr-4:2:2; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017;
```

6.1.2 1080i Format Group

6.1.2.1 Format 1080i50

```
width=1920; height=1080; exactframerate=25; interlace;
sampling=YCbCr-4:2:2; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017;
```

6.1.2.2 Format 1080i59.94

```
width=1920; height=1080; exactframerate=30000/1001;
interlace; sampling=YCbCr-4:2:2; depth=10; TCS=SDR;
colorimetry=BT709; PM=2110GPM; SSN=ST2110-20:2017;
```

6.1.3 1080p Format Group

6.1.3.1 Format 1080p50

```
width=1920; height=1080; exactframerate=50; sampling=YCbCr-
4:2:2; depth=10; TCS=SDR; colorimetry=BT709; PM=2110GPM;
SSN=ST2110-20:2017;
```

6.1.3.2 Format 1080p59.94

```
width=1920; height=1080; exactframerate=60000/1001;
sampling=YCbCr-4:2:2; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017;
```

6.1.3.3 Format 1080p23.98

```
width=1920; height=1080; exactframerate=24000/1001;
sampling=YCbCr-4:2:2; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017;
```

6.2 Ultra High Definition Category

The Ultra High Definition Category currently consists of one group.

6.2.1 UHD-1 SDR Format Group

6.2.1.1 Format 2160p50

```
width=3840; height=2160; exactframerate=50; sampling=YCbCr-
4:2:2; depth=10; TCS=SDR; colorimetry=BT709; PM=2110GPM;
SSN=ST2110-20:2017;
```

6.2.1.2 Format 2160p59.94

```
width=3840; height=2160; exactframerate=60000/1001;
sampling=YCbCr-4:2:2; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017;
```

6.3 Example Usage (Informative)

Consider a broadcaster that is seeking equipment for use in an environment that is only ever intended to work at frame/field rates of 59.94 and image sizes of 1920x1080 and 1280x720. This broadcaster could issue a requirements document that includes just their desired formats, as follows:

> Video Format Requirements as per VSF TR-05:2018:
> **720p59.94** Format
> **1080i59.94** Format
> **1080p59.94** Format

7. Equipment Planning Performance and Bandwidth (Informative)

In order to characterize equipment performance capabilities and bandwidth requirements, several formulas for calculating samples per second and megabits per second are defined in this section. By using these formulas, equipment performance capabilities can be stated in a clear, consistent manner.

7.1 Equipment Performance Level Formulas

These formulas calculate samples per second from SDP parameters (applicable parameter names are shown in **`bold monospaced font`** in the following paragraphs). These measurements will allow a vendor to specify the general performance capabilities of a piece of equipment. For example, if a piece of equipment can support 125,000,000 chroma/luma samples per second it has the performance to support 1080i59.94 video.

7.1.1 Image Samples (Pixels) per Second

The number of image samples per second can be calculated from the information provided in the SDP by multiplying the image **`width`** by the image **`height`** by the **`exactframerate`**.

$$image\text{-}samples\text{-}per\text{-}second = width \times height \times exactframerate$$

7.1.2 Chroma/Luma Samples per Second

The number of Chroma and Luma samples per second can be found by multiplying the number of *image-samples-per-second* by a factor that is related to the color sub-sampling format.

$$chroma\text{-}luma\text{-}samples\text{-}per\text{-}second = image\text{-}samples\text{-}per\text{-}second \times subsampling\text{-}factor$$

The subsampling-factors for several commonly used formats are shown in Table 1:

Color sub-sampling format (SDP **`sampling`** parameter)	*subsampling-factor*
`RGB, XYZ, YCbCr-4:4:4, ICtCp-4:4:4`	3

YCbCr-4:2:2, ICtCp-4:2:2	2
YCbCr-4:2:0, ICtCp-4:2:0	1.5
key	1

Table 1: List of subsampling-factor values

Note: For chroma key signals, use the factor for the corresponding video chroma format.

7.2 Bandwidth Formulas

The formulas in this section may be used to calculate a stream data rate in megabits per second using SDP parameters (applicable parameter names are shown in **bold monospaced font** in the following paragraphs). This measurement allows an end user to better understand the bandwidth requirements of various formats for network planning.

7.2.1 Approximate Signal Bandwidth (ASB)

To calculate the ASB in Megabits per second for a signal using General Packing Mode, one needs to know the size of each pgroup (in bytes) and the number of pixels in each pgroup. These values are based on the SDP parameters **sampling** and **depth** which can then be used to locate the appropriate values from the tables provided in SMPTE ST 2110-20. Once the appropriate values of *pgroupsize* (in bytes) and *pgroupcoverage* (in pixels) have been obtained, the following formulas can be used to calculate the ASB:

$$packets\text{-}per\text{-}frame = 1 + INT\left(\frac{width \times height}{INT\left(\frac{1426}{pgroupsize}\right) \times pgroupcoverage}\right)$$

$$bits\text{-}per\text{-}packet = 8 \times \left(INT\left(\frac{1426}{pgroupsize}\right) \times pgroupsize + 94\right)$$

$$ASB = \frac{packets\text{-}per\text{-}frame \times bits\text{-}per\text{-}packet \times exactframerate}{1{,}000{,}000}$$

7.2.2 Example Calculation (Informative)

Consider a 1080p59.94 format video signal as defined in section 6.3.2 of this document:

From the SDP provided in paragraph 6.1.3.2, the following values are needed to do the calculation:

```
width=1920
height=1080
exactframerate=60000/1001
sampling=YCbCr-4:2:2
depth=10
PM=2110GPM
```

Step 1: Referring to SMPTE ST 2110-20:2017, section 6.2.4, the following values can be determined using the **`sampling`** and **`depth`** parameters:

pgroupsize = 5
pgroupcoverage = 2

Step 2: Plugging these values into the first equation gives 3,368 packets per video frame:

$$3{,}638 = 1 + INT\left(\frac{1920\times1080}{INT\left(\frac{1426}{5}\right)\times2}\right)$$

Step 3: The second equation results in 12,152 bits per packet:

$$12{,}152 = 8\times\left(INT\left(\frac{1426}{5}\right)\times5 + 94\right)$$

Step 4: Using the third equation, the ASB can be calculated as 2,650 Mbit/s:

$$2{,}650 = \frac{3{,}638\times12{,}152\times\left(\frac{60000}{1001}\right)}{1{,}000{,}000}$$

Appendix A (Informative) Type N, NL and W Video Sources

Per section 8.1 of SMPTE ST 2110-21, every valid SDP Media Type parameters list needs to include a TP parameter with one of the following values: 2110TPN, 2110TPNL or 2110TPW. This Appendix provides an explanation of how these terms are used.

The ST 2110-21 standard defines three types of senders: N (for Narrow), NL (for Narrow Linear) and W (for Wide), which are included in the SDP as the parameter TP. These types define limits for the amount of packet delay variation (i.e. the burstiness) that a sender is allowed to exhibit in its output stream.

A.1 Type NL

Type NL is the easiest to understand, and corresponds to a stream where all the packets that contain the image essence for a video signal are evenly spaced across a period of time that corresponds to the duration of a video frame. The SDP for this type of stream must include a parameter "TP=2110TPNL"

A.2 Type N

Type N is similar to Type NL, except that the sender doesn't send packets during the time that would correspond to the VBI (Vertical Blanking Interval) or VANC (Vertical Ancillary data space) of the corresponding traditional SDI video signal. Thus, a Type N sender would be able to send image essence packets in stream that would have a noticeable gap that occurs during each video frame period. For example, in a 720p signal running at 50 frames per second, and a VBI equal in duration to 30 lines of video, the sender would deliver packets for (720/750*20 milliseconds) 19.2 milliseconds out of every 20 milliseconds, and have a 0.8 millisecond gap when no packets are sent. Note that this would be the behavior that would be the easiest to implement if an incoming SDI signal was simply converted to ST2110 packets whenever active video samples arrive. The SDP for this type of stream must include a parameter "TP=2110TPN"

A.3 Type W

Type W senders are allowed to have a significantly greater burstiness than Type N or NL. This category is included in the ST 2110-21 standard to accommodate software-based senders, such as a graphics generator.

Type W streams will also tend to consume larger amounts of buffer space within network switches and other devices; applications with significant quantities of Type W senders will need to be implemented using network devices that have sufficiently large internal buffers. Signal bursts also should be taken into consideration during link capacity planning and bandwidth allocation, particularly in systems that have multiple frame-aligned Type W senders operating simultaneously. The SDP for this type of stream must include a parameter "TP=2110TPW"

A.4 Type N, W and A Receivers

Receivers are also categorized in ST 2110-21, but this information is not required to be transmitted with SDP.

Type N receivers can be expected to correctly receive a flow originating from either a Type N or a Type NL sender (but not a Type W), provided that the receiver is locked to the same clock as the sender.

Type W receivers should be able to receive a stream from an N, NL or W sender, provided the receiver is locked to the same clock as the sender.

Type A (for Asynchronous) receivers should be capable of receiving a stream from any type of sender, regardless of clock source or signal phase.

Appendix B (Informative) Regarding 2110-40 Ancillary Data Flows

The ANC (Ancillary) Data spaces of the SDI signal have been used for the packaging and transport of many different signal or data types. Some signals or data types are inappropriate to send via this method because there are more specific methods for their transport (i.e. embedded audio signals). Other signals are inappropriate to send via this method because they are specific to the SDI link layer, such as EDH (Error Detection and Handling) packets.

B.1 Permitted ANC packets

With the exception of specifically discouraged packets which are enumerated below, any valid SMPTE ST 291-1 ANC Data packet can be transported using SMPTE ST 2110-40.

B.2 Discouraged ANC packets

The following ANC Data packet types should not be transported using 2110-40:

Embedded Audio Packets
Embedded Audio Control Packets
EDH Packets (DID F4h)
Packets marked for Deletion (DID 80h)
SDI Payload Identifier (DID 41h / SDID 01h)

Note: the information about video format contained in the SDI Payload Identifier is already transported within the video metadata set defined in SMPTE ST 2110-20.

Appendix C (Informative) Packing Mode

The SMPTE ST 2110-20 standard defines two packing modes for video signals -- BPM (Block Packing Mode) and GPM (General Packing Mode). BPM is a restricted subset of GPM, and GPM is the more general case. The standard requires all receivers to receive GPM; therefore all receivers are also capable of receiving the BPM subset. It therefore follows that all formats can be defined as using GPM, and even if a specific sender emits a BPM, signal interoperability will still be achieved in all cases. Thus, there is no need to define specific formats that are BPM-only. However, per ST 2110-20, each source is required to indicate the type of packing mode employed by including either "PM=2110GPM" or "PM=2110BPM" in their associated SDP definition.

Appendix D (Informative) Audio Interoperability

Regarding audio interoperability, the SMPTE ST 2110-30 and AES67 standards include tables with profiles and levels for this. Both standards state that senders and receivers shall support a 1 millisecond packet time with 8 channels of audio as the common interoperability point.

Bibliography (Informative)

AES67-2015 AES standard for audio applications of networks – High-performance streaming audio-over-IP interoperability

SMPTE ST 259:2008 Television – SDTV Digital Signal/Data – Serial Digital Interface

SMPTE ST 291-1:2011 Ancillary Data Packet and Space Formatting

SMPTE ST 292-1:2012 1.5 Gb/s Signal / Data Serial Interface

SMPTE ST 425-1:2017 Source Image Format and Ancillary Data Mapping for the 3 Gb/s Serial Interface

SMPTE ST 2022-8:2018 (Anticipated) Professional Media over Managed IP Networks: Timing of ST 2022-6 streams in ST 2110-10 Systems

SMPTE ST 2059-1:2015 Generation and Alignment of Interface Signals to the SMPTE Epoch

SMPTE ST 2110-21:2017 Professional Media Over Managed IP Networks: Traffic Shaping and Delivery

SMPTE ST 2110-30:2017 Professional Media Over Managed IP Networks: PCM Digital Audio

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