Video Services Forum (VSF)
Technical Recommendation TR-10-2

Internet Protocol Media Experience (IPMX): Uncompressed Active Video

April 25, 2023

http://www.videoservicesforum.org

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THE FOREGOING NEGATION OF DAMAGES IS A FUNDAMENTAL ELEMENT OF THE USE OF THE CONTENTS HEREOF, AND THESE CONTENTS WOULD NOT BE PUBLISHED BY VSF WITHOUT SUCH LIMITATIONS.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

Internet Protocol Media Experience (IPMX) was created to foster the adoption of open standards-based protocols for interoperability over IP in the media and entertainment and professional audio/video industries. IPMX is based on the SMPTE ST 2110 standard and as such the VSF TR-10 suite of Technical Recommendations is a set of differences between SMPTE ST 2110 and IPMX.

This Technical Recommendation corresponds to the SMPTE ST 2110-20 document and describes the transport of uncompressed active video using RTP protocol in IPMX. It documents the differences between TR-10-2 and SMPTE ST 2110-20. Some of the subjects covered in this document include the payload format, Media Clock, RTP Clock, RTP Timestamps and the IPMX Info Block definition for uncompressed active video.

Table of Contents

1 Introduction (Informative)

IPMX, which stands for IP Media Experience, is based on two families of specifications. The SMPTE ST 2110 Professional Media Over Managed IP Networks suite of standards for the transport of video, audio, and ancillary/control signals over IP networks, and the NMOS REST APIs from AMWA, which provide discovery, connection management, and control.

IPMX is an accessible, open standard that meets the needs of professional and consumer video and audio users in a wide variety of contexts while giving manufacturers and developers what they need to build low-latency, interoperable, IP based audiovisual products or applications.

This Technical Recommendation (TR) covers the IPMX transport of uncompressed active video using the RTP protocol. Other parts of the TR-10 family of Technical Recommendations describe IPMX individual media essence types, along with their requirements, and defines other aspects of the IPMX system.

2 Contributors

The following individuals participated in the Video Services Forum IPMX working group that developed this Technical Recommendation.

Aaron Doughten (Sencore)	Charles Buysschaert (Intopix)	Jean-Baptiste Lorent (IntoPIX)	Phil Nguyen (Nextera)
Alain Bouchard (Matrox)	Chris Lapp (Cisco)	Jed Deame (Nextera Video)	Prinyar Boon (Phabrix)
Albert Faust (Arista)	Clark Williams (Christie Digital)	JJ Eynon (CNN)	Raul Diaz (Intel)
Andre Testa (Matrox)	Daniel BOUQUET (Analogway)	John Belstner (Intel)	Raymond Hermans (Adeas)
Andreas Hildebrand (ALC NetworX)	Danny Pierini (Matrox)	John Dale (Media Links)	Robert Welch (Arista)
Andrew Starks (Macnica)	David Chiappini (Matrox)	John Fletcher (BBC)	Ron Stites (Macnica)
Antoine Hermans (Adeas)	David Mitchinson (Appear TV)	Karl Johnson (Christie Digital)	Tadahiro Watanabe (Macnica)
Arnaud Germain (Intopix)	Gerard Phillips (Arista)	Karl Paulsen (Diversified)	Teiji Kubota (Macnica)
Ben Cope (Intel)	Greg Schlechter (Intel)	Marc Levy (Macnica)	Thomas True (NVIDIA)
Brad Gilmer (VSF)	Greg Stigall (Warner Media)	Mike Boucke (AJA)	Tim Bruylants (intopix)
Bob Ruhl (VSF)	Jack Douglass (PacketStorm)	Paulo Francisco (Evertz)	Wes Simpson (LearnIPvideo)
Cassidy Phillips (Imagine)	Jean Lapierre (Matrox)	Peter Brightwell (BBC)	

3 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.

4 Conformance Notation

Normative text describes elements of the design that are indispensable or contain the conformance language keywords: "shall," "should," or "may."

Informative text 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 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.

5 Normative References

- SMPTE ST 2110-20:2022 Professional Media Over Managed IP Networks: Uncompressed Active Video
- SMPTE ST 2110-10:2022 Professional Media over Managed IP Networks: System Timing and Definitions
- VSF TR-10-1 Internet Protocol Media Experience (IPMX): System Timing and Definitions

6 Definitions

For the purposes of this document, the terms, and definitions of VSF TR-10-1 and those of SMPTE ST 2110-20 apply.

7 General Provisions

All uncompressed video IPMX Senders and Receivers compliant with this TR shall comply with the following specifications:

SMPTE ST 2110-20 Sections 1-5, 6.1.2, 6.1.4, 6.1.5, 6.2, 6.3, and 7, subject to the constraints in this document

IPMX network interface requirements shall be in accordance with the provisions of SMPTE ST 2110-10 section 6, subject to the additional constraints in this document.

All IPMX Media streams shall have a UDP destination port value that is even and that is greater than 1024.

All IPMX Media streams should have a UDP destination port value that is greater than 5000.

Note: The interested reader can refer to RFC 3551 section 8 for a description of the selection of the above port number range.

The active sample arrays of the uncompressed video essence shall be transported using RTP (IETF RFC 3550), subject to the constraints and payload formats defined elsewhere in this document and in the applicable sections of ST 2110-20.

Unless otherwise noted, multi-octet fields within the RTP Header, RTP Payload Header, and RTP Payload shall be transmitted in Network Byte Order (most significant byte first). When

represented in bit-field diagrams such as Figure 1 or Figure 2 of SMPTE ST 2110-20, the most significant bits of multi-bit fields shall occupy the lowest-numbered bit index positions (left-most positions in the figures) and shall be transmitted first.

The UDP size of each RTP packet shall not exceed the Standard UDP Size Limit as specified in SMPTE ST 2110-10.

The image technical metadata necessary to receive and interpret the RTP stream shall be communicated via SDP as defined in section 7 of SMPTE ST 2110-20, subject also to the provisions of TR-10-1.

IPMX Senders shall make their SDP object available through the management programming interface of the device.

Traffic shaping and transmission timing of the RTP stream shall be in accordance with the section 8 (IPMX Sender Timing) of TR-10-1.

8 Payload Formats

Uncompressed video IPMX Receivers shall support media payload sample format of YCbCr 4:2:2 with a bit depth of 10 bits per component and using pixel groups constructed according to table 2 of SMPTE ST 2110-20, and RGB 4:4:4 with a bit depth of 8 bits per component and using pixel groups constructed according to table 1 of SMPTE ST 2110-20.

9 Media Clock, RTP Clock, and RTP Timestamps

The Media Clock and RTP Clock shall comply with the provisions of VSF TR-10-1. The Media Clock and RTP Clock rate for IPMX uncompressed video streams compliant with this TR shall be 90 kHz.

All RTP packets which are part of the same progressive frame shall contain the same RTP Timestamp value.

All RTP packets which are part of the same interlaced field shall contain the same RTP Timestamp value.

10 IPMX Info block definition for uncompressed active video

IPMX Senders shall send RTCP Sender Reports as outlined in TR-10-1. These RTCP Sender Reports shall include an IPMX Info Block extension containing an uncompressed active video Media Info Block.

The format of the IPMX Info Block extension, including an uncompressed active video Media Info Block, shall be as in Figure 1 below.

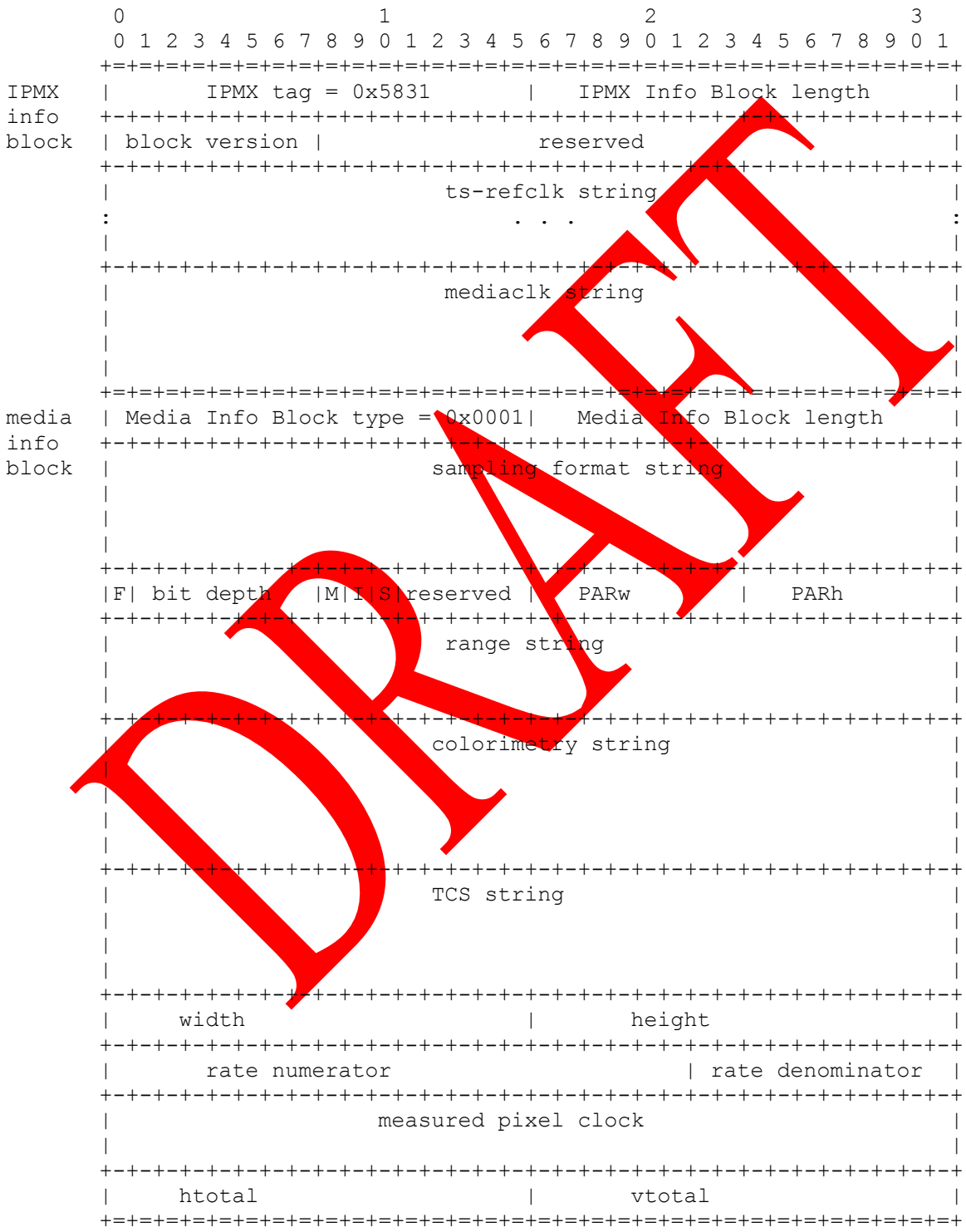


Figure 1 – IPMX Info Block extension for uncompressed active video media

IPMX tag: 16 bits
See TR-10-1 section 8.7 (RTCP Sender Report General Provision).

IPMX Info Block length: 16 bits
See TR-10-1 section 8.7 (RTCP Sender Report General Provision)

block version: 8bits
See TR-10-1 section 8.7 (RTCP Sender Report General Provision)

reserve: 24 bits
See TR-10-1 section 8.7 (RTCP Sender Report General Provision)

ts-refclk string: 64 bytes
See TR-10-1 section 8.7 (RTCP Sender Report General Provision)

mediaclk string: 12 bytes
See TR-10-1 section 8.7 (RTCP Sender Report General Provision)

Media Info Block type: 16 bits
Shall contain the constant 0x0001. This identifies the Media Info Block for Uncompressed Active Video.

Media Info Block length: 16 bits
The length of the Media Info Block, including the header, the Media Info Block content and any padding required to align the Media Info Block on a 32 bits boundary. The value of the Media Info Block length shall be the number of 32-bit words in the Media Info Block minus one.

sampling format string: 16 bytes
Shall correspond to the sample string value as defined in ST 2110-20 section 7.4.1 padded with 0x0 byte values

floating Point (F): 1 bit
Shall be 0 if the video media samples are integer values.
Shall be 1 if the video media samples are floating point values.

bit depth: 7 bits
Shall be set to the bit depth value of the video media samples.

packing mode(M): 1 bit
Shall be 0 if Block Packing Mode as defined in ST2110-20 section 6.3.3 is used.
Shall be 1 if General Packing Mode as defined in ST 2110-20 section 6.3.2 is used.

interlace (I): 1 bit
Shall be 0 if the signal is progressive.
Shall be 1 if the signal is interlaced or progressive segmented frame.

segmented(S): 1 bit
Shall be 1 if the signal is progressive segmented frame and shall be 0 otherwise.

PAR width (PARw): 8 bits
Shall be equal to the PAR width integer value specified in ST 2110-20 section 7.3. If the pixel aspect ratio width is unknown, a value of 1 shall be used.

PAR height (PARh): 8 bits
Shall be equal to the PAR height integer value specified in ST 2110-20 section 7.3. If the pixel aspect ratio height is unknown, a value of 1 shall be used.

range string: 12 bytes
Shall correspond to the range string value as defined in ST 2110-20 section 7.3 padded with 0x0 byte values. If the effective range of the signal is unknown, a value of "NARROW" shall be used.

colorimetry string: 20 bytes
Colorimetry string value as defined in ST 2110-20 section 7.5 padded with 0x0 byte values.

TCS string: 16 bytes
TCS string value as defined in ST 2110-20 section 7.6 padded with 0x0 byte values. If the exact transfer characteristic is unknown, a value of "SDR" shall be used.

width: 16 bits
Width of the active portion of the video

height: 16 bits
Height of the active portion of the video

rate numerator: 22 bits
Frame rate numerator

rate denominator: 10 bits
Frame rate denominator

measured pixel clock: 64 bits
Shall be equal to the measured pixel clock expressed in Hz.

vtotal: 16 bits
Shall be equal to the total number of lines in the video frame (2 fields for interlaced), including all active and blanking periods.

htotal: 16 bits
Shall be equal to the total number of luminance sample periods per video line, including all active and blanking periods.

11 RTCP Sender Report Example (Informative)

The following shows an example of a RTCP Sender Report for an IPMX Video Sender that has a synchronization source identifier (SSRC) value of 3254, sending the RTCP Sender Report for RTP timestamp 610164507 before it sends the first packet for that stream and having the following SDP file:

```
v=0
o=- 1618351493884125000 1618351539175204201 IN IP4 25.25.30.151
s=IP video OUT 1
t=0 0
m=video 10000 RTP/AVP 96
c=IN IP4 239.20.0.1/128
a=source-filter: incl IN IP4 239.20.0.1 25.25.30.151
a=rtpmap:96 raw/90000
a=fmtp:96 sampling=YCbCr-4:2:2; width=1920; height=1080;
exactframerate=60000/1001; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017; TP=2110TPN; IPMX;
measuredpixclk=1485501040; vtotal=1125; htotal=2200
a=ts-refclk:localmac=00-20-FC-32-2F-40
a=mediaclk:sender
```

Referring to Figure 1 from TR-10-1 the value for the different fields of the IPMX RTCP Sender Report corresponding to this example would be:

V = 2
P = 0
RC = 0
PT = 200
length = 50

Note: 50 is the size in 32-bit words of the packet payload - 1 as specified in RFC 3550 section 6.4.1. It is obtained by summing the length in Bytes of RTCP Sender report header (8) plus the length of the sender info (20) plus the length of the IPMX Info Block (2 bytes for IMPX tag, 2 for IPMX Info Block length, 1 for the IPMX Info Block version counter + 3 reserved bytes + 64 for ts-refclk string + 12 for mediaclk string) + the length of the Media Info Block (2 for the Media Info Block type + 2 for the Media Info Block length + 16 for sampling format string + 4 for (F, bit depth, M, I, S, reserved, PARw and PARh) + 12 for range string + 20 for colorimetry string + 16 for TCS string + 2 for width + 2 for height + 4 for rate numerator and denominator + 8 for measured pixel clock + 2 for vtotal + 2 for htotal). Thus, the total size in Bytes is 8+20+(2+2+1+3+64+12) + (2+2+16+4+12+20+16+2+2+4+8+2+2) = 204. To get the amount in 32-bit words we divide by 4 and get 51. And finally subtract 1 to get 50.

SSRC = 3254
NTP timestamp, most significant word = 1665165600
NTP timestamp, least significant word = 262167158

RTP timestamp = 610164507
sender’s packet count = 0
sender's octet count = 0
IPMX tag = 0x5831
IPMX Info Block length = 43

Note: 43 is the size in 32-bit words of the IPMX Info Block – 1 as specified in section 10 of this document. It is obtained by summing the length in Bytes as follows (2 for IMPX tag, 2 for IPMX Info Block length, 1 for the IPMX Info Block version counter + 3 reserved bytes + 64 for ts-refclk string + 12 for mediaclk string) + the length of the Media Info Block (2 for the Media Info Block type + 2 for the Media Info Block length + 16 for sampling format string + 4 for (F, bit depth, M, I, S reserved, PARw and PARh) + 12 for range string + 20 for colorimetry string + 16 for TCS string + 2 for width + 2 for height + 4 for rate numerator and denominator + 8 for measured pixel clock + 2 for vtotal + 2 for htotal). Thus, the total size in Bytes is (2+2+1+3+64+12) + (2+2+16+4+12+20+16+2+2+4+8+2+2) = 176. To get the amount in 32bit words we divide by 4 and get 44. And finally subtract 1 to get 43.

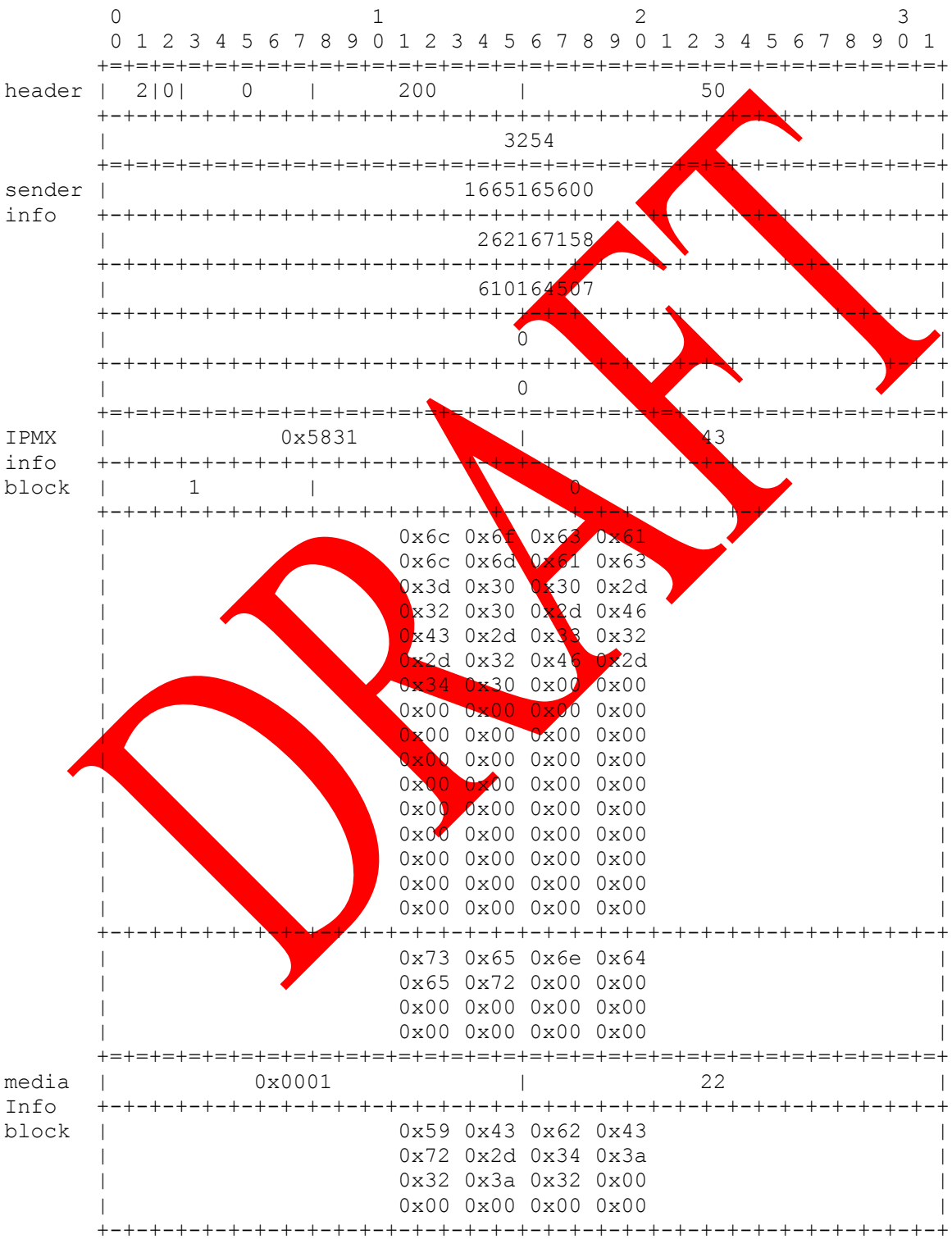
block version = 1
ts-refclk string = “localmac=00-20-FC-32-2F-40”
mediaclk string =”sender”
Media Info Block type = 0x0001
Media Info Block length: 22

Note: 22 is the size in 32-bit words of the Media Info Block – 1 as specified in section 10 of this document. It is obtained by summing the length in Bytes as follows 2 for the Media Info Block type + 2 for the Media Info Block length + 16 for sampling format string + 4 for (F, bit depth, M, I, S reserved, PARw and PARh) + 12 for range string + 20 for colorimetry string + 16 for TCS string + 2 for width + 2 for height + 4 for rate numerator and denominator + 8 for measured pixel clock + 2 for vtotal + 2 for htotal. Thus, the total size in Bytes is (2+2+16+4+12+20+16+2+2+4+8+2+2) = 92. To get the amount in 32bit words we divide by 4 and get 23. And finally subtract 1 to get 22.

sampling format string=” YCbCr-4:2:2”
floating Point (F) = 0
bit depth = 10
packing mode(M)=1
interlace (I)=0
segmented(S)=0
PARw=1
PARh=1
range string=”NARROW”
colorimetry string=”BT709”
TCS string=”SDR”
width=1920
height=1080
rate numerator=60000
rate denominator=1001

measured pixel clock=1485501040
vtotal=1125
htotal=2200

The following figure shows the corresponding RTCP Sender Report data.



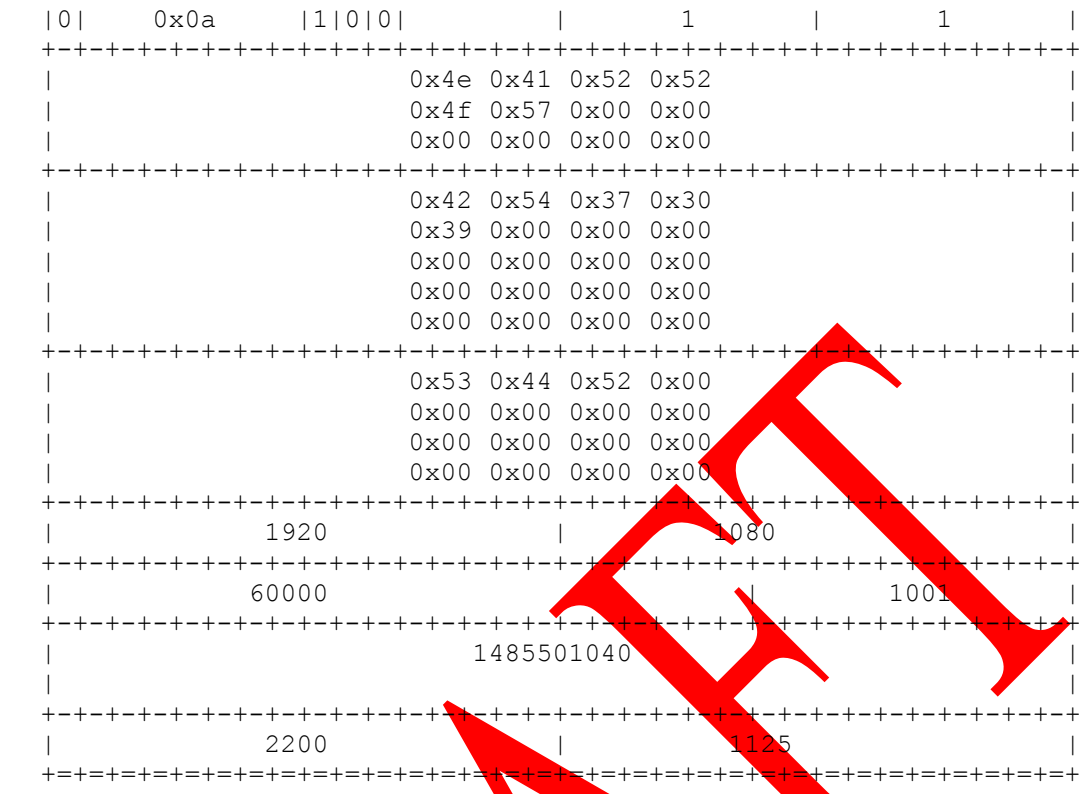


Figure 2 – Uncompressed active video IPMX RTCP Sender Report example