



VIDEO SERVICES FORUM

Video Services Forum (VSF) Technical Recommendation TR-10-15 Part 3

Internet Protocol Media Experience (IPMX):
H.264 Codec Requirements for Compressed Video



June 4, 2026

This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nd/4.0/>

or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.



<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 IMPLEMENTER 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THE FOREGOING NEGATION OF DAMAGES IS A FUNDAMENTAL ELEMENT OF THE USE OF THE CONTENTS HEREOF, AND THESE CONTENTS WOULD NOT BE PUBLISHED BY VSF WITHOUT SUCH LIMITATIONS.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 documents the minimum requirements for an IPMX device with respect to the use of the H.264 video codec in TR-10-7 streaming use-cases.

Table of Contents

1	Introduction (Informative)	4
2	Contributors	4
3	About the Video Services Forum.....	5
4	Conformance Notation.....	5
5	Normative References.....	7
6	Definitions.....	7
7	General Provisions	7
8	Picture	8
9	Packets	8
10	Timing.....	9
11	Random Access Point	9
12	IPMX H.264 Profile.....	10
13	Rate Control	10
14	Advanced Features.....	10
15	Sender Reports	11
16	H.264 Media Info Block	12
16.1	Example	14

1 Introduction (Informative)

IPMX, which stands for Internet Protocol 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 document covers the H.264 video codec requirements for IPMX. Other aspects of the IPMX system and their individual requirements are documented in other parts of this Technical Recommendation.

2 Contributors

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

Alain Bouchard (Matrox)	Jed Deame (Nextera)
Amanda Douglass (Macnica)	Jeroen Steendam (DekTec)
Andre Testa (Matrox)	Kai Hagseth (Bridge Technologies)
Andreas Hildebrand (RAVENNA)	Kenneth Klette Jonassen (Bridge Technologies)
Andrew Starks (Macnica)	Kyle Wilken (Cobalt Digital)
Daniel Bouquet (ANALOG WAY)	Leigh Whitcomb (Meinberg)
Danny Pierini (Matrox)	Phil Bernal (PB Technologies Group)
David Konz (plexusAV)	Phil Nguyen (Nextera)
Hung Lam (EvertzAV)	Raymond Hermans (Adeas)
Jack Douglass (PacketStorm)	Ryan Wallenberg (Cobalt)
Jean Lapierre (Matrox)	Steve Kolta (Meinberg)
Jean-Jacques Osytiguy (Matrox)	

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

- TR-10-1:2024 Internet Protocol Media Experience (IPMX): System Timing and Definitions
- TR-10-7:2024 Internet Protocol Media Experience (IPMX): Compressed Video Transport
- Internet Engineering Task Force (IETF) RFC 3550 RTP: A Transport Protocol for Real-Time Applications available at <https://www.ietf.org/rfc/rfc3550.txt>
- Internet Engineering Task Force (IETF) RFC 6184 RTP Payload Format for H.264 Video available at <https://datatracker.ietf.org/doc/html/rfc6184>
- BCP-004-01 NMOS Receiver Capabilities available at <https://specs.amwa.tv/bcp-004-01>
- BCP-004-02 NMOS Sender Capabilities available at <https://specs.amwa.tv/bcp-004-02>
- BCP-006-02 NMOS With H.264 available at <https://specs.amwa.tv/bcp-006-02>
- H.264 "Advanced video coding for generic audiovisual services" available at <https://www.itu.int/rec/T-REC-H.264>

6 Definitions

For the purposes of this document, the terms, and definitions of VSF TR-10-1 and the following apply.

H.264 Sender	An IPMX Sender configured to produce an H.264 coded stream.
H.264 Receiver	An IPMX Receiver capable of consuming an H.264 coded stream.
NMOS	AMWA Networked Media Open Specifications
CBR	H.264 HRD cbr_flag = 1, as per TR-10-7
VBR	H.264 HRD cbr_flag = 0, as per TR-10-7
SPS	H.264 Sequence Parameter Set
PPS	H.264 Picture Parameter Set
SEI	H.264 Supplemental Enhancement Information
Encoder CPB	An encoder transmit buffer used to queue coded access units prior to transmission for pacing/shaping purposes.

7 General Provisions

An IPMX Sender producing an H.264 coded stream shall comply with the VSF TR-10-7 Technical Recommendation.

An IPMX Sender producing an H.264 coded stream shall comply with the BCP-006-02 specification.

The H.264 coded bitstream produced by an IPMX Sender shall conform to the H.264 specification, as well as the requirements defined in BCP-006-02 and this Technical Recommendation.

An IPMX Receiver shall communicate its capabilities for the “video/H264” media type through BCP-004-01, Receiver Capabilities.

An IPMX Sender shall communicate its capabilities for the “video/H264” media type through BCP-004-02, Sender Capabilities. An IPMX Sender may publish partial capabilities.

Note: Partial capabilities can occur when certain parameters depend on variable external conditions (e.g., detected signal characteristics), and are therefore not fully determinable at all times.

8 Picture

The `vui_parameters()` shall be present in sequence parameter sets (SPS) and shall describe the uncompressed YCbCr video signal encoded in the H.264 coded stream, either directly or as a result of an RGB-to-YCbCr conversion.

The `vui_parameters()` shall have the following flags set to 1: `video_signal_type_present_flag`, `colour_description_present_flag`, `timing_info_present_flag`, and `nal_hrd_parameters_present_flag`. The `vui_parameters()` shall have the `pic_struct_present_flag` set to 1 for interlaced video and the `pic_timing` SEI message shall include a valid `pic_struct`.

A separate alpha channel may be supported. If supported, the alpha channel shall be transported as an independent monochrome bitstream using the KEY 4:0:0 sampling, ALPHA colorimetry and UNSPECIFIED transfer characteristics. Refer to the ST 2110-20 specification for a definition of the KEY, ALPHA and UNSPECIFIED values.

The video may be progressive or interlaced. The PAFF (PicAFF) feature shall not be used by encoders.

9 Packets

UDP/IP packets shall comply with the “RTP Payload Format” and “Packetization Rules” defined in RFC 6184, along with the additional restrictions specified in NMOS BCP-006-02.

A UDP/IP packet shall not contain more than one VCL NAL Unit.

10 Timing

H.264 coded video shall be transmitted and decoded using the HRD transmitter and decoder schedules.

The Virtual Receiver Buffer Model defined in ST 2110 does not apply; therefore, the concept of "gapped" transmission does not exist for with H.264. The traffic shaping mode shall be set to TP=2110TPW and explicitly declared in the fntp attribute of the SDP transport file.

The IPMX Network Compatibility Model applies to the burstiness of the transmission.

Buffering Period SEI messages shall be provided at each recovery point.

Picture Timing SEI messages shall be provided for each access unit. The pic_struct() shall be provided in the pic_timing() for interlaced video.

The timing_info_present_flag shall equal 1 and num_units_in_tick shall be equal to the frame rate denominator, time_scale shall be equal to twice the frame rate numerator.

The nominal removal time $tr,n(n)$ shall be an even number for a progressive byte stream.

max_num_reorder_frames should be 0. The coded sequence may use IDR, I, and P. It may also use B if only using references to pictures that come before the current picture in output order. Decode order shall equal output order.

The HRD parameters shall be specified for the bitrate of the base layer and HRD parameters shall be specified for the base layer. The cpb_cnt_minus1 value of hrd_parameters() shall be 0.

A coded bitstream shall conform to Type II HRD and nal_hrd_parameters_present_flag shall equal 1

The HRD parameters shall be specified in the vui_parameters of the SPS.

The low_delay flag may be set to 1, in which case an encoder may skip frames.

11 Random Access Point

A coded stream shall include a random access point at least once every 5 seconds. Each random access point shall provide the following NAL Units:

- H.264: (IDR) and SPS and PPS, or SEI recovery_point with information about the picture state.
- The use of GDR (Gradual Decoding Refresh) shall be signaled through the recovery point SEI recovery_frame_cnt attribute.

12 IPMX H.264 Profile

An H.264 Sender compliant with the IPMX H.264 Profile shall support producing a bitstream compliant with either a) the High profile using the YCbCr-420 color sampling, and the 8-bit sample depth, or b) the Main profile using the YCbCr-420 color sampling, and the 8-bit sample depth.

An H.264 Sender supporting a monochrome bitstream shall support producing a bitstream compliant with the HighPredictive-444 profile using 4:0:0 chroma subsampling, and the 8-bit sample depth.

An H.264 Sender may support additional profiles, color samplings, and bit depths beyond these minimum requirements.

The Extended profile shall not be used by an encoder. The ASO, FMO, RS, DP, SI and SP features shall not be used by an encoder.

An H.264 Receiver compliant with the IPMX H.264 Profile shall be capable of consuming a bitstream compliant with a) the High profile, using the YCbCr-420 color sampling and 8-bit sample depth, and b) the Main profile, using the YCbCr-420 color sampling and 8-bit sample depth.

An H.264 Receiver supporting a monochrome bitstream shall support consuming a bitstream compliant with the HighPredictive-444 profile, using 4:0:0 chroma subsampling and 8-bit sample depth.

An H.264 Receiver compliant with the IPMX H.264 Profile shall be capable of consuming a bitstream compliant with Level 4.2.

An H.264 Receiver may support additional profiles, levels, color samplings, and bit depths beyond these minimum requirements.

13 Rate Control

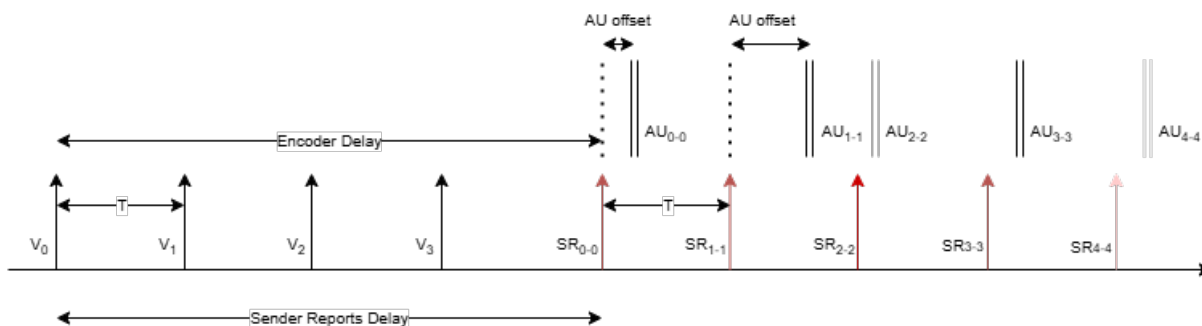
An encoder may produce either an H.264 CBR or VBR bitstream.

A decoder shall support consuming both H.264 CBR and VBR bitstreams.

14 Advanced Features

Annexes F, G, H, I and J of H.264 shall not be used.

15 Sender Reports



The Media Info Block of a sender report shall provide stream parameters that are compliant with the active VPS, SPS and PPS of the coded stream. In all cases, the information from the active VPS, SPS, and PPS prevails over the information from the Media Info Block. Information not provided by the coded stream active VPS, SPS, PPS shall be taken from the Media Info Block.

Section 8.8.2 of VSF_TR-10-1, titled "Video RTCP Sender Report schedule", is modified to replace the paragraph "The RTCP Sender Report shall be sent before the first video media packet of the associated frame or field but after the first video media packet of the previous frame or field." by this new one "The RTCP Sender Report shall be sent before the first video media packet of the associated frame or field, if any" with additional requirements as follow:

- Frames/Fields of video to be encoded are denoted V_i where i is the index in presentation order of the frame/field. For each video frame/field, there is an associated sender report SR_{i-j} where i is the associated video frame/field index in presentation order, and j is the associated access unit index in decoder order. For each video frame/field there is an associated access unit AU_{j-i} where j is the associated access unit index in decoder order and i is the associated video frame/field index in presentation order.
- The nominal period (T) between video frames/fields is a constant value in seconds.
 - The encoder is expected to transmit sender reports at this same nominal interval.
 - Coded access units are expected to be inserted into the encoder Coded Picture Buffer (Encoder CPB) for later transmission at this same nominal interval.
 - The encoder shall start inserting access units into the Encoder CPB after a constant delay ($encoder_delay$) from the time the frame/field is captured from baseband.
 - The $encoder_delay$ shall remain constant once streaming starts, such that the encoder continues to insert access units into the Encoder CPB after the same

constant delay (`encoder_delay`) from the time the frame/field is captured from baseband.

Note: The `encoder_delay` constant does not have to be the same for every stream transmitted by the Sender, nor be the same at each activation of the Sender. It is established when the encoder inserts the first access unit into the Encoder CPB.

- The encoder shall transmit the sender report for a frame/field no more than `encoder_delay` seconds after capturing the frame/field. This means the sender report can be transmitted either immediately upon capturing a frame/field or when the associated access unit is inserted into the Encoder CPB, but always before transmitting the coded access unit. The `sender_reports_delay` shall remain constant once streaming starts, such that the encoder continues to transmit sender reports after a constant delay (`sender_reports_delay`) from the time the frame/field is captured from baseband.

Note: The `sender_reports_delay` constant does not have to be the same for every stream transmitted by the Sender, nor be the same at each activation of the Sender. It is established when the encoder transmits the first sender report.

- Sender reports shall be transmitted in presentation order. If a frame/field is skipped by the encoder, it shall not skip the associated sender report.
- The AU offset is the time difference between when an access unit is inserted into the Encoder CPB and when it starts being transmitted, ranging from zero up to the Encoder CPB size divided by the maximum bitrate.

16 H.264 Media Info Block

An H.264 coded stream shall carry, in the Sender Report, an additional Media Info Block of type 0x000A, immediately following the Media Info Block of type 0x0005 described in TR-10-7.

Media Info Blocks of type 0x000A carry H.264 codec specific- information in the following binary format. Multi-bytes fields are expressed in network byte order (big-endian). The binary structure contains required and optional parameters described in this specification, and possibly additional vendor-specific ones. The parameters shall use the same syntax as in the 'a=fmtp' line of the associated SDP transport file.

Table 1 describes the size of each parameter of the binary structure and its associated flag in the FIELD-PRESENT-MASK which indicate when set to '1' that the parameter value is provided in

the binary structure, and when set to '0' that no value is provided. When a value is not provided for a parameter, the associated bytes in the binary structure shall be 0x00 and for N-byte parameters the length shall be set to 0. For N-byte parameters, the FIELD-PRESENT-MASK bit indicates if the associated length (N) has a valid value, indicating the number of bytes included at the end of the binary structure for the associated parameter. Unused bits of the FIELD-PRESENT-MASK shall be set to 0.

Parameter	Size	FIELD-PRESENT-MASK
profile-level-id	3-byte	bit 0 (least significant bit)
packetization-mode	1-byte	bit 1
sprop-max-don-diff	2-byte	bit 2
sprop-interleaving-depth	2-byte	bit 3
sprop-deint-buf-req	4-byte	bit 4
sprop-init-buf-time	4-byte	bit 5
sprop-parameter-sets	N-byte	bit 6
sprop-level-parameter-sets	N-byte	bit 7
extra bytes	N-byte	bit 8

Table 1

```

+++++
| Media Info Block type = 0x000A|   Media Info Block length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               FIELD-PRESENT-MASK            |
+-----+-----+-----+-----+-----+-----+-----+-----+
|           profile-level-id           |   pkt-mode           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|   sprop-max-don-diff           |   sprop-interleaving-depth   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|           sprop-deint-buf-req           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|           sprop-init-buf-time           |
+-----+-----+-----+-----+-----+-----+-----+-----+
| param-sets-N |l-param-sets-N |   extra-N   |   reserved   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|           sprop-parameter-sets           bytes           |
|   then sprop-level-parameter-sets bytes           |
|   then extra           bytes           |
|                               pad to 4 bytes boundary   |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Media Info Block type 0x000A

There shall be (param-sets-N + l-param-sets-N + extra-N) rounded to a multiple of 4 bytes following the 28th byte of the binary structure.

