SMPTE ST 2110-40:2023
Revision of SMPTE ST 2110-40:2018

ST SMPTE STANDARD

Professional Media over Managed IP Networks: SMPTE ST 291-1 Ancillary Data

Table of contents

445 Hamilton Ave., White Plains NY 10601, (914) 761-1100
Approved December 31, 2023

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in its Standards Operations Manual. This SMPTE Engineering Document was prepared by Technology Committee 32NF.

This revision updates the 2018 original publication, including updating the normative references to current revisions. In this revision, two timing models are defined – one compatible with the original 2018 publication, and another designed for use in lower-latency systems.

A previously published version of this document with the approval date of 2022-06-24 contained typos in the dates of certain normative references pointing at non-existent documents. These typos have been corrected in this version (approval date December 31, 2023).

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Engineering Document. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This clause is entirely informative and does not form an integral part of this Engineering Document.

The capability and capacity of IP networking equipment has improved steadily, enabling the use of IP switching and routing technology to transport and switch video, audio, and metadata essence within television facilities. Existing standards such as SMPTE ST 2022-6 have gained some amount of use in this application, but there was a desire in the industry to switch different essence elements separately.

This family of SMPTE engineering documents builds on the work of Video Services Forum (VSF) Technical Recommendations TR-03 and TR-04, and on AES67, documenting a system for transporting various essence streams over IP networks, capturing the timing relationships between those streams. The system is designed to be extensible to a variety of essence types.

SMPTE ST 2110-10 covers the system as a whole, the timing model, and common requirements across all essence types. Other documents cover specific media essence formats.

SMPTE ST 2110-40 documents the transport of SMPTE ST 291-1 Ancillary Data packets using RTP over an IP network.

1 Scope

This Standard specifies the real-time, RTP payload based transport of SMPTE ST 291-1 Ancillary (ANC) Data packets related to digital video data streams, over IP networks, referenced to a common reference clock.

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, any clause 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; then formal languages; then figures; and then any other language forms.

3 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this engineering document. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this engineering document are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE RP 168:2009 Definition of Vertical Interval Switching Point for Synchronous Video Switching

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

SMPTE RP 291-2:2013 Ancillary Data Space Use — 4:2:2 SDTV and HDTV Component Systems and 4:2:2 2048 × 1080 Production Image Formats

SMPTE ST 2110-10:2022 Professional Media over Managed IP Networks: System Timing and Definitions

SMPTE ST 2110-20:2022 Professional Media Over Managed IP Networks: Uncompressed Active Video

SMPTE ST 2110-21:2022 Professional Media Over Managed IP Networks: Traffic Shaping and Delivery Timing for Video

SMPTE ST 2059-1:2021 Generation and Alignment of Interface Signals to the SMPTE Epoch

IETF RFC 4566 Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", DOI 10.17487/RFC4566, July 2006, https://www.rfc-editor.org/info/rfc4566

IETF RFC 8331 Edwards, T., "RTP Payload for Society of Motion Picture and Television Engineers (SMPTE) ST 291-1 Ancillary Data", DOI 10.17487/RFC8331, February 2018, https://www.rfc-editor.org/info/rfc8331

4 Terms and Definitions

For the purposes of this document, the terms and definitions given in SMPTE ST 2110-10 and the following apply.

4.1
ANC Data Packet

ancillary data (ANC) packet as specified in SMPTE ST 291-1

4.2
Video Payload ID code
VPID_Code

code indicating the digital interface standard as specified in IETF RFC 8331

4.3
HANC

Horizontal ANC data space as specified in SMPTE RP 291-2

4.4
VANC

Vertical ANC data space as specified in SMPTE RP 291-2

4.5
Format Specific Parameter

parameter referred to as a “format specific parameter” in IETF RFC 4566 and signaled in the “a=fmtp” clause of the SDP

5 ANC Data Packet RTP Payload Format

5.1 Overview (Informative)

The television production and distribution ecosystem has developed a large and growing set of data items beyond simple video and audio essence. Rather than developing a custom physical transport for each of these data items, a practice has evolved in which the data items are mapped into ANC Data packets, and transported over the Serial Digital Interface (SDI) in either the HANC data space or the VANC data space. SMPTE ST 291-1 specifies the fundamental structures and properties which pertain to all ANC Data Packets. These Ancillary Data Packets often include time-related non-video essence elements (such as closed captions and subtitles) and also metadata about the video signals such as Active Format Description (AFD).

SMPTE maintains a registry of ANC packet types, which can be found on the SMPTE Registry for Data Identification Word Assignments for Registered DIDs at https://smpte-ra.org/smpte-ancillary-data-smpte-st-291/.

This standard leverages existing practices and definitions by mapping the contents of the ANC Data Packets into RTP, based on IETF RFC 8331.

This mechanism is not intended for the carriage of audio, since there are already well-established mechanisms for doing that; similarly this mechanism is not intended to carry EDH packets.

ANC data packets are contextually related to the video essence which appears in the same SDI signal. While not within the scope of this standard, the systemic implementation of facilities utilizing this transport method could require grouping or coupling of essence streams with ANC data streams to ensure that the contextual relationships are preserved.

5.2 RTP Payload Format

5.2.1 General Requirements

The contents of the ANC Data packets shall be directly mapped into RTP packets as specified in IETF RFC 8331 subject to the additional constraints and guidelines of this standard.

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

While embedded audio packets (including audio control packets) are valid SMPTE ST 291-1 ANC packets, they should not be transmitted using this method. Similarly, EDH packets should not be transported using this mechanism.

NOTE The DID, SDID or DBN, DC, UDW, and CS portions of the ANC packet (as specified in SMPTE ST 291-1) are included in the RTP payload. The Ancillary Data Flag (ADF) specified in SMPTE ST 291-1 is not included in the RTP payload. A single RTP packet can contain zero or more ANC data packets, subject to limits and conditions specified in IETF RFC 8331.

5.2.2 Indication of SDI Location Information

Senders of ANC Data packets may use the Line_Number, Horizontal_Offset, StreamNum, and C fields specified in IETF RFC 8331 to indicate the proposed SDI location information for the ANC Data Packet within a single- or multi-link SDI signal. The Sender may indicate an exact proposed location using these fields, or may specify constraints on the horizontal location or vertical location or both using the special Horizontal_Offset and Line_Number values defined for this purpose.

Senders of Ancillary Data Packets which propose exact SDI interface line number or stream number information (not the 0x7FE or 0x7FF Line_Number codes) shall signal the digital interface standard of the corresponding SDI format using the VPID_Code Format Specific Parameter as specified in IETF RFC 8331. Other Senders need not signal this.

Senders which signal proposed SDI location information shall ensure that the proposed locations of each ANC data packet are in increasing order as the ANC data packets within each frame are transmitted.

NOTE The special Line_Number value 0x7FE indicates a proposed location within the range from the second line after the line specified for switching, as defined in SMPTE RP 168, to the last line before active video, inclusive, while the Line_Number value of 0x7FF indicates that no specific proposal is made.

5.2.3 Receivers which construct SDI outputs

Receivers creating SDI output which contains Ancillary Data Packets shall meet the requirements of SMPTE ST 291-1 when constructing the HANC and VANC spaces. If the VPID_Code associated with the ANC Data packets RTP stream is consistent with the SDI format being created, these receivers should use the proposed SDI Line_Number, Horizontal_Offset, StreamNum, and C-bit information when available, to the extent that it meets the requirements of SMPTE ST 291-1.

When the proposed SDI location information is not specified exactly, or is not applicable to the output SDI format, devices constructing SDI should locate the ANC packets within the VANC space starting 2 lines after the SMPTE RP 168 switching point for the output SDI format, while maintaining compliance with SMPTE ST 291-1, and should comply with any constraints on the horizontal location or vertical location signaled using the special codes defined in IETF RFC 8331.

NOTE The Line_Number and StreamNum values in IETF RFC 8331 refer to digital interface line numbers and stream numbers in an SDI signal format, whereas the SRD Row numbers in SMPTE ST 2110-20 refer to video sample row numbers within the active sample array starting with zero. Any proposed SDI location information being conveyed within the RTP payload header relates to the location within the digital interface signal.

5.3 RTP Clock

The RTP Clock rate shall be 90 kHz. The RTP Clock Offset shall be as specified in SMPTE ST 2110-10.

5.4 RTP Timestamps

SMPTE ST 291-1 defines ANC Data Packets in the context of SDI video signals; as such, there exists a timing relationship between SMPTE ST 291-1 ANC data packets and frames of video essence.

As specified in IETF RFC 8331, the RTP timestamps of ANC Data mapped as specified in this standard are interpreted in a similar fashion to the RTP timestamps of frames of video essence.

Senders shall generate RTP timestamps for the ANC Data packets using the procedures specified for video signals in SMPTE ST 2110-10, such that the RTP Timestamp of the ANC Data is contemporaneous with the related field or frame of the video signal.

5.5 Keep Alive

Senders shall send at least one RTP packet corresponding to each video field, frame, or segment. In the event that no ANC packets are transmitted in relation to a frame (or interlaced field or PsF segment) of video, an RTP packet with an ANC_Count value of zero and a Marker bit set shall be transmitted.

6 Timing Model

6.1 General

IETF RFC 8331 suggests 1 millisecond as a reasonable upper bound between the time when an ANC packet appears in an input SDI signal, and when it is transmitted on the network. This clause defines the transmission timing requirements more exactly.

For the purpose of the models below, the following terms are as defined in SMPTE ST 2110-21:

T_{FRAME}, T_{LINE}, TR_{OFFSET}, $TRO_{DEFAULT}$

6.2 Definitions of Time Offsets

6.2.1 Location-Based Time Offset

For ANC Data Packets which signal proposed SDI locations, within progressive frames or the first field or segment of interlaced or segmented frames, the Location-Based Time Offset (T_{LBO}) shall be defined as the moment in time when the specified location is present on the SDI interface, relative to the moment in time when the most recent alignment point of the prevailing SDI format, as defined in SMPTE ST 2059-1, was present on the SDI interface.

For ANC Data Packets within progressive frames or the first field or segment of interlaced or segmented frames, which signal a proposed SDI line number but not a specific horizontal location, the T_{LBO} shall be taken to be the moment in time of the first sample of the HANC space or VANC space (whichever is signaled by the packet, presuming VANC if unspecified) on the specified line, relative to the moment in time when the most recent alignment point of the prevailing SDI format, as defined in SMPTE ST 2059-1, was present on the SDI interface.

For ANC Data Packets which do not signal a proposed SDI line number, the T_{LBO} is defined as the moment in time of the first sample of the HANC space or VANC space (whichever is signaled by the packet, presuming VANC if unspecified) of the location 2 lines after the SMPTE RP 168 switching point for the output SDI format, relative to the moment in time when the most recent alignment point of the prevailing SDI format, as defined in SMPTE ST 2059-1, was present on the SDI interface.

For the second field of interlaced frames, or the second segment of segmented frames, T_{LBO} shall be calculated as above, but reduced by the Second Field Offset T_{SFO} defined below:

$$T_{SFO} = \left(\frac{T_{FRAME}}{2}\right) + \left(\frac{T_{LINE}}{2}\right)$$

6.2.2 Earliest Packet Time Offset

Multiple ANC Data Packets can be consolidated into a single RTP Packet. The Earliest Packet Time Offset ($T_{EPO(j)}$) shall be defined for each RTP Packet (j) within the stream as the smallest T_{LBO} value of the ANC Data Packets within the RTP Packet.

For the case of RTP Packets containing no ANC Data Packets, for progressive frames or the first field or segment of interlaced or segmented frames, $T_{EPO(j)}$ is defined as the difference between the time of the beginning of the last line of the vertical ancillary data space in the first field if interlaced (first segment if segmented) relative to the moment in time when the most recent alignment point of the prevailing SDI format as defined in SMPTE ST 2059-1 was present on the SDI interface. For the second field or segment of interlaced or segmented frames, $T_{EPO(j)}$ is defined as the difference between the time of the beginning of the last line of the vertical ancillary data space of the second field (or segment) relative to the moment in time when the most recent alignment point of the prevailing SDI format, as defined in SMPTE ST 2059-1, was present on the SDI interface, minus the value T_{SFO} defined in 6.2.1.

6.3 Frame and Field Start Times

SMPTE ST 2110-21 defines the latest allowed moment of transmission of the first video packet of each video frame, for both gapped and linear Packet Read Schedules, as TPR_0 or T_{VD}. For interlaced and Progressive segmented Frame (PsF) signals, the equivalent value for the second field (or second segment) is also defined. Figure 1 shows the relationship between the timing parameters of this standard and the video frame timing using the frame timing values defined in ST 2110-21.

SMPTE ST 2110-21 defines the Format Specific Parameter `TROFF` for indicating the value of TR_{OFFSET} of a video stream. The transmission timing of the RTP Streams defined in this standard is also subject to a value $TR_{OFFSETANC}$ as described below, which is signaled as defined in Clause 7.

The timing definition for ANC data is interrelated with the timing defined for SDI signals, and as such the $TR_{OFFSETANC}$ value of an ANC stream shall be used to signal any non-default offset value of the ANC data relative to the alignment point of the prevailing format. Video and ANC streams from different sources (with potentially different TR_{OFFSET} and $TR_{OFFSETANC}$ values) can be routed to receivers which expect to compose them while reconstructing SDI. To avoid confusion, this standard uses the term $TR_{OFFSETANC}$ to indicate the value associated with (and signaled for) the ANC data stream.

Let Transmitted Frame Start Time (T_{FST}) for progressive frames (or the first field of interlaced frames) be:

$$T_{FST} = T_{AD} - TRO_{DEFAULT} \qquad where\ T_{AD} = (N \times T_{FRAME}) + TR_{OFFSETANC}$$

For the second field of interlaced frames, or the second segment of PsF frames, define the Transmitted Second Start Time (T_{SST}) using the Second Field Offset value defined in Clause 6.2.1:

$$T_{SST} = T_{FST} + T_{SFO}$$

NOTE While the diagram in Figure 1 uses $TRO_{DEFAULT}$ and other values from ST 2110-21 which are particular to a video format, many devices (such as closed caption or subtitle generators) generate ANC data RTP streams which do not have an accompanying video stream. RTP Timestamp values are used to time-align the contents of ANC data RTP streams and other streams.

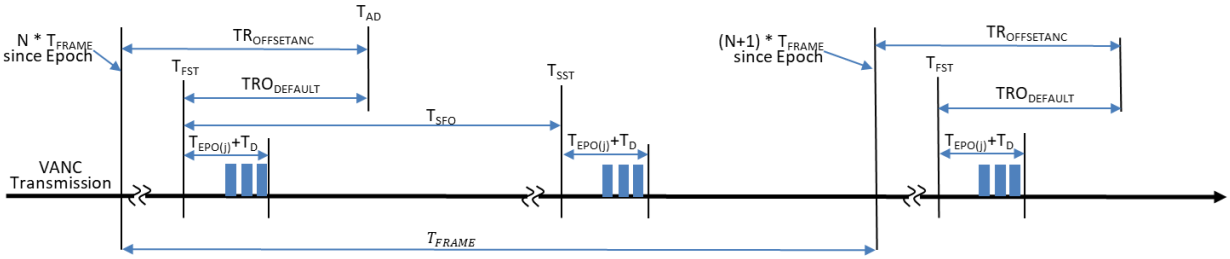

Figure 1 — Relationship of ANC transmission Frame Timing.

6.4 Low-Latency Transmission Model

The Low-Latency Transmission Model (LLTM) is intended for applications which require extremely tightly managed system latency, far tighter than the timing suggested in RFC 8331.

Senders conforming to the LLTM shall transmit each RTP Packet (j) at a time $T_{TRANSMIT(j)}$, no later than $T_{EPO(j)}$ + a format-specific delay value (T_D) after the relevant frame or field start time, or more formally:

$$T_{TRANSMIT(j)} \leq \left(T_{FST} + T_{EPO(j)} + T_D\right) \quad (for\ progressive\ frames\ and\ the\ first\ fields\ or\ segments)$$

$$T_{TRANSMIT(j)} \leq \left(T_{SST} + T_{EPO(j)} + T_D\right) \quad (for\ second\ fields\ or\ second\ segments)$$

T_D shall be determined based on the output SDI video format, as:

$$T_D = \frac{8}{(FrameRate \times TotalLines)}$$

Where *FrameRate* and *TotalLines* are as defined in SMPTE ST 2059-1:2021 Clauses 7.2 and 7.4 (as appropriate).

Senders conforming to the LLTM shall not transmit RTP packets earlier than one frame period before the requirement above.

All Receivers shall be able to receive LLTM streams.

6.5 Compatible Transmission Model (CTM)

The Compatible Transmission Model (CTM) is intended to codify the transmission timing suggested in IETF RFC 8331 and by reference in earlier versions of this standard.

Senders conforming to the CTM shall transmit each RTP Packet (j) at a time $T_{TRANSMIT(j)}$ no later than $T_{EPO(j)}$ + 1 millisecond after the T_{FST} *or* T_{SST}.

$$T_{TRANSMIT(j)} \leq \left(T_{FST} + T_{EPO(j)} + T_D\right) \quad (for\ progressive\ frames\ and\ the\ first\ fields\ or\ segments)$$

$$T_{TRANSMIT(j)} \leq \left(T_{SST} + T_{EPO(j)} + T_D\right) \quad (for\ second\ fields\ or\ second\ segments)$$

where:

$$T_D = (0.001\ seconds)$$

Senders conforming to the CTM shall not transmit RTP packets earlier than one frame period before the requirement above.

All Receivers shall be able to receive CTM streams.

7 Session Description Protocol (SDP)

The SDP object shall be constructed as described in IETF RFC 8331, subject also to the provisions of SMPTE ST 2110-10.

Section 4.1 of IETF RFC 8331 permits the use of Flow Identification ("FID") semantics to group streams within the SDP; such use is inconsistent with the "one SDP object per RTP Stream" provision of SMPTE ST 2110-10 and therefore Flow Identification ("FID") semantics shall not be used under this standard.

Senders implementing this standard shall signal a Format Specific Parameter `SSN` with the value `ST2110-40:2018` unless they are signaling Format Specific Parameter `TM`, in which case they shall signal the value `ST2110-40:2023`. Receivers shall consider a Format Specific Parameter SSN value of `ST2110-40:2021` as equivalent to a value of `ST2110-40:2023`.

Senders implementing the Low-Latency Transmission Model shall signal a Format Specific Parameter `TM` with the value `LLTM` in the SDP.

Senders implementing the Compatible Transmission Model may signal a Format Specific Parameter `TM` with the value `CTM` in the SDP.

If a Format Specific Parameter `TM` is not present in the SDP, Receivers should assume the value `CTM`.

All Senders shall signal the Format Specific Parameter `exactframerate` as defined in SMPTE ST 2110-20:2022 Clause 7.2 to indicate the frame rate related to the ANC data in the stream.

If a Sender is operating with a value of $TR_{OFFSETANC}$ that is different than the default value $TRO_{DEFAULT}$ for the prevailing video format, then the Format Specific Parameter `TROFF` shall be signaled using the same definition as SMPTE ST 2110-21.

Bibliography (informative)

Video Services Forum (VSF) TR-03 “Transport of Uncompressed Elementary Stream Media over IP”, November 2015, http://www.videoservicesforum.org/download/technical_recommendations/VSF_TR-03_2015-11-12.pdf

Video Services Forum (VSF) TR-04 “Utilization of ST 2022-6 Media Flows within a VSF TR-03 Environment”, November 2015, http://www.videoservicesforum.org/download/technical_recommendations/VSF_TR-04_2015-11-12.pdf

SMPTE ST 352:2013 Payload Identification Codes For Serial Digital Interfaces