SMPTE ST 2110-10:2022
Revision of SMPTE ST 2110-10:2017

SMPTE STANDARD

Professional Media over Managed IP Networks: System Timing and Definitions

Table of Contents **Page**

Approved March 28, 2022

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 extends and clarifies elements of the 2017 original publication, including updates to the normative references to reflect current revisions. In addition, this revision includes:

- Improvements to the definitions of the terms Media Clock and RTP Clock
- Making explicit the requirement of receivers to tolerate RTP header extensions
- Improvements to normative language based on the related PICS documents
- Additional definitions and considerations for streams which are not referenced to the Common Reference Clock, including related SDP signaling of the clock source
- Improvement of the definitions for RTP Timestamps of different essence types and origins
- New definitions of system timing concepts including Link Offset Delay and Transmission Delay, and timestamp preservation modes, including syntax for signaling these in the SDP, and a related new informative annex.

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 section 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. This new work encapsulates each production element separately into IP.

This family of SMPTE standards builds on the work of VSF TR-03 and TR-04, and of AES67, documenting a system for inter-networking various essence streams and capturing the timing relationships between those streams. The system is intended to be extensible to a variety of essence types.

This standard covers the system as a whole, the timing model, and common requirements across all essence types. Subsequent parts of this standard document the individual media essence types and their individual requirements as used within this system.

This family of SMPTE standards builds on the work of VSF TR-03 and TR-04, and of AES67, documenting a system for inter-networking various essence streams and capturing the timing relationships between those streams. The system is intended to be extensible to a variety of essence types. The AMWA has developed an interface specification, AMWA IS-05, for managing connections of the streams defined in this standard.

1 Scope

This standard is part of a family of engineering documents that define an extensible system of RTP-based essence streams referenced to a common reference clock, in a manner which specifies their timing relationships.

This standard specifies the system timing model and the requirements common to all of the essence streams, and defines timestamping methods for video streams and audio streams such that time alignment across essence is possible.

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 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; 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.

AES AES67-2018, “AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability”

IEEE Std 1588-2008 “IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", 24 July 2008, DOI: 10.1109/IEEESTD.2008.4579760

IETF RFC 768 Postel, J., "User Datagram Protocol", STD 6, DOI 10.17487/RFC0768, August 1980, https://www.rfc-editor.org/info/rfc768

IETF RFC 791 Postel, J., "Internet Protocol", STD 5, DOI 10.17487/RFC0791, September 1981, https://www.rfc-editor.org/info/rfc791

IETF RFC 2460 Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", December 1998, http://www.rfc-editor.org/info/rfc2460

IETF RFC 3550 Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, DOI 10.17487/RFC3550, July 2003, https://www.rfc-editor.org/info/rfc3550

IETF RFC 3551 Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, DOI 10.17487/RFC3551, July 2003, https://www.rfc-editor.org/info/rfc3551

IETF RFC 3376 Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", DOI 10.17487/RFC3376, October 2002, https://www.rfc-editor.org/info/rfc3376

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 4570 Quinn, B. and R. Finlayson, "Session Description Protocol (SDP) Source Filters", DOI 10.17487/RFC4570, July 2006, https://www.rfc-editor.org/info/rfc4570

IETF RFC 4604 Holbrook, H., Cain, B., and B. Haberman, "Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast", DOI 10.17487/RFC4604, August 2006, https://www.rfc-editor.org/info/rfc4604

IETF RFC 5760 Ott, J., Chesterfield, J., and E. Schooler, "RTP Control Protocol (RTCP) Extensions for Single-Source Multicast Sessions with Unicast Feedback", DOI 10.17487/RFC5760, February 2010, https://www.rfc-editor.org/info/rfc5760

IETF RFC 5771 Cotton, M., Vegoda, L., and D. Meyer, "IANA Guidelines for IPv4 Multicast Address Assignments", BCP 51, DOI 10.17487/RFC5771, March 2010, https://www.rfc-editor.org/info/rfc5771

IETF RFC 6128 Begen, A., "RTP Control Protocol (RTCP) Port for Source-Specific Multicast (SSM) Sessions", DOI 10.17487/RFC6128, February 2011, https://www.rfc-editor.org/info/rfc6128

IETF) RFC 7104 Begen, A., Cai, Y., and H. Ou, "Duplication Grouping Semantics in the Session Description Protocol", DOI 10.17487/RFC7104, January 2014, https://www.rfc-editor.org/info/rfc7104

IETF RFC 7273 Williams, A., Gross, K., van Brandenburg, R., and H. Stokking, "RTP Clock Source Signalling", DOI 10.17487/RFC7273, June 2014, https://www.rfc-editor.org/info/rfc7273

IETF RFC 8285 Singer, D., Desineni, H., and R. Even, Ed., "A General Mechanism for RTP Header Extensions", DOI 10.17487/RFC8285, October 2017, https://www.rfc-editor.org/info/rfc8285

SMPTE ST 272:2004 Formatting AES Audio and Auxiliary Data into Digital Video Ancillary Data Space

SMPTE ST 299-1:2009 24-Bit Digital Audio Format for SMPTE 292 Bit-Serial Interface

SMPTE ST 299-2:2010 Extension of the 24-Bit Digital Audio Format to 32 Channels for 3 Gb/s Bit-Serial Interfaces

SMPTE ST 2022-7:2019 Seamless Protection Switching of SMPTE ST 2022 IP Datagrams

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

SMPTE ST 2059-2:2021 SMPTE Profile for Use of IEEE-1588 Precision Time Protocol in Professional Broadcast Applications

4 Terms and Definitions

For the purposes of this standard, the following terms and definitions apply.

NOTE The terminology used in some of the normative references (particularly those of the IETF) differs from conventional usage within the SMPTE. Readers are reminded that the definitions of this section supersede any similar term in the normative references.

4.1 Device

hardware or software application that can include multiple Senders and Receivers

4.2 RTP Stream

sequence of IP Datagrams constructed in accordance with this family of standards, utilizing the Real-Time Transport Protocol as specified in IETF RFC 3550

4.3 Sender

subsystem within a device which originates one RTP stream (or redundant set of streams) into the Network

4.4 Receiver

subsystem within a device which terminates one RTP stream (or redundant set of streams) from the Network

4.5 Ordinary Clock

PTP (Precision Time Protocol) instance that has a single PTP port in a PTP domain and maintains the timescale used in the domain

4.6 Network

IP datagram transport mechanism with sufficient capacity to deliver the RTP stream from the sender to the Receiver

4.7 Big-Endian Byte Order

convention for the transport ordering of octets within values which require more than one octet to represent, such that the most numerically significant octet is transported first, followed by the remaining octets in order of decreasing significance

4.8 Standard UDP Size Limit

maximum limit on the size of the UDP portion of the IP datagrams when operating in the size range specified in section 6.3

4.9 Extended UDP Size Limit

maximum limit on the size of the UDP portion of the IP datagrams when operating in the size range specified in section 6.4

4.10 Image Sampling Instant

time instant representative of the scene capture time

4.11 RTP Timestamp

data field specified as “timestamp” in IETF RFC 3550 section 5.1 titled “RTP Fixed Header Fields”

4.12 Media Clock

timebase related to the sampling rate (or frame rate in the case of video) of the media within the stream, with a source specified by the `mediaclk` attribute in the SDP, used to advance the RTP timestamps included in RTP packets

4.13 RTP Clock

counter advanced by the Media Clock at the rate specified for the media type, and which is sampled to determine the timestamps included in RTP packets

4.14 Common Reference Clock

network-delivered clock that can be used for synchronization of stream contents across different Devices in the network

4.15 Timestamp Reference Clock

timebase specified with the `ts-refclk` attribute in the SDP, used for time messages inside RTCP sender reports

4.16 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 Textual Conventions

5.1 SDP Parameters and Values

The names and values of SDP Format Specific Parameters within the text of this standard are formatted using a monospaced font (such as Courier) except when they appear in section headings.

6 Network Interface Requirements

6.1 General Requirements

Devices compliant to this standard are interconnected at the data-plane by a Network. Two or more networks in parallel might be used in cases where a set of redundant connections are employed.

The network interfaces of Devices specified in this standard shall support IPv4, wherein streams are transported using IP version 4 as specified in IETF RFC 791. Devices should support IPv6 as specified in IETF RFC 2460.

6.2 Real-Time Transport Protocol (RTP)

All of the streams specified in this standard shall use the Real-time Transport Protocol (RTP) as specified in IETF RFC 3550, and shall conform to the RTP Profile specified in IETF RFC 3551, subject to the restrictions and additional requirements specified in this standard and in the specific media-format standards which define each of the media essence types.

All RTP Streams shall be transported on UDP as specified in IETF RFC 768. UDP does not guarantee reliable data transport and Receivers should be capable of receiving streams with occasional dropped, late or out-of-order packets.

NOTE The UDP header checksum is optional in IPv4, and many IPv4 senders populate zero instead of calculating the UDP checksum, as is permitted by IETF RFC 768. In IPv6, IETF RFC 2460 section 8.1 specifically requires the UDP checksum to be calculated, and updates the defined pseudo-header fields.

RTP Session Multiplexing on the same multicast group/port as specified in IETF RFC 3550 section 5.2 shall not be used in the context of this standard.

RTP Control Protocol (RTCP), as specified in IETF 3550 section 6, may be used in the context of this standard. Senders and Receivers may implement RTCP, and Receivers shall tolerate the presence of RTCP.

All RTP streams shall use dynamic payload types chosen in the range of 96 through 127, signaled as specified in section 6 of IETF RFC 4566, unless a fixed payload type designation exists for that RTP Stream within the IETF standard which specifies it.

RTP Header extensions as defined in IETF RFC3550 sections 3.1 & 3.5.1, and further detailed in IETF RFC 8285 may be used in the context of this standard. Senders and Receivers may use this mechanism, and Receivers shall tolerate the presence of an extended header.

Unless otherwise specified, all multi-octet numeric values expressed in the RTP Header, RTP Payload Headers, and Payloads shall be expressed in Big-Endian Byte Order.

When using redundant streams, the streams shall be generated using the method specified in SMPTE ST 2022-7 and as constrained in section 8.5 of this standard.

6.3 Standard UDP Size Limit

The Standard UDP Size Limit shall be 1460 octets. The UDP Size is reflected in the UDP header, and includes the length of the UDP header (8 octets) and also the RTP headers and data. Senders shall not generate IP Datagrams containing UDP packet sizes larger than this limit unless operating conformant to the optional Extended UDP Size Limit specified in section 6.4. Regardless of the presence or size of any RTP Header Extensions, Senders shall adhere to the UDP size constraints.

All Receivers shall be capable of receiving UDP packets up to the Standard UDP Size Limit.

Senders shall ensure that there are no fragmented IP packets in the egress interface of the Sender, notwithstanding the provisions of IETF RFC 791 which might allow them. Receivers need not reassemble fragmented IP datagrams.

NOTE Annex A provides additional background information about the Standard UDP Size Limit.

6.4 Extended UDP Size Limit

The Extended UDP Size Limit shall be 8960 octets.

Senders may transmit and Receivers may support reception of IP Datagrams up to the Extended UDP Size Limit, subject to the constraints of the specific essence transport standard in use.

Senders operating with UDP Sizes which exceed the Standard UDP Size Limit shall include a Format Specific Parameter `MAXUDP` as specified in section 8.6.

NOTE Annex A provides additional background information about the extended UDP size limit.

6.5 Unicast and Multicast

Senders and Receivers shall support IPv4 multicast transmission and reception (respectively) of streams including IGMP signaling as specified in IETF RFC 3376. Senders shall not transmit media signals on IPv4 multicast addresses within the “Local Network Control Block” nor the “Internetwork Control Block” specified in IETF RFC 5771.

Senders and Receivers shall support IPv4 unicast addressing of streams as specified in IETF RFC 791.

Senders and Receivers should support IPv6 multicast transmission and reception of streams as specified in IETF RFC 2460, including Multicast Listener Discovery Protocol version 2 as specified in IETF RFC 4604.

Senders and Receivers should support IPv6 unicast transmission and reception of streams as specified in IETF RFC 2460.

7 System Timing Model

7.1 Introduction to the System Timing Model (informative)

For streams which share the same clock source (typically the Common Reference Clock), inter-stream synchronization at a common destination relies on comparison of RTP Timestamp values in the RTP packet headers that are transmitted by various Senders.

The Common Reference Clock can be distributed to all participating Senders and Receivers via IEEE Std 1588-2008 Precision Time Protocol. If a Common Reference Clock is unavailable, devices can signal use of an alternative clock source, for example a local reference clock, such that streams which share the alternative clock source can still be synchronized with each other by a mutual receiving Device.

7.2 Distribution of the Common Reference Clock via PTP

A Common Reference Clock, potentially derived from a traceable time source, should be provided and distributed on the network using IEEE Std 1588-2008 Precision Time Protocol (PTP). If a Common Reference Clock is available, the Common Reference Clock should be used as the basis of the Timestamp Reference Clock.

The configurable PTP dataset member `defaultDS.slaveOnly` may be set to prevent an Ordinary Clock from entering the PTP LEAD state as defined in SMPTE ST 2059-2:2020. Ordinary Clocks which are not intended to become the PTP leader should be configured with `defaultDS.slaveOnly` set to `TRUE`. All Ordinary Clock devices containing Senders or Receivers shall have a method or control to allow a user to force the device into a `defaultDS.slaveOnly` equals TRUE state if the device is capable of operating with `defaultDS.slaveOnly` set to FALSE.

All Devices conforming to this standard shall support a Common Reference Clock delivered via IEEE Std 1588-2008 using any message rates allowed by the SMPTE ST 2059-2 PTP Profile. In any system in which there is an expectation to interchange audio streams with AES67 compliant devices, the message rates used in the distribution of the Common Reference Clock should be constrained to simultaneously meet the parametric limits of the Media Profile as specified in AES67.

NOTE AES has issued AES-R16-2016, a technical report regarding the compatibility of parameter ranges between the AES67 Media Profile and SMPTE ST 2059-2.

7.3 RTP Clock Offset

When the Media Clock is directly referenced to the Timestamp Reference Clock (`mediaclk:direct` in the SDP), the `offset` clause in the SDP indicates the RTP Clock value at the epoch (time of origin) of the Timestamp Reference Clock. The value of this `offset` is signaled as specified in section 8.3. In this standard, the `offset` value shall be zero.

NOTE 1 RFC 3550 recommends that the initial value of the RTP timestamp be random. In this standard, we override this with a requirement for a zero-offset relationship to the Timestamp Reference Clock (in the case when the Media Clock is directly referenced to the Timestamp Reference Clock). Receivers designed to maintain compatibility with other RTP implementations might need to comply with the RTP provisions in those RTP standards, specifically the possibility that the offset could be non-zero.

NOTE 2 The requirement of a zero offset value in this standard allows fast restoration of signals after Sender restarts, under the assumption that the sender resumes operation with the same stream parameters as before the re-start. Eliminating the random offset provision of IETF RFC 3550 allows the Receiver to attempt to make use of the signal as soon as the packet stream is restored, without waiting for the systemic propagation of a revised SDP object.

7.4 RTP Clock and Media Clock

The RTP Clock and Media Clock shall advance at uniform rates. The rates of the RTP Clock and Media Clock are as specified or implied by the media payload format.

7.5 RTP Timestamps – General Provisions

The RTP Timestamps of RTP streams are used to synchronize RTP streams from various Senders within an overall production environment. RTP packets contain RTP timestamps as specified for the specific media essence type.

The RTP Timestamps, as specified in IETF RFC 3550 clause 5.1, shall reflect the "sampling instant" of the essence samples contained within the RTP packet, subject to additional clarifications in the sections below.

7.6 RTP Timestamps for Video Streams

7.6.1 Video RTP Timestamps - General

The RTP Timestamps of successive video frames shall advance at regular increments based on the prevailing frame rate, truncating to integer values when necessary. When in conflict, this requirement supersedes the cases in the subsections below.

For interlaced video, the RTP Timestamps of the first field of successive frames shall advance at regular increments based on the prevailing video frame rate, and the RTP timestamp of the second field shall be offset from the RTP timestamp of the first field by one half of the prevailing frame period, truncating to integer values when necessary. For Progressive segmented Frame (PsF) signals, both segments shall have the same RTP Timestamp.

NOTE 1 The video RTP Timestamp is limited in temporal resolution to the values that the RTP Clock rate can convey. Not all frame periods will have an integer relationship with the rate of the video RTP Clock. The frame periods (difference between successive video RTP Timestamps) might not be exactly constant - for example 60/1.001 Hz frame periods effectively alternate between increments of 1501 and 1502 ticks of a 90 kHz clock.

NOTE 2 The "regular increments" provision of this section applies to each video stream individually. Different streams could have different timestamps and a discontinuity could occur when switching between different streams.

7.6.2 Video RTP Timestamps generated by Image Capture Devices

For progressive scan video essence, the RTP Timestamp should reflect the Image Sampling Instant of the progressive image whose samples are contained within the RTP packets. This clause applies even when the progressive essence is transported in a segmented manner.

For interlaced video essence, the RTP Timestamp of the first field should reflect the Image Sampling Instant of the first field whose samples are contained within the RTP packet. The RTP timestamp of the second field shall be as specified in section 7.6.1.

7.6.3 Video RTP Timestamps for synthetic essence including playback

In the case of video playback from storage devices, or generation of video essence via synthetic means, the RTP timestamp of a frame (or of the first interlaced field within a frame) should represent a point in time of $N \times T_{FRAME}$ unless there is a specific production intent to place it differently; in any case it shall not exceed +/- T_{FRAME} from the most recent $N \times T_{FRAME}$ time point. This provision is subject to the general provisions of section 7.6.1 including the "regular increments" provisions, and specifically applies where the media clock is directly referenced to the timestamping reference clock (mediaclk:direct). The RTP timestamp of the second field if applicable shall be as specified in section 7.6.1.

7.6.4 Video RTP Timestamps generated from a Serial Digital Interface (SDI) signal

In the case of encapsulation of the video contained within an SDI signal, the RTP timestamps shall be determined as follows:

- For a progressive video frame, or for the first field of an interlaced frame, the video RTP Timestamp shall be a sample of the value of the video RTP Clock at the Alignment Point of the SDI signal, as specified in SMPTE ST 2059-1 or the appropriate media reference standard. If applicable, the video RTP Timestamp of the second field shall be as specified in section 7.6.1.

- For Progressive segmented Frame (PsF) video signals on SDI, the segment data shall be treated as a single progressive image, and the RTP timestamp of the second segment shall be the same as the first segment.

7.7 RTP Timestamps for Audio Streams

7.7.1 Audio RTP Timestamps – General

The audio RTP Timestamps of successive audio packets shall advance at regular increments based on the audio RTP Clock rate and the audio packet time as specified in section 7.2 of AES67:2018. When in conflict, this requirement supersedes the cases in the subsections below.

7.7.2 Audio RTP Timestamps generated by Audio Capture Devices

For audio capture devices, the RTP Timestamp of audio RTP packets should reflect the sampling instant of the first sample of the audio signal within the audio RTP packet.

7.7.3 Audio RTP Timestamps for synthetic essence including playback

For the case of audio playback or synthetically generated audio essence, the audio RTP Timestamp should represent the time point at which the synthetic essence has the intended time relationship to other essences within the production.

7.7.4 Audio RTP Timestamps for audio signals derived from SDI

For audio essence embedded in SDI as specified in SMPTE ST 299-1, SMPTE ST 299-2, or SMPTE ST 272, the effective sampling instant for the first audio sample of each audio channel related to a frame of video shall be contemporaneous to the video frame RTP Timestamp as determined in section 7.6.4, offset by an amount determined during the de-embedding process. The effective sampling instant of subsequent audio samples for each audio channel shall increase monotonically with each sample. The audio RTP Timestamp of each audio RTP packet shall reflect the effective sampling instant of the first audio sample contained within the audio RTP packet.

NOTE SMPTE ST 299-1 for HD and SMPTE ST 299-2 for 3G signals describes the timing relationship between embedded audio and SDI including the phase offset information. For SD signals, SMPTE ST 272 applies, and some additional skill-of-art is needed to infer the phase information.

7.7.5 Audio RTP Timestamps for audio signals derived from AES3

The RTP Timestamp of an audio packet shall be a sample of the RTP Clock at the X or Z preamble of the first audio sample in the packet.

7.8 Link Offset Delay

Figure 1 shows the demarcation points for calculation of the Link Offset Delay value. For Receivers, a packet j with an RTP Timestamp value equivalent to time $T_{RTP(j)}$ is said to be *Reconstructed* when the media is reconstituted from the received packet and thus available for further use. Let time $T_{REC(j)}$ indicate the time that packet j is reconstructed. Where multiple packets have the same RTP timestamp e.g. in the case of video the expressions below apply to the first packet with that RTP Timestamp.

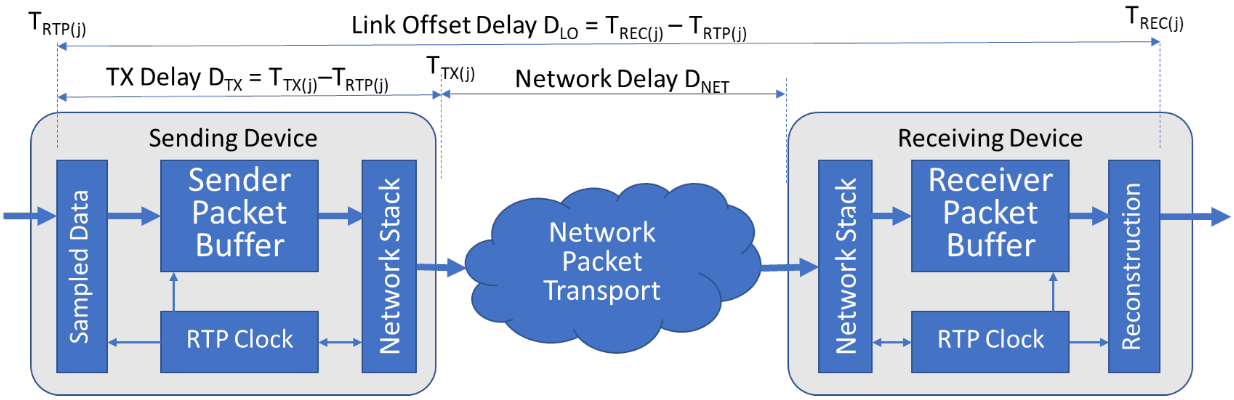

Figure 1 — Link Offset Delay

From a system perspective, let the Link Offset Delay D_{LO} of the Receiver be defined as

$$D_{LO} == T_{REC(j)} - T_{RTP(j)}$$

For any Receiver, the Link Offset Delay D_{LO} should be a constant value, designed or configured into the Receiver. Receivers should document their Link Offset Delay, and provide a means to configure it, through their management API or user interface. Another property of the Link Offset Delay D_{LO} is that at time T_{NOW}, a Receiver reconstructs the packet whose RTP timestamp is equivalent to ($T_{NOW} - D_{LO}$) where T_{NOW} is the time of reconstruction of the first packet with that particular RTP Timestamp as noted above.

The network delay D_{NET} is a function of network topology, queueing, and transport delays, and in some cases can be very small. Likewise the transmission delay D_{TX} is a function of the internal architecture of the Sender and can be very small. For successful reconstruction, the packet j must be received by the receiving device no later than time $T_{REC(j)}$. Since the Transmission Delay D_{TX} and Network Delay D_{NET} values can both be very small, the packet containing media sample j could arrive at the Receiver as early as $T_{RTP(j)}$. By implication, the Receiver packet buffer needs to be able to hold at least the number of packets that could arrive during the time period D_{LO}.

As illustrated in Figure 1, the instant of transmission of packet j from the sender is denoted $T_{TX(j)}$. The Transmission Delay D_{TX} of a sender is defined as the typical delay between the RTP timestamp of a packet and its actual transmission instant, where the packet in question is the first packet with that RTP Timestamp.

$$D_{TX} = T_{TX(j)} - T_{RTP(j)}$$

The Transmission Delay value of a Sender should be signaled in the SDP as defined in section 8.7.

NOTE When the RTP Timestamp is equal to the sampling instant, the Link Offset Delay definition above is equivalent to the Link Offset definition in AES67.

7.9 RTP Timestamps of Derived Signals

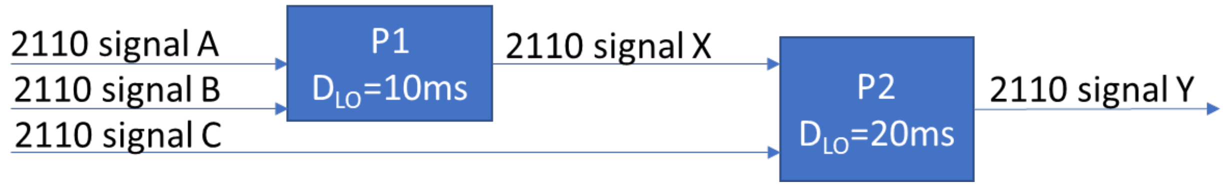

Figure 2 — Example of signal processing workflow

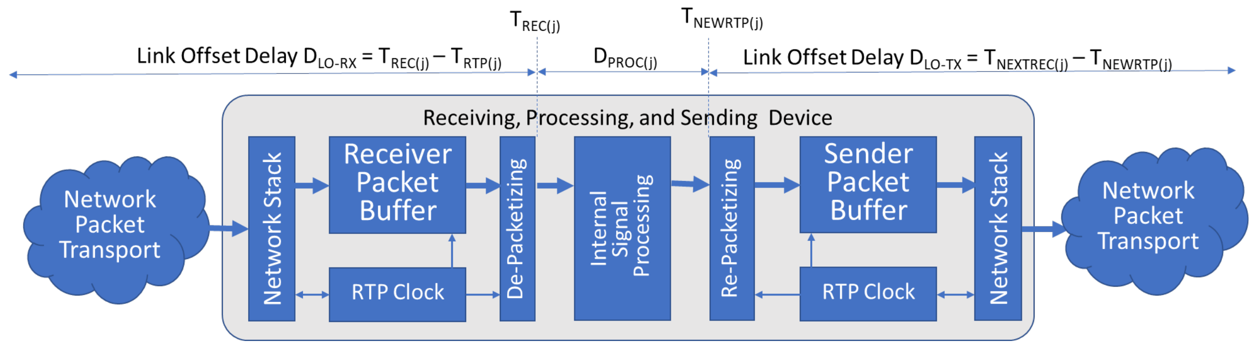

Figure 3— Example of a Receiver-Processing-Sender Device

As noted in previously, the RTP Timestamp generally indicates the sampling time of the input signal in a live environment. The RTP timestamp values are the means of time-aligning signals which were contemporaneously sampled when further signal processing operations require time alignment between signals. Figure 2 illustrates an example where signal X is created by a time-aligning processing function applied to input signals A and B, and a subsequent processing function is applied which needs to time-align signal X with original signal C. As defined in section 0, the Link Offset Delays D_{LO} of receivers P1-A and P1-B are quasi-constant. Sample j of signal X depends on samples j of signals A and B, such that each sample $X_{(j)}$ can be calculated no earlier than time ($T_{RTP(j)} + D_{LO(P1)}$).

Figure 3 details the Receiver, processing, and Sender aspects of a device such as P1 and P2 from the signal processing chain of Figure 2 for the purposes of illustrating the stream timestamp determination.

In the event that the input stream of a processing function is known to be timestamped in a manner representative of the original sampling time (as signaled by `TSMODE=SAMP` in the incoming SDP) it is advantageous to preserve the time-aligning characteristic of the RTP timestamp from input to output – in which case inline processing devices such as P1 and P2 in Figure 2 with relatively small D_{LO} and D_{PROC} values should mark the samples of their derived output signal with the same RTP timestamp as the incoming signal(s) which compose into that output media sample, so that subsequent processing can time-align the processed signal with contemporaneous samples of other signals. In the terms of the figure, $T_{NEWRTP(j)} = T_{RTP(j)}$. Pursuant to section 8.7 of this document, such devices shall signal `TSMODE=SAMP` in their SDP; the Transmission Delay D_{TX} value of such a device reflects the incurred delay of the signal, which includes D_{LO} of the Receiver section, plus D_{PROC} of any internal processing, plus any intrinsic transmission delay of the sender subsystem. Devices which signal `TSMODE=SAMP` shall also signal their Transmission Delay value in the SDP as indicated in section 8.7.

Alternatively, if the D_{LO} or D_{PROC} value of the processor is impractically long, or the input signals cannot be time-aligned or arrive later than the D_{LO} value of the processor, or the input signal is not marked with `TSMODE=SAMP` in the incoming SDP, then the sender shall determine the RTP Timestamps of the samples of the derived output signal as if it were a new signal "sampled" at the current time ($T_{NEWRTP(j)} = T_{NOW}$) and should indicate the intrinsic sender Transmission Delay D_{TX} in the SDP, along with `TSMODE=NEW`.

Notwithstanding the foregoing provisions of this section, there are cases where a control system or a user may have additional information about the RTP Timestamping behavior of signals within an overall system. In order to facilitate incorporation of this additional information into the overall behavior of devices, devices should have a capability to be directed through their management API or user interface to override the incoming SDP information about `TSMODE` (or the lack of information leading to a `TSMODE=NEW` presumption) and be forced to treat an incoming signal as if it had a specific value of `TSMODE` or `TSDELAY`.

Senders which preserve the RTP timestamp values from their input to output as described in this section shall be referred to as *time-preserving*, while senders which create new RTP timestamps shall be referred to as *time-resetting*.

In all cases, the "regular increments" requirements of section 7.6.1 or 7.7.1 shall take precedence over the behaviors described in this section.

8 Session Description Protocol (SDP)

8.1 General

Devices which contain one or more Senders shall construct one SDP session description as specified in IETF RFC 4566 for each RTP stream (or redundant pair of streams) that is originated by that device. These SDP objects shall be made available through a management API of the device.

Devices which contain one or more receivers shall provide a capability to ingest and act upon an SDP object created in accordance with the requirements of this standard. These SDP objects shall be communicated through a management API of the device.

An example SDP is shown in Annex B.

8.2 Timestamp Reference Clock Signaling

All stream descriptions shall have a `ts-refclk` attribute as specified in IETF RFC 7273 section 4. Devices which are referenced to IEEE Std 1588-2008 shall use the `ts-refclk:ptp` form, signaling either the grandmaster clockIdentity and domain number, or signaling that the PTP is traceable.

Devices shall signal that the PTP is traceable in the SDP if the following conditions are all met:

- the PTP timescale is in use
- the current grandmaster timeTraceable field indicates traceable
- the current grandmaster clockAccuracy field indicates accuracy to within 250ns or better

Devices which are not referenced to IEEE Std 1588-2008 shall use an appropriate `ts-refclk` format as specified in IETF RFC 7272 or the extended form shown below, indicating the MAC address of the Sender using the token `localmac`. Receivers may assume that different streams which signal the same value for `localmac` are using the same Timestamp Reference Clock.

```
a=ts-refclk:localmac=<Ethernet MAC address of sender>
```


The following examples show the PTP forms and the localmac form:

```
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:37

a=ts-refclk:ptp=IEEE1588-2008:traceable

a=ts-refclk:localmac=7C-E9-D3-1B-9A-AF
```

NOTE 1 The first PTP example above signals that the Sender is using a PTP clock conforming to IEEE Std 1588-2008, the clockIdentity of the grandmaster is `39-A7-94-FF-FE-07-CB-D0`, and the domain number is 37. ClockIdentity is expressed in EUI-64 format, which is a sequence of hexadecimal values, while the PTP domain number is expressed as a decimal number. The second PTP example indicates that the PTP source is traceable as specified in IETF RFC 7273 section 4.7. The third example shows the case of a local timebase signaled by an Ethernet MAC address of the Sender in EUI-48-format.

NOTE 2 The 2017 version of this document erroneously indicated a syntax for declaring traceable PTP which excluded the IEEE Std 1588-2008 clause. While this revision corrects the error, receivers might tolerate the 2017 construction.

Additional requirements of the SDP object can be specified in the media essence-specific documents.

8.3 Media Clock Signaling

All stream descriptions shall have a media-level `mediaclk` attribute as per IETF RFC 7273 section 5. If the Media Clock is directly derived from the Timestamp Reference Clock, the `direct` reference shall be used and the offset shall be included. In this case, as per clause 7.3 of this document, the offset will be zero. For example:

```
a=mediaclk:direct=0
```

If the Media Clock is asynchronous with respect to the Timestamp Reference Clock, for example if the input media stream for a Sender is not locked to the Common Reference Clock, the following form shall be used:

```
a=mediaclk:sender
```

NOTE If RTCP Sender Reports are being sent, they will provide paired samples of Timestamp Reference Clock and RTP Clock which could be used to estimate relationship between the two clocks.

8.4 Source Address Signaling

Senders should indicate the source address information for streams within the SDP in order to support source-specific multicast sessions by use of an inclusive source filter line as defined in IETF RFC 4570, for example:

```
a=source-filter: incl IN IP4 233.252.0.2 198.51.100.1
```

If RTCP is in use within a source-specific multicast session, the SDP should also follow the provisions of RFC 5760 and RFC 6128.

8.5 Signaling for Duplicate RTP streams

Duplicate RTP streams meeting the requirements of SMPTE ST 2022-7 may be used for redundant transmission to achieve higher system availability. Senders which transmit these duplicated RTP streams, using the mechanisms of Separate Source Addresses (IETF RFC 7104, Section 4.1) or Separate Destination Addresses (IETF RFC 7104, section 4.2) shall signal the RTP duplication using the session level `group` attribute of IETF RFC 5888 and the duplication grouping `DUP` semantics of IETF RFC 7104.

Redundant streams shall not use both identical source addresses and identical destination addresses at the same time.

NOTE SMPTE ST 2022-7 allows the above methods of specifying duplicate RTP streams, but also allows for RTP streams with duplicated source and destination addresses (on separate physical networks); such a mechanism cannot be represented with SDP, and therefore the use of duplicate source and destination addresses is not supported by this Standard.

8.6 UDP Datagram Size

As specified in sections 6.3 and 6.4, Senders may choose to operate within the Standard UDP Size Limit or the Extended UDP Size Limit.

Senders operating with UDP Sizes which exceed the Standard UDP Size Limit shall include a Format Specific Parameter `MAXUDP` with a decimal value indicating the largest UDP Datagram Size (in octets) that might be present in the stream.

If the `MAXUDP` parameter is not present, Receivers shall assume the Standard UDP Size Limit specified in section 6.3.

NOTE The Format Specific Parameter refers to the maximum UDP datagram size – the Ethernet MTU size also includes the IP header in addition to the UDP packet itself.

8.7 RTP Timestamp Mode and Delay

Format Specific Parameter `TSMODE` is defined to indicate the relationship of the RTP timestamps to the content sampling instant or production timeline. If the `TSMODE` parameter is not present, the receiver shall presume a value of `TSMODE=NEW`. Allowed values are:

`TSMODE=SAMP` The RTP timestamp indicates the effective sampling instant of the media (or an equivalent value) and is suitable for time-alignment purposes across multiple essence flows.

`TSMODE=NEW` The RTP timestamp has been created anew at the egress of this sender based on the contemporaneous value of the sender's RTP Clock.

`TSMODE=PRES` The RTP timestamp has been preserved from an input signal, however the input signal did not indicate a value of `TSMODE=SAMP`.

Format Specific Parameter `TSDELAY` is defined to signal the Transmission Delay D_{TX} of senders as defined in section 0. The time value is represented as a decimal positive integer number of microseconds. If the `TSDELAY` parameter is not present, the receiver shall take a receiver-dependent action.

Senders which produce signals in accordance with sections 7.6.2, 7.6.3, 7.6.4, 7.7.2, 7.7.3, 7.7.4, or 7.7.5 should include `TSMODE` with a value of `SAMP` and also should include `TSDELAY` with a value representative of their transmission delay D_{TX} in their SDP if (and only if) their RTP timestamps actually reflect the sampling instant or equivalent as defined in the relevant section.

Senders specified in section 7.9 as *time-preserving*, when acting upon input signals with a value of `TSMODE=SAMP`, shall include in their SDP the `TSMODE` parameter with a value of `SAMP`, and `TSDELAY` with a value representative of their inclusive transmission delay D_{TX} in the SDP. In any other case, such a time-preserving sender shall include `TSMODE=PRES` and an appropriate `TSDELAY` value.

Senders specified in section 7.9 as *time-resetting* should include `TSMODE` with a value of `NEW`, and should include `TSDELAY` with a value representative of their transmission delay D_{TX} in their SDP.

Additional information regarding the `TSMODE` and `TSDELAY` parameters can be found in Annex C.

Annex A Datagram Size Limits (Informative)

This family of documents defines RTP essence formats, and mandates their transport over UDP and IP. The most common method for transporting IP datagrams within facilities is Ethernet, as specified in the IEEE 802.3 family of standards.

The length of an IP datagram is limited only by the representable values within the IPv4 or IPv6 header fields. However in practice the underlying transport mechanism imposes more significant limits on the datagram sizes.

For technical reasons relating to certain older variants of the Ethernet system, the payload of the Ethernet Frame is officially limited to a maximum of 1500 octets within the IEEE 802.3 family of standards. When IP datagrams are transported over Ethernet, this limits the size of the IP datagram, including the IP, UDP, and RTP headers and data, to a total of 1500 octets.

The IPv4 standard header is 20 octets long, while the standard IPv6 header is 40 octets long. In order to accommodate either standard, and to simplify the case of in-network mappings between the two IP standards, the larger of the two values is assumed. Thus the “standard” UDP datagram size limit specified in section 6.3 of this standard is 1500-40 = 1460 octets.

While not strictly allowed within the IEEE 802.3 family of standards, support for so-called “Jumbo” Ethernet frames have been a common feature of Ethernet networking equipment for many years, with an industry consensus value of 9000 octets as the maximum payload length within the Ethernet “Jumbo” frame.

The 8960 octet limit in section 6.4 is based on a 9000 octet Ethernet “Jumbo” frame payload size, and accommodates IPv4 or IPv6 headers.

Ethernet frames have a minimum payload size of 46 octets, and in the rare case of an IP datagram smaller than 46 octets the payload is zero-padded in the process of mapping IP datagrams into Ethernet frames. There is no need to pad up the IP datagram to 46 octets in the RTP protocol.

When contemplating the potential throughput of network interfaces, it is important to remember that the nominal bit-rates of the interfaces, such as “Ten Gigabit Ethernet” or “Twenty-Five Gigabit Ethernet” refer to the bit rate of the Ethernet Frames.

Each Ethernet Frame contains the following “overhead” which must be accounted for in any calculation of the IP throughput of an Ethernet connection:

18 octets	Layer-2 Ethernet frame header and FCS (without an 802.1Q VLAN tag)
4 octets	Optional 802.1Q VLAN tag if present
20 octets	Fixed preamble, start-of-frame delimiter, and minimum inter-packet gap

As an example, the throughput of Ethernet payload data for maximum “standard” sized UDP datagrams specified in section 6.3 (assuming IPv4 and an 801.2Q tag) on a 10 Gigabit Ethernet link is:

Ethernet-Frames/sec = 10,000,000,000 / (8 * (1460 + 20 + 18 + 4 + 20)) = 821,287.7792

IP bits/sec = Ethernet-Frames/sec * IP bits/frame = 821,287.7792 * (1480 * 8) = 9,724,047,306

The above calculation assumes minimum inter-packet gap, an 801.2Q tag, and assumes that all of the IPv4 datagrams are 1480 octets (1460 payload + 20 IPv4 header).

Annex B SDP Example (Informative)

The following example SDP for a video stream including the fields described in this standard.

```
v=0
o=- 123456 11 IN IP4 192.168.100.2
s=Example of a SMPTE ST2110-20 signal
i=this example is for 720p video at 59.94
t=0 0
a=recvonly
a=group:DUP primary secondary
m=video 50000 RTP/AVP 112
c=IN IP4 239.100.9.10/32
a=source-filter: incl IN IP4 239.100.9.10 192.168.100.2
a=rtpmap:112 raw/90000
a=fmtp:112 sampling=YCbCr-4:2:2; width=1280; height=720;
exactframerate=60000/1001; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017; TSMODE=SAMP; TSDELAY=0
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:37
a=mediaclk:direct=0
a=mid:primary
m=video 50020 RTP/AVP 112
c=IN IP4 239.101.9.10/32
a=source-filter: incl IN IP4 239.101.9.10 192.168.101.2
a=rtpmap:112 raw/90000
a=fmtp:112 sampling=YCbCr-4:2:2; width=1280; height=720;
exactframerate=60000/1001; depth=10; TCS=SDR; colorimetry=BT709;
PM=2110GPM; SSN=ST2110-20:2017; TSMODE=SAMP; TSDELAY=0
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:37
a=mediaclk:direct=0
a=mid:secondary
```

This SDP reflects a session ID of 123456 and a session version of 11. The session name is indicated in the s= clause, with additional session information in the i= clause. The source addresses of the primary and secondary streams are indicated in the a=source-filter clauses for each stream.

The t=0 0 clause indicates that this session is permanent (has no begin or end time).

The a=group:DUP clause is as specified in section 8.4 of this standard, indicating two RTP Streams are sent, tagged primary and secondary inside this SDP.

The first m= section describes the primary RTP Stream, which is transmitted via IPv4 on group 239.100.9.10 and UDP port 50000 from source address 192.168.100.2. Each m= line signals the start of a “media-specific section” within the SDP.

The a=fmtp clause contains a number of Format Specific Parameters specified in the media-specific document.

The a=ts-refclk clause is as specified in section 8.2 of this standard. The a=mediaclk:direct=0 clause signals that the media clock is directly referenced to the clock in the ts-refclk clause, and the offset of 0 is as mandated in section 7.3 of this standard. The a=mid:primary section tags this media section as the “primary” stream within the grouping semantics.

The second media section (starting with the second m= line) documents the same information for the secondary stream. This secondary stream is on multicast group 239.101.9.10, UDP port 50020, and originates from source address 192.168.101.2.

Note that when utilizing SMPTE ST 2022-7 hitless reconstruction, the RTP headers and RTP payloads need to be identical between the two RTP streams, because packets can be selected from either stream at any time.

Annex C Timestamp Methodology Notes (Informative)

This revision introduces signaling within the SDP to indicate to the control system information about the creation of the RTP timestamps. In particular, the `TSMODE` and `TSDELAY` parameters are added to the SDP. Both of the new parameters have default values that are meant to be consistent with the 2017 original version of this standard. Section 8.7 provides the normative definitions of these new parameters and some requirements around their use.

The conceptual workflow of an ST 2110-10 production environment, when using these new parameters is as follows:

- Media Sampling devices such as audio A/D converters and video cameras mark their output samples as `TSMODE=SAMP`. This indicates that they approximate the actual sampling instant, and that the timestamp values can be used across signals similarly marked when mixing audio or video together.

- Media Playback devices such as disk recorders or graphics stores can mark their output samples as `TSMODE=SAMP`, again indicating that these samples can be freely mixed with live audio and video samples directly from sampling devices.

- Media Conversion devices which encapsulate already-sampled essence sources (such as SDI or AES3) often mark these encapsulated samples as `TSMODE=NEW` (the default value if unspecified); however such an encapsulator could support precompensation for upstream delays (configured manually or through a control system) and if so could mark the samples as `TSMODE=SAMP` if the compensated delay value approximates the sampling instant.

- Media Processing devices which derive output RTP Streams from ST 2110 input RTP Streams can mark the output signals as `TSMODE=SAMP`if the input signals are marked as `TSMODE=SAMP` provided the introduced delay is small.

- Media processing devices which need to time-correlate across multiple input RTP Streams can do so based on the RTP Timestamp values, for inputs which are marked `TSMODE=SAMP`. If the inputs are unmarked, or marked `TSMODE=NEW` or `TSMODE=PRES`, then the control system might need to provide additional information about the accumulated delay of the incoming streams in order for a device to perfectly correlate the incoming signals together.

In addition to the `TSMODE` parameter, the new parameter `TSDELAY` is introduced. This parameter, along with the Link Offset Delay valued defined in section 7.8, can be used by control systems to infer the delays accumulated along signal processing paths.

Bibliography (Informative)

AMWA IS-05 NMOS Device Connection Management Specification (version 1.0) https://specs.amwa.tv/is-05/releases/v1.0.0/

AES AES-R16-2016 "AES Standards Report - PTP parameters for AES67 and SMPTE ST 2059-2 interoperability"

SMPTE ST 2022-6:2012 Transport of High Bit Rate Media Signals over IP Networks (HBRMT)

VSF TR-03 "Transport of Uncompressed Elementary Stream Media over IP", November 12, 2015, http://www.videoservicesforum.org/download/technical_recommendations/VSF_TR-03_2015-11-12.pdf

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

