SMPTE STANDARD

Professional Media Over Managed IP Networks: Timing of ST 2022-6 Streams in ST 2110-10 Systems

Table of Contents

Approved February 22, 2019

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.

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.

The SMPTE ST 2110 family of engineering documents builds on the work of Video Services Forum (VSF) Technical Recommendations TR03 and TR04, 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.

This specification documents the use and constraints of use of SMPTE ST 2022-6 signals within SMPTE ST 2110-10 systems. In particular it constrains the use of ST 2022-6 such that the timestamps of ST 2022-6 signals can be related to the timing architecture of ST 2110-10, so that video, audio, and metadata can be embedded and dis-embedded from the ST 2022-6 signal while maintaining timing alignment.

1 Scope

This standard specifies the use and constraints for SMPTE ST 2022-6 streams in conjunction with the timing model of SMPTE ST 2110-10, including definition of a Synchronizing Timestamp derived in the ST 2022-8 receiver to relate the two standards. It also documents the SDP announcement of SMPTE ST 2022-5 FEC and SMPTE ST 2022-7 redundancy for these streams.

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.

SMPTE ST 2110-10:2017 Professional Media over Managed IP Networks: System Timing and Definitions

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

SMPTE ST 2022-5:2013 Forward Error Correction for Transport of High Bit Rate Media Signals over IP Networks (HBRMT)

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

SMPTE ST 2022-7:2013 Seamless Protection Switching of SMPTE ST 2022 IP Datagrams

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

Internet Engineering Task Force (IETF) RFC 6364 Session Description Protocol Elements for the Forward Error Correction (FEC) Framework Multicast [online, viewed 2017-11-20] Available at https://www.ietf.org/rfc/rfc6364.txt

4 Terms and Definitions

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

4.1 Synchronizing Timestamp

Derived timestamp value used for the purpose of synchronization of the video essence within the SMPTE ST 2022-6 payload with other essence streams

5 Synchronizing SMPTE ST 2022-6 Essence within a SMPTE ST 2110-10 System

5.1 Overview (Informative)

SMPTE ST 2110-10 specifies a timing architecture in which various essence flows can be synchronized to a common network clock. This document defines and constrains SMPTE ST 2022-6 streams such that their video, audio, and Ancillary data payloads can be synchronized with video, audio, and ancillary data streams from SMPTE ST 2110-10 systems.

ST 2022-6 streams contain a different RTP timestamp in each packet of the stream, which differs from ST 2110-20 in which there is a single RTP timestamp for the entire frame (or field) of video. In order to relate the timestamps of the ST 2022-8 signals with essence in a ST 2110-10 system, this document defines a single reference value for each frame (or field), called the Synchronizing Timestamp. For the purposes of synchronization to other essences, this Synchronizing Timestamp represents the “time” of the video frame (or field).

5.2 General Requirements

Streams shall comply with the provisions of SMPTE ST 2022-6, as further constrained herein.

All SMPTE ST 2022-6 streams compliant to this standard shall comply with the Real-time Transport Protocol provisions of SMPTE ST 2110-10.

5.3 Synchronizing Timestamps, Media Clock, and RTP Clock

For SMPTE ST 2022-6 streams compliant to this standard, the Media Clock and RTP Clock shall comply with SMPTE ST 2110-10.

In accordance with SMPTE ST 2022-6, the Media Clock and RTP Clock rates shall be 27.0 MHz.

The SDI signal encapsulated within the SMPTE ST 2022-6 signal shall be synchronized to the Media Clock.

The last ST 2022-6 RTP packet of a frame has the "M" bit (Marker bit) set; the First ST 2022-6 RTP Packet of a frame is the packet immediately following the last ST 2022-6 RTP packet of the previous frame. The Synchronizing Timestamp of a progressive-scan SDI frame shall be equal to the RTP timestamp of the First ST 2022-6 RTP Packet of that frame, plus the offset specified below. The Synchronizing Timestamp of the first field of an interlaced SDI frame shall be calculated in the same manner as a progressive frame.

The First ST 2022-6 RTP Packet of a frame contains the EAV(s) at the start of the first line(s) of the frame. The Synchronizing Timestamp of the frame (or first field in the case of interlace) shall be equal to the RTP timestamp of the first packet of the frame plus an offset equal to the difference between the time of the EAV(s) at the beginning of the Frame, and the time of the alignment point as defined in ST 2059-1 or the defining SDI document for SDI formats not captured in ST 2059-1. Using the values of P, HA, and SR from the relevant row of Table 2 or Table 4 of SMPTE ST 2059-1 (or of the defining SDI document):

$$Timestamp_{SYNCHRONIZING} = Timestamp_{RTPFirstOfFrame} + INT\left(((P - HA)/SR) \times 27{,}000{,}000\right)$$

Annex A contains additional information about the relationship between the alignment point and the First ST 2022-6 RTP packet of the frame.

The Synchronizing Timestamp of the second field of an interlaced frame shall be calculated as the mid-point between the Synchronizing Timestamp values of the surrounding first fields, rounding down when necessary.

All of the samples of a progressive video frame or of an interlaced field in the case of interlace video shall have the same Synchronizing Timestamp value, even though they are delivered across many packets with increasing RTP Timestamp values.

The Synchronizing Timestamp shall be used as the temporal reference point for the field or frame of video samples when aligning the video essence contained in the SMPTE ST 2022-6 stream with other essences in a SMPTE ST 2110-10 system, compensating as required for the differences in Media Clock rate between the essences.

For audio essence samples contained as embedded audio within the payload of the ST 2022-6 stream, the effective sampling instant of each audio sample (for synchronization purposes) shall be determined as specified in ST 2110-10 section 7.5.4.

Note 1: This Synchronizing Timestamp is different from the RTP Timestamp transmitted in each RTP packet of the SMPTE ST 2022-6 stream; it is constructed by the receiver as specified above when needed.

Note 2: SMPTE ST 2022-6 does not explicitly require that the RTP clock be synchronized to the underlying SDI clock, though many implementations do so. The provisions above force this requirement.

Note 3: References to frames in the text above refer to video image structures on the SDI interface, even in the case of SDI mappings in which the interface frame rate differs from the source frame rate.

5.4 Well-Formed SDI

The SDI signals encapsulated within the SMPTE ST 2022-6 flows shall be well-formed, meaning that they comply strictly with the structures defined in the relevant SMPTE documents for the encapsulated SDI signals.

Note: SMPTE RP 168 switching operations that are upstream of an IP encapsulator might cause one or more lines of an SDI video signal to contain an incorrect number of samples (i.e. not in strict conformance with the relevant specification for the SDI signal) in the vicinity of the switching point. When such lines are encountered by an encapsulator, video sample data might need to be added or removed from these lines prior the encapsulation process, thereby ensuring that every line within an ST 2022-6 encapsulated stream contains the correct number of video samples.

6 Network Compatibility and Transmission Traffic Shape Models

Streams from senders compliant to this standard shall comply with the Network Compatibility Model and the Virtual Receiver Buffer Model specified in SMPTE ST 2110-21, using a value of $N_{PACKETS}$ equal to the number of packets in a frame as specified in SMPTE ST 2022-6 for the prevailing SDI format. Senders shall comply with either the “Narrow Linear Senders (Type NL)” requirements, or the “Wide Senders (Type W)” requirements.

The $TRO_{DEFAULT}$ value for streams under this standard shall be

$$TRO_{DEFAULT} = MAX\left(INT\left(\frac{1500 \times 8}{MAXIP}\right),\ INT\left(\frac{N_{PACKETS}}{27000 \times T_{FRAME}}\right)\right) \times \left(\frac{T_{FRAME}}{N_{PACKETS}}\right) \quad seconds$$

Senders shall signal Media Type Parameters `TP` and `TROFF` as specified in ST 2110-21.

7 Session Description Requirements

7.1 Session Description of SMPTE ST 2022-6 Streams

SMPTE ST 2022-6 senders compliant to this standard shall comply with the Session Description Protocol as specified in SMPTE ST 2110-10, with the exception that the requirement of separate SDP objects for each RTP stream is relaxed for the case of SMPTE ST 2022-5 FEC which shall be as described below.

Mapping of the 2022-6 media type into SDP shall use the media type definition defined in section 8.1.

Video RTP Senders whose streams utilize a value of TRO_{OFFSET} which differs from $TRO_{DEFAULT}$ shall signal the prevailing value of TRO_{OFFSET} in the Session Description Protocol (SDP) with a Media Type Parameter `TROFF` of the prevailing TRO_{OFFSET} value, in microseconds, expressed as a positive integer decimal value (rounded if necessary). If this parameter is not present, receivers shall assume the default value $TRO_{DEFAULT}$.

7.2 Session Description of SMPTE ST 2022-5 FEC Streams

If SMPTE ST 2022-5 FEC packets are being transmitted by the sender, then the provisions of IETF RFC 6364 shall apply. An additional "a=group" clause shall be included in the SDP session header, referencing the ST 2022-6 stream and the ST 2022-5 FEC packet stream using symbolic tags as follows:

```
a=group:FEC-FR <vidtag> <fectag>                    (in the session header)
```

The media section for the SMPTE ST 2022-6 stream shall contain an additional "a=" clause as follows:

```
a=mid:<vidtag>                                      (name ties to a=group attribute)
```

The ST 2022-5 FEC flow shall have its own media section, referencing the media types defined in section 8.2 for the FEC stream, containing the following "a=" clauses:

```
a=rtpmap:<FPT> SMPTE2022-5-FEC/27000000
a=fec-repair-flow: encoding-id=10
a=mid:<fectag>                                      (name ties to a=group attribute)
```

where <FPT> is the dynamically assigned Payload Type for the FEC stream. The values of the symbolic tags `<vidtag>` and `<fectag>` shall meet the syntactic requirements of IETF RFC 6364.

7.3 Session Description of SMPTE ST 2022-7

The use of SMPTE ST 2022-7 redundant streams shall be signaled using the methods defined in SMPTE ST 2110-10, adapted as described in sections 7.1 and 7.2 of this document for the cases of SMPTE ST 2022-6 and SMPTE ST 2022-5 streams, respectively.

8 IANA Considerations

The following items related to this document require registration with IANA.

8.1 SMPTE 2022-6 Media Type Registration Details

```
IANA Type Name:  video

IANA Subtype Name: SMPTE2022-6

IANA Required Parameters: none

IANA Optional Parameters: TROFF
```

Note: the use of large UDP packets is restricted by SMPTE ST 2022-6, therefore the MAXUDP parameter defined in SMPTE ST 2110-10 is not used in the context of this subtype.

8.2 SMPTE 2022-5 FEC Media Type Registration Details

```
IANA Type Name:  application

IANA Subtype Name: SMPTE2022-5-FEC

IANA Required Parameters: none

IANA optional Parameters: none
```

8.3 SMPTE 2022-5 FEC Framework (FECFRAME) FEC Encoding Registry ID

```
IANA FECFRAME ID:  10
```

Annex A (Informative)

SMPTE ST 2022-6 defines the first RTP packet of the frame as beginning with the EAV sequence at the beginning of the first line of the frame. ST 2110-10 defines the RTP timestamp of a video frame derived from SDI, based on the alignment point as defined in ST 2059-1 or (for SDI formats not captured in ST 2059-1) the defining SDI document. Figure 1 shows the temporal relationship between the two, in terms of the format-specific timing parameters defined in ST 2059-1, for an SDI signal aligned according to ST 2059-1.

The Alignment point of ST 2059-1 is not coincident with the EAV of SDI; instead it is at a format-dependent offset before the first active pixel of the SDI line. In order to have correspondence between the ST 2110-10 video frame RTP timestamp and the Synchronizing Timestamp, it is necessary to add this small offset to the RTP timestamp of the first ST 2022-6 packet in order to compensate for this difference.

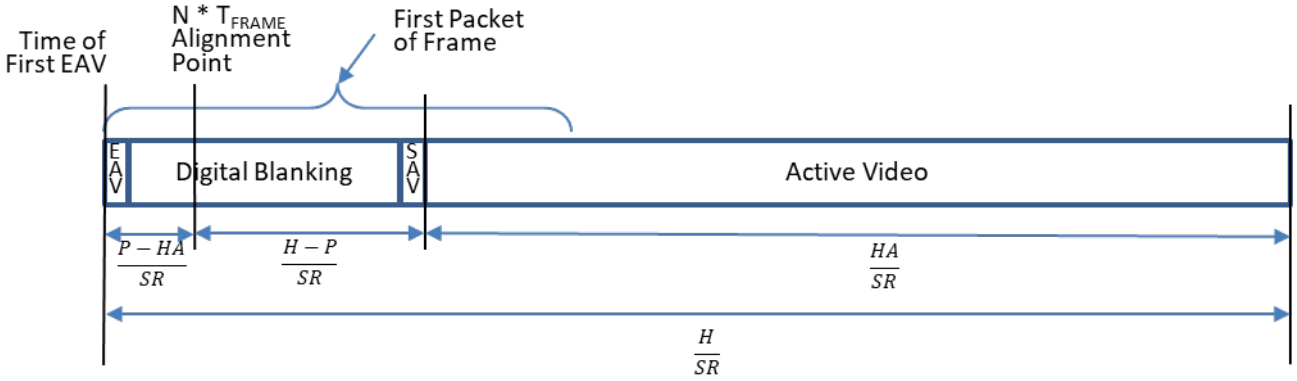

Figure 1 – Relationship between the EAV and the Alignment Point

Bibliography (Informative)

Video Services Forum (VSF) TR-03 Transport of Uncompressed Elementary Stream Media over IP [online, viewed 2017-11-20] Available at 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 [online, viewed 2017-11-20] Available at http://www.videoservicesforum.org/download/technical_recommendations/VSF_TR-04_2015-11-12.pdf

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