SMPTE STANDARD

Professional Media Over Managed IP Networks: Traffic Shaping and Delivery Timing for Video

Table of Contents

Approved November 22, 2017

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. 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 TR03 and TR04, 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 .

Part 10 covers the system as a whole, the timing model, and common requirements across all essence types. Other documents will cover specific media essence formats. Part 20 documents the transport of uncompressed active video in such systems, using an RTP format based on IETF RFC 4175.

Part 21 (this part) specifies the timing model for senders and receivers of video RTP streams.

1 Scope

This standard specifies a timing model for SMPTE ST 2110-10 video RTP streams as measured leaving the RTP sender, and defines the sender SDP parameters used to signal the timing properties of such 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.

Recommendation ITU-R BT.656-5 Interfaces for digital component video signals in 525-line and 625-line television systems operating at the 4:2:2 level of Recommendation ITU-R BT.601

Recommendation ITU-R BT.709-6 Parameter values for the HDTV standards for production and international programme exchange

Recommendation ITU-R BT.1543-1 1280×720, 16:9 progressively-captured image format for production and international programme exchange in the 60 Hz environment

Recommendation ITU-R BT.1847-1 1280×720, 16:9 progressively-captured image format for production and international programme exchange in the 50 Hz environment

Recommendation ITU-R BT.2020-2 (10/2015) Parameter values for ultra-high definition television systems for production and international programme exchange

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

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

4 Terms and Definitions

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

5 Textual Conventions

5.1 Mathematical Functions

The following functions are defined for one or more numerical arguments.

Maximum	MAX(a, b) shall return the largest value of a and b.
Integer	INT(a) shall return the largest integer not greater than a.

6 Stream Timing Characteristics

6.1 General

This section specifies two parametric models for the packet delivery timing characteristics of frame-based video RTP streams as they leave the transmission interface of the Sender. A network compatibility model is defined in section 6.6.1, which regulates the burst characteristics of senders in order to promote compatibility with a wide variety of switches with varied buffer sizes. Additionally, a Virtual Receiver Buffer model is defined in section 6.6.2.

In the Virtual Receiver Buffer model, packets are deposited at the actual moment of transmission, and removed on a specific schedule. Two fundamental Packet Read Schedules (PRS) are defined, a Gapped schedule and a Linear schedule. The number of outstanding packets in the Virtual Receiver Buffer cannot exceed a parametric maximum for each sender class.

This Virtual Receiver Buffer model provides important information about the stream characteristics useful in the design of receivers – however a practical receiver needs to also accommodate the network-induced packet jitter and latency which accumulates along the path from sender to receiver. Design of receivers is outside the scope of this standard.

6.2 Virtual Receiver Buffer Packet Read Schedule (PRS) Parameters

The following parameters, illustrated for the case of the gapped model in Figure 1 and the linear model in Figure 2, are used in specifying the read schedule for the Virtual Receiver Buffer model of section 6.6. The model defines a Packet Read Time instant TPR_j for each packet j within the video frame of an RTP stream, relative to the Video Transmission Datum T_{VD}. TPR_j is the time at which the packet j will be removed from the Virtual Receiver Buffer in the model.

T_{FRAME} is the time period between consecutive frames of video at the prevailing frame rate

$N_{PACKETS}$ is the number of packets per frame of video (depends on mapping details)

T_{VD} a time point given by $(N \times T_{FRAME}) + TR_{OFFSET}$, where N is an integer and the time scale has its origin at the SMPTE Epoch as defined in SMPTE ST 2059-1.

TR_{OFFSET} is the difference between the most recent integer multiple of T_{FRAME} and T_{VD}. TR_{OFFSET} shall be a positive number or zero, such that $T_{VD} = (N \times T_{FRAME}) + TR_{OFFSET}$ for each frame

$TRO_{DEFAULT}$ is the model-specific default value for the TR_{OFFSET} parameter.

T_{RS} is the time between removing adjacent packets from the Virtual Receiver Buffer during the frame/field (Time-Read-Spacing). Packet removal shall be modeled as an instantaneous event with zero time duration.

TPR_0 is the time when the first packet of the frame is removed from the Virtual Receiver Buffer. TPR_0 is coincident with T_{VD}. $(TPR_0 = T_{VD})$ (Time-Packet-Read-Zero).

TPR_j is the time when packet j will be removed from the Virtual Receiver Buffer (Time-Packet-Read-j).

Video RTP Senders whose streams utilize a value of TR_{OFFSET} which differs from $TRO_{DEFAULT}$ shall signal the prevailing value of TR_{OFFSET} in the Session Description Protocol (SDP) with a Media Type Parameter `TROFF` of the prevailing TR_{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 $TRO_{DEFAULT}$.

6.3 Gapped Packet Read Schedule (PRS)

6.3.1 Overview

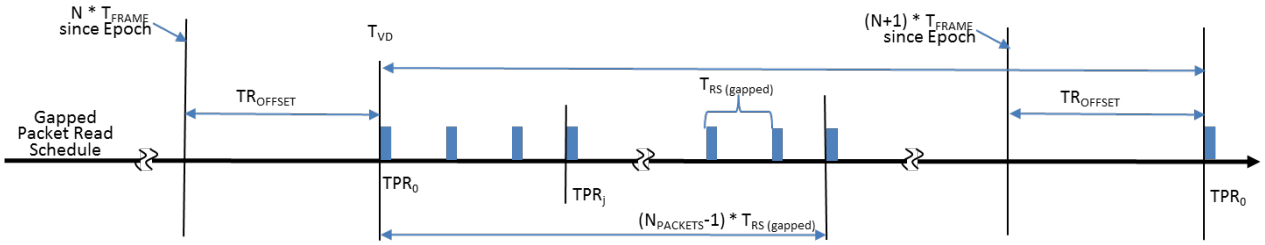

Figure 1 -- Packet Read Schedule (Gapped)

The gapped PRS, illustrated in Figure 1, consists of a sequence of TPR_j values which loosely approximates the delivery of samples in the SDI signal, including a gap in the transmission corresponding to the vertical blanking time of SDI. The TPR_j instants are spaced uniformly throughout the active field (or segment) or frame interval. The progressive scan signal has one gap after each frame, and an interlaced scan signal (and PsF signal) has two gaps, one after each field (or segment in the case of PsF). The following variable is used in the equations below:

R_{ACTIVE} is the ratio of active time to total time within the frame period.

Although RTP puts no specific constraints on the image height, width, or frame rate, the gapped PRS shall only apply to streams with image dimensions and frame rates specified in Recommendation ITU-R BT.656-5, Recommendation ITU-R BT.1543-1, Recommendation ITU-R BT.1847-1, Recommendation ITU-R BT.709-6, or Recommendation ITU-R BT.2020-2.

6.3.2 Gapped PRS – Progressive Images

For progressively scanned images (but excluding Progressive segmented Frame (PsF) images) using the gapped model, TPR_j shall be defined as follows:

$$R_{ACTIVE} = (1080/1125)$$

$$T_{RS} = (T_{FRAME} \times R_{ACTIVE})/N_{PACKETS}$$

$$T_{VD} = (N \times T_{FRAME}) + TR_{OFFSET} \qquad (for\ an\ integer\ value\ of\ N)$$

$$TPR_j = (j \times T_{RS}) + T_{VD}$$

$$TRO_{DEFAULT} = \begin{cases} (43/1125) \times T_{FRAME} & \text{Image Height} \geq 1080 \text{ lines} \\ (28/\ 750\ ) \times T_{FRAME} & \text{Image Height} < 1080 \text{ lines} \end{cases}$$

Note: this gapped model defines the inter-frame gap and TR_{OFFSET} value for all (non-PsF) progressive formats, including 720p, 1080p, 2160p, and 4320p. The $TRO_{DEFAULT}$ values are chosen slightly later than the beginning of the active video in order to allow sufficient time to assemble packets and buffer the samples for transmission when working from a timed SDI signal and packing into the Standard UDP Size Limit.