SMPTE STANDARD

Professional Media Over Managed IP Networks: AES3 Transparent Transport

Table of Contents **Page**

Copyright © 2018 by THE SOCIETY OF MOTION PICTURE AND TELEVISION ENGINEERS
445 Hamilton Avenue., White Plains, NY 10601
(914) 761-1100

Approved July 27, 2018

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 of 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 will cover specific media essence formats.

SMPTE ST 2110-31 (this part) documents the transport of AES3 signals in transparent manner, preserving the V, U, C, and P bits.

1 Scope

This Standard specifies the real-time, RTP-based transport of AES3 signals over IP networks, referenced to a network 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 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.

AES3-3-2009 AES standard for digital audio — Digital input-output interfacing — Serial transmission format for two-channel linearly represented digital audio data, Part 3: Transport

Internet Engineering Task Force (IETF) RFC 3550 RTP: A Transport Protocol for Real-Time Applications [online, viewed 2017-11-07] Available at https://www.ietf.org/rfc/rfc3550.txt

Internet Engineering Task Force (IETF) RFC 3551 RTP Profile for Audio and Video Conferences with Minimal Control [online, viewed 2017-11-07] Available at https://www.ietf.org/rfc/rfc3551.txt

Internet Engineering Task Force (IETF) RFC 4566 SDP: Session Description Protocol [online, viewed 2017-11-07] Available at https://www.ietf.org/rfc/rfc4566.txt

Internet Engineering Task Force (IETF) RFC 8285 A General Mechanism for RTP Header Extensions [online, viewed 2017-11-07] Available at https://www.ietf.org/rfc/rfc8285.txt

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

SMPTE ST2110-30:2017 Professional Media over Managed IP Networks -- PCM Audio

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 AES3 Block

Block as defined in AES3-3

4.2 AES3 Frame

Frame as defined in AES3-3

4.3 AES3 Subframe

Subframe as defined in AES3-3

5 AES3 Transport RTP Format

5.1 Historical Overview (Informative)

SMPTE ST 2110-30 defines a method of transporting PCM audio data based upon AES67. While PCM audio and integration with AES67 are important in many television production environments, there also exist cases where transparent transport of entire AES3 signals (including the V, P and U bits and the channel status information) is beneficial. The modern television ecosystem has leveraged the prevalence of AES3 signal transport to encapsulate many different data items – SMPTE ST 337 defines a general method for the encapsulation of various payloads into the AES3 transport and SMPTE ST 338 manages the growing namespace of such payloads.

This standard defines an RTP payload format for the transport of AES3 signals over IP.

5.2 RTP Encapsulation

The sequence of AES3 Subframes inside the AES3 signals shall be transported using RTP as specified in IETF RFC 3550, subject to the constraints and payload definition below.

The technical metadata necessary to receive and interpret the RTP stream shall be communicated via SDP as defined in Section 6.

5.3 RTP Header Definitions

The fields of the RTP packet header shall be as specified in IETF RFC 3550 section 5.1 (RTP Fixed Header Fields) and shown in Figure 1.
Fields not listed below are exactly as specified in IETF RFC 3550. The following additional constraints shall apply:

Payload Type (PT): 7 bits	Dynamically allocated payload type. The payload type shall be dynamically allocated in accordance with IETF RFC 3551
CSRC Count (CC): 4 bits	Set to zero (0)
Timestamp: 32 bits	RTP Timestamp as specified in SMPTE ST 2110-10 for PCM audio signals. For AES3 transparent Subframes, the RTP timestamp indicates the time of the AES3 Subframe as presented to the encapsulator on the AES3 interface, or the equivalent value for an AES3 signal embedded into SDI.
Marker bit (M): 1 bit	Set to zero (0)
Extension bit (X): 1 bit	When this bit is set, an RTP header extension formatted as specified in IETF RFC 8285 is present immediately following the SSRC field.

Unless otherwise specified, all multi-octet numeric values expressed in the RTP Header and in the RTP Payload Headers shall be expressed in Network Byte Order (traditionally referred to as “big-endian”).

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X|  CC   |M|     PT      |       sequence number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           synchronization source (SSRC) identifier            |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|                           AM824 Subframe                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           AM824 Subframe                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
...
...
|                           AM824 Subframe                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1 – RTP Header with RTP Payload

5.4 RTP Payload Definition

The RTP Payload shall consist of an interleaved set of sequences of AM824 Subframes. The contents of each AES3 Subframe shall be transported in one AM824 Subframe. The fields of the AM824 Subframe are shown in Figure 2 and defined below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0|B|F|  PCUV |                    DATA24                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2 -- AM824 Subframe

Block Start (B): 1 bit	When set to 1, indicates the first Subframe of an AES3 Block. If this bit is set to 1, the Frame Start (F) bit shall also be set to 1.
Frame Start (F): 1 bit	When set to 1, indicates the first Subframe of an AES3 Frame
PCUV: 4 bits	The P, C, U and V bits of the AES3 Subframe, in order. The P bit is copied from timeslot 31 of the AES3 Subframe. The C bit is copied from timeslot 30, the U bit from timeslot 29 and the V bit from timeslot 28 of the AES3 Subframe.
DATA24: 24 bits	24 bits originating from time slots 4 -27 of the AES3 Subframe. AES3 timeslot 27 is conveyed in AM824 Subframe bit position 8.

AES3 Subframes 1 and 2 of each AES3 Frame shall be sequentially interleaved. If multiple AES3 signals are transmitted within the same RTP stream, then the AES3 Subframes from each AES3 signal are sequentially interleaved. Every packet within a stream shall contain data from the same number of AES3 signals. Each packet within a stream shall contain AES3 Subframes from the same number of periods of the Media Clock. The time period corresponding to each packet within the stream shall be signaled using the `ptime` attribute in the SDP, as defined in section 6.1, using one of the permitted values from Table 1.

Note 1: The Block Start (B) field set to 1 corresponds to the "Z" preamble defined in AES3-3 and requires that the Frame Start (F) bit is also set to 1. When the Block Start (B) field is set to 0, the Frame Start (F) bit is set to 1 in correspondence to the "X" preamble, and set to 0 in correspondence to the "Y" preamble.

Note 2: The AES3 Subframe in AES3-3 is divided into timeslots numbered 0...31 where timeslot 0 is transmitted first, and multi-bit quantities such as the audio sample value are organized with the lowest-numbered bit in the bitfield carrying the least significant bit in the value. This RTP payload format is composed of 32 bit AM824 Subframes in which the most significant bit of a multi-bit quantity is actually the lowest-numbered bit in the big-endian representation shown.

5.5 Media Clock

Streams in scope of this standard shall use one Media Clock rate of 44.1kHz, 48kHz, or 96kHz. Senders and Receivers under this standard shall support a Media Clock and RTP clock rate of 48kHz, and rates of 44.1kHz or 96kHz may be supported. Devices which support multiple streams are not required to support multiple Media Clock rates simultaneously.

The offset between the Media Clock and the RTP Clock shall be zero as specified in SMPTE ST 2110-10. Other provisions of the Media Clock and RTP Clock shall be as specified in SMPTE ST 2110-10 section 7.

6 Session Description

6.1 General

Streams under this standard shall be signaled in the SDP using the media type "audio" and the media subtype "AM824".

The number of AES3 Subframe sequences multiplexed within the payload shall be signaled in the SDP object on the a=rtpmap line, using the syntactic field which typically communicates the number of channels in an audio signal, as shown below:

```
a=rtpmap:<pt> AM824/<clock-rate>/<nchan>
```

where

`<clock-rate>`	is the Media Clock Rate (sampling rate)
`<nchan>`	is the number of AES3 Subframes multiplexed together within the payload
`<pt>`	is the dynamically assigned RTP Payload Type found in the RTP header

The `<clock-rate>` parameter shall take one of the values 44100, 48000 or 96000.

Since this standard transports AES3 signals, and each AES3 signal contains two sequences of AES3 Subframes, the number of AES3 Subframe sequences `<nchan>` expressed in the SDP object shall always be an even number.

Senders under this standard shall signal a `ptime` attribute in the SDP, as defined in IETF RFC 4566 and shown as `<packet-time>` below.

```
a=ptime:<packet-time>
```

For the avoidance of doubt, the `<packet-time>` value in the context of this RTP Payload specification indicates the amount of time required (in milliseconds) for the transmission of the underlying AES3 Subframes when formatted into AES3-3 signals at the signaled `<clock-rate>`.

The `<packet-time>` parameter shall take one of the values from the Table 1, based on the prevailing sampling rate and the desired number of AES3 Subframe sequences interleaved together within the RTP stream.

Table 1 – Permitted Values of Packet Time for each Media Clock rate

<packet-time>	<clock-rate> Media Clock (Hz)	Number of periods of the Media Clock per packet
1	48000	48
0.12	48000	6
0.08	48000	4
1	96000	96
0.12	96000	12
0.08	96000	8
1.09	44100	48
0.14	44100	6
0.09	44100	4

6.2 Channel Order Signaling

For AES3 Subframes containing PCM Audio, Senders may signal the channel order in the SDP using the Channel Order Convention specified in SMPTE ST 2110-30. Because this standard may also transport non-PCM audio signals, the additional Channel Grouping Symbol listed in Table 2 may be used in addition to those specified in ST 2110-30.

Table 2 – Additional Channel Order Convention Grouping Symbols

Channel Grouping Symbol	Quantity of AES3 Subframe sequences in group	Description of group
`AES3`	2	Either Non-PCM signals or PCM signal, or one Subframe of each.

7 Conformance Levels

All receivers under this standard shall implement the Level A requirements below. Senders and Receivers conforming to this standard should specify the Table 3 Conformance Levels which are supported when indicating conformance to this standard.

Table 3 - Conformance Levels

Level	Receiver shall support			
	Media Clock Rate (Hz)	**<Packet-Time>**	**Maximum Number of AES3 Subframe Sequences per packet**	**Maximum Number of AES3 signals**
A	48000	1	6	3
AX	48000	1	6	3
	44100	1.09	6	3
	96000	1	2	1
B	48000	1	6	3
	48000	0.12	8	4
BX	48000	1	6	3
	48000	0.12	8	4
	44100	1.09	6	3
	44100	0.14	8	4
	96000	1	2	1
	96000	0.12	4	2
C	48000	1	6	3
	48000	0.12	60	30
CX	48000	1	6	3
	48000	0.12	60	30
	44100	1.09	6	3
	44100	0.14	60	30
	96000	1	2	1
	96000	0.12	30	15
D	48000	1	6	3
	48000	0.12	60	30
	48000	0.08	80	40
DX	48000	1	6	3
	48000	0.12	60	30
	48000	0.08	80	40
	44100	1.09	6	3

	44100 44100 96000 96000 96000	0.14 0.09 1 0.12 0.08	60 80 2 30 40	30 40 1 15 20

Bibliography (Informative)

AES67-2015 AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability

Video Services Forum (VSF) TR-03 Transport of Uncompressed Elementary Stream Media over IP [online, viewed 2018-03-06] Available at http://www.videoservicesforum.org/download/technical_recommendations/VSF_TR-03_2015-11-12.pdf

SMPTE ST 337:2015 Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

SMPTE ST 338:2015 Format for Non-PCM Audio and Data in AES3 — Data Types