SMPTE ST 340:2015
Revision of
SMPTE 340-2008

SMPTE STANDARD

Format for Non-PCM Audio and Data in AES3 — ATSC A/52 Digital Audio Compression Standard for AC-3 and Enhanced AC-3 Data Types

Table of Contents **Page**

Approved
March 27, 2015

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.

SMPTE ST 340 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.

This standard describes how both AC-3 and E-AC-3 ("Enhanced AC-3") data rate reduced (non PCM) audio streams conforming to ATSC A/52 Annex E (and to ETSI TS 102 366) are packed into an AES3 data stream, following the methods described by SMPTE ST 337. The AC-3 and E-AC-3 data types are identified by a data type number listed in SMPTE ST 338 and carried in the burst information word of the data burst preamble. The standard describes how the AC-3 data is packed into the AES3 data stream, then extends the technique to E-AC-3.

1 Scope

This standard specifies data type specific format requirements for both AC-3 and E-AC-3 (“Enhanced AC-3”) data bursts carried within an AES3 interface according to SMPTE ST 337.

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; followed by 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 recommended practice. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this recommended practice are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

ATSC A/52:2012, Digital Audio Compression (AC-3, E-AC-3) Standard, 17 December 2012

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

SMPTE ST 339:2015, Format for Non-PCM Audio and Data in AES3 — Generic Data Types

4 AC-3

4.1 Overview

As defined in A/52, an AC-3 synchronization frame (or as used hereafter, “Sync Frame”) is a unit of the AC-3 serial bit stream capable of being fully decoded. A Sync Frame begins with a sync word and contains 1536 coded audio samples. The coded data representing 1536 samples of the audio from all the audio channels of one program forms a Burst Payload. A burst preamble (defined by SMPTE ST 337) containing information about the payload is prepended to the Burst Payload and enough padding words are appended to it to make the resulting data burst duration exactly match that of the 1536 samples of baseband (PCM) audio that it represents. The resulting data bursts are placed in the audio sample word/aux data fields of AES3 subframes at regular intervals in either the frame or subframe mode (see SMPTE ST 337, section 5). Multiple streams of data may be multiplexed into the AES3 transport per SMPTE ST 337. Data bursts are placed in the AES3 transport using either 16, 20, or 24 bits of the available data space. While the 24-bit mode allows more efficient use of the AES3 capacity, the 16- and 20-bit modes allow use with existing equipment limited to 16- or 20-bit operation.

A single AC-3 Sync Frame forms the Burst Payload, as shown in Figure 1. Each AC-3 Sync Frame contains 6 coded audio blocks (AB), each of which represent 256 new audio samples per channel. A synchronization word at the beginning of each Sync Frame contains information needed to acquire and maintain synchronization. A bit stream information (BSI) header follows the sync word, and contains parameters describing the coded audio service. The coded audio blocks may be followed by an auxiliary data (Aux) field. At the end of each frame is an error check field that includes a CRC word for error detection.

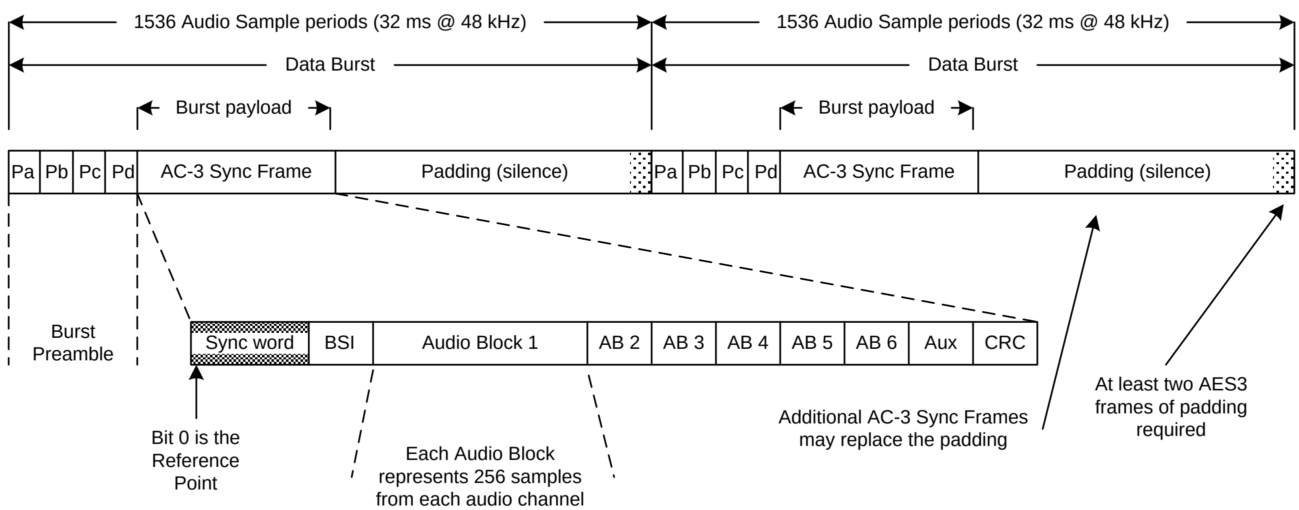

Figure 1 – Structure of an AC-3 Data Burst

4.2 burst_preamble

The Pc word (burst_info value) of the burst_preamble carries the data_type identifier, the data_type_dependent and the data_stream_number information (see SMPTE ST 337, Table 7).

4.2.1 data_type identifier

The data_type identifier shall be set to 1.

4.2.2 data_type_dependent

The values of the data_type_dependent bits shall be as shown in Table 1, and as described below.

Table 1 – Values of data_type_dependent field for AC-3 data type

data_type_dependent bit number	Meaning
0 - 2	Reserved, must be set to 000
3	Repetition rate flag (rep_rate_flag)
4	Not full service flag (not_full_svc)

rep_rate_flag - The repetition rate flag shall be set to 0 if the AC-3 data burst is placed in the AES3 interface such that the reference point of the data burst (see section 3.5) occurs at the AC-3 standard repetition rate (see section 3.6). The flag shall be set to 1 if the reference point does not occur at the AC-3 standard repetition rate. This flag is intended to be set to the same state for all data bursts of a given AC-3 data stream to indicate whether data bursts for the stream occur at the standard repetition rate.

not_full_svc - This is a 1-bit field which indicates whether or not this audio service is a full service suitable for presentation, or whether it is only a partial service which must be combined with another audio service before presentation. This bit shall be set to 0 if the audio service is suitable for presentation and shall be set to 1 if it is not.

4.2.3 data_stream_number

The data_stream_number shall be set to any number from 0h to 6h. 7h is a reserved value. See SMPTE ST 337.

4.3 AC-3 Burst Payload

The AC-3 encoder produces a stream of AC-3 Sync Frames that are defined by ATSC A/52. Each AC-3 Sync Frame contains the data that represents 1536 audio samples of all the audio channels in a program plus a sync word, a check sum etc. The length of the AC-3 Burst Payload will depend on the encoded bit rate.

The Burst Payload of each AC-3 data burst shall contain one complete AC-3 Sync Frame.

4.4 AES3 Frame Rate (Sampling Frequency)

The frame rate of the AES3 stream used to transport the AC-3 Sync Frames shall be the same as the rate at which the encoded audio was sampled.

If multiple AC-3 streams are transported in the same AES3 stream, all the audio represented by these Burst Payloads must be sampled at the same rate.

4.5 AC-3 Reference Point

The Reference Point of an AC-3 Burst Payload shall be bit 0 of the Burst Payload, as shown in Figure 1.

4.6 AC-3 Standard Repetition Rate

AC-3 Burst Payloads occur at the standard Repetition Rate if the Reference points for consecutive data bursts (in the same data stream number) occur 1536 AES3 frames apart.

4.7 AC-3 Decode Latency (Professional)

The AC-3 decode latency for professional applications shall be 1536 AES3 frame periods of the rate at which the baseband audio was sampled (e.g., 32 ms at 48-kHz sampling rate) measured from the reference point of the data burst. A reference decoder must output the first PCM sample exactly 1536 sample periods after the first bit of the data burst is received by the decoder.

4.8 AC-3 Reference Position

The Reference Position of a Burst Payload is defined by the relationship of the decoded audio to an associated video signal. A Burst Payload is in the Reference Position when the decoded audio from that Burst Payload is in sync with the associated video.

The Reference Point of the Burst Payload carried in an AES3 stream that is locked to the associated video signal must therefore precede the video sync point by 1536 AES3 frame periods.

4.9 AC-3 Burst Payload Timing

It may not be possible to carry Burst Payloads in the Reference position of an AES3 stream. SMPTE ST 339 defines a Time Stamp data type whose Burst Payload carries either the offset (in AES3 frame periods) between the AC-3 Reference Position and the Reference Point of the following AC-3 Burst Payload or a SMPTE ST 12-1 time code associated with the decoded audio from the following Burst Payload. Either method of time stamping the Burst Payload allows the carriage of multiple AC-3 Burst Payloads in a single data burst. If an AC-3 Burst Payload does occupy the reference position, the delay shall be 0.

Burst Preambles and their associated Burst Payloads may be placed randomly in AC-3 data bursts. In this case, the Burst Preamble rep_rate_flag must be set to 1 to indicate that the Burst Payloads do not occur at the standard Repetition Rate. Time stamps shall be used to indicate when to present the decoded audio samples.

4.10 Use of Pause Date Bursts Between AC-3 Data Streams

It is recommended that Pause data-bursts (ref SMPTE ST 339) are used to fill gaps in the AC-3 data stream. The Pa word of the first Pause data burst transmitted after the last valid AC-3 data burst shall occur 1536 AES3 frames after the Pa word of the preceding AC-3 data burst. The pause data-bursts shall be transmitted with a repetition period of 3 AES3 frames, unless other repetition periods are necessary to precisely fill the stream gap (whose length may not be a multiple of 3 AES3 frames), subject to the zero padding requirements of SMPTE ST 337, Section 7.4, “Burst Spacing”.

When it is possible to control the length of the gap in an AC-3 data stream, it is recommended that the gap length be an integer multiple of 256 AES3 frames, again respecting SMPTE ST 337, Section 7.3. This allows many AC-3 decoders to optimize the gap concealment process.

The sequence of Pause data bursts may be interrupted to allow other data bursts to be multiplexed into the AES3 transport. See SMPTE ST 339 for details of the Pause data type and its use.

5 E-AC-3

5.1 Overview

An E-AC-3 bitstream is constructed from one or more substreams. A single E-AC-3 substream can be considered equivalent to a complete AC-3 bitstream — it can carry up to 5.1 channels of audio, and is constructed from a sequence of Sync Frames. While an AC-3 bitstream is limited to 5.1 channels, E-AC-3 bitstreams can include multiple substreams, allowing delivery of a single program with more than 5.1 channels, multiple independent programs, or a combination of both. The substreams are time-multiplexed into the Burst Payload, as shown in Figure 2.

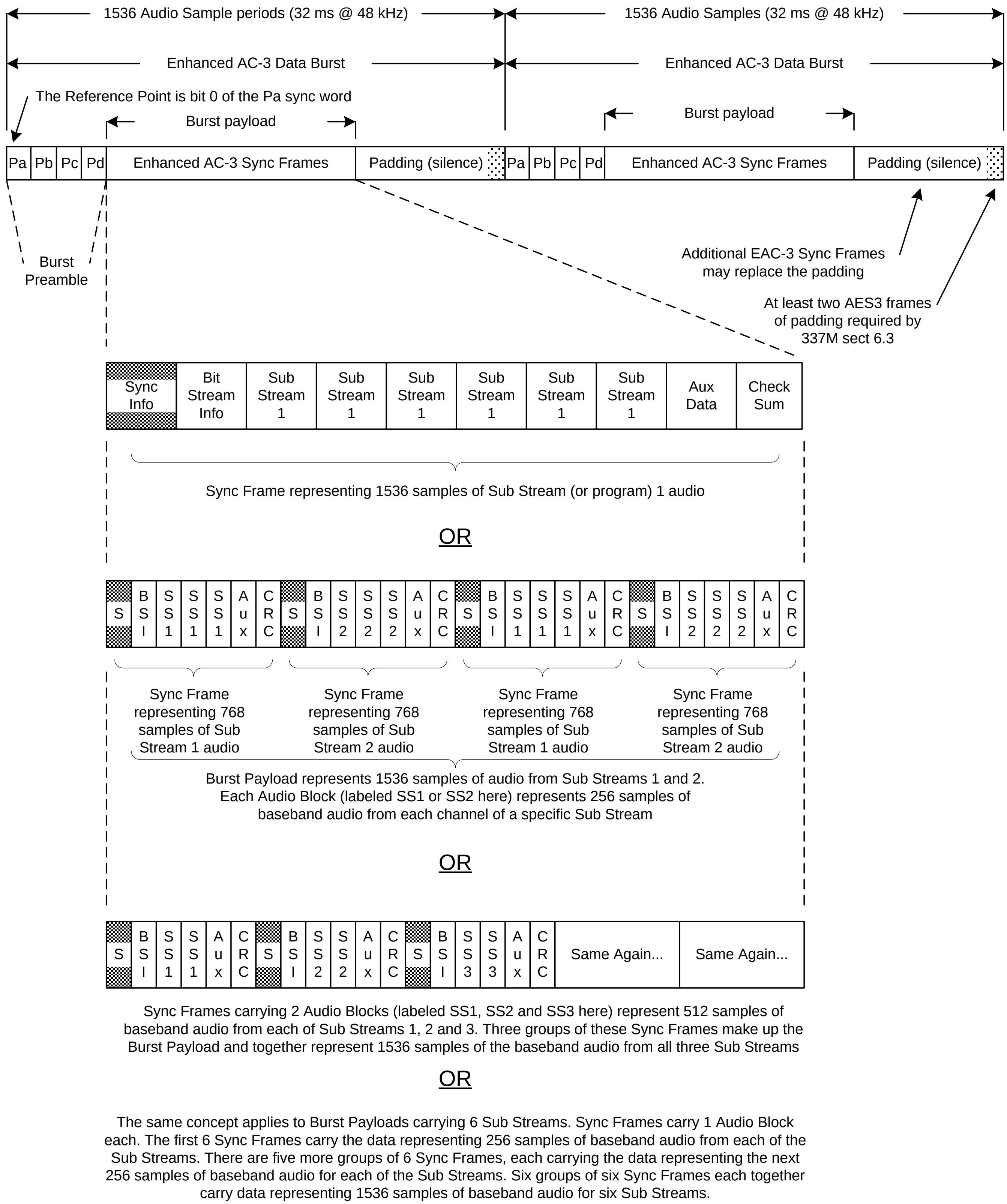

Figure 2 – Structure of an E-AC-3 Data Burst

The Sync Frames carrying the data for each substream are made up of a sync word, Bit Stream information (BSI), 1, 2, 3, or 6 Audio Blocks, each of which represents 256 samples from each channel of one of the (possibly multichannel) substreams being coded, some auxiliary (Aux) data and a check sum (CRC). The Audio Blocks shown in Figure 2 are labeled SS1, SS2, etc. to indicate which substream they belong to, and to show the order in which they are multiplexed into the Burst Payload. Every Sync Frame carried in an E-AC-3 bitstream must contain the same number of Audio Blocks.

Multiple streams of data may be multiplexed into the AES3 transport per SMPTE ST 337. Data bursts are placed in the AES3 transport using either 16, 20, or 24 bits of the available data space. While the 24-bit mode allows more efficient use of the AES3 capacity, the 16- and 20-bit modes allow use with existing equipment limited to 16- or 20-bit operation.

5.2 burst_preamble

The Pc word (burst_info value) of the burst_preamble carries the data_type identifier, the data_type_dependent and the data_stream_number information (see SMPTE ST 337, Table 7)

5.2.1 data_type identifier

The data_type identifier shall be set to 16.

5.2.2 data_type_dependent

The values of the data_type_dependent bits shall be as shown in Table 2, and as described below.

Table 2 – Values of data_type_dependent field for E-AC-3 data type

data_type_dependent bit numbers	Meaning
0 - 2	Reserved, must be set to 000
3	Repetition rate flag (rep_rate_flag)
4	Not full service flag (not_full_svc)

rep_rate_flag – The repetition rate flag shall be set to 0 if the E-AC-3 data burst is placed in the AES3 interface such that the Reference Point of the data burst (see Section 4.5) occurs at the E-AC-3 standard repetition rate (see Section 4.7). The flag shall be set to 1 if the reference point does not occur at the E-AC-3 standard repetition rate. This flag is intended to be set to the same state for all data bursts of a given E-AC-3 data stream to indicate whether data bursts for the stream occur at the standard repetition rate.

not_full_svc – This is a 1-bit field which indicates whether or not this audio service is a full service suitable for presentation, or whether it is only a partial service which must be combined with another audio service before presentation. This bit shall be set to a 0 if the audio service is suitable for presentation and shall be set to 1 if it is not.

5.2.3 data_stream_number

The data_stream_number shall be set to any number from 0h to 6h. 7h is a reserved value. See SMPTE ST 337.

5.3 E-AC-3 Burst Payload

The E-AC-3 Burst Payload shall contain one or more Sync Frames as shown in Figure 2. Each Sync Frame is defined by ATSC A/52, and shall consist of a sync word, Bit Stream information (BSI), 1, 2, 3, or 6 Audio Blocks, each of which shall represent 256 samples from each channel of one of the (possibly multichannel) substreams being coded, optional auxiliary data (Aux) and a check sum (CRC). If the Burst Payload is constructed from Sync Frames containing less than 6 Audio Blocks (i.e. when the burst payload is carrying audio from multiple substreams) then the Burst Payload shall contain the number of Sync Frames needed to deliver six blocks of audio data from all substreams being carried in the bitstream. Additionally, the transmitting device shall ensure that the first Sync Frame in the data-burst is always the Sync Frame whose “Converter Synchronization Flag” flag is set to “1”

5.4 AES3 Frame Rate (Sampling Frequency)

The frame rate of the AES3 stream used to transport the E-AC-3 Sync Frames shall be the same as the rate at which the baseband audio to be encoded was sampled.

If multiple E-AC-3 Burst Payloads are transported in the same AES3 stream, all the audio represented by these Burst Payloads must be sampled at the same rate. Bits 24-27 of the channel status word of the AES3 transport stream shall indicate the sampling frequency.

The following table shows the maximum data rate of the E-AC-3 bitstream that can be supported for each data word bit depth and sampling frequency of the AES3 interface. The data rates shown in this table refer to an AES3 interface operating in frame mode. Subframe mode shall not be used when transmitting E-AC-3 bit streams.

Table 3 – AES3 data bit depth and maximum available data rates for E-AC-3 data

E-AC-3 sampling frequency	Maximum data rate of E-AC-3 bit stream (kbps)		
	16-bit mode	20-bit mode	24-bit mode
32 kHz	1021.333	1276.667	1532
44.1 kHz	1407.525	1759.406	2111.288
48 kHz	1532	1915	2298

The maximum data rates shown in this table take into account the two AES3 frames of padding bits between data bursts that are required by Section 7.4, ”Burst Spacing” of SMPTE ST 337.

5.5 E-AC-3 Reference Point

The reference point of an E-AC-3 data burst shall be bit 0 of the Pa sync word, as shown in Figure 2.

5.6 E-AC-3 Standard Repetition Rate

The Standard Repetition Rate for E-AC-3 shall be such that Data Bursts carrying the same data stream number occur at 1536 AES3 frame intervals.

5.7 E-AC-3 Standard Decode Latency (Professional)

The E-AC-3 reference decode latency for professional applications shall be 3324 frame periods of the AES3 stream used to transport the E-AC-3 data (i.e., 69.25 ms at a 48-kHz frame rate).

The 3324 AES3 frame period is the sum of the time required to receive the Burst Payload (the E-AC-3 Sync Frame) and the time required to decode it.

5.8 E-AC-3 Reference Position

The Reference Position of a Burst Payload is defined by the relationship of the decoded audio to the associated video signal. A Burst Payload is in the Reference Position when the decoded audio from that Burst Payload is in sync with the associated video.

The Reference Point of the Burst Payload carried in an AES3 stream that is locked to the associated video signal must therefore precede the video sync point by 3324 AES3 frame periods.

5.9 Burst Payload Timing

It may not be possible to carry E-AC-3 Burst Payloads in the Reference position of an AES3 stream. SMPTE ST 339 defines a Time Stamp data type whose own Burst Payload carries either the offset (in AES3 frame periods) between the E-AC-3 Reference position and the Reference point of the E-AC-3 Burst Payload immediately following the Time Stamp, or a SMPTE ST 12-1 time code that applies to the decoded audio from the following E-AC-3 Burst Payload. Either method of time stamping the E-AC-3 Burst Payload allows the carriage of multiple E-AC-3 Burst Payloads in a single data stream. If an E-AC-3 Burst Payload does occupy the reference position, the delay shall be 0.

Burst Preambles and their associated Burst Payloads may be placed randomly in E-AC-3 data bursts. In this case, the Burst Preamble rep_rate_flag must be set to 1 to indicate that the Burst Payloads do not occur at the standard Repetition Rate. Time stamps shall be used to indicate when to present the decoded audio samples.

5.10 Use of Pause Data Bursts Between E-AC-3 Bit Streams

When a stream gap in an E-AC-3 stream is filled by a sequence of pause data-bursts, the Pa of the first pause data-burst shall be placed 1536 AES3 frames after the Pa word of the preceding E-AC-3 frame. The pause data-bursts shall be transmitted with a repetition period of 3 AES3 frames, unless other repetition periods are necessary to precisely fill the stream gap (whose length may not be a multiple of 3 AES3 frames), subject to the zero padding requirements of SMPTE ST 337, Section 7.4, “Burst Spacing”.

When it is possible to control the length of the gap in an E-AC-3 data stream, it is recommended that the gap length be an integer multiple of 256 AES3 frames, again respecting SMPTE ST 337, Section 7.4, “Burst Spacing”. This allows many E-AC-3 decoders to optimize the gap concealment process.

The sequence of Pause data bursts may be interrupted to allow other data bursts to be multiplexed into the AES3 transport. See SMPTE ST 339 for details of the Pause data type and its use.

Annex A Bibliography (Informative)

AES3-2009, AES Standard for Digital Audio — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data (a multi-part document)

SMPTE ST 12-1:2014, Time and Control Code

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

ITU-R BS.1196-1 (04/01), Audio Coding for Digital Terrestrial Television Broadcasting