SMPTE STANDARD

Format for Non-PCM Audio and Data in AES3 — AC-4 Data Type

Table of Contents

Page

Approved
November 6, 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 Standards Operating Manual.

SMPTE ST 2021 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.

1 Scope

This standard describes how AC-4 data rate reduced (non PCM) audio streams conforming to ETSI TS 103 190 are packed into an AES3 data stream, following the methods described by SMPTE ST 337. The AC-4 data type is identified by a data type number listed in SMPTE ST 338 and carried in the burst information word of the data burst preamble.

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

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 274:2008) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 274M-2008). Documents with the same root number (e.g. 274) and publication year (e.g. 2008) are functionally identical.

The following standards contain provisions that, 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.

ETSI TS 103 190 v.1.1.1 (2014-04), Digital Audio Compression (AC-4) Standard

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

SMPTE ST 274:2008, Television — 1920 × 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

SMPTE ST 296:2012, 1280 × 720 Progressive Image 4:2:2 and 4:4:4 Sample Structure — Analog and Digital Representation and Analog Interface

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

SMPTE ST 2036-1:2014, Ultra High Definition Television — Image Parameter Values for Program Production

4 Definitions and Acronyms

4.1 Definitions

4.1.1 AC-4 frame

An AC-4 sync frame, as specified in Annex A.

4.1.2 AC-4 I-frame

An independently decodable AC-4 frame, as specified in ETSI TS 103 190.

4.1.3 Audio Frame Rate

The number of AC-4 audio frames per second, indicated by the value of the frame_rate_index bitstream parameter as specified in ETSI TS 103 190.

4.1.4 Base sampling frequency

The sampling frequency of the AC-4 bitstream, indicated by the value of the fs_index bitstream parameter as specified in ETSI TS 103 190.

4.1.5 Fractional frame rates

AC-4 supports a number of fractional audio frame rates. These frame rates are written in shorthand notation, as specified in Table 1.

Table 1 – Shorthand notation for fractional frame rates

Fractional AC-4 audio frame rate (fps)	Shorthand version
$24 \times {}^{1\,000}/_{1\,001}$	23.976
$30 \times {}^{1\,000}/_{1\,001}$	29.97
$48 \times {}^{1\,000}/_{1\,001}$	47.952
$60 \times {}^{1\,000}/_{1\,001}$	59.94
$120 \times {}^{1\,000}/_{1\,001}$	119.88
12 000 / 512	23.438
11 025 / 512	21.533

4.1.6 Latency

Delay time of an external audio decoder to decode an AC-4 data burst, defined as the sum of two values; the receiving delay time and the decoding delay time.

4.2 Acronyms

4.2.1 ETSI

European Telecommunication Standards Institute

4.2.2 fps

frames per second

4.2.3 UIMSBF

Unsigned integer, most significant bit first

5 AC-4

5.1 Overview

AC-4 coded audio shall be transported in an AES3 data stream as a series of Data Bursts. Each Data Burst shall start with a Burst Preamble as defined by SMPTE ST 337, containing information about the Burst Payload, which shall follow the Burst Preamble. The Burst Payload shall consist of an AC-4 Frame. The Burst Payload shall be followed by enough zero padding words to make the resulting Data Burst duration exactly match the duration in samples of baseband (PCM) audio that the AC-4 coded audio represents.

The resulting Data Bursts shall be 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). Data Bursts shall be 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-4 Frame shall form the Burst Payload, as shown in Figure 1. As specified by ETSI TS 103 190 and Annex A of this document, each AC-4 Frame begins with a sync word and a frame size field that indicates the size of the following raw AC-4 frame. If the value of the sync word is AC41h, the raw AC-4 frame is followed by an error check (crc_word) word.

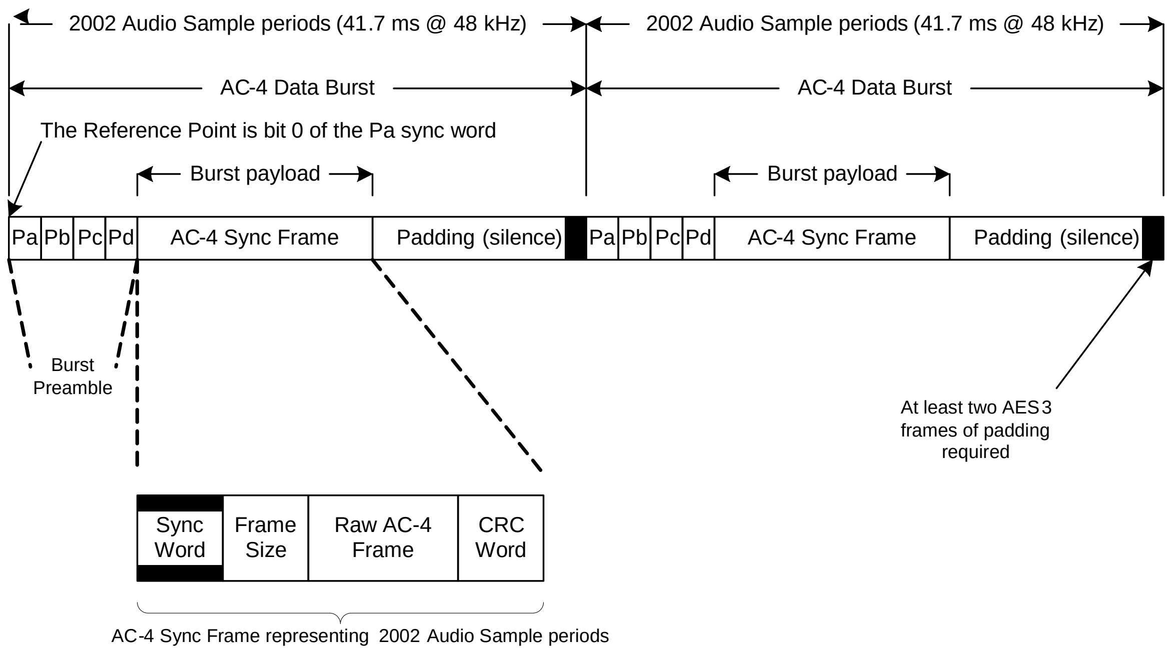

Figure 1 – Structure of an AC-4 data burst (AC-4 frame rate = 23.976 fps)

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 24 (see SMPTE ST 338).

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 AC-4 data type

data_type_dependent bit number	Value	Meaning
0 – 3	See Table 3 and Table 4	Repetition rate of the data-burst in AES3 frames
4	Repetition rate flag (rep_rate_flag)	

The value of bits 0 to 3 of the data_type_dependent field shall indicate the repetition rate of the AC-4 data-burst. The interpretation of these bits is dependent on the base sampling frequency of the AC-4 bitstream and the AES3 frame rate, as specified in Table 3 and Table 4.

Table 3 – Values of data_type_dependent field for AC-4 at an AES3 frame rate of 48 kHz

data_type_dependent bit numbers	Value	Meaning	
0 – 3	0	Repetition rate of the data-burst in AES3 frames	2002
	1		2000
	2		1920
	3		1601 / 1602 (see Table 7
	4		1600
	5		1001
	6		1000
	7		960
	8		800 / 801 (see Table 7)
	9		800
	10		480
	11		400 / 401 (see Table 7)
	12		400
	13		2048
	14 – 15	Reserved	
4	Repetition rate flag (rep_rate_flag)		

Table 4 – Values of data_type_dependent field for AC-4 at an AES3 frame rate of 44.1 kHz

data_type_dependent bit numbers	Value	Contents	
0 – 3	0 – 12	Reserved	
	13	Repetition rate of the data-burst in AES3 frames	2048
	14, 15	Reserved	
4	Repetition rate flag (rep_rate_flag)		

rep_rate_flag – The repetition rate flag shall be set to 0 if the AC-4 data burst is placed in the AES3 interface such that the Reference Point of the data burst (see Section 5.5) occurs at the AC-4 standard repetition rate (see Section 5.6). The flag shall be set to 1 if the reference point does not occur at the AC-4 standard repetition rate. This flag is intended to be set to the same state for all data bursts of a given AC-4 data stream to indicate whether data bursts for the stream occur at the standard repetition rate.

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 AC-4 Burst Payload

An AC-4 bitstream consists of a sequence of AC-4 frames. The AC-4 burst-payload shall consist of a single AC-4 frame. The length of the AC-4 data-burst will depend on the encoded bit rate (which determines the AC-4 frame length). The specification for the AC-4 bitstream may be found in ETSI document TS 103 190.

5.4 AES3 Frame Rate (Sampling Frequency)

The frame rate of the AES3 stream used to transport the AC-4 frames shall be the same as the rate at which the encoded audio was sampled.

If multiple AC-4 streams are transported in the same AES3 stream, all the audio represented by these Burst Payloads must be sampled at the same rate.

The units of burst-length shall be in bits. The maximum size of an AC-4 burst payload is dependent on the repetition rate of the data-burst, and is specified.

The maximum data rates shown in Table 5 take into account the two AES3 frames of padding bits between data bursts that are required by SMPTE ST 337, Section 7.3 (burst spacing), but does not take into account the use of the AC-4 guard band (specified in Section 5.9).

Table 5 – AES3 data bit depth and maximum available data rates for AC-4 data

Repetition rate of the AC-4 data-burst in AES3 frames	AC-4 base sampling frequency	Maximum data rate of AC-4 bitstream (kbit/s)		
		16-bit	20-bit	24-bit
2002	48 kHz	1532.931	1916.164	2299.397
2000		1532.928	1916.160	2299.392
1920		1532.800	1916	2299.200
1601/1602		1532.164	1915.205	2298.246
1600		1532.160	1915.200	2298.240
1001		1529.862	1912.328	2294.793
1000		1529.856	1912.320	2294.784
960		1529.600	1912	2294.400
800/801		1527.944	1909.930	2291.916
800		1528.320	1910.400	2292.480
480		1523.200	1904	2284.800
400/401		1521.423	1901.778	2282.134
400		1520.640	1900.800	2280.960
2048		1533	1916.250	2299.500
2048	44.1 kHz	1408.444	1760.555	2112.666

5.5 AC-4 Reference Point

The reference point of an AC-4 data-burst shall be bit 0 of the Pa sync word, as shown in Figure 1.

5.6 AC-4 Standard Repetition Rate

The Standard Repetition Rate of the AC-4 data-burst shall be defined by the base sampling frequency and frame rate of the AC-4 audio bitstream. As AC-4 supports multiple audio frame rates that match commonly used video frame rates, specifically RP 168, ST 274, ST 296 and ST 2036-1 as appropriate, the transmission device shall ensure that the selected AC-4 data-burst repetition rate is equal to the duration of the AC-4 frame, as specified in Table 6.

Table 6 – AC-4 base sampling frequency, AC-4 audio frame rate and corresponding AC-4 data-burst repetition rate

AC-4 base sampling frequency	AC-4 audio frame rate (fps)	AC-4 data-burst repetition rate in AES3 frames
48 kHz	23.976	2 002
	24	2 000
	25	1 920
	29.97	1 601 / 1 602 (see Table 7)
	30	1 600
	47.952	1 001
	48	1 000
	50	960
	59.94	800 / 801 (see Table 7)
	60	800
	100	480
	119.88	400 / 401 (see Table 7)
	120	400
	23.438	2 048
44,1 kHz	21.533	2 048

For AC-4 audio frame rates of 29.97, 59.94 and 119.88 fps, the duration of an AC-4 audio frame does not correspond to an integer number of AES3 frames. To ensure that precise time alignment is maintained between the AC-4 data-burst and the AC-4 audio frames at these frame rates, the repetition rate of data-bursts varies so that over a sequence of 5 data bursts, the AC-4 data-bursts are time-aligned with the corresponding 5 audio frames, as specified in Table 7.

Table 7 – AC-4 data-burst sequence and repetition rate variance to ensure AC-4 audio frame alignment at 29.97, 59.94 and 119.88 fps

AC-4 audio frame rate	Repetition rates for AC-4 data-burst sequence in AES3 frames				
	Data-burst 0	Data-burst 1	Data-burst 2	Data-burst 3	Data-burst 4
29.97 fps	1 602	1 601	1 602	1 601	1 602
59.94 fps	801	801	800	801	801
119.88 fps	400	401	400	401	400

For AC-4 bitstream operations (e.g. splice, edit, switch) requiring the 5-frame sequence(s) outlined in Table 7 to be maintained, detection of a specific data burst in the 5-frame sequence may be implemented utilizing the sequence_counter field defined in ETSI TS 103 190. The sequence_counter value modulo 5 can be used to determine the number of samples the decoder will produce for a particular input frame to maintain the 5-frame sequence.

5.7 AC-4 Standard Decode Latency

The latency of an AC-4 decoder is defined as the sum of the receiving delay time and decoding delay time.

The receiving delay time is the time taken to receive the complete AC-4 burst payload, and is dependent on the encoded bitrate and audio frame rate of the AC-4 bitstream. For the purposes of maintaining synchronization with video, it is recommended that a constant value of receiving delay time for each audio frame rate and corresponding repetition rate be maintained. This value is calculated based on the maximum possible size of an AC-4 burst-payload, and is equal to the time occupied by the duration of the AC-4 data-burst in AES3 frames at the AES3 frame rate.

The decoding delay time is equal to the time occupied by 3195 PCM samples at the AC-4 base sampling frequency.

5.8 AC-4 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.

For AC-4 audio frame rates of 23.438 and 21.533 fps that do not match video frame rates, 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 the number of AES3 frame periods equal to the duration of one AC-4 frame at the AC-4 frame rate, plus 3195 AES3 frames.

For AC-4 audio frame rates that match video frame rates, the AC-4 reference position is defined in relation to the video frame that corresponds to the AC-4-encoded audio samples.

An AC-4 data burst is defined as occupying the reference position in the AES3 stream if at least one of the following requirements is met:

1. Ideal reference position: The ideal reference position occurs in the AES3 stream at the first AES3 frame after the start of the defined video line of a video reference signal that corresponds to the AC-4 frame. The ideal AC-4 reference line position is specified in the “Ideal Position ±80µS” column of Table 8.

2. Acceptable reference position: the AC-4 data burst is considered to be in an acceptable reference position if the reference position is located ±80 microseconds from the ideal position. The acceptable reference positions are specified in the "Ideal Position -80µS" and "Ideal Position +80µS" column of Table 8.

3. Safe reference position: the AC-4 data burst is considered to be in a safe position (meaning no AC-4 frame data will be corrupted at a switch) if the reference point is within the following range:

   a. Earliest valid position: the reference point is immediately after the defined video switch point, as specified in the Earliest Valid Position column of Table 8.

   b. Latest valid position: the reference point ensures that the AC-4 frame data ends before the end of the video frame, as specified in the Latest Valid Position column of Table 8.

Table 8 – AC-4 reference position per video format

Video Format	AC-4 Frame Rate (fps)	Total Number of Lines	Earliest Valid Position		Ideal Position						Latest Valid Position	
					-80µS		±80µS		+80µS			
			Line	µS	Line	µS	Line	µS	Line	µS	Line	µS
625/50/i	25	625	8	450	11	650	12	730	13	810	30	1860
525/59.94/i	29.97	525	12	510	13	610	14	690	16	770	26	1400
1920x1080/60/i	30	1125	18	510	21	610	24	690	26	770	52	1530
1920x1080/59.94/i	29.97	1125	18	510	21	610	24	690	26	770	48	1400
1920x1080/50/i	25	1125	13	450	19	650	21	730	23	810	53	1860
1920x1080/48/i	24	1125	11	390	25	890	27	970	29	1050	98	3610
1920x1080/47.97/i	23.98	1125	11	390	25	890	27	970	29	1050	98	3610
1920x1080/60/p	60	1125	35	510	42	610	47	690	52	770	104	1530
1920x1080/59.94/p	59.94	1125	35	510	42	610	47	690	52	770	95	1400
1920x1080/50/p	50	1125	26	450	37	650	42	730	46	810	105	1860
1920x1080/30/p	30	1125	18	510	21	610	24	690	26	770	52	1530
1920x1080/29.97/p	29.97	1125	18	510	21	610	24	690	26	770	48	1400
1920x1080/25/p	25	1125	13	450	19	650	21	730	23	810	53	1860
1920x1080/24/p	24	1125	11	390	25	890	27	970	29	1050	98	3610
1920x1080/23.98/p	23.98	1125	11	390	25	890	27	970	29	1050	98	3610
1280x720/60/p	60	750	23	510	28	610	32	690	35	770	69	1530
1280x720/59.94/p	59.94	750	23	510	28	610	32	690	35	770	63	1400
1280x720/50/p	50	750	17	450	25	650	28	730	31	810	70	1860
1280x720/30/p	30	750	12	510	14	610	16	690	18	770	35	1530
1280x720/29.97/p	29.97	750	12	510	14	610	16	690	18	770	32	1400
1280x720/25/p	25	750	9	450	13	650	14	730	16	810	35	1860
1280x720/24/p	24	750	8	390	17	890	18	970	19	1050	65	3610
1280x720/23.98/p	23.98	750	8	390	17	890	18	970	19	1050	65	3610
µS = Microseconds												

The AC-4 data burst shall be coincident in time with the video frame that corresponds to the audio samples coded within the AC-4 frame.

Note: Requirements 1 and 2 reference a defined video reference signal. Requirement 3 references a content video signal that might be the video reference signal. If not, it is assumed that the AC-4 burst is locked to the video reference signal.

Note: Requirements 1 and 2 meet the basic AC-4 phase synchronization requirement for AES3 transport applications. Requirement 3 defines the reference location for lip sync.

5.9 AC-4 Guard Band

To allow a video stream to be cut-edited or switched while ensuring that the accompanying AC-4 data remains error-free, a period of AES3 null data precedes the AC-4 data burst. This null data is referred to as the AC-4 guard band. The guard-band location and duration is chosen so that AC-4 data is not present on the AES3 interface during video lines in which switches or edits will occur, and so that some delay (a few lines) can be tolerated in the audio path.

The portion of the guard band that precedes the AC-4 data burst begins at the AES3 sample that is aligned to the SMPTE RP 168, ST 274, ST 296 and ST 2036-1 as appropriate reference point, and ends at the start of the AC-4 data burst at the AC-4 reference position (see Table 8). The size and location of the guard band are determined by the synchronization requirements specified in Section 5.7, and the size of the encoded AC-4 frame. The total duration of the AC-4 guard band for each frame rate is shown in Table 9.

Table 9 – AC-4 guard band duration

AC-4 Frame Rate (fps)	Guard band duration (AES3 frames)
23.976	100
24	100
25	96
29.97	80
30	80
47.952	50
48	50
50	48
59.94	40
60	40
100	24
119.88	20
120	20

Figure 2 shows the alignment of a AC-4 data burst from a 25 fps AC-4 stream, and a 625 line, 25 fps video stream. In this case, the AC-4 reference position is at the Ideal Position as specified in Table 8.

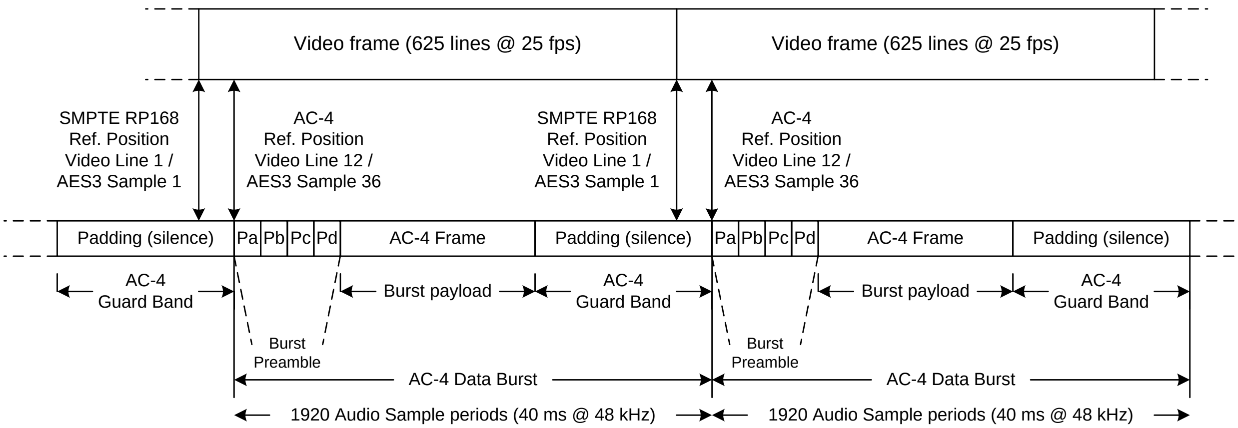

Figure 2 – 25 fps AC-4 data burst alignment with 625 line video @ 25 fps

Note: The AC-4 guard band applies to AC-4 data bursts only. Other data burst types can appear during the AC-4 guard-band interval. For instance, captioning data, timecode, or other types of low-bit-rate data can be carried in the space not used by AC-4. In general, this data has a higher chance of being corrupted during a switch or cut edit.

5.10 Burst Payload Timing

It might not be possible to carry AC-4 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 AC-4 Reference position and the Reference point of the AC-4 Burst Payload immediately following the Time Stamp, or a SMPTE ST 12-1 time code that applies to the decoded audio from the following AC-4 Burst Payload. Either method of time stamping the AC-4 Burst Payload allows the carriage of multiple AC-4 Burst Payloads in a single data stream. If an AC-4 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-4 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.

5.11 Use of Pause Data Bursts Between AC-4 Data Bursts

When a stream gap in an AC-4 stream is filled by a sequence of pause data-bursts, the Pa of the first pause data-burst shall be located one frame repetition rate following the Pa of the previous AC-4 frame. The pause data-bursts shall be transmitted with a repetition rate of 3 AES3 frames, unless other repetition rates 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 7.3.

When it is possible to control the length of the gap in an AC-4 data stream, it is recommended that the gap length be an integer multiple of the AC-4 audio frame duration (2 002, 2 000, 1 920, etc. AES3 frames), again respecting SMPTE ST 337, Section 7.3. This allows AC-4 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 AC-4 Sync Frame

Figure A.1 shows how synchronization, frame size and CRC fields are added to a raw AC-4 frame, as defined in ETSI TS 103 190, to create an AC-4 sync frame. Each AC-4 burst-payload contains a single AC-4 sync frame.

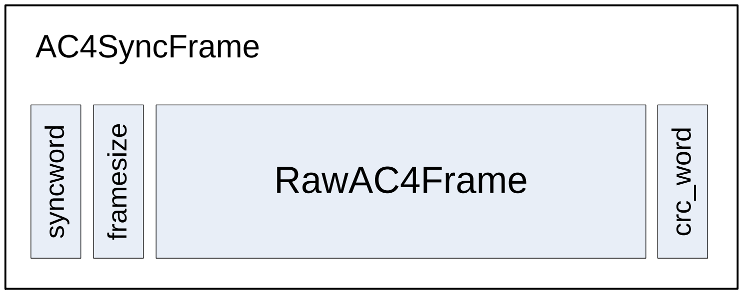

Figure A.1 – An AC-4 sync frame

The syntax of the AC-4 sync frame shall be as specified in Table A.1.

Table A.1 – AC-4 sync frame syntax

Syntax	No. of Bits	Identifier
`ac4syncframe() {`		
`syncword`	16	uimsbf
`framesize()`		
`RawAC4Frame()`		
`if (syncword == '0xAC41') {`		
`crc_word`	16	uimsbf
`}`		
`}`		

The syncword shall be from Table A.2 according to the presence of the CRC word.

Table A.2 – AC-4 sync word values

Value	Description
0xAC41	CRC word is present
0xAC40	CRC word is not present

The value of the framesize field shall be equal to the size of the RawAC4Frame in bytes. The function for deriving the frame size of the RawAC4Frame in bytes is given in Table A.3.

Table A.3 – framesize()

Syntax	No. of Bits	Identifier
`frameSize() {`		
`size`	16	uimsbf
`if (size == '0xFFFF') {`		
`size`	24	uimsbf
`}`		
`return size`		
`}`		

The 16-bit CRC word is optional for AC-4. Its presence shall be determined by the value of the syncword (see Table A.2).

The following CRC generator polynomial shall be used to generate the CRC word: $x^{16} + x^{15} + x^2 + 1$.

The CRC calculation on the decoder side may be implemented by one of several standard techniques, e.g. by a linear feedback shift register (LFSR). A CRC-protected AC-4 sync frame is considered valid if the register contains all zeros after the complete data of the AC-4 sync frame, excluding the syncword (but including the frameSize), has been shifted in.

Annex B 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