SMPTE 339-2008
Revision of
SMPTE 339M-2004

SMPTE STANDARD

Format for Non-PCM Audio and Data in AES3 — Generic Data Types

Table of Contents

Copyright © 2008 by THE SOCIETY OF
MOTION PICTURE AND TELEVISION ENGINEERS
3 Barker Avenue, White Plains, NY 10601
(914) 761-1100

Approved
October 23, 2008

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 Part XIII of its Administrative Practices. This SMPTE Engineering Document was prepared by Technology Committee N26.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. 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 specifies data type specific format requirements for null data, time stamp data, user defined data, pause data and utility data which may include video frame synchronization information that may be carried within an AES3 interface according to SMPTE 337.

1 Scope

This standard specifies data type specific format requirements for several types of data bursts that may be carried within an AES3 interface according to SMPTE 337. Included are descriptions of the data type (defined in SMPTE 338), the format of the burst_payload for the data type, the coding of data type dependent fields in the burst_preamble, and additional data burst and bitstream formatting requirements not defined in SMPTE 337. This includes specific synchronization methods which may affect formatting.

This standard covers generic data types which include null data, time stamp data, user defined data, pause data, and utility data which may include video frame synchronization information.

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 standard. 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.

AES3-2003, AES Standard for Digital Audio Engineering — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data (Revision of AES3-1992)

SMPTE 12M-1-2008, Television — Time and Control Code

SMPTE 337-2008, Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

SMPTE 338-2008, Format for Non-PCM Audio and Data in AES3 — Data Types

SMPTE RP 168-2002, Definition of Vertical Interval Switching Point for Synchronous Video Switching

4 Null Data (data_type = 0)

The null data type is provided so that the preamble sync codes may be intentionally inserted into the data stream. Null data bursts may be inserted in an AES3 transport stream to help receivers detect the fact that the stream is carrying non-PCM audio data. Null data bursts shall obey the burst spacing requirements specified by SMPTE 337, § 7.3. The Null data burst should not be used as a sync run in sequence. The Pause data type should be used instead.

4.1 Null Data burst_preamble

The burst_preamble for a null data burst shall set the length_code, error_flag, and data_type_dependent values to 0. The data_stream_number shall be set to any valid number other than 0x7. The data_type shall be set to 0 as defined in SMPTE 338.

4.2 Null Data burst_payload

The null data burst shall not contain a burst_payload (a payload of length 0).

5 Time Stamp (data_type = 2)

The time stamp data type is used to convey time synchronization information associated with a specific data burst. This information shall include SMPTE 12M-1 time code information and/or specific delay information. The use of time stamp bursts is optional. When present, the time stamp information shall apply to the data burst which immediately follows in the AES3 interface. When time code information is present, the time stamp may contain date and time zone information coded in the time code binary groups according to SMPTE 309M.

5.1 Time Stamp burst_preamble

The burst_preamble for a time stamp data burst shall include a data_type_dependent field set to a value of 0. The data_stream_number shall be set to 0x7. The data_mode parameter shall be set to 0. The data_type shall be set to 2 as defined in SMPTE 338.

NOTE: When time code information is conveyed within time stamp data bursts, data stream number 7 should not be interpreted as a single data stream of time code information. Time stamp data bursts are independent and apply only to the data burst that immediately follows in the AES3 interface; therefore time code information from multiple time code sources may be contained within data stream number 7. Within an individual data burst time code information from at most one time code source will be present.

5.2 Time Stamp burst_payload

The time stamp burst_payload shall always be packed in the 16 bit data mode and shall have a length of six to nine 16-bit words. The burst_payload shall be in the format shown in Table 1.

The flag bit f2 in word 5 shall indicate the status of all SMPTE 12M-1 LTC time code fields in words 0-5. If this bit is set to 0, 12M-1 time code information is present. If this bit is set to 1, SMPTE 12M-1 time code information is not present and all bit fields in words 0 to 5 other than bit f2 are not defined.

In Table 1, numbers in brackets (e.g., [63]) shall refer to the corresponding bit numbers of the SMPTE 12M-1 LTC time code words. The time code hours field is denoted by bit fields H20-H1, which corresponding to the

SMPTE 12M-1 BCD coded hours bits (e.g., H20 corresponds to 20s of hours). The minutes, seconds, and frames address fields are indicated in a similar fashion by the bit fields M40-M1, S40-S1, and F20-F1 respectively.

The flag bit f1 in word 5 shall indicate the status of the binary group and flag bit fields in words 0 to 3. If this bit is set to a 1 the binary group and flag bit fields in words 0 to 3 have been copied from a source of SMPTE 12M-1 time code. If this information has not been provided, flag bit f1 shall be set to a 0, and these binary group and flag bit fields shall be set to 0.

Bit fields a3-a0 of word 5 carry a 4 bit frame rate code. The frame rate code shall be encoded as shown in Table 2. Bit 0 of word 5 shall carry the drop-frame flag bit from bit 10 of the SMPTE 12M-1 time code word. If the f1 bit (word 5, bit 1) is set to 1, bit 6 of word 3 shall also carry the drop-frame flag. Reserved bits (R) in word 5 shall not be used and shall be set to 0.

NOTE: The interpretation of the 12M-1 time code flag bits is dependent on the time code frame rate and whether or not the codeword is LTC or VITC.

Word 4 of the time stamp shall be used to associate the time code carried in words 0 to 3 with a specific sample of the linearly coded audio recovered from the data burst carrying (low bit rate) encoded audio immediately following the time stamp burst. The sample number is carried as an unsigned integer. For example, if an audio data burst represents 1536 linear PCM samples and the sample number is set to 1007, then the SMPTE 12M-1 time code applies to sample 1007 of the 1536 samples of the linearly coded audio.

If the sample number contained in a time stamp is greater than the number of samples represented by the corresponding encoded audio data burst, then the time code shall be associated with a sample of the linearly coded audio recovered from the following data burst having the same stream number. As an example of this case, and again using an audio data burst that represents 1536 linear PCM samples, if word 4 is set to 1546, then the SMPTE 12M-1 time code applies to sample 10 of the 1536 samples recovered from the next data burst with the same stream number. In such cases, the same SMPTE 12M-1 time code information may be repeated more than once with different sample numbers in each case. In all cases, the sample number shall be accurate to within $\pm$ 0.5 ms of the ideal value.

When the time stamp is followed by a data burst that does not contain encoded audio, word 4 is not used, and shall be set to a value of 0.

Table 1 – Time stamp burst_payload

Time Stamp		MSB		Bit Number												LSB	
Payload Word		15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	Usr8, Usr7, flags, hours	[63]	[62]	[61]	[60]	[55]	[54]	[53]	[52]	[59]	[58]	H20	H10	H8	H4	H2	H1
1	Usr6, Usr5, flag, minutes	[47]	[46]	[45]	[44]	[39]	[38]	[37]	[36]	[43]	M40	M20	M10	M8	M4	M2	M1
2	Usr4, Usr3, flag, seconds	[31]	[30]	[29]	[28]	[23]	[22]	[21]	[20]	[27]	S40	S20	S10	S8	S4	S2	S1
3	Usr2, Usr1, cf, df, frames	[15]	[14]	[13]	[12]	[7]	[6]	[5]	[4]	[11]	[10]	F20	F10	F8	F4	F2	F1
4	Sample number	s15	s14	s13	s12	s11	s10	s9	s8	s7	S6	s5	s4	s3	s2	s1	s0
5	Reserved, flags	R	R	R	R	R	R	R	R	R	f2	a3	a2	a1	a0	f1	[10]
6	User private (optional)	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U
7	User private (optional)	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U
8	Delay (optional)	D15	D14	D13	D12	D11	D10	D9	D8	D7	D6	D5	D4	D3	D2	D1	D0

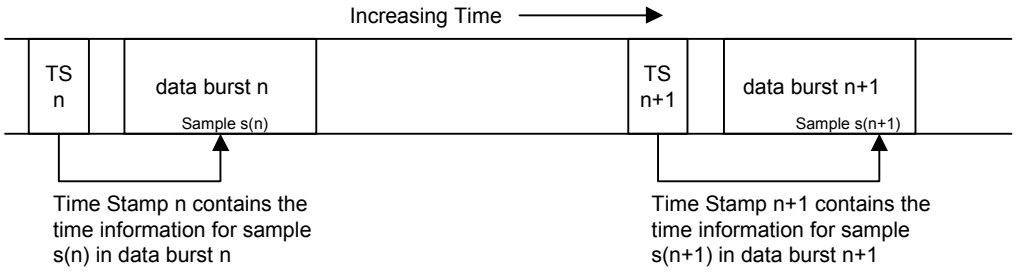


Figure 1 – Sample number indication in time stamp data bursts

Table 2 – Frame rate code

frame rate code				frame rate
a3	a2	a1	a0	
0	0	0	0	not indicated
0	0	0	1	24 ÷ 1.001 (23.98)
0	0	1	0	24
0	0	1	1	25
0	1	0	0	30 ÷ 1.001 (29.97)
0	1	0	1	30
0	1	1	0	50
0	1	1	1	60 ÷ 1.001 (59.94)
1	0	0	0	60
-	-	-	-	reserved
1	1	1	1	reserved

Words 6 and 7 are optional words that may contain user private data; however their presence is required when the optional delay field (word 8) is present.

Word 8 is optional. If present, it shall contain a delay indication denoted by bit fields D15-D0. The delay field is a signed integer (in 2's complement format) that indicates the offset, in terms of AES3 frames, of the reference point of the corresponding data burst from the defined reference position for that data burst. A positive value shall indicate the reference point is advanced (in time) from the reference position, while a negative value shall indicate the reference point is delayed (in time) from the reference position. For instance a delay setting of +2037 indicates that the reference point of the corresponding data burst is present in the AES3 bitstream 2037 frames ahead of the reference position for the burst. An exception is the value 0x8000 which shall be used to indicate no delay information is provided even though word 8 is present. This gives a range of ± 32767 AES3 frames, equivalent to approximately ± 682 msec with an AES3 reference sample rate of 48 kHz.

The definitions of the reference point and reference position are relative to the data_type setting of the corresponding data burst. The definitions may vary between data types and may not exist for some data types, in which case the meaning of the delay field is undefined.

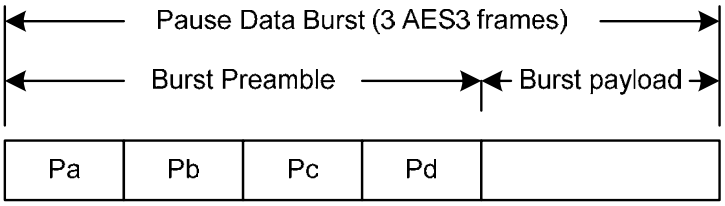
6 Pause (data_type = 3)

6.1 General

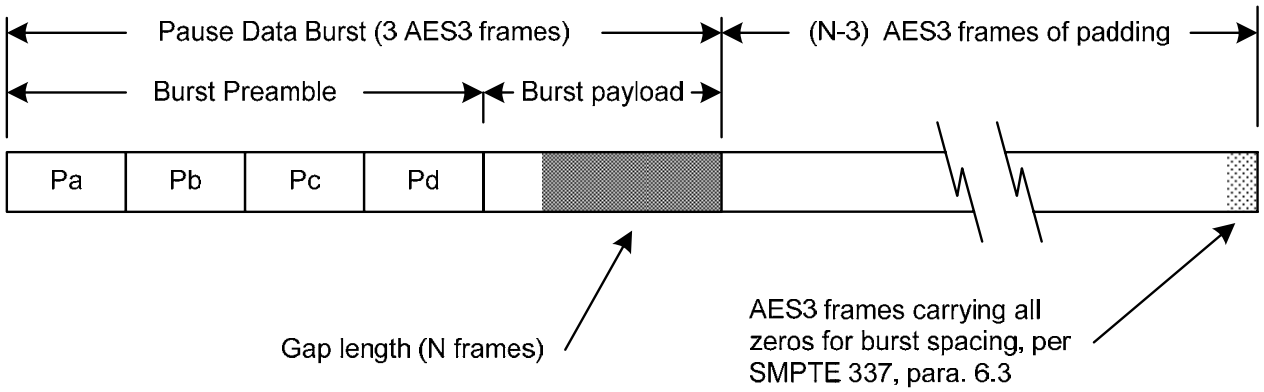
The Pause data type is provided to fill gaps in streams of other types of data, such as the non-PCM audio data described by SMPTE 340M. The Pause data type may be used to provide a consistent stream of data to downstream equipment to keep it operating properly, to signal the duration of a known gap in a data stream, or as a synchronizing “run-in sequence”, for the benefit of downstream equipment. SMPTE 340M, § 4.10 and § 5.10, specify the use of the Pause data type in AES3 streams carrying ATSC A/52B data types.

The Pause data burst shall consist of a burst_preamble as specified by SMPTE 337, followed by a one AES3 frame long burst payload that may carry gap length information. The Pause data burst shall operate only in the frame mode, thus the preamble and payload occupy a total of three AES3 frames. If the burst payload calls for a gap length of (N) AES3 frames, then that Pause data burst shall be followed by (N-3) AES3 frames of padding data (silence). See § 6.5 for use of the Pause data burst with multiplexed data streams.

Note that the Pause data burst is most commonly found in consumer (SP-DIF format) data streams per IEC 61937. Because some professional equipment may use consumer style interface chips, the Pause data burst can be found in data streams carried by the professional AES3 transport.



The Pause data burst without gap length information in the payload



The Pause data burst with gap length information in the payload

Figure 2 – Structure of the Pause data burst and Pause data burst followed by padding frames Pause data burst_preamble

6.2 The burst_preamble

The burst_preamble format is specified by SMPTE 337, Tables 6 and 7.

The data_type shall be set to 3 as specified by SMPTE 338.

The data in the burst payload of the Pause data burst shall be carried in the same 16, 20 or 24 bit mode as the data in the audio stream containing the gap to be filled. The data_mode bits of the burst_ preamble shall reflect this.

The data_type_dependent bit values of the pause data-bursts are given in Table 3.

Table 3 – Values of data_type_dependent info of the pause data-burst

data_type_dependent bit numbers	Value	Interpretation
0-4	0 1 2 – 31	General use * Stop, frame sequence discontinued ** Reserved
* - A pause data_burst with data_type_dependent info set to 'General use' is used to fill a gap or as a sync run-in sequence preceding an encoded audio bitstream. ** - Transmitters may use the STOP value to indicate that the transmission of the associated encoded audio bitstream has ceased. The AES3 transport itself shall remain active.		

The pause data_burst shall carry the same data_stream_number as the data_stream_number of the audio data stream which contains the stream gap to be filled with the pause data_bursts, or for which it is acting as a synchronization run-in sequence. In the case where a main audio service data stream and one or more associated audio service data streams are interleaved on the interface, the pause data-bursts shall have the same data_stream_number as the main audio service. The pause data burst shall only be used to fill the stream gaps between data bursts of the main audio service data streams.

6.3 Pause data_burst_payload

The pause data_burst_payload is 32, 40 or 48 bits long, depending on the data_mode value . The first 16-bits of the payload contain the optional audio_gap_length parameter. The remaining bits are reserved and shall all be set to '0'. When used, the value of the audio_gap_length parameter shall be equal to the number of PCM audio samples which would be missing from one channel of the decoded output signal.

Table 4 – Burst Payload Permitted values

Bits of payload LSB..MSB	Value	Interpretation	Remarks
0 – 15	0	Not specified	Mandatory when data-type-dependent info = 1.
	1	Reserved	
	2	Reserved	
	3 – 65 535	Gap-length	Gap length measured in number of AES3 frames (optional)
16 – 31	0	Reserved	All '0'
32 – 39	0	Reserved	All '0' when data_mode > 0
40 – 47	0	Reserved	All '0' when data_mode > 1

A gap shall be filled with either a sequence of Pause data_bursts or a single Pause data_burst carrying the audio gap-length, followed by the required amount of padding. AES3 frames carrying an all zero payload shall be used for padding, unless a second data stream is multiplexed into the AES3 transport, as detailed in § 6.5. The data_stream_number of the Pause data burst shall be the same as the data_stream_number of the paused data stream. Figure 3, parts A and B shows the use of the Pause data burst without and with the gap length carried in the burst payload.

If the data source does not have the information about the audio gap length at the time the stream gap begins, the pause data burst may carry an initial value for gap-length. If the gap is longer than the initial indication, another pause data burst carrying a new gap length value may be sent, or a sequence of pause data bursts alone may be sent until the stream of encoded audio data bursts resumes.

6.4 Burst Spacing

Pause data_bursts shall be spaced as specified by SMPTE 337, § 7.3.

6.5 Multiplexed Data Streams

SMPTE 337 allows the carriage of several data streams, multiplexed into a single AES3 transport.

When it is necessary to multiplex one or more additional data streams into the same AES3 transport, and to transmit Pause data bursts because one of those streams has a gap in it, a three frame Pause data burst carrying an audio_gap_length parameter, but not followed by any padding, shall be transmitted immediately before the additional multiplexed data stream. The audio_gap_length parameter of that Pause data burst shall be set equal to the length of the additional data stream(s) multiplexed into the AES3 transport, measured in AES3 frames, as shown in Figure 3, part C.

The data_stream_number of the Pause data burst shall be the same as the data_stream_number of the paused data stream.

The data streams to be muliplexed into the AES3 transport may be any of the types identified in SMPTE 338.

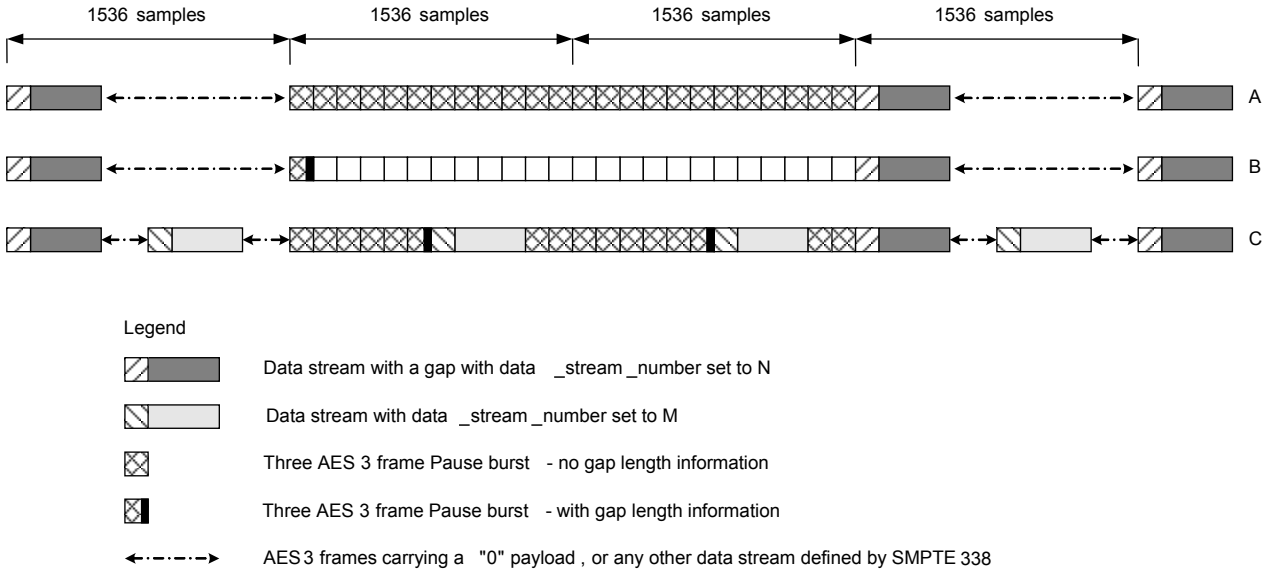


Figure 3 – Use of Pause data bursts in an AES3 transport carrying one or

more data streams, one of which has a gap

7 Utility (data_type = 26)

The utility data type provides additional functions that support synchronization and reception of SMPTE 337 formatted data.

7.1 Utility Data burst_preamble

The data_type shall be set to 26 as defined in SMPTE 338. The data_stream_number shall be set to any number other than 0x7. The data_mode parameter shall be set to a value of 0 (16-bit mode). Bit 4 of the data_type_dependent field shall be set to indicate the specific utility function as defined in Table 5. Utility functions are described in § 7.2. When bit 4 is set to a 1, the value of bits 3-0 of the data_type_dependent field are undefined.

Table 5 – Value of data_type_dependent bit 4 utility data

data_type_dependent bit 4	Utility function
0	V-sync
1	Undefined

7.2 Utility Data burst_payload

The contents of the burst_payload and data burst formatting requirements are defined by the utility function as specified in the data_type_dependent field.

7.2.1 V-sync function

The V-sync function is intended to allow identification of an alignment point between an AES3 stream containing SMPTE 337 formatted data and a corresponding video raster. Knowledge of the alignment point may be used to aid in synchronizing the AES3 signal with the associated video signal. The V-sync function may also be used to define a set of AES3 frames containing SMPTE 337 data that are associated with a specific video frame/field.

7.2.1.1 V-sync data_type_dependent

V-sync data bursts shall set bits 3-0 of the data_type_dependent field as shown in Table 6 to indicate the frame/field rate of the video signal with which the AES3 signal is associated. When associated with a 25, 29.97, or 30 frame per second (fps) video signal, the V-sync frame rate may be indicated as either the frame rate or twice the frame rate (i.e., 50, 59.94, or 60 fps) of the video signal. When associated with to a 50, 59.94, or 60 fps video signal, the frame rate may be indicated as either the frame rate or half the frame rate (i.e., 25, 29.97, or 30 fps). Bits 3-0 of the data_type_dependent field shall be set to 0 (frame rate not indicated) in all other cases.

Table 6 – V-sync data_type_dependent field bits 3-0 (frame rate code)

data_type_dependent bits 3-0				V-sync frame rate
b3	b2	b1	b0	
0	0	0	0	Not indicated
0	0	0	1	24 ÷ 1.001 (23.98)
0	0	1	0	24
0	0	1	1	25
0	1	0	0	30 ÷ 1.001 (29.97)
0	1	0	1	30
0	1	1	0	50
0	1	1	1	60 ÷ 1.001 (59.94)
1	0	0	0	60
–	–	–	–	reserved
1	1	1	1	reserved

7.2.1.2 V-sync burst_payload

The V-sync data burst carries no payload and shall set the burst_preamble length_code to 0.

7.2.1.3 V-sync reference position

V-sync data bursts are intended to occur at regular intervals within the AES3 signal synchronous to a related video signal. The reference point of a V-sync data burst is defined as the burst preamble word Pa. A V-sync data burst is defined as being in the reference position when the reference point of the data burst (preamble word Pa) is placed in the first AES3 frame occurring after the vertical sync reference point of the associated video signal. The vertical sync reference point is the beginning of the first line of the vertical sync interval as shown in SMPTE RP 168. V-sync data bursts may be coded in SMPTE 337 frame mode or subframe mode. In the frame mode, the V-sync preamble word Pa reference position shall always occur in subframe 1 (Ch1) of an AES3 frame (following an AES3 X or Z preamble). In the subframe mode, the V-sync Pa word reference position may occur in either AES3 subframe.

The location of V-sync data bursts may be used to define a set of SMPTE 337 data bursts that are associated with a specific video frame/field. Downstream equipment shall assume that the AES3 frames containing the V-sync burst and all AES3 frames up to but not including the subsequent V-sync burst are associated with the video frame whose vertical sync reference point occurs closest in time to the V-sync burst. In the event that the AES3 data between V-sync bursts contains a time code value, the linkage may be deduced in the event of multi-frame offsets.

7.2.1.4 V-sync usage

The precise alignment between AES3 signals and the corresponding video signal is often lost as these two signals travel through video facilities. Identification of a specific AES3 frame that is intended to line up with the reference point of an associated video signal allows downstream equipment to match a set of AES3 frames to a corresponding video frame. One example would be to determine which AES3 frames of data to include in an access unit of SMPTE 302M.

NOTE: The time stamp data bursts may also be used for synchronization to an associated video reference signal. However, time stamp data bursts must occur immediately prior to data bursts to which they correspond. V-sync data

bursts are not associated with other data bursts and may occur even in the absence of other SMPTE 337 data, which, as with null data bursts, may be useful for maintaining receiver synchronization.

NOTE: The reference location of the V-sync data bursts will vary by up to one AES3 frame with respect to the video vertical sync reference point, and at some frame rates (e.g., 29.97 fps) the location of the V-sync data burst will vary between video frames. Receivers should allow for the V-sync data burst location to vary by at least ±1 AES3 frame from the ideal location. In practice, V-sync data bursts may vary significantly from the defined reference position due to equipment processing or other system delays. See SMPTE RP 168 for a more complete discussion of vertical timing under different video formats.

7.2.1.5 V-sync example

Figure 4 and Figure 5 show an example of V-sync data burst placement with a SMPTE 337 data signal in a 48-kHz AES3 stream associated with a 29.97 (NTSC) interlaced video signal. In this example the V-sync data burst is coded as a 29.97 frame rate in the SMPTE 337 frame mode. The first preamble word Pa is, therefore, located in subframe 1 (Ch 1) of the first AES3 frame following the odd field vertical sync reference point. In this example, the actual number of AES3 frames associated with each video frame will vary between 1601 and 1602 frames.

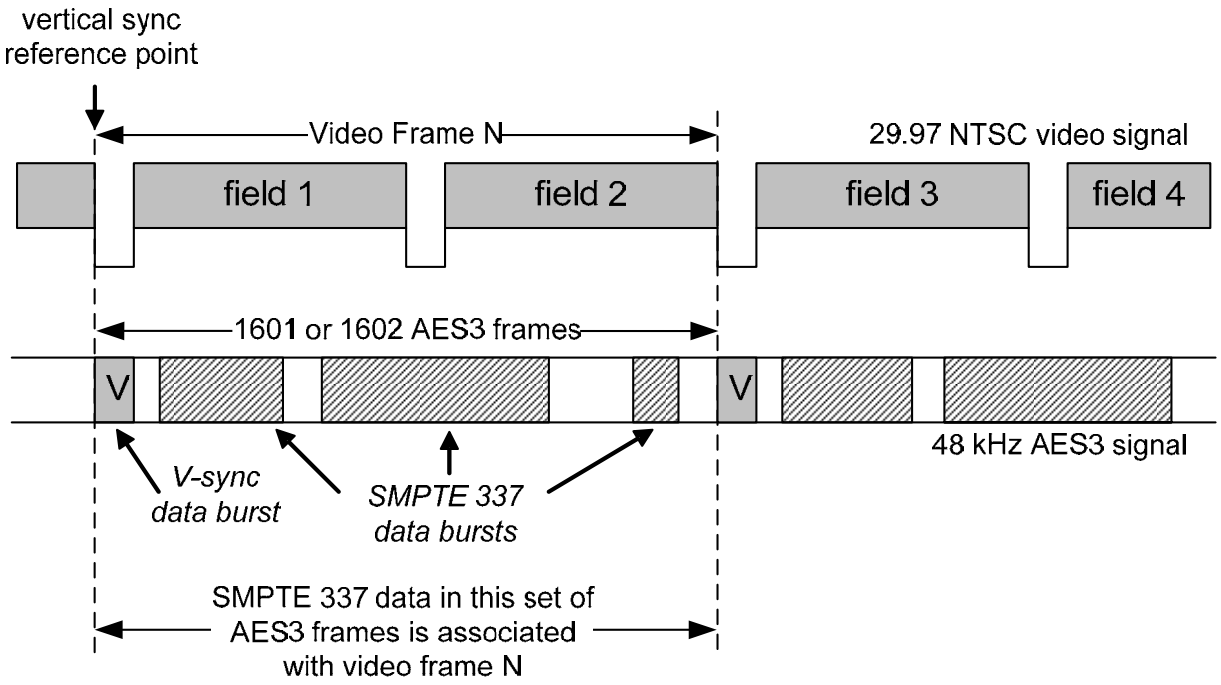


Figure 4 – V-sync data burst example (29.97 frame rate with interlaced video signal)

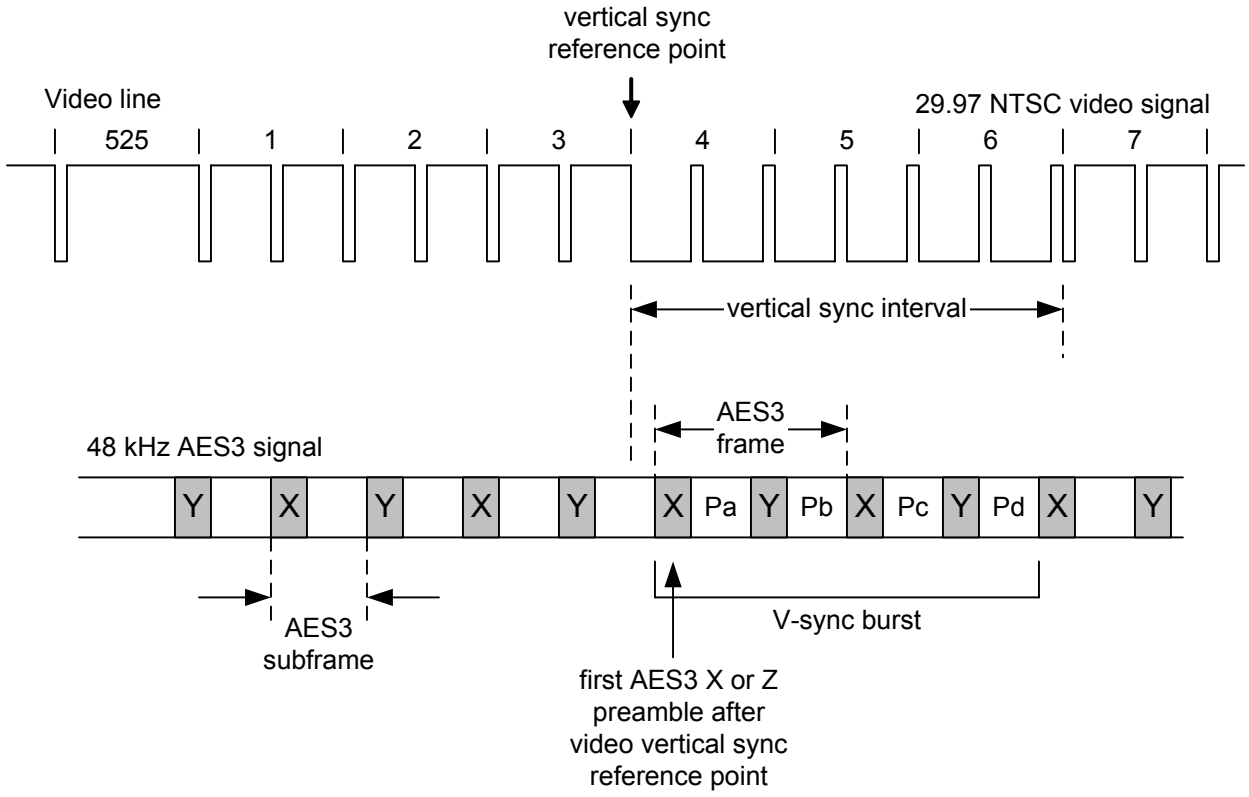


Figure 5 – V-sync reference position example (29.97 frame rate with interlaced video signal)

8 User Defined (data_type = 30)

The user data type is provided for the transmission of arbitrary user data.

8.1 User Data burst_preamble

The data_type_dependent field is undefined for the user data type. The data_stream_number shall be set to any valid number other than 0x7. The data_type shall be set to 30 as defined in SMPTE 338.

8.2 User Data burst_payload

The user data burst may be of any length and data mode. The contents of the burst_payload are undefined.

Annex A (Informative)
Bibliography

ATSC A/52B: Digital Audio Compression (AC-3) (E-AC-3) Standard, Revision B (June 2005)

IEC 61937:2000, Digital Audio — Interface for Non-Linear PCM Encoded Audio Bitstreams Applying IEC 60958

SMPTE 302M-2007, Television — Mapping of AES3 Data into MPEG-2 Transport Stream

SMPTE 309M-1999, Television — Transmission of Data and Time Zone Information in Binary Groups of Time and Control Code

SMPTE 340-2008, Format for Non-PCM Audio and Data in AES3 — ATSC A/52B Digital Audio Compression Standard for AC-3 and Enhanced AC-3 Data Types

SMPTE RP 168-2002, Definition of Vertical Interval Switching Point for Synchronous Video Switching