PROPOSED SMPTE STANDARD

for Television —

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

SMPTE 339M
Revision of
SMPTE 339M-2000

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 337M. Included are descriptions of the data type (defined in SMPTE 338M), 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 337M. 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, and utility data which may include video frame synchronization information.

2 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 standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

AES3-2003, AES Standard for Digital Audio — Digital Input-Output Interfacing — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data (AES3) *anticipating satisfactory completion of the current Call for Comment*

SMPTE 12M-1999, Television, Audio and Film — Time and Control Code

SMPTE 337M-2000, Television — Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

SMPTE 338M-2000, Television — 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

3 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 of use in receiving devices to potentially enhance reliable autodetection of whether or not the subframe contains PCM audio or non-PCM data.

3.1 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 338M.

3.2 Null data burst_payload

The null data burst shall contain no burst_payload (a payload of length 0).

4 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 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.

4.1 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 338M.

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.

4.2 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 time code fields in words 0-5. If this bit is set to 0, 12M time code information is present. If this bit is set to 1, SMPTE 12M 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]) refer to the corresponding bit number of the SMPTE 12M time code word (the bit numbers reflect the LTC bit assignment). These bits shall correspond to binary groups (Usr8-Usr1), color frame flag (cf), drop frame flag (df), and other flag bits from a SMPTE 12M time code word. The SMPTE 12M time code hours address field is denoted by the bit fields H20-H1 corresponding to the SMPTE 12M BCD coded hours bits (e.g., H20 corresponds to 20s of hours). The minutes, seconds, and frames address fields are denoted 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 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.

In word 5, bit fields a3-a0 denote a 4 bit frame rate code. The frame rate code shall be encoded as shown in table 2. Bit 0 of word 5 shall contain the drop-frame flag bit (bit 10 of the SMPTE 12M time code word). If the f1 bit is set to 1, bit 6 of word 3 will also contain 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 time code flag bits is dependent on the time code frame rate. The function of the phase bit / field flag bit is different for LTC and VITC time codes.

Word 4 carries a value designated sample number. This word is provided for use when the time stamp is followed by (and applies to) a data burst containing low bit rate encoded audio. The sample number field allows the time code value to be associated with a particular audio sample that is represented in a burst of encoded audio. When the time stamp is followed by a data burst that does not contain encoded audio, this field is not used but must still be present. In this case it shall be set to a value of 0.

Table 1 – Time stamp burst_payload

Time Stamp Payload Word		MSB Bit Number															LSB
		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

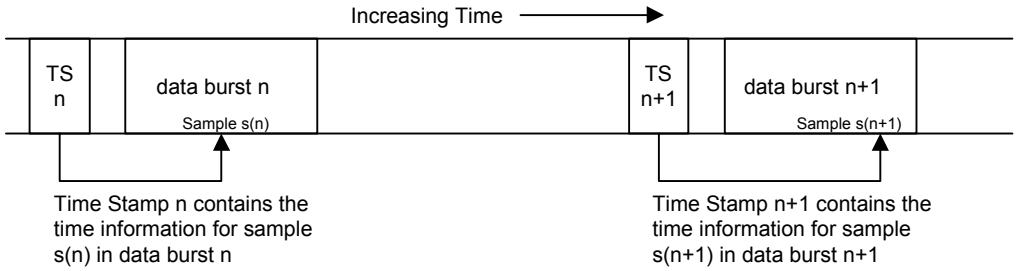

Figure 1 – Sample number indication in time stamp data bursts

In word 4, bit fields s15-s0 denote a sample number to which the SMPTE 12M time code applies. The sample number is an unsigned integer that shall indicate the linear PCM sample to which the time code information applies. When the time stamp data burst is associated with a data burst containing coded audio, the sample number shall be relative to the linear PCM samples represented by the coded audio data burst as illustrated in figure 1. For example, if an audio data burst represents 1536 linear PCM samples and the sample number is set to 1007, then the SMPTE 12M time code information applies to sample 1007 of the 1536 samples. If the sample number contained in a time stamp is greater than the number of samples represented by the corresponding audio data burst, then the sample number shall correspond to a sample represented by a subsequent data burst of the same stream number. For example, if an audio data burst represents 1536 linear PCM samples and the sample number is set to 1546, then the SMPTE 12M time code information applies to sample 10 of the 1536 samples represented by the next data burst of the same stream number. In such cases, the same SMPTE 12M 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.

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 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.

5 User defined (data_type = 30)

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

5.1 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 338M.

5.2 User data burst_payload

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

6 Utility (data_type = 26)

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

6.1 Utility data burst_preamble

The data_type shall be set to 26 as defined in SMPTE 338M. 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 3. Utility functions are described in clause 6.2. When bit 4 is set to a 1, the value of bits 3-0 of the data_type_dependent field are undefined.

Table 3 – Value of data_type_dependent bit 4 utility data

data_type_dependent bit 4	Utility function
0	V-sync
1	Undefined

6.2 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.

6.2.1 V-sync function

The V-sync function is intended to allow identification of an alignment point between an AES3 stream containing SMPTE 337M 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 337M data that are associated with a specific video frame/field.

6.2.1.1 V-sync data_type_dependent

V-sync data bursts shall set bits 3-0 of the data_type_dependent field as shown in table 4 to indicate the frame/field rate of the video signal to which the AES3 signal is associated. When related to a 25, 29.97, or 30 frame per second (fps) video signal the V-sync frame rate may be coded as either the frame rate or twice the frame rate (i.e., 50, 59.94, or 60 fps) of the video signal. When related to a 50, 59.94, or 60 fps video signal, the frame rate may be coded at 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 4 – 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

6.2.1.2 V-sync burst_payload

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

6.2.1.3 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 337M frame mode or subframe mode. In the frame mode, the V-sync preamble word Pa reference position will 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 337M 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.

6.2.1.4 V-sync usage

The alignment between AES3 signals and corresponding video signals is often not precise as these two signals travel through video facilities. Identification of a specific AES3 frame that corresponds to the reference point in the video raster 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 that time stamp data bursts may also be used for synchronization to an associated video reference signal. However, time stamp data bursts are restricted to 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 337M data, which, as with null data bursts, may be useful for maintaining receiver synchronization.

Note that 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.

6.2.1.5 V-sync example

Figure 2 and figure 3 show an example of V-sync data burst placement with a SMPTE 337M 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 337M 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. Note that in this example the actual number of AES3 frames associated with each video frame will vary between 1601 and 1602 frames.

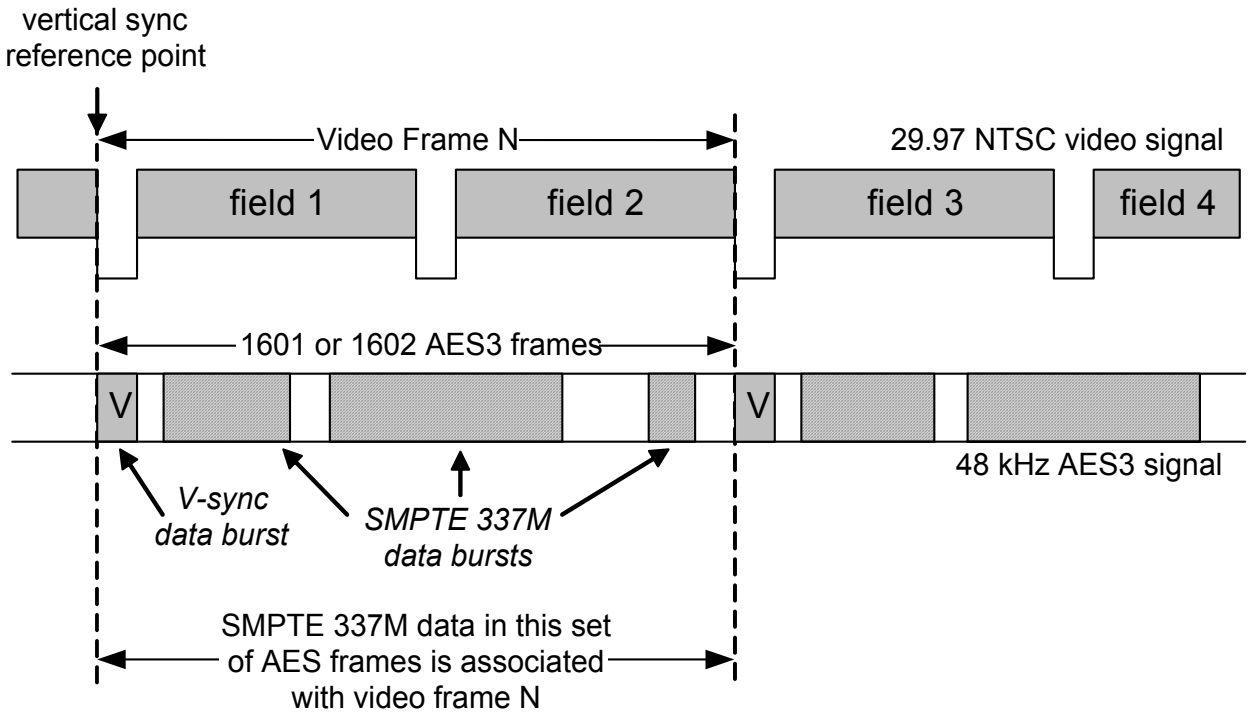

Figure 2 – V-sync data burst example (29.97 frame rate with interlaced video signal)

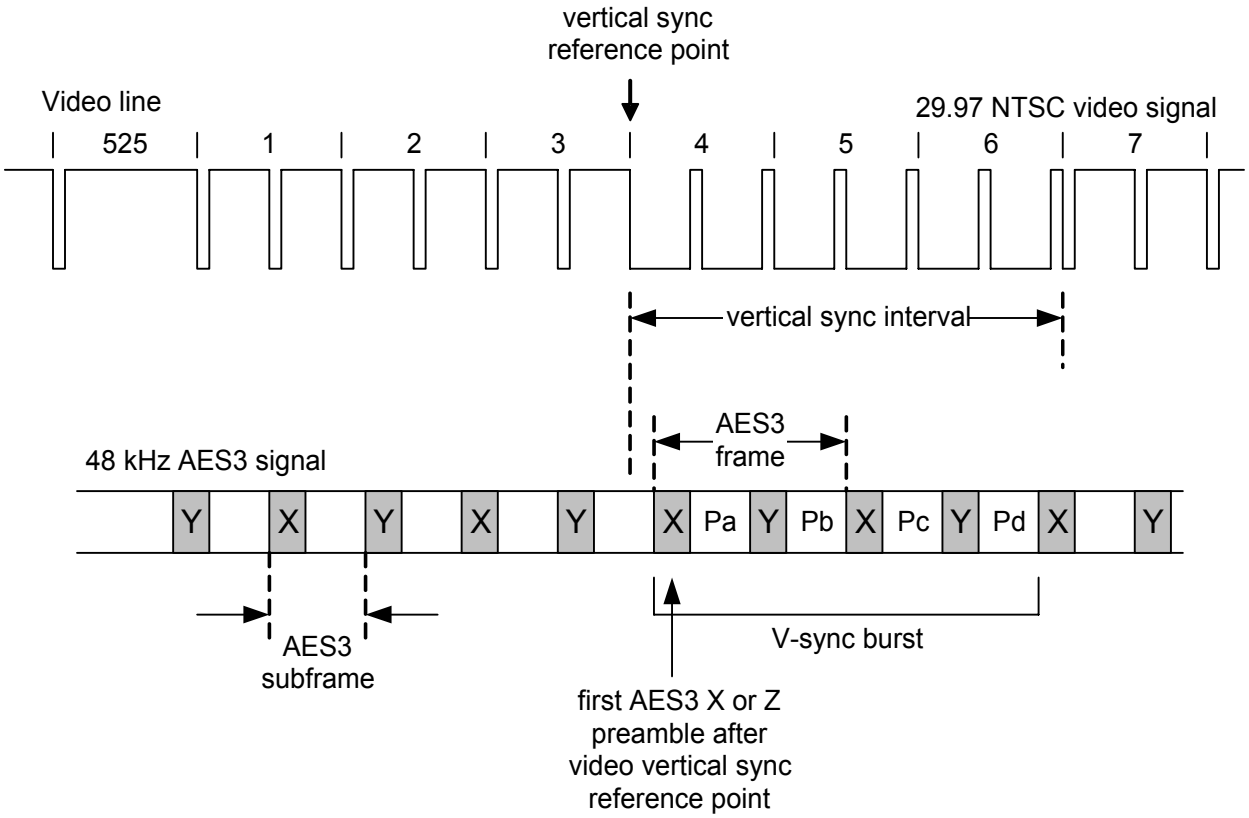

Figure 3 – V-sync reference position example (29.97 frame rate with interlaced video signal)

Annex A (informative)
Bibliography

SMPTE 302M-1998, Television — Mapping of AES3 Data into MPEG-2 Transport Streams

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