SMPTE 324M-2004

SMPTE STANDARD

for Television —
Serial Interface for Multiplexing Eight AES3 Data Streams

Table of contents

1 Scope

1.1 This standard defines a method of mapping up to eight synchronous streams of AES3-compliant data into a self-clocking serial interface. When used for PCM digital audio, the interface can carry up to 16 channels of linearly encoded audio and auxiliary data. Alternatively or concurrently (within the eight stream limit), the interface may be used to carry streams of compressed audio or other data that is packaged into an AES3-compliant stream. The interface is transparent to any data type that is packaged in an AES3-compliant stream that is synchronous with the multiplex.

1.2 PCM audio sampled at 48 kHz and clock-locked to video is the preferred linear baseband audio representation for studio applications and should be used whenever baseband audio is associated with video that is running at standard frame rates. However, this interface supports any frequency of operation supported by AES3, provided that all of the audio channels and/or all the non-audio (see 3.15) data streams are synchronized to a common clock. Ideally all the audio channels should be audio-synchronous (see 3.5) for guaranteed audio phase coherence.

1.3 This standard is intended to provide a reliable method of distributing multiple co-phased channels of digital audio around a studio without losing the initial relative sample-phase relationship. A mechanism is provided to allow multiple 16-channel streams to be re-aligned after a relative misalignment of up to ± 8 samples.

1.4 This interface is intended to be compatible with the complete range of digital television scanning standards and standard film rates.

Approved
May 10, 2004

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, Digital Audio Engineering — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data

AES-3id-2001, AES Information Document for Digital Audio Engineering — Transmission of AES3 Formatted Data by Unbalanced Coaxial Cable

AES11-2003, Digital Audio Engineering — Synchronization of Digital Audio Equipment in Studio Operations

ANSI/SMPTE 276M-1995, Television — Transmission of AES/EBU Digital Audio Signals Over Coaxial Cable.

IEC 60169-8, Radio-Frequency Connectors. Part 8: R.F. Coaxial Connectors with Inner Diameter of Outer Conductor 6.5 mm (0.256 in) with Bayonet Lock — Characteristic Impedance 50 ohms (Type BNC)

3 Definition of terms

3.1 AES audio: All of the data, audio and auxiliary, associated with one AES digital stream as defined in AES3.

3.2 AES frame: Two AES subframes, one with audio data for channel a (subframe 1) followed by one with audio data for channel b (subframe 2).

3.3 AES subframe: All data associated with one AES audio sample for one channel in a channel pair **.**

3.4 audio data: 28 bits: 24 bits of audio associated with one audio sample plus the following 4 bits: sample validity (V bit), channel status (C bit), user data (U bit), and parity (P bit). Alternatively, there may be 20 bits of audio, 4 auxiliary bits, and the V, C, U, and P bits per the AES3 standard.

3.5 audio synchronous: One audio channel is defined as being synchronous with another when the two channels are running from the same clock and the analog inputs are concurrently sampled as described in AES11.

3.6 auxiliary data: Four bits of AES audio associated with one sample defined as auxiliary data by AES3. The four bits may be used to extend the resolution of the audio sample.

3.7 carrier data: Available data bits in the header data that may be used for metadata transmission or other user applications.

3.8 carrier data block start bit (CB bit): An optional bit provided to allow the formation of a carrier data block that spans multiple AES frames.

3.9 channel pair: Two digital audio channels derived from the same AES audio source.

3.10 header data: All the data in the multiplex header.

3.11 multichannel phasing flag (MC bit): A dedicated bit that identifies the start of a 16 frame period with a 1 state. This bit immediately precedes the A1 bit and has a 1 state in the first frame following the frame containing the Z preamble in the AES11 timing reference.

3.12 multiplex frame: All the data associated with each sample of audio in each AES frame for each of 8 AES inputs and the multiplex header associated with that sample period.

3.13 multiplex frame sync: A bi-phase code violation at the beginning of the multiplex header that identifies the start of each multiplex frame

3.14 multiplex header: A 64-bit data packet at the start of each frame that includes a bi-phase code violation for synchronization, 32 bits of carrier data, an optional carrier block flag (CB bit), a parity bit for carrier data (CP bit), a multi-channel phasing flag (MC bit), sixteen channel active bits (A bits), eight AES block start bits (Z bits), and a parity bit for the MC, A, and Z bits (HP bit).

3.15 non-audio data: Any data payload in an AES3 stream that requires byte 0, bit 1 of the channel status data to be set to 1 per the AES3 standard. This encompasses any data that is not linearly coded PCM audio, so “non-audio data” may actually be compressed audio or non-linearly coded audio as well as metadata, compressed video, or other data that is truly not audio.

3.16 video synchronous: Audio is defined as being clock-synchronous with video if the sampling rate of the audio is such that the number of audio samples occurring within an integer number of video frames is itself a constant integer number.

4 Overview

Audio data derived from eight synchronous AES frames are multiplexed together in a single, bi-phase-mark-encoded serial data stream at exactly eight times the bit rate of each incoming AES signal. Each AES subframe is organized in time sequence in the multiplexed frame. All the data in each AES subframe is included with the exception of the preamble. The four bit-times from each preamble are moved to the front of the multiplexed word to produce a 64-bit header. Four of the header bits are used for a code violation that provides frame synchronization. One bit is used for a sixteen frame flag that allows multiple streams that have been subjected to differential delay to be re-phased with minimum latency. Sixteen A bits are used to identify active channels. Eight bits are assigned as Z bits to identify the block start frame for each of the eight AES channel pairs. One bit is assigned to a parity bit that covers the MC bit and the Z bits. Thirty-two bits are allocated to carrier data. One bit is reserved for an optional block start reference for carrier data. The last bit is used to ensure even parity across carrier data bits, the carrier block bit and the reserved bits.

5 Multiplex carrier channel-coding

The data in the multiplex carrier is coded bi-phase mark. Data zeroes are identified by a transition only at the beginning of the bit period. Data ones are identified by transitions at the beginning and at the middle of the bit period as shown in figure 1. The receiver shall be able to detect the bi-phase mark data in either normal or reverse polarity so that it can operate correctly if the data has passed through an odd number of inversions.

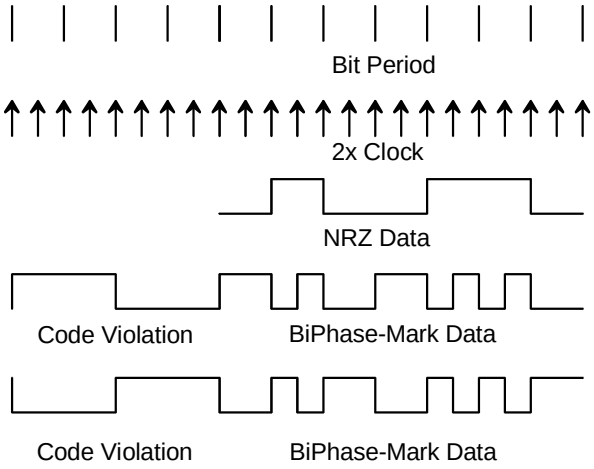

Figure 1 – Channel coding

6 Synchronization and audio phasing

Multiplex frame synchronization is achieved by use of a code violation at the start of each multiplex frame. The code violation consists of a transition at the beginning of the first 2-bit period and a second transition at the beginning of the second 2-bit period. This produces the maximum violation time achievable in a 4-bit period for an even-parity violation. A multichannel flag bit (MC bit) is provided to allow re-phasing of multiplex streams that have suffered differential delay up to ± 8 samples. The MC bit is a “1” in the sample frame following the frame that contains the Z preamble in the AES11 system sync reference. Subsequently, the MC bit is a “1” every 16 frames, and has a value of 0 in all other frames. This bit must always be passed with the identical delay as the audio data.

7 AES data mapping

AES frame data from AES inputs are rigidly organized in the multiplex carrier. AES input 1 follows the multiplex header with subframe 1 (AES A channel) preceding subframe 2 (AES B channel). AES inputs 2 through 8 follow in the same fashion, as shown in figure 2. With the exception of the preamble, all data in each AES subframe is replicated in order in the multiplex carrier. Each subframe 1 preamble is replaced by a single bit which is zero at all times except during a Z preamble at which time it is a one. These single Z bits immediately follow the A bits in the multiplex header. They are arranged in the same order as the AES inputs. The subframes are numbered 1 through 16 in the multiplex according to the above mapping scheme. The fixed mapping of the subframes of each AES signal pair into the multiplex is the key to maintaining the identity of the individual signals. The labels on the connectors at the inputs to the multiplexer and the outputs of the demultiplexer shall clearly indicate the identity of each of the eight AES signal pairs and the sequence in which these signal pairs are inserted into the multiplex.

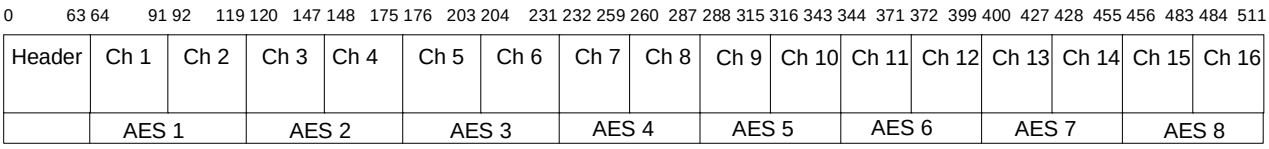

Figure 2 – AES data mapping

8 Active channels

Sixteen A bits are assigned to identify which channels in the multiplex are in use. The A bits immediately follow the MC bit and are arranged in sequential order. The A bit representing a channel must be set to one to indicate that channel as active (in use). The A bits representing inactive (unused) channels must be set to zero. When a channel is inactive, the audio data word for that channel must be set two's complement zero (silence) and the Z bit must be set correctly for any associated active channel; e.g., if channel 8 is active and channel 7 is inactive, the Z 7-8 bit must be set correctly for proper decoding of channel 8. Given the nature of the inputs, the A bits will typically be set in pairs, but the flexibility is provided to allow individual channel usage. In some situations, this could allow a downstream processing device to better utilize its resources.

9 Multiplex header

Each incoming AES subframe includes a 4-bit preamble from which only the block start information (Z bit) is necessary in a rigid multiplex. The four bit periods from each AES preamble are assigned to the multiplex header. The header is thus 64 bits long.

Bits 0 through 3 in the header are assigned to multiplex frame sync (a code violation sequence).

Bits 4 through 35 are assigned to carrier data bits and organized as four bytes with MSB first.

Bit 36 is assigned to an optional carrier data block start bit (CB bit).

Bit 37 is assigned to a parity bit (CP) which ensures even parity across the carrier data bits and the CB bit.

Bit 38 is assigned to the multi-channel flag bit (MC).

Bits 39 through 54 (A bits) are assigned in sequential order.

Bit 55 is assigned to the block start (Z 1-2 bit) for the first AES pair (channels 1 and 2).

Bit 56 is assigned to the block start (Z 3-4 bit) for the second AES pair (channels 3 and 4).

Bit 57 is assigned to the block start (Z 5-6 bit) for the third AES pair (channels 5 and 6).

Bit 58 is assigned to the block start (Z 7-8 bit) for the fourth AES pair (channels 7 and 8).

Bit 59 is assigned to the block start (Z 9-10 bit) for the fifth AES pair (channels 9 and 10).

Bit 60 is assigned to the block start (Z 11-12 bit) for the sixth AES pair (channels 11 and 12).

Bit 61 is assigned to the block start (Z 13-14 bit) for the seventh AES pair (channels 13 and 14).

Bit 62 is assigned to the block start (Z 15-16 bit) for the eighth AES pair (channels 15 and 16).

Each Z bit is set to one in the frame coinciding with the Z preamble in the corresponding AES input. At all other times, the Z bits are set to zero.

Bit 63 is assigned to a parity bit (HP bit) that ensures even parity across the MC bit and Z bits. Figure 3 shows the header detail.

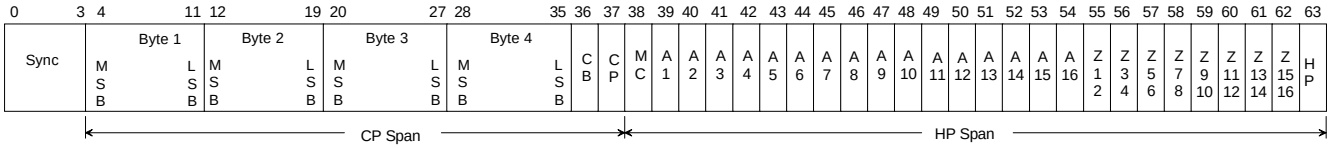

Figure 3 – Header detail

10 Electrical characteristics

10.1 The output interface of the multiplexer is a 1-volt peak-to-peak, 75-ohm coaxial, source-terminated signal, centered on ground.

10.1.1 Output level tolerance shall be ± 10% and the output return loss shall be better than 25 dB from 300 kHz to 50 MHz and better than 15 dB from 50 MHz to 100 MHz.

10.1.2 The rise and fall times, determined between the 10% and 90% amplitude points and measured across a 75-ohm resistive load through 1 meter of high quality coaxial cable shall be between 5% and 30% of a half-cycle of the highest signaling frequency; i.e., the period of either the high or low state of a data one, or mark.

NOTE – For example, a 2.0 ns rise time is equal to 10% of the shortest symbol for 48 kHz operation. This results in 7% of the symbol period at 32 kHz and 20% at 96 kHz. Therefore equipment with an output rise time of 2.0 ns (actually a window of 1.5 ns to 3 ns) will meet this specification for operation from 32 kHz to 96 kHz.

10.1.3 Data jitter shall be less than 0.2 unit interval peak to peak.

NOTE – For example, at 48-kHz operation, the shortest symbol period is 20 ns. At this operating frequency, the jitter shall not exceed 4 ns peak to peak. As the operating frequency is increased, the allowable jitter is proportionately less.

10.2 The receiver shall have a nominal input impedance of 75 ohms.

10.2.1 The receiver shall have a return loss better than 25 dB from 300 kHz to 50 MHz and better than 15 dB from 50 MHz to 100 MHz.

10.2.2 The receiver shall correctly interpret the data with an input level from 200 mV to 1.2 V peak to peak with a data jitter up to plus or minus 0.25 unit interval peak.

11 Equalization

Equalization for transmission cable is not required for typical distances. When equalization is used, it shall correct link cable losses over the range from 100 kHz to 100 MHz.

NOTE – Typical coaxial cables exhibit loss, but very little differential group delay from 5 MHz to 100 MHz and equalizers, if used, must have complementary response to avoid introducing group delay.

12 Cable

The interconnecting coaxial cable shall have nominal characteristic impedance of 75 ohms.

13 Connector type

The connector shall have mechanical characteristics conforming to the 50-ohm BNC type. Mechanical dimensions of the connector may produce either a nominal 50-ohm or a nominal 75-ohm impedance. However, the electrical characteristics of the connector and its associated interface circuitry shall provide a nominally resistive impedance of 75 ohms and meet the specified 75-ohm return loss. Where a 75-ohm connector is used, its mechanical characteristics must reliably interface with the nominal 50-ohm BNC type defined by IEC 60169-8.

Annex A (informative)
Additional data

A.1 Abstract

This document describes a method for multiplexing up to eight AES3-compliant data streams into one serial data stream. The input and output signal formats of the multiplexer and demultiplexer meet the AES3, AES3-ID, or SMPTE 276M specification. The aggregate transmitted data rate is eight times that of a single AES input. The inputs may consist of any combination of PCM audio and non-audio data (see 3.15) streams. In post-mix environments, audio inputs should be inserted starting at 1 and organized per SMPTE 320M or SMPTE EG 26. Inputs must at a minimum be synchronized in frequency, and ideally audio inputs should be phase aligned.

A.2 Background

The advent of DTV, be it HDTV, resolution enhanced SDTV, multicast, or multilingual broadcast, has generated an increased desire to improve audio quality. In fact, results from a number of ATV user tests indicate that improving audio quality provides at least as much enhancement to the viewing experience as increasing picture aspect ratio and resolution. The surround sound or theatre sound environment can be implemented and coded in many forms. AC-3, DTS, SDDS, and MPEG Audio are some examples. Each of these multi-channel coding methods requires original audio source material that is image accurate both during production and through distribution in order to maintain the original spatial perception of the program. With multiple diverse audio paths it is difficult to provide precisely matched path delays for multiple channels of audio in a typical video production environment due to the multiple opportunities for sample slips from the beginning to the end of a production path.

Several proposals have been made for distribution of audio using compression. While compression is very useful for recording, long haul transmission, and final delivery, it creates some problems as a distribution method for production. Production requires that multiple sources be available to a common processing point, and that these sources be in a common editable format. Since there are a number of compression formats available, the point of production must be “multilingual”, and must also be able to manage the different time delays associated with each format, and the variable time delay with a given format. This is a complex process to distribute throughout a production facility.

The audio channels of current DVTRs, disk recorders, and other AES paths through the typical plant are not transparent data channels in spite of the fact that digital audio is PCM data. There are several audio signal impairment issues including phase uncertainty between AES audio and video sync. Error concealment algorithms, sample rate converters, and audio level processing are operations that can be found in many machines. If compressed data is stored on the machine, the effects of these processing steps, if not successfully bypassed, will typically render the data and the subsequent decompressed audio output useless.

Different machines may well require different coding algorithms to insure transparency within the recorder. As a minimum, a common, or standard, compression algorithm would need to allow for full error correction given the statistics of every

transport, fixed compression delay across all equipment in the plant, and a coding scheme that would allow a cut edit. It is most likely that future machines will employ proprietary compression to implement the feature sets desired by individual machine manufacturers. Furthermore, it is probable that any commercially successful DVTR (or disk recorder) will have full bandwidth outputs for both video and audio, since this is the level at which the signals can be edited.

This standard represents a straightforward intra-plant distribution solution based on multiplexed 16-channel full-bandwidth distribution and switching between equipment with CODECs (encoder-decoder pairs), such as recorders or STL interface points, where compression may be required for bandwidth preservation. This topology insures that digital audio signals are coherent and that image accurate audio may be moved throughout the facility with confidence. It also allows for optimal CODEC design to match the bandwidth and error behavior of specific storage and transmission channels.

Eight AES signals may be easily transported in a single path for only a small increase in complexity over a single AES signal. The complexity of the multiplexer and demultiplexer is minimal when compared with separate distribution CODECs. Given the proliferation of 8-track audio recorders, the presence of 8-, 10-, and 12-channel HDTV DVTRs, and the economies associated with multi-channel transport, a 16-channel interface standard will dramatically simplify in-plant distribution of multi-channel audio.

A.3 Implementation

A.3.1 Multiplexer inputs and format

There are eight AES inputs which will be referred to as AES 1, 2, 3, 4, 5, 6, 7, and 8; each of these has two channels of audio, A (subframe 1) and B (subframe 2), consistent with the AES3 standard. Each channel of audio contains 24 bits of data, data, 4 bits of overhead information (C, V, U and P), and a 4 bit preamble. The bit usage and position in the frame, as well as the preamble codes are all identical to AES3 specifications. For simplicity in the multichannel environment, the AES channels (subframes) are assigned numbers. For example, AES 1 channels A and B are labeled channels 1 and 2, and AES 2 channels A and B are labeled channels 3 and 4, up to AES 8 channels A and B being labeled channels 15 and 16. In the case of non-audio data (see 3.15), an AES channel pair may be a single grouping of data that is not functionally separable into channels.

A.3.2 Demultiplexer outputs and format

There are eight AES outputs which will be referred to as AES 1, 2, 3 4, 5, 6, 7, and 8; each of these has two channels of audio, A (subframe 1) and B (subframe 2), consistent with the AES3 standard. Each channel of audio contains 24 bits of data, 4 bits of overhead information (C, V, U and P) and a 4-bit preamble. The bit usage and position in the frame, as well as the preamble codes are all identical to AES3 specifications. For simplicity in the multichannel environment, the AES channels (subframes) are assigned numbers. For example AES 1 channels A and B are labeled channels 1 and 2, and AES 2 channels A and B are labeled channels 3 and 4, up to AES 8 channels A and B being labeled channels 15 and 16. In the case of non-audio data (see 3.15) an AES channel pair may be a single grouping of data that is not functionally separable into channels.

A.3.3 Input/output electrical specifications

The signals should be electrically compatible with AES3, AES-3ID, and/or SMPTE 276M for both multiplexer inputs and demultiplexer outputs.

A.3.4 Multiplexer transmission format

The multiplexer serial data format calls for transmitting consecutive data frames composed of 16 AES data packets and one header packet, in the rigid order shown (see figure 2). Note that the AES inputs or outputs of a multiplexer or demultiplexer have a rigid position in the transmission frame format. This is a critical economic and operational advantage for manufacturers and users.

Each AES data packet is a truncated version of the 32-bit AES sub-frame as shown. This packet is 28 bits long, with complete preservation of the entire AES subframe data payload. The block start information is moved to the multiplex header for data efficiency. The header packet contains 64 bits (see figure 3). A four-bit preamble is used for multiplexer framing. A single MC bit is used for alignment of multiple data streams ensuring synchronous performance and preservation of audio image across parallel machines with differential latencies of several AES frames.

The header contains 4 bytes of channel data and an optional channel block bit. A parity bit (CP) sets this group to even parity. The last data group in the header contains the MC bit, an A bit for each channel to identify channel usage, the Z bits corresponding to each of the AES inputs, and a second parity bit (HP bit) which sets the last data group to even parity.

The Z preamble from each pair of AES subframes is saved as a Z bit for decoding channel status. The Z bit will be 1 at the start of the AES block and 0 for the remaining 191 sub-frames in the standard AES block. This allows accurate recovery of all the channel status information. The Z bit is coincident in time for each AES A/B pair since this is a requirement of the AES3 specification. Equipment is required to pass channel status transparently. Also, equipment that processes audio and reinitializes the channel status bits must re-stripe the Z framing bit in accordance with maintaining the channel pair correlation.

The complete multiplexed frame of data is 512 bits long generating a bit rate of 512 x F_S (the audio sampling frequency). For example, 48 kHz audio is transmitted at 24.576 MHz.

A.3.5 Channel code

The multiplexed data stream utilizes bi-phase mark coding (see figure 1). Synchronization is provided by a code violation consisting of 4 baud-periods low followed by 4 baud-periods high (or the inverse). The advantages of this coding scheme are that the signal is insensitive to polarity, the clock is always contained in the data, and the clock is easily recoverable with low jitter. The code violation allows for economical frame location within a period of 1 frame for rapid synchronization.

A.3.6 Electrical interface

The electrical interface is coaxial with a characteristic impedance of 75 Ohms. Connections are made with BNC type connectors, and signal lines are terminated with 75 Ohms. The output signal level is 1.0 V peak to peak $\pm$10%. Rise time is 5% to 30% of the baud interval. A 2-ns rise time will stay within specification for sample rates from 32 kHz to 96 kHz. This signal is very robust. Since the energy band is largely above the RC turnover frequency in typical coaxial cable, the need for cable equalization is reduced or eliminated in most situations, and jitter introduction is minimal without equalizers.

Annex B (informative)
Bibliography

ANSI/SMPTE 125M-1995, Television — Component Video Signal 4:2:2 Bit-Parallel Interface

SMPTE 260M-1999, Television — 1125/60 High-Definition Production System — Digital Representation and Bit-Parallel Interface

SMPTE 274M-2003, Television — 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

SMPTE 293M-2003, Television — 720 x 483 Active Line at 59.94-Hz Progressive Scan Production — Digital Representation

SMPTE 296M-2001, Television — 1280 x 720 Progressive Image Sample Structure — Analog and Digital Representation and Analog Interface

SMPTE 320M-1999, Television — Channel Assignment and Levels on Multichannel Audio Media

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

SMPTE EG 26-1995, Audio Channel Assignments for Digital Television Tape Recorders with AES/EBU Digital Audio Inputs