AES3 (AES/EBU) Professional Interface

The Audio Engineering Society (AES) has established a standard interconnection generally known as the AES3 or AES/EBU digital interface. It is a serial transmission format for linearly represented digital audio data. It permits transmission of two-channel digital audio information, including both audio and nonaudio data, from one professional audio device to another. The specification provides flexibility within the defined standard for specialized applications; for example, it also supports multichannel audio and higher sampling frequencies. The format has been codified as the AES3-1992 standard; this is a revised version of the original AES3-1985 standard. In addition, other standards organizations have published substantially similar interface specifications. The International Electrotechnical Commission (IEC) developed the IEC-60958 professional or broadcast use (known as type I) format. The International Radio Consultative Committee (CCIR) provides Rec. 647 (1990). The Electronic Industries Association of Japan specifies an EIAJ CP-340-type I format. The American National Standards Institute has ratified the ANSI S4.40-1985 standard. The European Broadcasting Union has established EBU Tech. 3250-E.

The AES3 format establishes a standard for conveying nominally two channels of periodically sampled and uniformly quantized audio signals on a single twisted wire pair. The format is intended to convey data over distances of up to 100 meters without equalization. Longer distances are possible with equalization. Left and right audio channels are multiplexed, and the channel is self-clocking and self-synchronizing. Because it is independent of sampling frequency, the format can be used with any sampling frequency. A sampling frequency of 48 kHz $\pm$ 10 parts per million is often used but frequencies of 32, 44.1, 48 and 96 kHz are all recognized as standard sampling frequencies by the AES for PCM applications in standards document AES5-1998. Moreover, AES3 has provisions for sampling frequencies of 22.05, 24, 88.2, 96, 176.4 and 192 kHz. Sixty-four bits are conveyed in one sampling period; the period is thus 22.7 μs with a 44.1-kHz sampling frequency. AES3 alleviates polarity shift between channels, channel imbalances, absolute polarity inversion, gain shifts, as well as analog transmission problems such as hum and noise pickups, and high frequency loss. Furthermore, an AES3 data stream can identify mono/stereo, use of pre-emphasis, and the sampling frequency of the signal.

The biphase mark code, a self-clocking code, is the binary frequency modulation channel code used to convey data over the AES3 interconnection. There is always a transition (high to low, or low to high) at the beginning of a bit interval; a binary 1 places another transition in the center of the interval; a binary 0 has no transition in the center. A transition at the start of every bit ensures that the bit clock rate can be recovered by the receiver. The code also minimizes low-frequency content, and is polarity free (information lies in the timing of transitions, not their direction). All information is contained in the code's transitions. Using the code, a properly encoded data stream will have no transition lengths greater than one data period (two cells), and no transition lengths shorter than one-half coding period (one cell). This kind of differential code can tolerate about twice as much noise as channels using threshold detection. However, its bandwidth is large, limiting channel rate; logical 1 message bits cause the channel frequency to equal the message bit rate. The overall bit rate is 64 times the sampling frequency; for example, it is 3.072 Mbps at a 48-kHz sampling frequency. Channel codes are discussed in chapter 3.

The AES3 specification defines a number of terms. An audio sample is a signal that has been periodically sampled, quantized, and digitally represented in a two's complement manner. A subframe is a set of audio sample data with other auxiliary information. Two subframes, one for each channel, are transmitted within the sampling period; the first subframe is labeled 1, and second is labeled 2. A frame is a sequence of two subframes; the rate of transmission of frames corresponds exactly to the sampling rate of the source. With stereo transmissions, subframe 1 contains left A-channel data and subframe 2 contains right B-channel data, as shown in Fig. 13.2*A*. For mono, the rate remains at the 2-channel rate, and audio data is placed in subframe 1. A block is a group of channel status data bits and an optional group of user bits, one per subframe, collected over 192 source sample periods. A subframe preamble designates the starts of subframes and channel status blocks, and synchronizes and identifies audio channels. There are three types of preambles. Preamble Z identifies the start of subframe 1 and frame 0, which is also the start of a channel status block. Preamble X identifies the start of subframe 1 otherwise, and Preamble Y identifies the start of subframe 2. Preambles occupy four bits; they are formed by violating the biphase mark coding in specific ways.

The format specifies that a subframe has a length of 32 bits, with fields that are defined as shown in Fig. 13.2*B*. Audio data might occupy 24 bits. Data is linearly represented in two's complement form, with the least significant bit (LSB) transmitted first. If the audio data does not require 24 bits, then the first four bits can be used as an auxiliary data sample, as defined in the channel status data. When devices use 16-bit words, the last 16 bits in the data field are used, with the others set to 0. Four bits conclude the subframe: 1) An audio sample validity bit is 0 if the transmitted audio sample

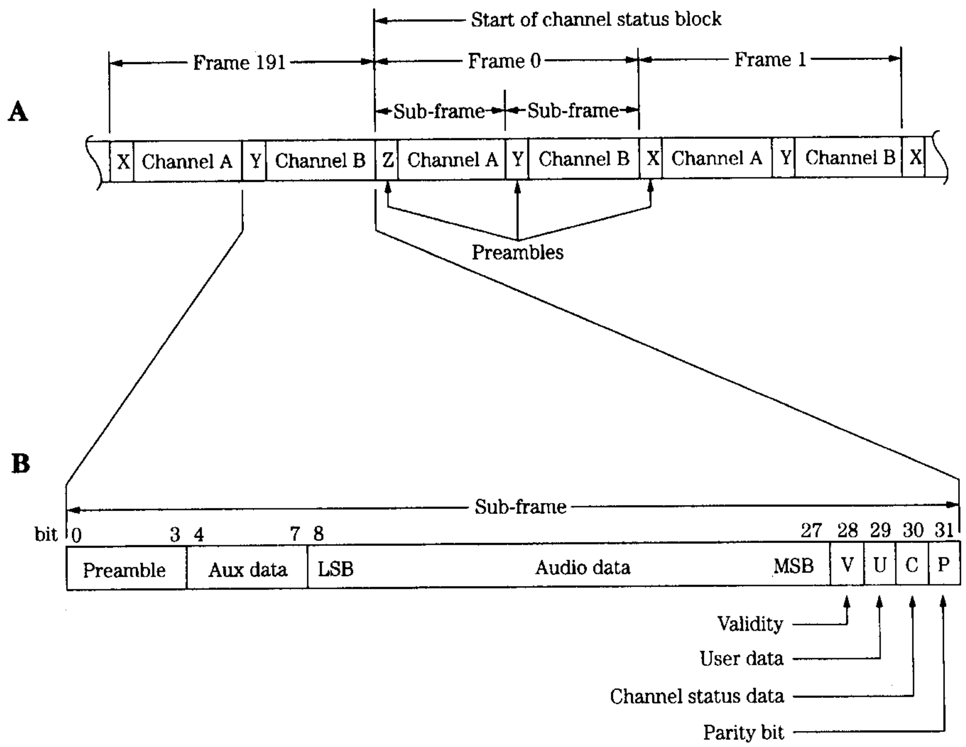

Figure 13.2 The AES3 interface is structured in frames and subframes as well as channel status blocks formed over 192 frames. A. There are two subframes per frame; each subframe is identified with a preamble. B. The interface uses a subframe of 32 bits.

is error free, and 1 if the sample is defective and not suitable for conversion to an analog signal. 2) A user data bit can optionally be used to convey blocks of user data. A recommended format for user data is defined in the AES18-1992 standard, as described below. 3) A channel status bit is used to form blocks describing information about the interconnection channel and other system parameters, as described below. 4) A subframe parity bit provides even parity for the subframe; the bit can detect when an odd number of errors have occurred in the transmission.

Channel status block

The audio channel status bit is used to convey a block of data 192 bits in length. An overview of the block is shown in Fig. 13.3. Received blocks of channel status data are accumulated from each of the subframes to yield two independent channel status data blocks, one for each channel. At a sampling

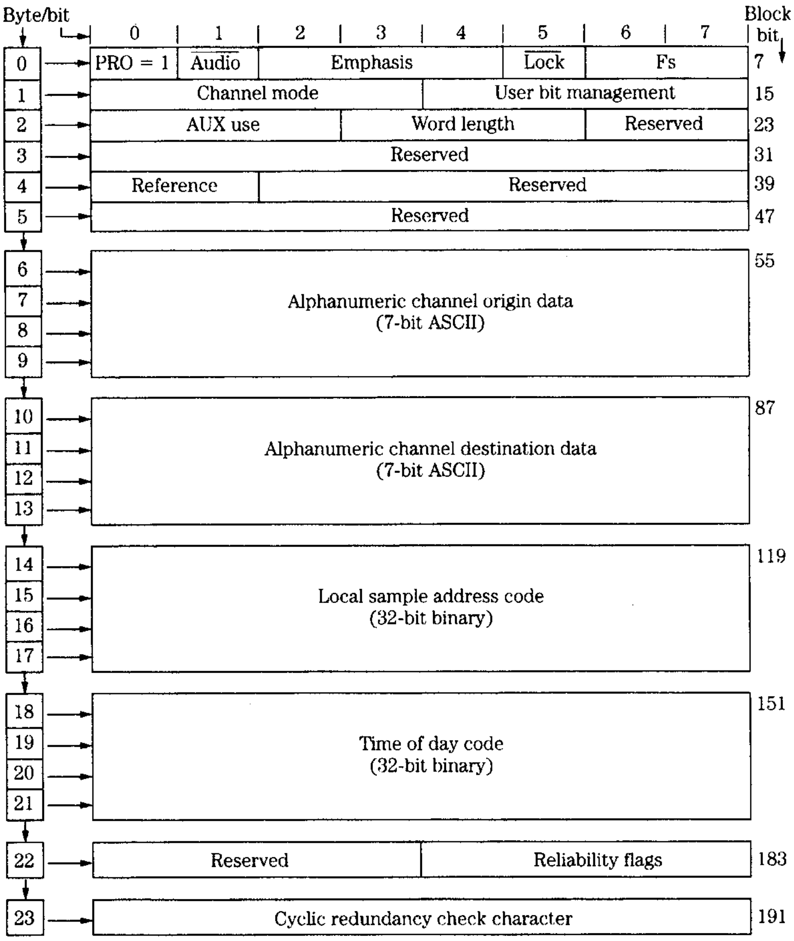

Figure 13.3 Specification of the 24-byte channel status block used in the professional AES3 serial interface.

frequency of 48 kHz, the blocks repeat at 4-ms intervals. Each channel status data block consists of 192 bits of data in 24 bytes, transmitted as one bit in each subframe, and collected from 192 successive frames. The block rate is 250 Hz at a 48-kHz sampling frequency. The channel status block is synchronized by the alternate subframe preamble occurring every 192 blocks.

There are 24 bytes of channel status data. The first six bytes (outlined at the top of Fig. 13.3) are detailed in Fig. 13.4. Byte 0 of the channel status block contains information that identifies the data for professional use, as well as information on sampling rate and use of pre-emphasis; with any AES3 communication, bit 0 in byte 0 must be set to 1 to signify professional use of the channel status block. Byte 1 specifies signal mode such as stereo, mono or multichannel. Byte 2 specifies maximum audio word length and number of bits used in the word; an auxiliary coordination signal can be specified. Byte

BYTE 0	Bit 0	PRO = 1 (Professional)
	0	Consumer use of channel status block
	1	Professional use of channel status block
	Bit 1	$\overline{\text{Audio}}$
	0	Digital audio
	1	Non-audio
	Bit 234	Emphasis
	000	Emphasis not indicated
	100	No emphasis
	110	50/15 μs emphasis
	111	CCITT J.17 emphasis
	Bit 5	$\overline{\text{Lock}}$
	0	Sampling frequency locked
	1	Sampling frequency unlocked
	Bit 67	Encoded sampling frequency
	00	Not indicated
	01	48 kHz
	10	44.1 kHz
	11	32 kHz

BYTE 1	Bit 0123	Encoded channel mode
	0000	Not indicated
	0001	Two-channel
	0010	Single-channel
	0011	Primary/Secondary
	0100	Stereophonic
	0101	Reserved
	0110	Reserved
	0111	Single-channel double sampling frequency
	1000	Single-channel double sampling frequency
	1001	Single-channel double sampling frequency
	1111	Multichannel mode. Vector to byte 3 for channel identification
	Bit 4567	User bits management
	0000	Default
	0001	192-bit block
	0010	AES18
	0011	User defined

Figure 13.4 Description of the data contained in bytes 0–5 in the 24-byte channel status block used in the professional AES3 serial interface.

BYTE 2	Bit 012	Auxiliary sample bits	
	000	Maximum 20 bits	
	001	Maximum 24 bits	
	010	Maximum 20 bits (single coordination)	
	011	Reserved	
	Bit 345	Sample wordlength (24)	Sample wordlength (20)
	000	Default	Default
	001	23 bits	19 bits
	010	22 bits	18 bits
	011	21 bits	17 bits
	100	20 bits	16 bits
	101	24 bits	20 bits
	Bit 67	Reserved	

3	Bit 0-7	Vector from byte 1 (multichannel modes)
BYTE 4	Bit 01	Digital audio reference signal (AES11)
	00	Default
	01	Grade 1
	10	Grade 2
	11	Reserved
	Bit 2	Reserved
	Bit 3456	Sampling frequency
	0000	Not indicated (default)
	1000	24 kHz
	0100	96 kHz
	1100	192 kHz
	0010	Reserved
	1010	Reserved
	0110	Reserved
	1110	Reserved
	0001	Reserved
	1001	22.05 kHz
	0101	88.2 kHz
	1101	176.4 kHz
	0011	Reserved
	1011	Reserved
	0111	Reserved
	1111	User defined
	Bit 7	Sampling frequency scaling flag
	0	No scaling
	1	Sampling frequency is 1/1.001 times
5	Bit 0-7	Reserved

Figure 13.4 *(Continued)*

3 is reserved for multichannel functions. Byte 4 identifies multichannel modes, type of digital audio reference signal (Grade 1 or 2) and alternative sampling frequencies. Byte 5 is reserved. Bytes 6 through 9 contain alphanumeric channel origin code, and bytes 10 through 13 contain alphanumeric destination code; these can be used to route a data stream to a destination, then display its origin at the receiver. Bytes 14 through 17 specify a 32-bit sample address. Bytes 18 through 21 specify a 32-bit time-of-day timecode with 4-ms intervals at 48-kHz sampling frequency; this timecode can be divided to obtain video frames. Byte 22 contains data reliability flags for the channel status block, and indicates when an incomplete block is transmitted. The final byte, byte 23, contains a CRCC codeword with generation polynomial $x^8 + x^4 + x^3 + x^2 + 1$ across the channel status block for error detection.

Three levels of channel status implementation are defined in the AES3 standard: minimum, standard, and enhanced; these establish the nature of the data directed to receiving units. With the minimum level, the first bit of the channel status block is set to 1 to indicate professional status, and all other channel status bits are set to 0. With standard implementation, all channel status bits in bytes 0, 1, and 2 (used for sampling frequency, pre-emphasis, mono/stereo, audio resolution, etc.) and CRCC data in byte 23 must be transmitted; this level is the most commonly used. With enhanced implementation, all channel status bits are used.

As noted, audio data can occupy 24 bits per sample. Because most audio data occupies 20 bits or less, the four remaining bits can be optionally used as an auxiliary speech-quality coordination channel, providing a path so that verbal communication can accompany the audio data signal. Such a channel could use a sampling rate that is ⅓ that of the main data rate, and use 12-bit coding; one 4-bit nibble is transmitted in each subframe. Complete words would be collected over three frames, providing two independent speech channels. The resolution of the main audio data must be identified by information in byte 2 of the channel status block.