

SMPTE STANDARD

NOTE: This document previously numbered
SMPTE 299-2009; Content Unchanged

24-Bit Digital Audio Format for
SMPTE 292 Bit-Serial Interface



Table of Contents	Page
Foreword	2
Intellectual Property	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	4
4 Definition of Terms	4
5 Overview	6
6 Audio Data Packet	9
7 Audio Control Packet	19
Annex A Recommendations for Handling of SMPTE 337 Non-PCM Data (Informative)	25
Annex B Recommendations for Handling Legacy Implementations (Informative)	26
Annex C Bibliography (Informative)	27
Annex D Additional Sampling Frequency (Informative)	28

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.

SMPTE Standard 299-1 was prepared by Technology Committee 32NF.

Intellectual Property

SMPTE draws attention to the fact that it is claimed that compliance with this Standard may involve the use of one or more patents or other intellectual property rights (collectively, "IPR"). The Society takes no position concerning the evidence, validity, or scope of this IPR.

Each holder of claimed IPR has assured the Society that it is willing to License all IPR it owns, and any third party IPR it has the right to sublicense, that is essential to the implementation of this Standard to those (Members and non-Members alike) desiring to implement this Standard under reasonable terms and conditions, demonstrably free of discrimination. Each holder of claimed IPR has filed a statement to such effect with SMPTE. Information may be obtained from the Director, Standards & Engineering at SMPTE Headquarters.

Attention is also drawn to the possibility that elements of this Standard may be subject to IPR other than those identified above. The Society shall not be responsible for identifying any or all such IPR.

1 Scope

1.1 This standard defines the mapping of 24-bit AES digital audio data and associated control information into the ancillary data space of a serial digital video conforming to SMPTE 292. The audio data are derived from AES3, hereafter referred to as AES audio. The AES audio data may contain linear PCM audio or non-PCM data formatted according to SMPTE 337. Not all implementations compliant with this standard may support all sampling frequencies. Manufacturers are encouraged to indicate which sampling frequencies are supported (see Annex D).

1.2 Audio sampled at a clock frequency of either 48 kHz or 96 kHz locked (synchronous) to video, is the preferred implementation for intrastudio applications. As an option, this standard supports AES audio at synchronous or asynchronous sampling rates from 32 kHz to 48 kHz.

1.3 Audio channels are transmitted in groups of four, up to a maximum of 16 audio channels in the case of 32 kHz, 44.1 kHz or 48 kHz sampling, and up to a maximum of 8 audio channels in case of 96-kHz sampling. Each group is identified by a unique ancillary data ID.

1.4 Audio data packets are multiplexed (embedded) into the horizontal ancillary data space of the C_b/C_r data stream, and audio control packets are multiplexed into the horizontal ancillary data space of the Y data stream.

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.

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 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)

SMPTE 291M-2006, Television — Ancillary Data Packet and Space Formatting

SMPTE 292-2008, 1.5 Gb/s Signal/Data Serial Interface

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

4 Definition of Terms

4.1 AES audio: All the VUCP (sample validity bit (V), channel status bit (C), user data bit (U), even parity bit (P)) data, audio data and auxiliary data, associated with one AES digital stream as defined in AES3.

4.2 AES frame: Two AES subframes: in the case of the 32 kHz to 48 kHz sampling subframes, one and two carry AES audio channel 1 and 2, respectively. In the case of 96 kHz sampling subframes, one and two carry successive samples of the same AES audio signal which is mandatory for 96 kHz application.

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

4.4 Ancillary data packet: A data packet as defined by SMPTE 291M.

4.5 audio clock phase data: Audio clock phase is indicated by the number of video clocks between the first word of EAV and the video sample appearing at the same time as the audio sample at the input to the formatter.

4.6 audio control packet: An ancillary data packet occurring once a field in an interlaced system and once a frame in a progressive system and containing data used in the process of decoding the audio data stream.

4.7 audio data: 29 bits: 24 bits of AES audio associated with one audio sample, including AES auxiliary data, plus VUCP bits and the Z flag which is derived from the preamble of AES3 stream. The Z bit is common to the two channels of an AES channel pair.

4.8 audio data packet: An ancillary data packet containing audio clock phase data, audio data for 2 channel pairs (4 channels) and error correction code. An audio data packet shall contain audio data of one sample associated with each audio channel.

4.9 audio frame number: A number, starting at 1, for each frame within the audio frame sequence. For the example in § 4.9, 48 kHz sampling at 30.00/1.001 frames/s system, the frame numbers would be 1, 2, 3, 4 and 5.

4.10 audio frame sequence: The number of video frames required for an integer number of audio samples in synchronous operation. As an example, the audio frame sequence for synchronous 48 kHz sampling at 30.00/1.001 frames/s system is 5 frames.

4.11 audio group: Consists of two channel pairs that are contained in one ancillary data packet. Each audio group has a unique ID as defined in § 6.1 and § 7.1. Audio groups are numbered 1 through 4.

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

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

4.14 data ID: A word in the ancillary data packet that identifies the use of the data therein.

4.15 error correction code: BCH (31, 25) code (an error-correction method) in each bit sequence of b0-b7. Errors between the first word of ancillary data flag (ADF) through the last word of audio data of channel 4 (CH4) in user data words (UDW) can be corrected or detected within the capability of this code.

4.16 horizontal ancillary data block: An ancillary data space located in the digital line blanking interval of one television line.

4.17 synchronous audio: Audio is defined as being clock synchronous with video if the sampling rate of audio is such that the number of audio samples occurring within an integer number of video frames is itself a constant integer number. Examples are shown in Table 1.

Table 1 – Examples of audio samples per frame for synchronous audio

Audio sampling rate	Samples/frame				
	30.00 frame/s	30.00/1.001 frame/s	25.00 frames/s	24.00 frames/s	24.00/1.001 frames/s
96.0 kHz	3200/1	16016/5	3840/1	4000/1	4004/1
48.0 kHz	1600/1	8008/5	1920/1	2000/1	2002/1
44.1 kHz	1470/1	147147/100	1764/1	3675/2	147147/80
32.0 kHz	3200/3	16016/15	1280/1	4000/3	4004/3

AES11 provides specific recommendations for audio and video synchronization.

Note: Implementations of this standard may achieve synchronous or asynchronous operation through the use of sample rate converters. In the context of this standard, synchronous audio applies to the AES audio stream that is directly mapped into the ancillary data space, which may or may not be the AES audio stream present on device interfaces. It is recommended that product manufacturers clearly state when sample rate conversion is used to support multiple sample rates and/or asynchronous operation. It is also recommended that the use of sample rate conversion be user selectable. For example, when the AES audio data contains SMPTE 337 formatted data the use of sample rate conversion will corrupt the SMPTE 337 data (see Annex A). This recommendation applies to both multiplexing (embedding) and demultiplexing (receiving) devices.

5 Overview

5.1 The modes of transmission carried in an audio data packet defined in § 6 shall be the TWO-CHANNEL MODE at all sampling frequencies from 32 kHz to 48 kHz, and the SINGLE CHANNEL DOUBLE SAMPLING FREQUENCY MODE at the sampling frequency of 96 kHz as defined in AES3. Audio data channels 1~4 (CH1~ CH4) defined in § 6.2.2 carry two AES audio channel pairs (AES1 channel 1 and 2 and AES2 channel 1 and 2) in the case of 32 kHz to 48 kHz sampling. For 96 kHz sampling, two successive samples of two AES audio channels (AES1 channel 1 1st and 2nd sample and AES2 channel 1 1st and 2nd sample) shall be carried.

Note: AES3 document allows for different audio samples configurations of the digital audio. Single sample mode or successive sample modes are permitted in AES3 without regard to used sampling frequency.

5.2 The 32 kHz, 44.1 kHz or 48 kHz sampling audio data derived from two channel pairs shall be configured in an audio data packet as shown in Figure 1. Both channels of a channel pair are derived from the same AES audio source. The number of samples per channel used for one audio data packet shall be constant and is equal to one. The number of audio data packets in a given group shall be less than or equal to N_a in a horizontal ancillary data block. The definition and examples of N_a are described in § 6.3.3.

5.3 Figure 2 shows the audio data packet at the sampling rate of 96 kHz. AES subframes 1 and 2 carry successive samples of the same AES audio signal. Both channels shall be derived from the same AES audio source. The number of samples per channel used for one audio data packet shall be constant and equal to two. The number of audio data packets in a given group is less than or equal to $N_a/2$ in a horizontal ancillary data block. The definition and examples of N_a are described in § 6.3.3.

5.4 Two types of ancillary data packets carrying AES audio information are defined and formatted per SMPTE 292. Each audio data packet shall carry all of the information in the AES bit stream as defined by AES3. The audio data packet shall be located in the horizontal ancillary data space of the C_b/C_r data stream. An audio control packet shall be transmitted once per field in an interlaced system and once per frame in a progressive system in the horizontal ancillary data space of the second line after the switching point of the Y data stream.

5.5 Data ID shall be defined for four separate packets of each packet type. This allows for up to eight channel pairs. In this standard, the audio groups are numbered 1 through 4 and the channels are numbered 1 through 16. Channels 1 through 4 are in group 1, channels 5 through 8 are in group 2, and so on. Table 2 defines the relationship between CH1~CH4 (UDW2~UDW17) in the audio data packet and the channel/sample number for 32 kHz to 48 kHz sampling and 96-kHz sampling respectively.

5.6 The audio data packet and audio control packet shall be located in SMPTE 292 transport HANC space that is equal to 268 clock pulses at 30 Hz video frame rate.

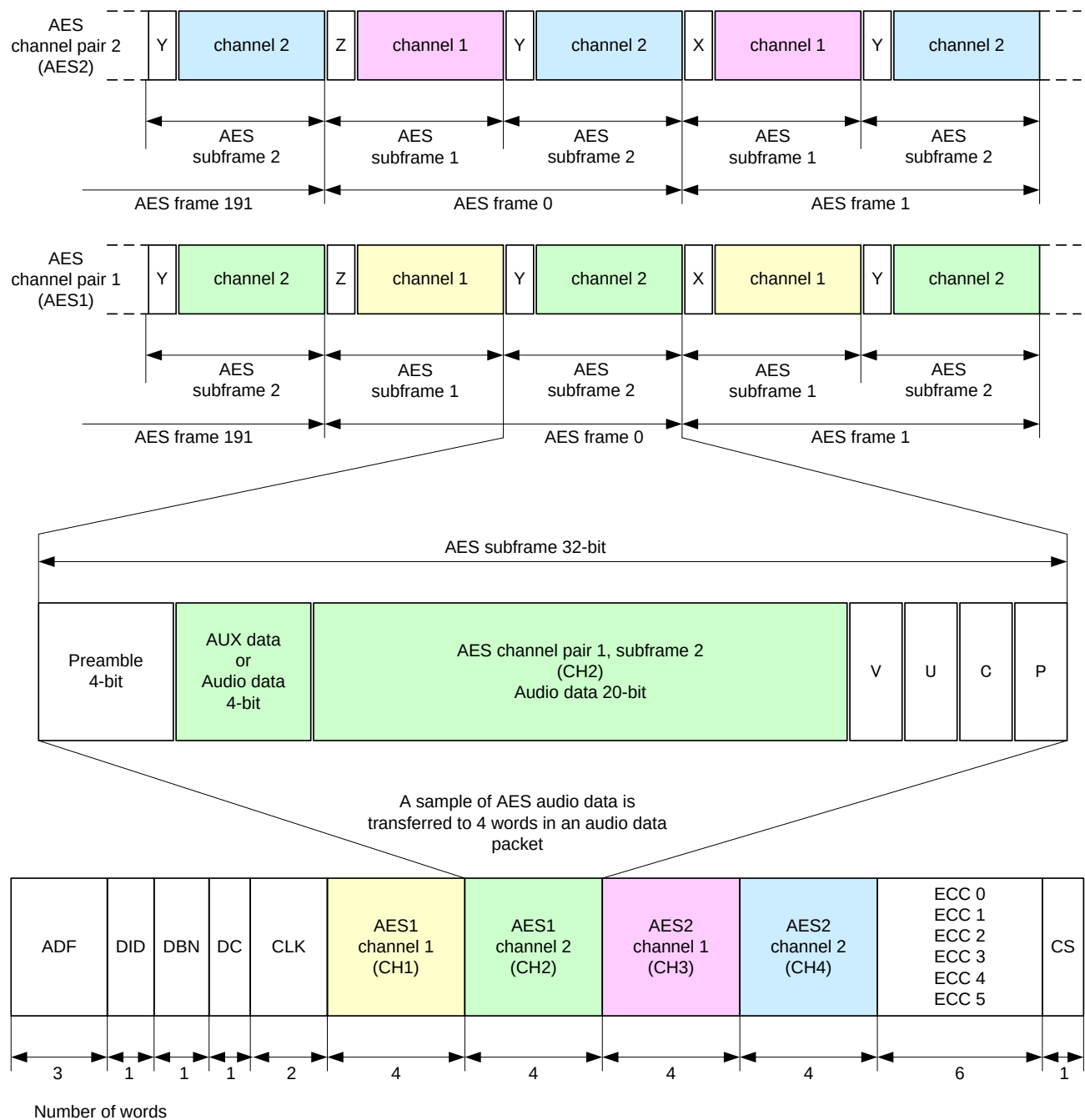


Figure 1 – Relationship between AES audio and audio data packets at sampling rates of 32 kHz, 44.1 kHz or 48 kHz

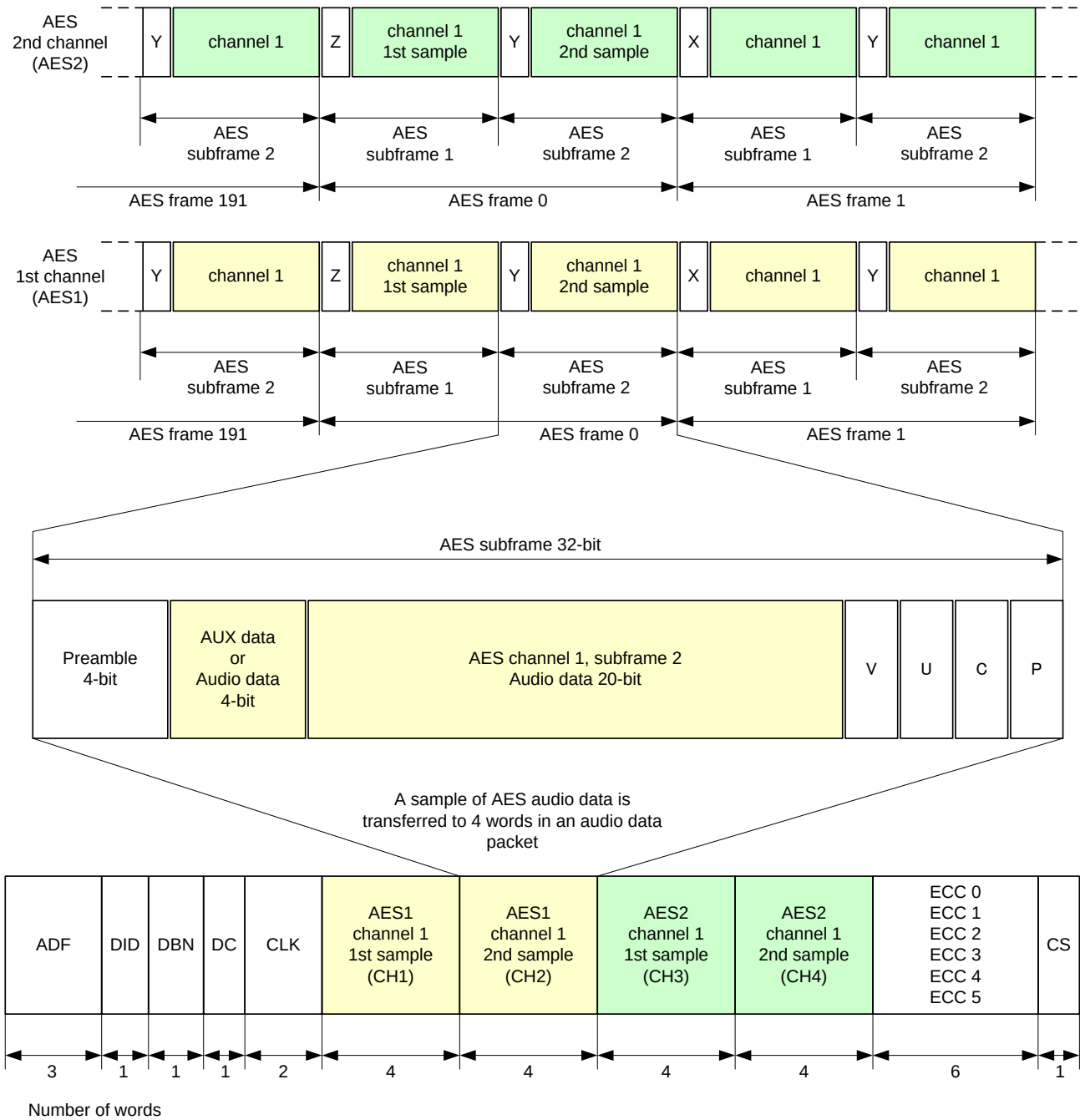


Figure 2 – Relationship between AES audio and audio data packets at a sampling rate of 96 kHz

Table 2 – Relationship between audio data packets and the channel/sample number of 32 kHz to 48 kHz and 96 kHz sampling

Audio sampling rate	Audio Group 1			
	UDW2~UDW5 CH1	UDW6~UDW9 CH2	UDW10~UDW13 CH3	UDW14~UDW17 CH4
32.0 kHz, 44.1 kHz or 48.0 kHz	AES1 channel 1	AES1 channel 2	AES2 channel 1	AES2 channel 2
96.0 kHz	AES1 channel 1 1st sample	AES1 channel 1 2nd sample	AES2 channel 1 1st sample	AES2 channel 1 2nd sample

6 Audio Data Packet

6.1 Structure of Audio Data Packet

6.1.1 The structure of the audio data packet shall be as shown in Figure 3. Audio data packets shall be formatted according to the requirements of SMPTE 291M and shall include ancillary data flag (ADF), data identification (DID), data block number (DBN), data count (DC), user data words (UDW) and checksum (CS) fields as specified in SMPTE 291M. DC is always 218_h.

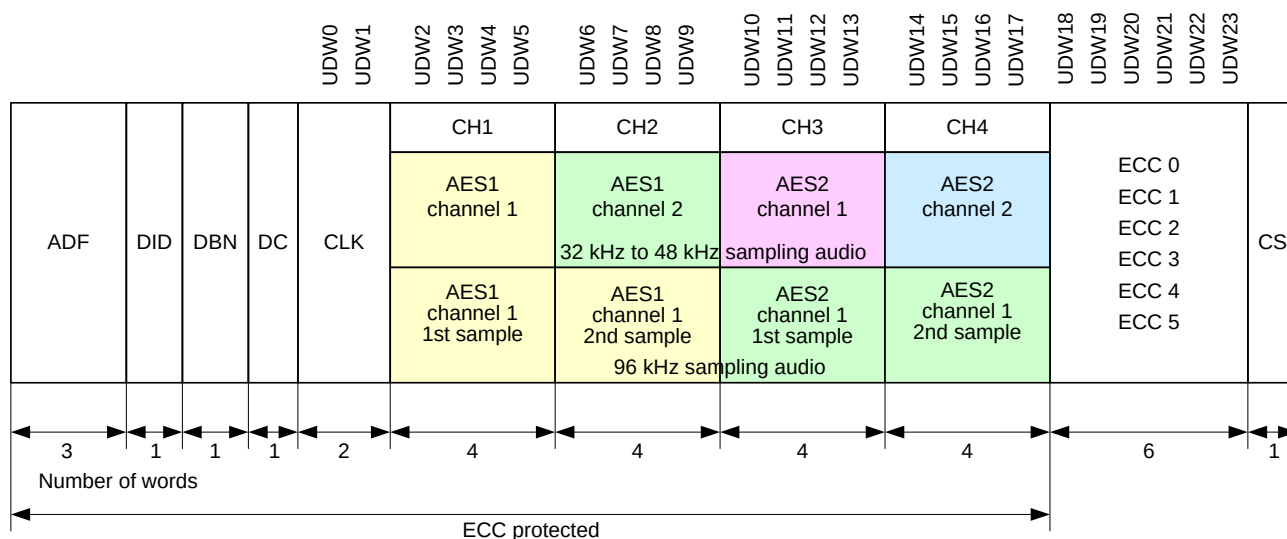


Figure 3 – Structure of audio data packets

6.1.2 The DID value shall be defined as 2E7_h for audio group 1 (channel 1~4), 1E6_h for audio group 2 (channel 5~8), 1E5_h for audio group 3 (channel 9~12) and 2E4_h for audio group 4 (channel 13~16), respectively.