International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Recommendation ITU-R BS.2143-0

(01/2022)

Transport method for non-Pulse-Code Modulation audio signals and data over digital audio interfaces for programme production and exchange

BS Series

Broadcasting service (sound)

International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Recommendations

(Also available online at http://www.itu.int/publ/R-REC/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	**Broadcasting service (sound)**
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management
SNG	Satellite news gathering
TF	Time signals and frequency standards emissions
V	Vocabulary and related subjects

***Note**: This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2022

RECOMMENDATION ITU-R BS.2143-0

Transport method for non-Pulse-Code Modulation audio signals and data over digital audio interfaces for programme production and exchange

(Question ITU-R 130-3/6)

(2022)

Scope

This Recommendation specifies a transport method for non-Pulse-Code Modulation (non-PCM) audio signals and data including the Serial Audio Definition Model (S-ADM) metadata over the digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647 (AES3) for programme production and exchange.

Keywords

Non-PCM audio signal, Audio Definition Model (ADM), Serial ADM (S-ADM), advanced sound system, digital audio interface, AES3

Abbreviations/Glossary

ADM: Audio Definition Model, metadata set specified in Recommendation ITU-R BS.2076

S-ADM: Serial representation of the Audio Definition Model, format of metadata based on the ADM segmented into a time series of frames specified in Recommendation ITU-R BS.2125.

The ITU Radiocommunication Assembly,

considering

a) that the advanced sound systems require a set of audio-related metadata to play back an advanced sound programme;

b) that the advanced sound systems in live or real-time workflows of programme production and exchange require real-time interfaces to transport metadata synchronized with audio signals,

recognizing

a) that Recommendation ITU-R BS.2051 – Advanced sound system for programme production, specifies the next-generation audio systems that require audio-related metadata for programme production;

b) that Recommendation ITU-R BS.2076 – Audio Definition Model, specifies the set of audio-related metadata for the file-based production of advanced sound systems;

c) that Recommendation ITU-R BS.2125 – A serial representation of the Audio Definition Model, specifies the format of metadata based on the Audio Definition Model segmented into a time series of frames for linear workflows such as live or real-time production for broadcasting and streaming applications;

d) that Recommendation ITU-R BS.647 – A digital audio interface for broadcasting studios, specifies a digital audio interface of pairwise audio channels and its compatible interface AES3 that are used worldwide;

e) that Recommendation ITU-R BS.1873 – Serial multichannel audio digital interface for broadcasting studios, specifies a multichannel digital audio interface (MADI) of 56 or 64 channels based on the two-channel audio signal specified in Recommendation ITU-R BS.647;

f) that Recommendation ITU-R BT.1365 – 24-bit digital audio format as ancillary data signals in HDTV serial interfaces, specifies the mapping of 24-bit digital audio data conforming to Recommendation ITU-R BS.647 and associated control information into the ancillary data space of serial digital video interfaces conforming to Recommendation ITU-R BT.1120;

g) that SMPTE ST 2110-31 "Professional Media Over Managed IP Networks: AES3 Transparent Transport" specifies the real-time, RTP-based transport of AES3 signals over IP networks;

h) that SMPTE ST 337 "Format for Non-PCM audio and data in an AES3 serial digital audio interface" specifies the transport method of non-PCM audio signals and data based on the two-channel audio signal specified in Recommendation ITU-R BS.647;

i) that SMPTE ST 2116 "Format for Non-PCM audio and data in an AES3 – Carriage of Metadata of Serial ADM (Audio Definition Model)" specifies the transport method to convey Serial ADM metadata with synchronized audio signals in professional applications using the AES3 serial digital audio interface,

noting

that most of the digital audio interfaces for programme production and exchange are compatible with the signal format specified in Recommendation ITU-R BS.647 (AES3),

recommends

1 that the specification described in Annex 1 should be used to carry non-PCM audio signals and data over the digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647 (AES3) for programme production and exchange;

2 that the specification described in Annex 2 should be used to carry the S-ADM metadata defined in Recommendation ITU-R BS.2125 over the digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647 (AES3) for programme production and exchange using the transport method described in Annex 1.

Annex 1

Transport method for non-PCM audio signals and data over digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647 (AES3)

1 Introduction

The digital audio interface specified in Recommendation ITU-R BS.647, also known as AES3, is widely used to convey linear PCM audio signals. Many of the other digital audio interfaces are compatible with the AES3 signal format, including:

– the multichannel digital audio interface (MADI) specified in Recommendation ITU-R BS.1873;

– serial digital interfaces for HDTV and UHDTV specified in Recommendations ITU-R BT.1120 and BT.2077, which can convey multiple audio channels using the ancillary data space as per Recommendation ITU-R BT.1365;
– IP interfaces for audio signals specified in SMPTE ST 2110-31.

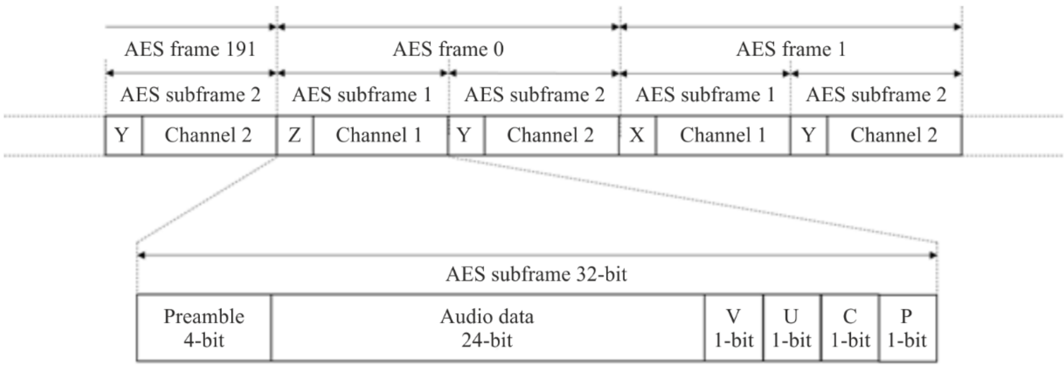
The AES3 digital audio interface can also convey non-PCM audio signals and data by the transport method specified in SMPTE ST 347. This Annex describes the method for non-PCM audio signals and data over digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647.

2 Overview

Recommendation ITU-R BS.647 specifies the transport method of audio signals compatible with the AES3 digital audio interface. The AES3 digital audio interface consists of a sequence of subframes as shown in Fig. 1. Each subframe is intended to convey one linear PCM sample, and contains 32-bit time slots, each of which can carry a single bit of information (V, U, C and P). A pair of subframes, each containing the PCM word of one audio channel, make up an AES3 frame containing two PCM words. A sequence of 192 frames makes up a block. The 192 channel status bits for each channel during a block make up the 192-bit (24-byte) channel status word for that channel.

FIGURE 1

Structure of audio subframes specified in Recommendation ITU-R BS.647

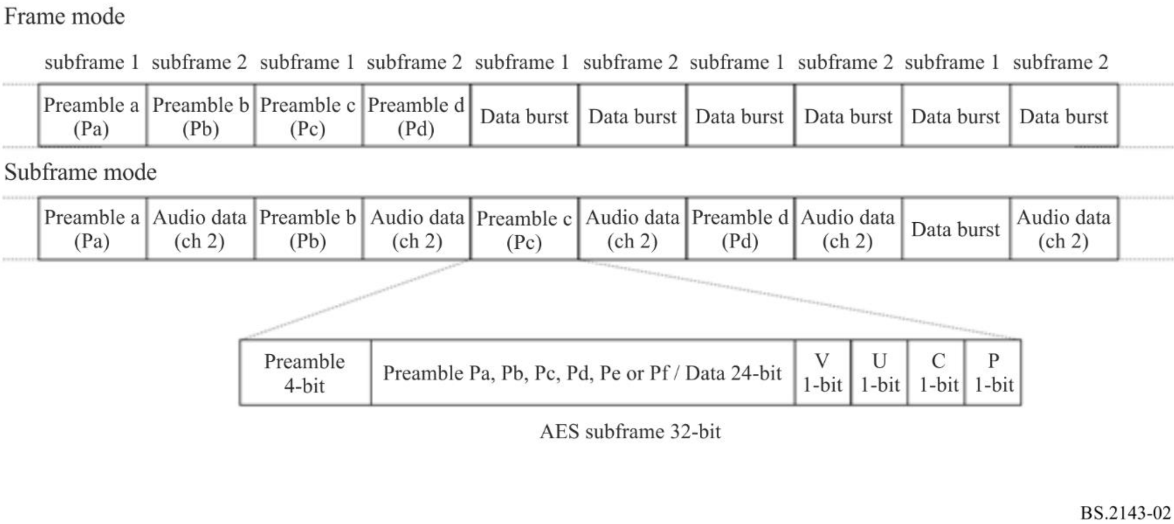


BS.2143-01

When non-PCM audio signals and data are to be transported using the interface compatible with Recommendation ITU-R BS.647, non-PCM audio signals and data are mapped to 24-bit time slots instead of audio data. The non-PCM data streams to be conveyed are formed into data burst, each consisting of a preamble containing information about the burst followed by a data payload. The data bursts are placed in the audio sample word of subframes in one of two modes as shown in Fig. 2. In the frame mode, the data space from each subframe within an AES frame is combined to allow up to 48 bits of data to be placed in each frame. In the subframe mode, each subframe contains either linear PCM audio signals or non-PCM audio signals and data.

FIGURE 2

Structure of data burst to convey non-PCM and data



Data bursts are tagged with a number indicating to which data stream they belong. Up to seven different streams of non-PCM audio signals and data can be time-multiplexed together to form a set of data bit streams.

3 Interface format to convey non-PCM audio signals and data

The 32-bit time slots are specified as follows.

TABLE 1

Bit field of a subframe for non-PCM data

Bit locations	Definitions
0-3	Sync preamble as per Rec. ITU-R BS.647 (AES3)
4-27	Non-PCM audio signals and data
28	Validity bit as per Rec. ITU-R BS.647 (AES3)
29	User data bit as per Rec. ITU-R BS.647 (AES3)
30	Channel status bit – bytes 0, 1, 2, 23 specified in § 3.1. Other bytes are undefined.
31	Parity bit as per Rec. ITU-R BS.647 (AES3)

3.1 Channel status word

A sequence of 192 frames makes up a block. The 192 channel status bits for each channel during a block make up the 192-bit (24-byte) channel status word. For AES3 channels that convey non-PCM audio signals and data, byte 0, byte 1, byte 2 and byte 23 of the channel status word are set as shown in Tables 2 to 5, respectively, and the other bytes are set to ‘0’.

TABLE 2

Channel status bits in byte 0

Bits	Value	Definitions
0	1	Professional use of channel block
1	1	Non-PCM audio mode
2-4	000	Emphasis not indicated
5	-	Frame frequency lock status
6-7	-	Indicates frame frequency per Rec. ITU-R BS.647 (AES3)

TABLE 3

Channel status bits in byte 1

Bits	Value	Definitions
0-3	0000	Encoded channel mode not indicated
4-7	-	Encoded user bit management as per Rec. ITU-R BS.647 (AES3)

TABLE 4

Channel status bits in byte 2

Bits	Value	Definitions
0-2	-	Auxiliary sample bit usage per Rec. ITU-R BS.647（AES3)
3-5	-	Non-PCM data word length per Rec. ITU-R BS.647 (AES3)
6-7	00	Reserved

TABLE 5

Channel status bits in byte 23

Bits	Value	Definitions
0-7	-	CRCC word per Rec. ITU-R BS.647 (AES3)

3.2 Sample rate synchronization

There is no requirement on synchronization between the audio digital interface rate and sample rates of the audio coded within non-PCM audio signal.

4 Data burst format to convey non-PCM audio signals and data

The non-PCM audio signals and data streams to be conveyed are formed into data bursts consisting of data words in a continuous sequence of audio subframes. Each data burst consists of a **burst_preamble** followed by a **burst_payload**. When multiple streams are present, bursts from each stream are placed in the AES3 stream in a time-division multiplexed fashion.

4.1 burst_preamble

The **burst_preamble** occurs at the beginning of each data burst and is followed by the **burst_payload**. When non-PCM audio signals and data stream are conveyed, the preamble of a four- or six-subframe form, which consists of words designated as **Pa** to **Pd** or **Pa** to **Pf**, is used. The four-subframe version only has a 5-bit capacity for defining the **data_type** of the payload and the six-subframe version provides extended code points for additional **data_type** definitions. The six-subframe version is used when the **data_type** has the value 31, otherwise, the four-subframe version is used.

TABLE 6

Preamble words

Preamble word	Contents
Pa	Sync word 1. **Pa** = 0x96F872 (24-bit mode)
Pb	Sync word 1. **Pb** = 0xA54E1F (24-bit mode)
Pc	**burst_info** value.
Pd	**length_code**, equal to the number of data bits in the **burst_payload**. When using a six-subframe preamble, **Pe** and **Pf** are counted as payload bytes.
Pe	**extended_data_type**. Table 10 shows the data type of the extended data.
Pf	Reserved. **Pf** = 0x0000.

4.1.1 Frame mode

Four-subframe version

The four preamble words are contained in two sequential frames. The frame beginning the data burst contains the preamble words **Pa** in the Ch1 subframe and **Pb** in the Ch2 subframe.

Six-subframe version

The six preamble words are contained in three sequential frames. The frame beginning the data burst contains the preamble words **Pa** in the Ch1 subframe and **Pb** in the Ch2 subframe. When using the six-subframe version, the preamble words **Pe** and **Pf** are counted as payload bytes.

4.1.2 Subframe mode

Four-subframe version

The four preamble words are contained in four sequential subframes of the individual channel (Ch1 or Ch2) being employed to convey the non-PCM audio signals and data. The subframe of the channel being used at the beginning the data burst contains the preamble word **Pa**.

Six-subframe version

The six preamble words are contained in six sequential subframes of the individual channel (Ch1 or Ch2) being employed to convey the non-PCM audio signals and data. The subframe of the channel being used at the beginning the data burst contains the preamble word **Pa**. When using the six-subframe version, the preamble words **Pe** and **Pf** are counted as payload bytes.

4.2 burst_info (Pc)

The **burst_info** contains information about the contents of the **burst_payload** as specified in Table 7. Bit 23 of the **burst_info** is to be considered the MSB and is located in time slot 27 of an audio subframe.

TABLE 7

burst_info

Bits	Definitions
0-7	Reserved
8-12	The **data_type** (5-bit unsigned integer) indicates the type of data contained in the **burst_payload**. Supported data types and the mapping of **data_type** values to specific data types are as defined in Table 8. The value 31 indicates that the **data_type** in the **extended_data_type** (the **burst_preamble** word **Pe**) is used. The MSB of the **data_type** is placed in bit 12 and is located in time slot 16 of an audio subframe.
13-14	**data_mode**. The 2-bit **data_mode** indicates the mode in which the data for the **burst_payload** is placed in audio subframes. The **data_mode** is set to ‘2’ to indicate the 24-bit mode. The MSB of the **data_mode** is placed in bit 14 and is located in time slot 18 of an audio subframe.
15	**error_flag**. The 1-bit **error_flag** provides an error indication for the data in the **burst_payload**. If the data in the **burst_payload** is known to be error free or if it is unknown whether the data contains errors, then the value of this bit is set to ‘0’. If the data in the **burst_payload** is known to contain errors, this bit may be set to ‘1’. The **error_flag** bit is located in time slot 19 of an audio subframe.
16-20	**data_type_dependent**. The meaning of the 5-bit **data_type_dependent** is dependent on the value of data_type.
21-23	**data_stream_number**. The 3-bit **data_stream_number** indicates the number of the data stream to which the burst belongs. Each independent data stream uses a unique value for **data_stream_number**. Eight data stream numbers (0-7) are available as defined in Table 10. The MSB of the **data_stream_number** is placed in bit 23 and is located in time slot 27 of an audio subframe. In the subframe mode, each AES3 channel is treated independently and the requirement for unique data stream numbers for each data stream applies only within a given AES3 channel. In this mode, up to 14 independent data streams (7 in each channel) may be time multiplexed in the AES3 interface. Individual time stamp data bursts apply to specific data bursts of other data types. Although all time stamp data bursts are identified as data stream number 7, they should not be considered as a single stream of related time stamp values. When time code information is carried within time code data bursts, multiple time code streams may be conveyed within data bursts identified as data stream number 7.

TABLE 8

data_type in preamble Pc

data_type value	Definitions
0	Null data
1-30	Not defined in this Recommendation
31	Data type of **extended_data_type** in preamble **Pe** as defined in Table 10

TABLE 9

data_stream_number in preamble Pc

data_stream_number value	Definitions
0	Main audio service
1-6	All data types except the time stamp data type
7	Time stamp data

TABLE 10

extended_data_type in preamble Pe

data_type value	Definitions
0x0000	Reserved
0x0001	Metadata of S-ADM
0x0002 – 0xFFFF	Reserved

4.3 length_code (Pd)

The **length_code** indicates the length of the **burst_payload** in bits. The **length_code** MSB is always located in time slot 27 of an audio subframe. The **burst_payload** field is limited from 0 to 16 777 215 bits in the 24-bit mode. The sizes of the **burst_preamble** words **Pa - Pd** are not counted in the value of the **length_code**.

4.4 burst_payload

The **burst_payload** is segmented into data words and placed in a continuous sequence of audio subframes.

4.4.1 Frame mode

In the frame mode, both AES3 channels are utilized to carry one set of non-PCM data streams. The available data space from each subframe within an audio frame is combined when packing data bursts into a continuous sequence of frames. This mode allows up to 48 data bits to be placed in a single audio frame.

The burst_payload is considered as a serial stream of bits; the first bit of the first data word of the payload in a burst occupies the MSB bit position of subframe 1 (time slot 27) and the last bit of the first data word occupies the LSB bit position of subframe 2. The last data bits of the burst_payload may occupy only a fraction of the last frame. Any unused bits in the last frame are set to ‘0’.

4.4.2 Subframe mode

In the subframe mode, each AES3 channel is utilized independently to carry either one set of non-PCM data streams or linear PCM audio signals. The subframe from each AES3 channel within a frame is considered independently when packing data bursts into a continuous sequence of frames. This mode allows up to 24 data bits per channel to be placed in a single audio frame.

The burst_payload is considered as a serial stream of bits; the first bit of the first data word of the payload in a burst occupies the MSB bit position of the subframe (time slot 27) and the last bit of the first data word occupies the LSB bit position of the subframe. The last data bits of the burst_payload may occupy only a fraction of the last frame. Any unused bits in the last frame are set to ‘0’.

The channel status words for each channel are treated independently in the subframe mode.

4.5 Burst spacing

There should not be a sequence of 4 096 or more audio frames (in frame mode) or subframes (in subframe mode) which contain as least one data burst, without the beginning of at least one of the data bursts preceded by four audio subframes which have subframes contents in time slots 8 - 27 of all 0s. This requirement ensures that there are occurrences of an extended sync code of 0, 0, 0, 0, **Pa**, **Pb**.

Data bursts from a given non-PCM data stream are placed in the AES3 interface in consecutive order. If multiple non-PCM data streams are placed in the AES3 interface (or in an individual channel in the subframe mode), the data bursts from each stream are interleaved in a time multiplexed manner.

4.6 Data type dependent fields

The format of data contained within the **data_type_specific** and **burst_payload** fields is dependent on the **data_type** field. Specific coding for these fields is described in other Annexes.

Annex 2

Transport method for the Serial ADM metadata over digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647

1 Introduction

Serial ADM (S-ADM) metadata is one type of the non-PCM data that can be transported over digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647 (AES3) using the method described in Annex 1 with additional constraints and data mapping specified in SMPTE ST 2116. This Annex describes the method of the Serial ADM (S-ADM) metadata over digital audio interfaces compatible with the signal format specified in Recommendation ITU-R BS.647.

2 Data burst format to convey the S-ADM metadata

The S-ADM metadata stream to be conveyed is formed into data bursts consisting of data words in a continuous sequence of audio subframes. Each data burst consists of a **burst_preamble** followed by a **burst_payload**.

2.1 burst_preamble (see § 4.1 in Annex 1)

The **burst_preamble** occurs at the beginning of each data burst and is followed by the **burst_payload**. When the S-ADM metadata stream is conveyed, the preamble of a six-subframe form, which consists of words designated **Pa** to **Pf**, is used. The six preamble words are contained in six sequential subframes of the individual channel employed to convey the S-ADM metadata. The preamble words **Pe** and **Pf** are counted as payload bytes. This ensures compatibility with equipment that does not support six-subframe operation.

TABLE 11

Preamble words

Preamble word	Contents
Pa	Sync word 1. **Pa** is set to '0x96F872' (24-bit mode).
Pb	Sync word 1. **Pb** is set to '0xA54E1F' (24-bit mode).
Pc	**burst_info** value (see Table 12).
Pd	**length_code**, equal to the number of data bits in the **burst_payload**. The preamble words **Pe** and **Pf** are counted as payload bytes.
Pe	**extended_data_type**. **Pe** is set to '0x0001' for S-ADM metadata.
Pf	Reserved. **Pf** is set to '0x0000'.

2.2 burst_info (Pc) (see § 4.2 in Annex 1)

The **burst_info** contains information about the contents of the **burst_payload** as specified in Table 12.

TABLE 12

burst_info

Bits	Definitions
0-7	Reserved
8-12	The **data_type** is set to 31, indicating that the **data_type** in the **extended_data_type** (the burst_preamble word "**Pe**") is used.
13-14	**data_mode**. The 2-bit **data_mode** indicates the mode in which the data for the **burst_payload**. The **data_mode** is set to '2' to indicate the 24-bit mode.
15	**error_flag**. 1 indicates that the data burst may contain errors. 0 indicates that the data may be valid.

TABLE 12 (*end*)

Bits	Definitions
16	**changedMetadata_flag**. 0 indicates that the Serial ADM metadata do not have any difference between the previous and current ADM metadata frames. 1 indicates that the Serial ADM metadata have a difference between the previous and current ADM metadata frames.
17	**assemble_flag**. 0 indicates that the **assemble_info** word is not present. The Serial ADM metadata are conveyed by a single data burst. 1 indicates that the **assemble_info** word is present. The Serial ADM metadata are conveyed by multiple data bursts.
18	**format_flag**. 0 (default) indicates that the **format_info** word is not present. The Serial ADM metadata are encoded as UTF-8. 1 indicates that the **format_info** word is present. The Serial ADM metadata are coded with an optional encoding format type.
19-20	**multiple_chunk_flag**. 00 indicates that a single chunk is used in the frame type of ‘header’, ‘full, ‘intermediate’ or ‘all’ to convey the Serial ADM metadata. 01 indicates the last chunk in the frame type of ‘divided’. 10 indicates the intermediate chunks in the frame type of ‘divided’. 11 indicates the first chunk in the frame type of ‘divided’.
21-23	**data_stream_number**. The 3-bit **data_stream_number** indicates the number of the data stream to which the burst belongs. The **data_stream_number** is set to the same number when multiple data bursts convey the Serial ADM metadata.

2.3 length_code (Pd) (see § 4.3 in Annex 1)

The **length_code** indicates the length of the **burst_payload** in bits. The **length_code** MSB is always located in time slot 27 of an audio subframe. The **burst_payload** field is limited from 0 to 16 777 215 bits in the 24-bit mode. The sizes of the **burst_preamble** words **Pa - Pd** are not counted in the value of the **length_code**.

2.4 extended_data_type (Pe) (see § 4.3 in Annex 1)

The **extended_data_type** indicates an additional **data_type**. The **data_type** value in the **extended_data_type** is set to 0x0001.

2.5 burst_payload (see § 4.4 in Annex 1)

The **burst_payload** is segmented into data words and placed in a continuous sequence of audio subframes. Each audio subframe (audio channel) is utilized independently to carry one set of S-ADM metadata. The first bit of the first data word of the payload in a burst occupies the MSB bit position of the audio subframe (time slot 27) and the last bit of the first data word occupies the LSB bit position of the audio subframe. The last data bits of the **burst_payload** may occupy only a fraction of the last audio subframe. Any unused bits in the last frame are set to ‘0’.

The **burst_payload** contains the **assemble_info**, **format_info** and **SADM_metadata_container**. The **assemble_info** and **format_info** are placed before the **SADM_metadata_container** if one or both are required.

2.5.1 assemble_info

The **assemble_info** consists of one word of the audio subframe and is positioned in the first word of the **burst_payload**. The **assemble_info** value is shown in Table 13.

TABLE 13

assemble_info

Bits	Definitions
0-7	Reserved
8, 9	**in_timeline_flag**. 00 indicates that the multiple in-timeline mode is not used. 01 indicates the last data burst in the multiple in-timeline mode. 10 indicates the intermediate data bursts in the multiple in-timeline mode. 11 indicates the first data burst in the multiple in-timeline mode.
10-15	**track_numbers** (over_track_flag). 6-bit unsigned integer = 0 to 63. The **track_numbers** with 1 added indicates the total number of tracks conveying the Serial ADM metadata. 0 indicates that the multiple over-track mode is not used. Not 0 indicates that the multiple over-track mode is used.
16-21	**Track_ID**. 6-bit unsigned integer = 0 to 63. Index of tracks conveying the Serial ADM metadata.
22, 23	Reserved

2.5.2 format_info

The **format_info** consists of one word of the audio subframe. When the **assemble_info** is used, the **format_info** is positioned in the second word of the **burst_payload**. When the **assemble_info** is not used, the **format_info** is positioned in the first word of the **burst_payload**. The **format_info** value is as shown in Table 14.

TABLE 14

format_info

Bits	Definitions
0-7	Reserved
8-11	**format_type**. The **format_type** indicates the encoding format type of the S-ADM metadata as defined in Table 15.
12-23	Reserved

TABLE 15

format_type

Value	Definitions
0000	UTF-8 (8-bit text)
0001	UTF-8 compressed with gzip as specified in RFC 1952
0010 to 1111	Reserved

2.5.3 SADM_metadata_container

The **SADM_metadata_container** contains the Serial ADM metadata.

The **SADM_metadata_container** word is filled with 24-bit data. The encoded Serial ADM metadata are separated into 24-bit data fields starting from the first data sample. When the **format_flag** is set to 0 or the **format_type** is set to 0000, the Serial ADM metadata encoded as 8-bit characters with UTF-8 are packed as shown in Table 16. One word can convey data of three characters.

TABLE 16

SADM_metadata_container values for UTF-8 text

Bits	Value
0-7	First character of trio
8-15	Second character of trio
16-23	Third character of trio

When the **format_flag** is set to 1 and the **format_type** is set to 0001, the S-ADM metadata encoded with UTF-8 are compressed with gzip (as specified in RFC 1952). The compressed data are divided into 24-bit blocks for packing in the **SADM_metadata_container** words.

In the multiple over-track mode, the series of the **SADM_metadata_container** words are divided into multiple tracks (see § 3.4).

2.6 Burst spacing (see § 4.5 in Annex 1)

As per § 4.5 in Annex 1. Since a single AES3 track conveys a stream of the Serial ADM metadata, time slots 8-27 of the AES3 subframe in Burst spacing are filled by ‘0’.

3 Mapping of the data burst to convey the S-ADM metadata

The data burst to convey the S-ADM metadata has the structure shown below.

Data_burst

{

burst_preamble (**Pa** … **Pf**)

If **assemble_flag** == 0 and **format_flag** == 0

burst_payload (**SADM_metadata_container**)

else if **assemble_flag** == 1 and **format_flag** == 0

burst_payload (**assemble_info**, **SADM_metadata_container**) else if **assemble_flag** == 0 and **format_flag** == 1

burst_payload (**format_info**, **SADM_metadata_container**)

else

burst_payload (**assemble_info**, **format_info**, **SADM_metadata_container**)

end

}

The S-ADM metadata is conveyed by multiple data bursts according to the multiple in-timeline mode (see § 3.3), the multiple over-track mode (see § 3.4) or both modes (see § 3.5).

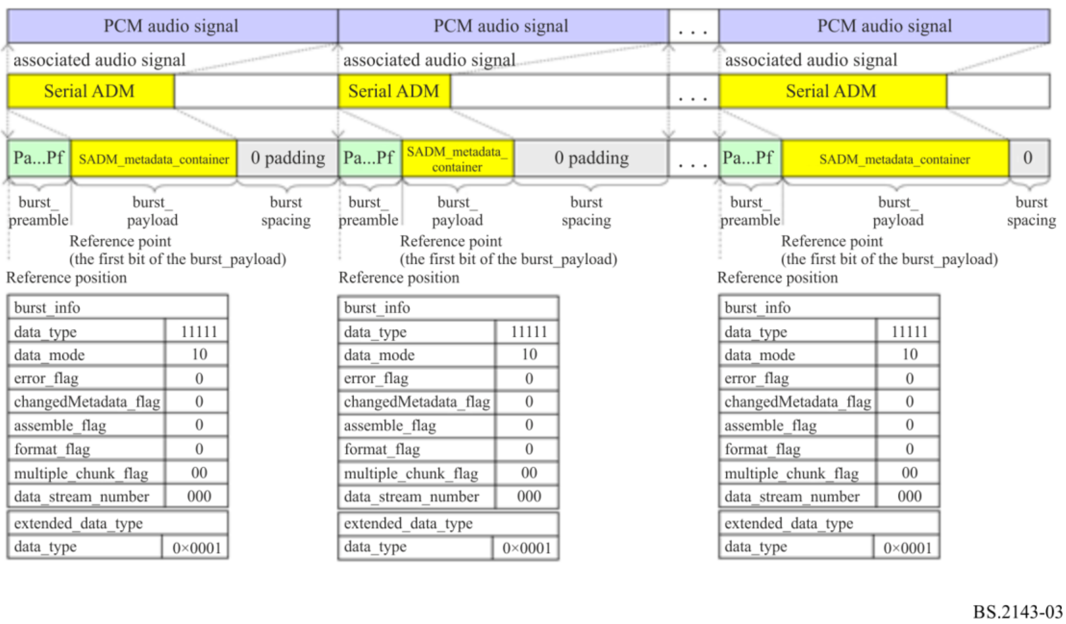
3.1 Fundamental structure of the data burst

Figure 3 shows the fundamental structure of the data burst sequence conveying the S-ADM metadata using the subframe mode (see § 4.4.2 in Annex 1).

The S-ADM metadata are contained in a single **SADM_metadata_container**. The first sample of PCM audio signals associated with the S-ADM metadata is synchronized with the first **burst_preamble** word **Pa** in each data burst.

FIGURE 3

Fundamental structure of the data burst sequence conveying the S-ADM metadata

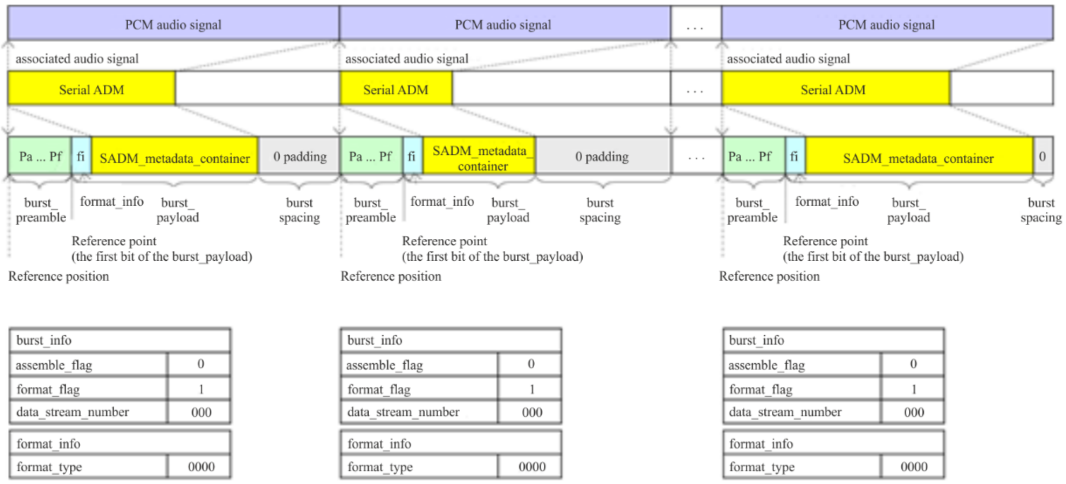


3.2 Structure of the data burst with the format_info

When the S-ADM metadata are coded with an optional encoding format type, the **format_flag** is set to ‘1’. Figure 4 shows the structure of the data burst sequence conveying the S-ADM metadata. In this case, the **format_info** is conveyed by the first word of the **burst_payload**.

FIGURE 4

Structure of the data burst sequence with the format_info conveying the S-ADM metadata



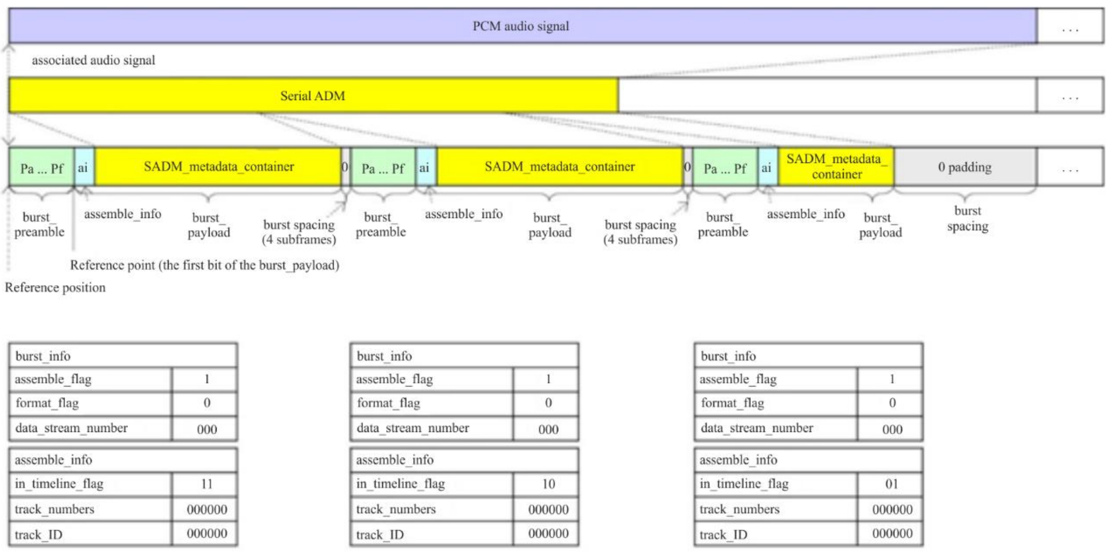
BS.2143-04

3.3 Structure of the data bursts in the multiple in-timeline mode

The multiple in-timeline mode is used to convey the S-ADM metadata over multiple continuous data bursts. Figure 5 shows an example of the data burst structure. In this case, the S-ADM metadata are divided into three continuous data bursts in the same track. The **in_timeline_flags** in the **assemble_info** words of the first, second and third data bursts are set to 11, 10 and 01, respectively. The **data_stream_numbers** in the preamble word **Pc** of both data bursts are set to the same value '000'. The **track_numbers** (over_track_flags) in the **assemble_info** words of all data bursts are also set to the same value '000000'. The first and second data bursts have the last audio subframes in which time slots 8-27 are filled with '0'.

FIGURE 5

Example of the data burst structure in the multiple in-time mode



BS.2143-05

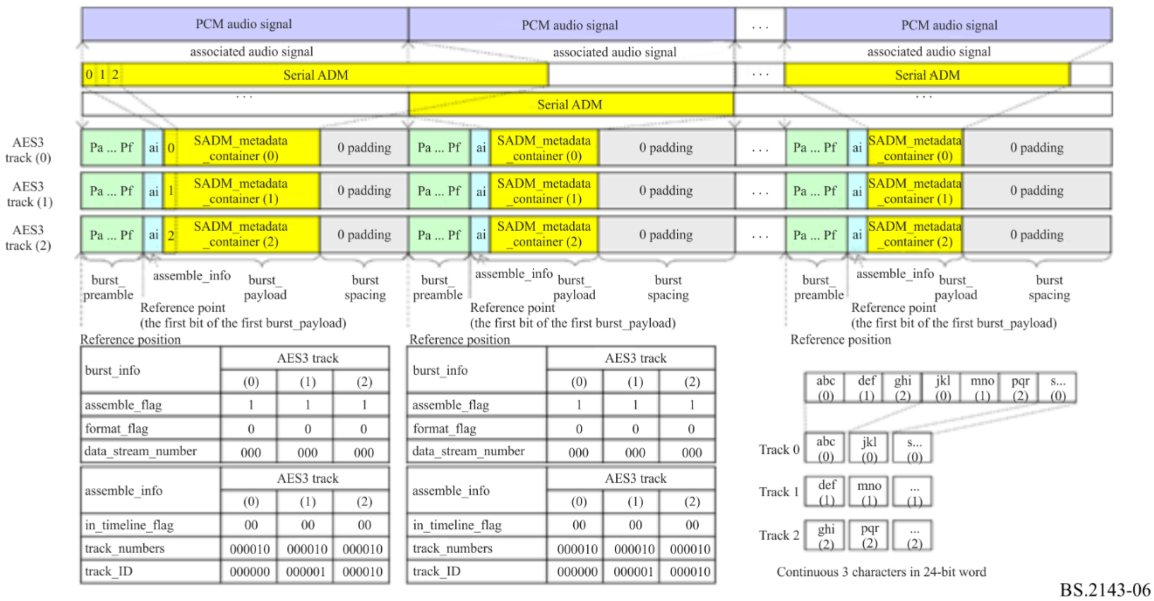
NOTE – In this case, the S-ADM metadata are divided into two continuous data bursts in the same track.

3.4 Structure of the data bursts in the multiple over-track mode

The multiple over-track mode is used to convey the S-ADM metadata over multiple simultaneous data bursts. The synchronized **burst_payload**s with a continuous **track_ID** and the same **data_stream_number** are combined. Figure 6 shows an example of the data burst structure. In this case, the S-ADM metadata are divided into three simultaneous data bursts in different tracks. The **track_numbers** (over_track_flag) in the **assemble_info** of each data burst is set to the same value ‘000010’. The **track_ID**s in the **assemble_info** of the first, second and third data bursts are ‘000000’, ‘000001’ and ‘000010’, respectively. The **data_stream_number** in the **burst_preamble** word ‘**Pc**’ of each data burst is set to the same value ‘000’.

FIGURE 6

Example of the data burst structure in the multiple over-track mode



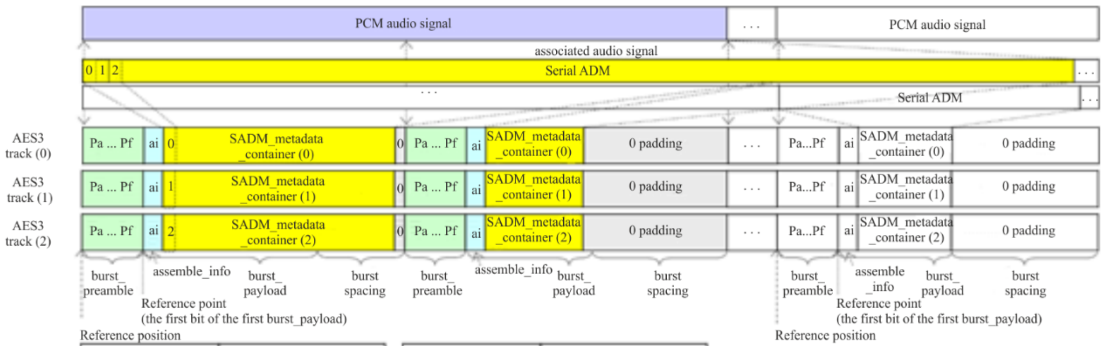
NOTE – In this case, the S-ADM metadata are divided into three simultaneous data bursts over different tracks.

3.5 Structure of the data bursts in the multiple over-track and multiple in-timeline modes

Both the multiple over-track and multiple in-timeline modes can be used at the same time. Figure 7 shows an example of the data burst structure. In this case, the S-ADM metadata are divided into a continuous pair of three simultaneous data bursts. The **track_numbers** (over_track_flag) in the **assemble_info** of each data burst are set to the same value ‘000010’. The **track_IDs** in the **assemble_info** of the first / fourth, second / fifth and third / sixth data bursts are ‘000000’, ‘000001’ and ‘000010’, respectively. The **in_timeline_flags** in the **assemble_info** words of the first three and second three data bursts are set to ‘11’ and ‘01’, respectively. The first three data bursts have the last four audio subframes in which time slots 8-27 are filled with ‘0’. The **data_stream_number** in the **burst_preamble** word **Pc** of each data burst is set to the same value ‘000’.

FIGURE 7

Example of the data burst structure in the multiple over-track and in-timeline modes



burst_info	AES3 track (0)	AES3 track (1)	AES3 track (2)
assemble_flag	1	1	1
format_flag	0	0	0
data_stream_number	000	000	000

assemble_info	AES3 track (0)	AES3 track (1)	AES3 track (2)
in_timeline_flag	11	11	11
track_numbers	000010	000010	000010
track_ID	000000	000001	000010

burst_info	AES3 track (0)	AES3 track (1)	AES3 track (2)
assemble_flag	1	1	1
format_flag	0	0	0
data_stream_number	000	000	000

assemble_info	AES3 track (0)	AES3 track (1)	AES3 track (2)
in_timeline_flag	01	01	01
track_numbers	000010	000010	000010
track_ID	000000	000001	000010

BS.2143-07

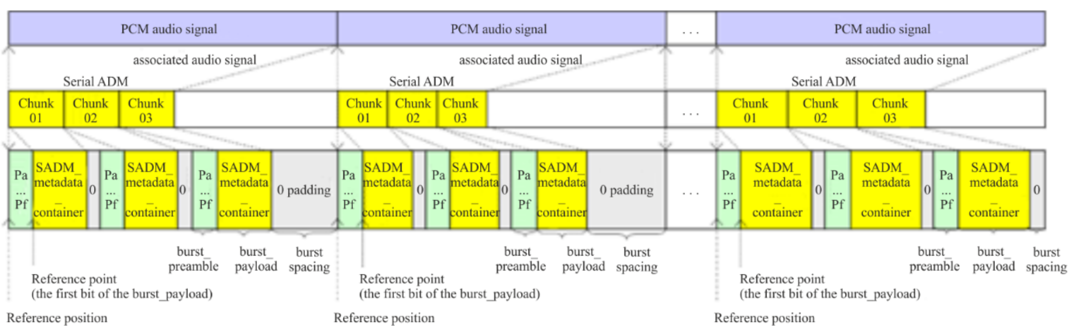
NOTE – In this case, the S-ADM metadata are divided into a continuous pair of three simultaneous data bursts.

3.6 Structure of the data burst for multiple chunks

When the S-ADM metadata are divided into multiple chunks, multiple continuous data bursts are used. Figure 8 shows an example of the data burst structure for multiple chunks. In this case, the S-ADM metadata are divided into three chunks in each ADM metadata frame. The **multiple_chunk_flags** in the **burst_info** words of the first, second and third data bursts are set to '11', '10' and '01', respectively.

FIGURE 8

Structure of the data burst sequence for multiple chunks



burst_info for the first chunk	
data_type	11111
data_mode	10
error_flag	0
changedMetadata_flag	0
assemble_flag	0
format_flag	0
multiple_chunk_flag	11
data_stream_number	000
extended_data_type	
data_type	0×0001

burst_info for the intermediate chunks	
data_type	11111
data_mode	10
error_flag	0
changedMetadata_flag	0
assemble_flag	0
format_flag	0
multiple_chunk_flag	10
data_stream_number	000
extended_data_type	
data_type	0×0001

burst_info for the last chunk	
data_type	11111
data_mode	10
error_flag	0
changedMetadata_flag	0
assemble_flag	0
format_flag	0
multiple_chunk_flag	01
data_stream_number	000
extended_data_type	
data_type	0×0001

BS.2143-08

3.7 Reference point of the Serial ADM metadata

The reference point of the Serial ADM metadata is the first bit of the **burst_payload** (the **burst_payload** that has the **track_ID** '000000' and the **in_timeline_flag** '00' or '11' when multiple **burst_payloads** are used) following the **burst_preamble** as shown in Fig. 3.

3.8 Reference position

The **SADM_metadata_container** is defined to be in the reference position when the reference point of the **burst_payload** is aligned with the first sample of the segmented audio signals associated with the S-ADM metadata in the **SADM_metadata_container**.

4 Types of interfaces conforming to the transport method for S-ADM

4.1 Single AES interface (two subframes) for real-time applications

The length of data bursts and the numbers of tracks for the multiple over-track mode and continuous data bursts for the multiple in-timeline mode are limited depending on the use case. Multiple over-track mode requires synchronized AES3 interfaces. Multiple AES3 interfaces, however, might not be synchronized with each other, sample by sample. Therefore, at least one of the single AES3 interfaces with the parameters of data burst defined in Table 17 is used to convey the S-ADM metadata for real-time applications.

TABLE 17

Parameters for single AES3 interface for real-time applications

System parameters		**Value**		
		A1	**B2**	**C2**
Length of data burst		Up to 3 200 samples	Up to 3 200 samples	Up to 4 096 samples
assemble_info				
	multiple over-track mode	N/A (1 track)	Up to 2 tracks (1)	Up to 2 tracks (1)
	multiple in-timeline mode	N/A (2)	Up to 2 continuous data bursts	Up to 3 continuous data bursts
format_info				
	format type	N/A (3)	N/A (3)	N/A (3)
Bit depth (bits)		24	24	24
Maximum latency for 48 000 Hz (ms)		66.7 ms	133.3 ms with 2 data bursts	256 ms with 3 data bursts

(1) A pair of subframes within a single AES3 interface is used to transport the S-ADM metadata using the multiple over-track mode.

(2) The multiple in-timeline mode is not supported for real-time applications because low latencies are required.

(3) The format type '0000' is used in this case. The S-ADM metadata are encoded as UTF-8.

4.2 Multiple AES3 interfaces

The S-ADM metadata are conveyed by multiple AES3 interfaces using the parameters of data burst as shown in Table 18. Maximum numbers of multiple over-track mode are 2, 4/8/16 and 64 audio tracks for AES3, SDI and MADI, respectively. Typical size of the Serial ADM is at most 100 kbytes or so. 16 audio tracks, therefore, are enough to transport the S-ADM.

TABLE 18

Parameters for multiple AES3 interfaces

System parameters		Value		
		Real-time applications		Non-real-time applications
		A	B	D
Length of data burst		Up to 3 200 samples	Up to 3 200 samples	Up to 4 096 samples
assemble_info				
	multiple over-track mode	Up to 4 tracks (A4) Up to 8 tracks (A8) Up to 16 tracks (A16)	Up to 4 tracks (B4) Up to 8 tracks (B8) Up to 16 tracks (B16)	Up to 4 tracks (D4) Up to 8 tracks (D8) Up to 16 tracks (D16)
	multiple in-timeline mode	N/A (1)	Up to two continuous data bursts	Up to six continuous data bursts
format_info				
	format type	N/A (2)	N/A (2)	N/A (2)
Bit depth (bits)		24	24	24
Maximum latency for 48 000 Hz (ms)		66.7 ms (3)	133.3 ms with 2 data bursts (3) 66.7 ms with a data burst (3)	512 ms with 6 data bursts 85.3 ms with a data burst

(1) The multiple in-timeline mode is not supported for real-time applications because low latencies are required.

(2) The format type “0000” is used in this case. The S-ADM metadata are encoded as UTF-8.

(3) 66.7 and 133.3 ms correspond to two and four video frames of the 60i video format, respectively.

4.3 Multiple AES3 interfaces using compression tool

The compressed S-ADM metadata are conveyed using the parameters of the data burst shown in Table 19.

TABLE 19

Parameters for multiple AES3 interfaces using compression tool

System parameters		Value		
		Real-time applications		Non-real-time applications
		A	B	D
Length of data burst		**Up to 3 200 samples**	**Up to 3 200 samples**	**Up to 4 096 samples**
assemble_info				
	multiple over-track mode	N/A (1 track) (AX1) Up to 2 tracks (AX2) Up to 4 tracks (AX4)	NA (1 track) (BX1) Up to 2 tracks (BX2) Up to 4 tracks (BX4)	NA (1 track) (DX1) Up to 2 tracks (DX2) Up to 4 tracks (DX4)
	multiple in-timeline mode	N/A	Up to 2 continuous data bursts	Up to 6 continuous data bursts
format_info				
	format type	0001 (gzip)	0001 (gzip)	0001 (gzip)
Bit depth (bits)		24	24	24
Maximum latency for 48 000 Hz (ms)		66.7 ms with a data burst	133.3 ms with 2 data bursts 66.7 ms with a data burst	512 ms with 6 data bursts 85.3 ms with a data burst

4.4 Synchronization with video frames

The S-ADM metadata are conveyed using the parameters of the data burst shown in Table 20 to synchronize with video frames.

TABLE 20

Parameters for synchronization with video frames

System parameters		Value			
		50 Hz		60 Hz	
Length of data burst		Up to 960 samples	Up to 1 920 samples	Up to 800 samples	Up to 1 600 samples
assemble_info					
	multiple over-track mode	N/A (1 track) (V50X-1) Up to 2 tracks (V50X-2) Up to 4 tracks (V50X-4)	N/A (1 track) (V25X-1) Up to 2 tracks (V25X-2) Up to 4 tracks (V25X-4)	N/A (1 track) (V60X-1) Up to 2 tracks (V60X-2) Up to 4 tracks (V60X-4)	N/A (1 track) (V30X-1) Up to 2 tracks (V30X-2) Up to 4 tracks (V30X-4)
	multiple in-timeline mode	N/A			

TABLE 20 (*end*)

System parameters	**Value**			
	50 Hz		**60 Hz**	
Length of data burst	**Up to 960 samples**	**Up to 1 920 samples**	**Up to 800 samples**	**Up to 1 600 samples**
format_info				
format type	0001 (gzip)			
Bit depth (bits)	24			
Maximum latency for 48 000 Hz (ms)	20 ms with a data burst	40 ms with a data burst	16.67 ms with a data burst	33.33 ms with a data burst

4.5 Channel allocations to the transport method for S-ADM

The S-ADM metadata are allocated to channels of AES3-based interfaces shown in Table 21.

TABLE 21

Channel allocations

Number of tracks to convey S-ADM	**Channel allocations**		
	AES3	**SDI**	**MADI**
1	2	16	64
2	1 to 2	15 to 16	63 to 64
4	N/A	13 to 16	61 to 64
8	N/A	9 to 16	57 to 64
16	N/A	1 to 16	49 to 64

Annex 3

Bibliography

[1] AES3-2009 (r2019), AES Standard for Digital Audio Engineering – Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data

[2] SMPTE ST 337:2015, Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

[3] SMPTE ST 338:2016, Format for Non-PCM Audio and Data in AES3 – Data Types, Amendment 1:2019 to SMPTE ST 338:2016

[4] SMPTE ST 2116:2019, Format for Non-PCM Audio and Data in AES3 – Carriage of Metadata of Serial ADM (Audio Definition Model)

[5] Recommendation ITU-R BS.2125 (01/2019) – A serial representation of the Audio Definition Model, https://www.itu.int/rec/R-REC-BS.2125/en

[6] Internet Engineering Task Force (IETF) RFC 1952 (05/1996), GZIP file format specification version 4.3 [online, viewed 2018-12-04], http://tools.ietf.org/html/rfc1952