SMPTE STANDARD

Format for Non-PCM Audio and Data in AES3 — Audio Metadata

Table of Contents

Approved July 9, 2019

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 its Standards Operations Manual. This SMPTE Engineering Document was prepared by Technology Committee 32NF.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Engineering Document. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

This standard describes an open and extensible framework for transport of Audio Metadata over an AES3 interface. This framework is based on the KLV data encoding protocol (SMPTE ST 336) and the SMPTE ST 337 format for Non-PCM Audio and Data in AES3.

Audio Metadata is defined as metadata that is associated with audio essence which is present in other AES3 channels, either as one or more PCM audio channels, or as coded audio essence (in ST 337 format). This audio metadata might require time synchronization with the audio essence, in some cases down to a resolution of an audio sample. This standard includes signaling that supports this synchronization.

1 Scope

This standard specifies format requirements for real-time Audio Metadata carried within an AES3 interface according to SMPTE ST 337. Audio Metadata is grouped into *payloads* for transport. This standard covers wrapping and transport of the Audio Metadata payloads only. The contents of audio metadata payloads are defined elsewhere and are beyond the scope of this standard, however Utility payloads that contain information that supports carriage of the Audio Metadata payloads are within scope and are defined in this standard. This standard specifies KLV encoding for wrapping and transport of Audio Metadata payloads, but places no requirement on the format of the metadata carried within Audio Metadata payloads.

This standard applies to Audio Metadata payloads only. Transport of payloads other than Audio Metadata payloads and associated Utility payloads is beyond the scope of this standard.

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.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; then formal languages; then figures; and then any other language forms.

3 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this engineering document. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this engineering document are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

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

SMPTE ST 355:2001, Format for Non-PCM Audio and Data in AES3 - KLV Data Type

SMPTE ST 336:2017, Data Encoding Protocol Using Key Length Value

ISO/IEC 8825-1:2008 Information Technology — ASN.1 Encoding Rules: Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER) and Distinguished Encoding Rules (DER)

ETSI EN 300 468 v1.15.1 - Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems

4 Terms and Definitions

For the purposes of this document, the following terms and definitions apply.

4.1 Section

group of payloads within the Audio Metadata Pack

4.2 Payload

item consisting of a payload tag, payload length, and payload value

4.3 Utility Payload

payload whose value field contains information other than audio metadata

4.4 Audio Metadata Payload

payload whose value field contains audio metadata.

4.5 Audio Metadata Pack

ST 336 Defined-Length Pack, as defined in this document, with a single element, ST 2109 Audio Metadata, consisting of Utility Payloads and/or Audio Metadata Payloads

4.6 Static payload tag

ASN.1 OID BER SubId whose UL mapping is permanently assigned via the ST 2109 Administrative Register and is valid for all Audio Metadata Packs

4.7 Dynamic payload tag

ASN.1 OID BER SubId whose UL mapping is assigned within an Audio Metadata Pack, and whose assignment scope is limited to that pack

4.8 ASN.1 OID BER SubId

BER encoding of subidentifiers as defined in clause 8.1.2.4.2 of ISO/IEC 8825-1. The first 127 nonzero values are encoded in a single byte, values 128 to 16383 are encoded in 2 bytes, greater values ere encoded in 3 or more bytes as necessary.

4.9 BER

Basic Encoding Rules, defined in ISO/IEC 8825-1

4.10 BER short form

BER encoding as defined in clause 8.1.3.4 of ISO/IEC 8825-1

4.11 BER long form

BER encoding as defined in clause 8.1.3.5 of ISO/IEC 8825-1

4.12 UL

SMPTE administered Universal Label as defined in SMPTE ST 298

5 Notations

Table columns shall have the meanings as defined in Table 1.

Table 1 - Definition of Table Headings

Table heading	Definition
Item Name	Name of this pack or metadata item as indicated in the defining document
Type	The data type of the entry
Len	Length of the value field for this entry, in bytes, according to the data type
Payload Tag	Value of the payload tag that identifies this entry
Req	Req indicates that the item is always present. This indicator maps to the isOptional entry in the SMPTE Groups Register, where isOptional corresponds to table entry != "Req". Opt indicates that the item can be present or omitted.

Table heading	Definition
	Opt* indicates that this item is present or omitted depending on the presence or value of other items.
Definition	Description of this entry
Symbol	The registry symbol that is associated with this item
Kind	The kind of the registry entry
Item UL	The SMPTE Universal Label that shall be associated with this item

6 Audio Metadata

In this standard, Audio Metadata is segmented into *payloads* that are intended to combine related audio metadata together (e.g. audio loudness metadata, object position metadata, etc.). Multiple payloads, grouped into sections, can be combined together into a single *Audio Metadata Pack*. For AES3 transport Audio Metadata Packs are wrapped within SMPTE ST 337 data bursts with no more than one Audio Metadata Pack per data burst. In addition to audio metadata payloads, packs can include one or more utility payloads that contain metadata that support the transport format. These payloads can include information that supports synchronization with associated audio essence, or identification of associated audio essence, for example.

Audio Metadata Packs are encoded in KLV format per SMPTE ST 336. Payloads are not required to be KLV encoded. The pack encoding is defined as a KLV Defined-Length Pack as specified in this standard and identified by a SMPTE Universal Label (Audio Metadata Pack). The Audio Metadata Pack consists of a single element, ST 2109 Audio Metadata, which is a defined as a sequence of payloads. Each type of payload within an Audio Metadata Pack is identified by a unique SMPTE Universal Label (UL). The ST 2109 Administrative Register defines the mapping of payload ULs to static payload tags. Dynamic mapping of payload ULs to payload tags is also supported for payload types not defined in the ST 2109 Administrative Register. Consistent with the SMPTE metadata dictionary, Audio Metadata Packs can contain payloads with keys defined as private use ULs. Data within payloads identified as private use does not have to be publicly disclosed.

Figure 1 shows an overview of the Audio Metadata Pack structure.

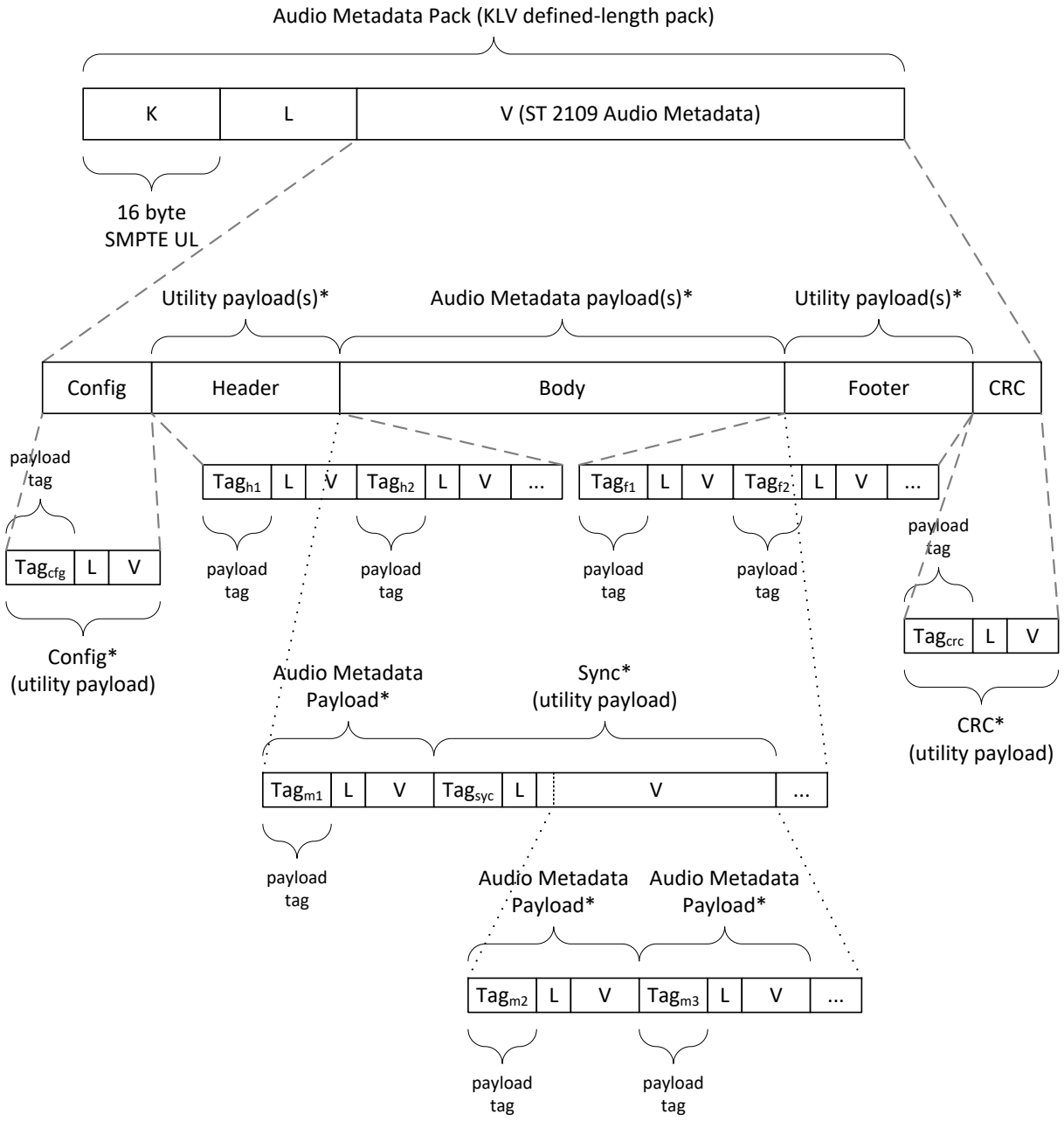

Figure 1 - Audio Metadata Pack structure

* All payloads are optional

7 Encoding of Audio Metadata Pack

7.1 Introduction

This clause specifies the encoding of the Audio Metadata Pack using the Key-Length-Value data encoding protocol (KLV), defined in SMPTE ST 336. Each Audio Metadata Pack shall be a KLV Defined-Length pack, as defined in ST 336. The Audio Metadata Pack is composed a single item, ST 2109 Audio Metadata, which may include one or more payloads.

7.2 Audio Metadata Pack Key

The Key of the Audio Metadata Pack shall be a SMPTE Administered Universal Label as per SMPTE ST 336.

The group node and the pack key shall have the ULs defined in Table 2. These are listed in the SMPTE Groups Register.

The structure of the Audio Metadata Pack shall be as shown in Table 3. The Audio Metadata Pack shall include a single element, ST 2109 Audio Metadata, which is composed of a sequence of payloads encoded according to clause 7.4.

Table 2 – Group Node and Pack Key for Audio Metadata Packs

Item Name	Symbol	Kind	Item UL
Audio Metadata	AudioMetadata	NODE	06.0E.2B.34.02.7F.01.01.0C.04.00.00.00.00.00.00
Audio Metadata Pack	AudioMetadataPack	LEAF	06.0E.2B.34.02.7F.01.01.0C.04.01.00.00.00.00.00

Table 3 – Audio Metadata Pack

Item Name	Type	Len	Req	Definition
Audio Metadata Pack	Pack key	16	Req	Identifies an Audio Metadata Pack
Length	BER length	Var	Req	Pack length
ST 2109 Audio Metadata	ST 2109 Payload Series	Var	Opt	Sequence of ST 2109 payloads

7.3 Audio Metadata Pack Length

The Length of the Audio Metadata Pack shall be coded as BER length, long or short form as required.

7.4 Audio Metadata Pack Value

The Value of the Audio Metadata Pack shall be the single element, ST 2109 Audio Metadata, whose total length is given by the Length field.

The ST 2109 Audio Metadata element and its associated node shall have the ULs defined in Table 4. These are listed in the SMPTE Elements Register.

Table 4 – ST 2109 Audio Metadata Node and Element

Item Name	Symbol	Kind	Item UL
Audio Metadata Elements	AudioMetadataElements	NODE	06.0E.2B.34.01.01.01.0E.04.02.09.00.00.00.00.00
ST 2109 Audio Metadata	ST2109AudioMetadata	LEAF	06.0E.2B.34.01.01.01.0E.04.02.09.01.00.00.00.00

ST 2109 Audio Metadata shall have the type ST 2109 Payload Series with the UL as defined in Table 5. This is listed in the SMPTE Types Register. The ST 2109 Payload Series type shall be a sequence of ST 2109 payloads as defined in this clause.

Table 5 – ST 2109 Payload Series

Item Name	Symbol	Kind	Item UL
ST 2109 Payload Series	ST2109PayloadSeries	LEAF	06.0E.2B.34.01.04.01.01.04.10.0B.00.00.00.00.00

ST 2109 Audio Metadata shall consist of a sequence of payloads grouped into sections. The payloads in ST 2109 Audio Metadata shall be constructed as defined in this clause and shall consist of Utility Payloads and Audio Metadata Payloads. Utility Payloads and Audio Metadata Payloads may be public payloads or private payloads. All payloads, whether public or private, shall be identified by SMPTE Universal Labels.

ST 2109 Payloads consist of three consecutive fields: Payload Tag, Payload Length, and Payload Value. The Payload Tag shall be ASN.1 OID BER coded and shall be associated with a SMPTE UL as described in clause 7.5. The Payload Length shall be ASN.1 BER short or long encoded and shall indicate the length of the Payload Value field. The contents of the Payload Value field are defined by the SMPTE UL associated with the Payload Tag. The ST 2109 Payload structure is illustrated in Figure 2.

NOTE: The Value of the ST 2109 Audio Metadata Element (the series of ST 2109 Payloads) constitutes the Value of the Audio Metadata Pack in its entirety. In this document references to the Value or contents of the Audio Metadata Pack are equivalent to references to the ST 2109 Audio Metadata Element.

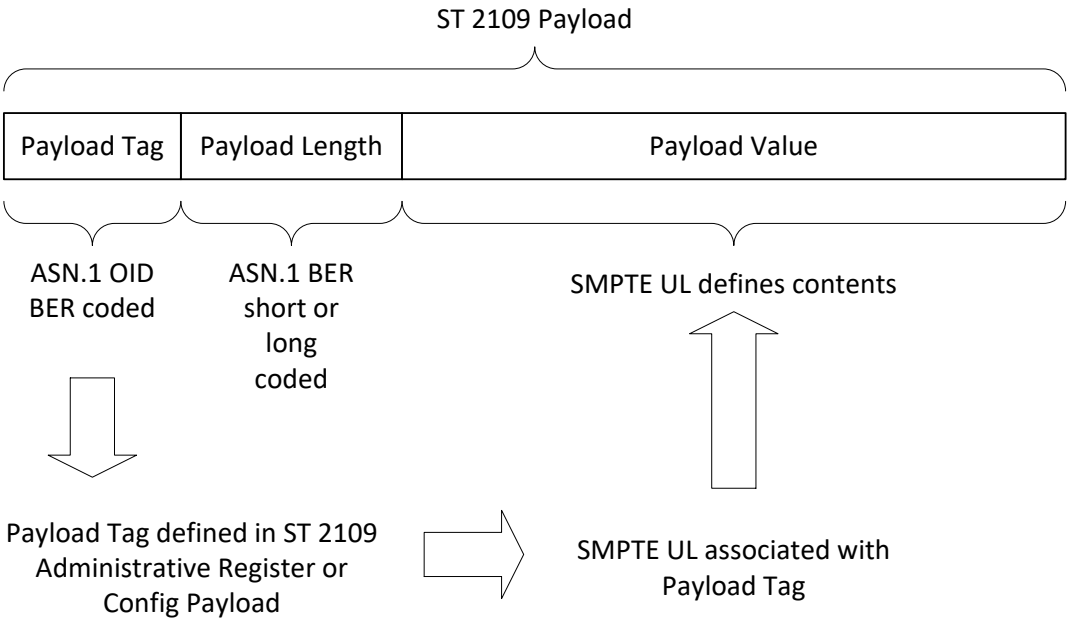

Figure 2 – ST 2109 Payload structure

The Value of the Audio Metadata Pack shall be constructed such that the Config Payload, if present, shall be the first payload of the Value field. Any Utility Payloads defined as "Header Payloads" shall occur in sequence following the Config Payload (if present) and prior to any Audio Metadata Payloads. Any Utility Payloads defined as "Footer Payloads" shall occur in sequence following Audio Metadata Payloads. Audio Metadata Payloads may be present in the Value field after any Header Payloads and prior to any Footer Payloads. The CRC Payload, if present, shall be the final payload at the end of the Value field.

The Config Payload defines the Config section of the Audio Metadata Pack. The sequence of Header Payloads within an Audio Metadata Pack defines the Header section of the Audio Metadata Pack. The sequence of Footer Payloads within an Audio Metadata Pack defines the Footer section of the Audio Metadata Pack. The sequence, within an Audio Metadata Pack, of Audio Metadata Payloads and other Utility Payloads not defined as Header Payloads or Footer Payloads defines the Body section of the Audio Metadata Pack. The CRC Payload defines the CRC section of the Audio Metadata Pack. This is illustrated in Figure 1.

Audio Metadata Payloads may be contained within a Sync Payload. The information contained in a Sync Payload shall apply solely to the Audio Metadata Payloads that are contained within the Sync Payload Value. Header and Footer Utility Payloads should contain information that applies to the entire Audio Metadata Pack including all Audio Metadata Payloads in the pack.

Other standards or recommendations may specify the order in which specific payloads are encoded within each section. For example, a standard defining audio metadata payloads may specify ordering requirements for these payloads.

Utility and Audio Metadata Payloads are described in Sections 8 and 9. Specific payloads may be included in the pack by addition to the ST 2109 Administrative Register described in Section 10. All payloads included in the ST 2109 Administrative Register will have static payload tags as assigned in the register. The Audio

Metadata Pack can include payloads not assigned in the ST 2109 Administrative Register by the Dynamic Tag mapping assignment described in 7.5.

7.5 Audio Metadata Pack Payload Tag Assignment

All payloads included in the Audio Metadata Pack shall be identified by payload tags associated with a SMPTE UL. Payload tags can be statically or dynamically defined. Any payloads included in an Audio Metadata Pack that do not have statically defined tags shall have a dynamic tag that is assigned in the Config Payload. The dynamic tag assignment shall apply only to payload tags in the Audio Metadata Pack that includes the Config Payload with the assignments. All payload tags in an Audio Metadata Pack that do not have dynamic tag assignments shall have static payload tags as assigned in the ST 2109 Administrative Register. The Static Payload Tag of each Payload shall be listed in the ST 2109 Administrative Register specified in clause 10.

Dynamic tags may override statically defined tags, but if so the payload type identified by the payload tag that is overridden shall not be included in the Audio Metadata Pack unless an alternate dynamic tag is assigned to that payload. Exceptions are the Config Payload, Sync Payload, and CRC payload which shall always have a static local tag.

In practical usage it is anticipated that commonly used payloads will have static local tags assigned, as use of dynamic tags requires repetition of the tag assignment in each Audio Metadata Pack. Dynamic tag assignments are intended for payloads that are not commonly used, or for cases where static payload tags are awaiting assignment in the ST 2109 Administrative Register. However some applications might benefit from the use of dynamic tags and this standard places no restriction on their use.

8 Utility Payloads

8.1 Introduction

This clause defines Utility Payloads that may be present in an Audio Metadata Pack. Utility Payloads shall be encoded in the Audio Metadata Pack as defined in this standard. Utility payloads defined in this document, and their associated node, shall have SMPTE ULs as defined in Table 6. These are listed in the SMPTE Labels Register.

Table 6 – Utility Payload Node and Label ULs

Item Name	Symbol	Kind	Item UL
Audio Metadata	AudioMetadata	NODE	06.0E.2B.34.04.01.01.0D.04.04.02.00.00.00.00.00
Config Payload	ConfigPayload	LEAF	06.0E.2B.34.04.01.01.0D.04.04.02.01.00.00.00.00
Sync Payload	SyncPayload	LEAF	06.0E.2B.34.04.01.01.0D.04.04.02.02.00.00.00.00
CRC Payload	CRCPayload	LEAF	06.0E.2B.34.04.01.01.0D.04.04.02.03.00.00.00.00

8.2 Config Payload

8.2.1 Config Payload Overview

The Config Payload is a Utility Payload that shall be encoded in the Audio Metadata Pack in the Config section. The Config Payload is optional for all Audio Metadata Packs. If present it shall be the first payload at the beginning of the Audio Metadata Pack Value section. The Config Payload contains information describing

the Audio Metadata Pack including pack synchronization information and optional dynamic payload tag mapping. Figure 3 illustrates the structure of the Config Payload.

8.2.2 Config Payload Tag

The Payload Tag of the Config Payload shall be a static payload tag with the tag corresponding to a 16-byte SMPTE Universal Label as defined in Table 6.

8.2.3 Config Payload Length

The Payload Length of the Config Payload shall be coded as BER length, long or short form, and shall indicate the size of the Config Payload Value field.

8.2.4 Config Payload Value

The first two bytes of the Payload Value field shall contain the pack_sample_offset field. pack_sample_offset shall be coded as a 16-bit signed integer, most significant byte first, and shall indicate the offset, in units of PCM audio samples, of the Audio Metadata Pack Reference Position from the default Reference Position. The Audio Metadata Pack Reference Position shall be the default Reference Position plus the pack_sample_offset. The value may indicate a positive or negative offset relative to the default Reference Position. A value of '0' shall indicate no modification to the default Audio Metadata Pack Reference Position.

NOTE: The Reference Position of the Audio Metadata Pack is defined in terms of the transport interface. The Reference Position for AES3 transport is defined in clause 11.4.

When the Length field indicates data is present beyond the Pack Sample Offset (Length > 2) it shall indicate that mapping information for dynamic payload tags is present. The mapping is indicated by a payload tag followed by a 16 SMPTE universal label (UL). Multiple dynamic tags are defined by a sequence of tag/UL mappings.

The scope of the dynamic tag mapping shall be the current Audio Metadata Pack only. The dynamic tag mapping information shall be present in all Audio Metadata Packs that contain the payload identified in the mapping. All payloads that do not have statically defined payload tags shall include a dynamic tag mapping.

The dynamic payload tag mapping shall override statically defined payload tags. Dynamic mapping of a UL that has been assigned a static tag shall override the static tag mapping. Dynamic mapping to a payload tag that has been assigned as a static tag shall override the SMPTE UL mapping defined in the static tag mapping. SyncPayload, CRCpayload, and the ConfigPayload shall have a permanently assigned payload tag that shall not be overridden (the Payload UL shall not be mapped to a dynamic tag, nor shall the static payload tag be dynamically assigned to another payload UL).

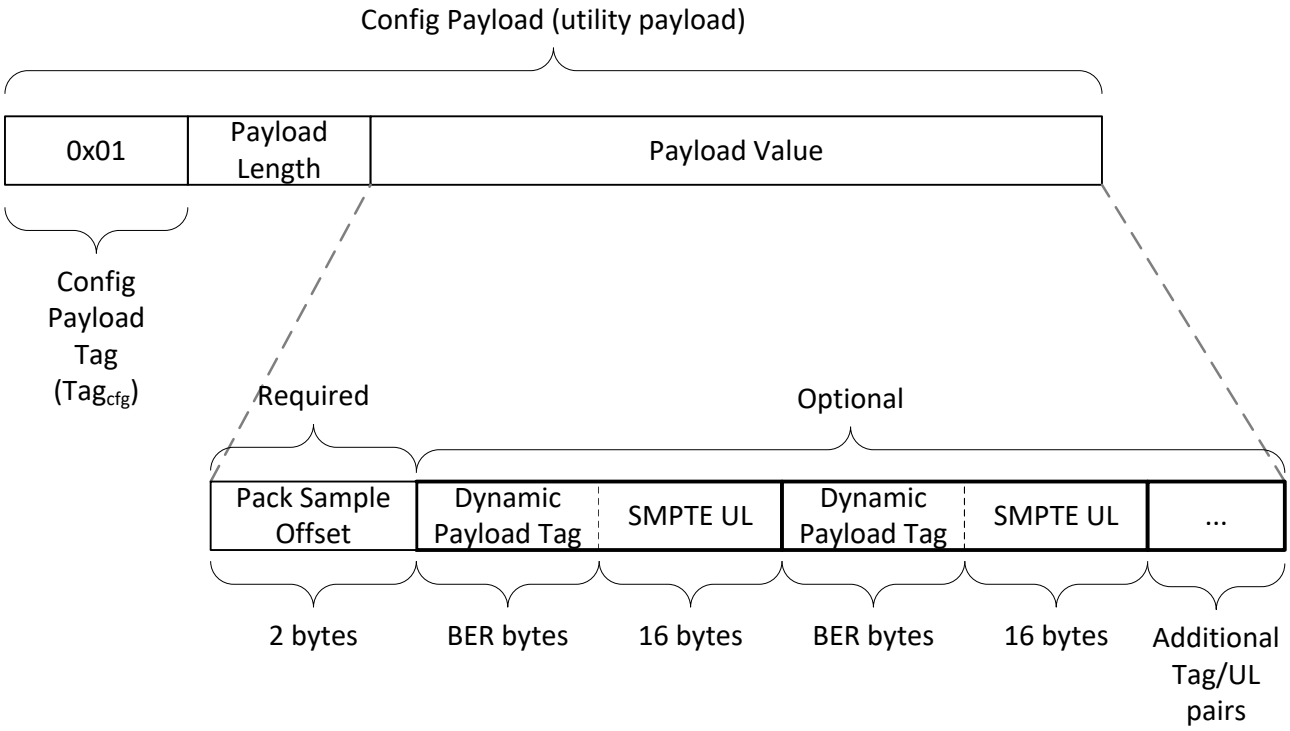

Figure 3 –Config Utility Payload

8.3 Sync Payload

8.3.1 Sync Payload Overview

The Sync Payload is a Utility Payload that shall be encoded in the Audio Metadata Pack as part of the Audio Metadata payload section. Sync Payloads are optional. When present, a Sync Payload shall contain one or more Audio Metadata Payloads in its value. The information contained in a Sync Payload shall apply only to the Audio Metadata Payloads that are present in its value.

8.3.2 Sync Payload Tag

The Payload Tag of the Sync Payload shall be a payload tag, with the tag corresponding to a 16-byte SMPTE Universal Label as defined in Table 6.

8.3.3 Sync Payload Length

The Payload Length of the Sync Payload shall be coded as BER length, long or short form, and shall indicate the size of the Sync Payload Value field (the Sync Payload content).

8.3.4 Sync Payload Value

The Payload Value of the Sync Payload shall be the following:

```
sync_payload_value()
{
        sample_offset..................................................................... 16 bits
        duration.......................................................................... 16 bits
        Audio Metadata Payloads.......................................................... var bits
}
```

sample_offset - payload sample offset - 16 bits (signed integer)

The 16-bit sample_offset field shall indicate the offset, in units of PCM audio samples, from the first sample of the audio essence to which it associated (as defined by the Reference Position of the Audio Metadata Pack) to the first PCM audio sample that the data in the Audio Metadata Payload shall apply to. If the current Audio Metadata Payload applies to the first sample of the audio essence to which it associated (as defined by the Reference Position of the Audio Metadata Pack), this field shall be set to '0'. sample_offset shall be coded as a 16 bit signed integer that can indicate positive of negative offsets of the payload relative to the Audio Metadata Pack Reference Position. A positive offset corresponds to a Reference Position later in time than the default Reference Position.

The default Reference Position of an Audio Metadata Pack may be modified by the pack_sample_offset in the Config Payload. For Audio Metadata Payloads contained in a Sync Payload, both pack_sample_offset and sample_offset jointly act to modify the offset of these Audio Metadata Payloads. The total offset is the set_sample_offset plus the sample_offset.

NOTE: The pack Reference Position for AES3 transport is as defined in clause 11.4.

duration - payload duration data – 16 bits (unsigned integer)

The 16-bit duration field shall indicate the time period in units of PCM audio samples of the associated audio essence that the Audio Metadata Payloads in the Sync Payload applies to. If the current payload applies to all PCM audio samples of the associated audio essence until a subsequent Audio Metadata Payload of the same type is received, this field shall be set to '0'. duration shall be coded as a 16 bit unsigned integer.

sample_offset and duration shall be coded most significant byte first, in the Sync Payload Value field.

The structure of the Sync Payload is illustrated in Figure 4.

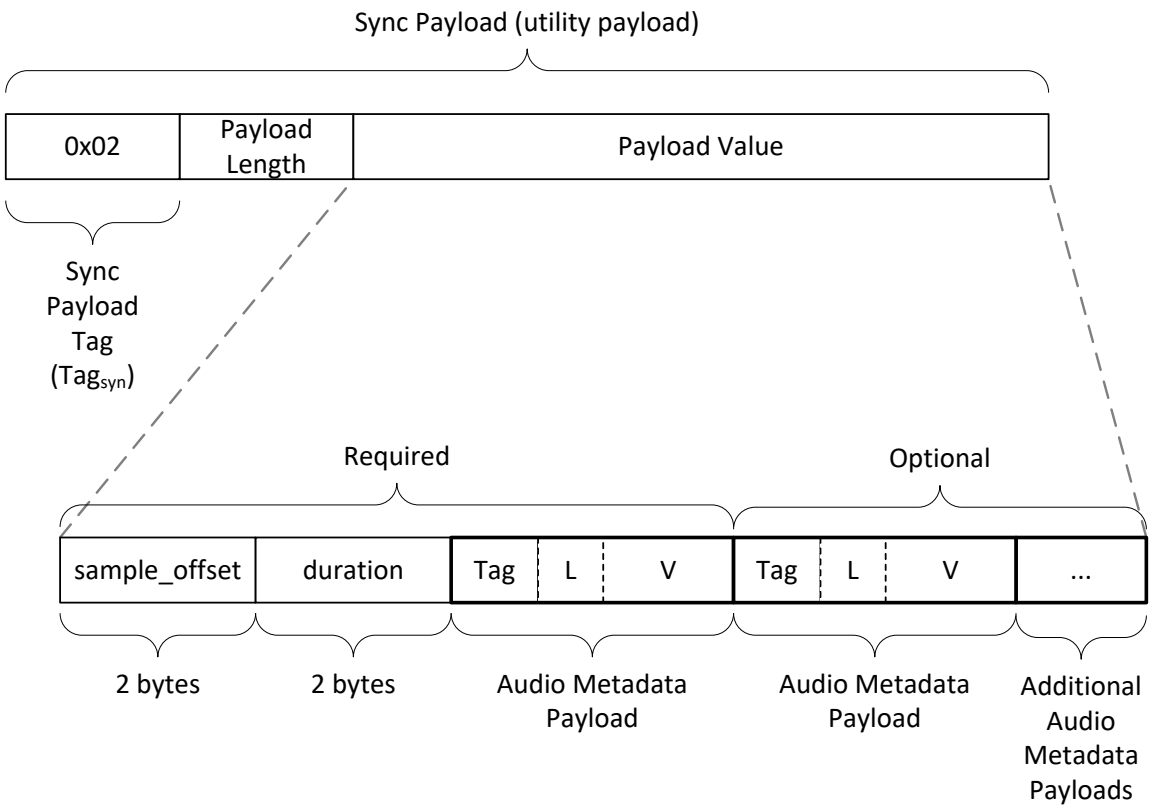

Figure 4 –Sync Utility Payload

8.4 CRC Payload

8.4.1 CRC Payload Overview

The CRC Payload is a Utility Payload that shall be encoded in the Audio Metadata Pack in the CRC section. CRC Payloads are optional. When present, a CRC Payload shall occur as the last payload in the pack.

8.4.2 CRC Payload Tag

The Payload Tag of the CRC Payload shall be a payload tag, with the tag corresponding to a 16-byte SMPTE Universal Label as defined in Table 6.

8.4.3 CRC Payload Length

The Payload Length of the CRC Payload shall be coded as BER length, long or short form, and shall indicate 4 bytes, the size of the CRC Payload Value field (the CRC Payload content).

8.4.4 CRC Payload Value

The Payload Value of the CRC Payload shall be a 4-byte (32-bit) Cyclic Redundancy Check (CRC) code, as specified in ETSI EN 300 468 v1.15.1, Annex B. The CRC is computed over the entire contents of the Audio Metadata Pack, including the pack Key and Length, but excluding the CRC Payload Value (the CRC itself).

9 Audio Metadata Payloads

9.1 Audio Metadata Payload Overview

This clause defines the format of Audio Metadata Payloads that may be present in an Audio Metadata Pack. Audio Metadata Payloads shall be encoded in the Body section of the Audio Metadata Pack as defined in this standard.

An Audio Metadata Payload shall consist of one payload of audio metadata. Audio Metadata Payloads are defined elsewhere and may contain public or private data. All Audio Metadata Payloads, whether public or private, shall have a SMPTE registered Universal Label and shall be mapped to an Audio Metadata Pack payload tag.

9.2 Audio Metadata Payload Tag

The Payload Tag of the Audio Metadata Payload shall be a payload tag, with the tag corresponding to a SMPTE Universal Label. The Universal Label may correspond to public or private data. All Audio Metadata Payloads encoded according to this standard shall have a valid SMPTE Universal Label that is mapped to a payload tag, either as a static payload tag as per the ST 2109 Administrative Register described in clause 10, or as a payload tag that is dynamically assigned in the Config Payload, as described in clause 8.2.

9.3 Audio Metadata Payload Length

The Payload Length of the Audio Metadata Payload shall be coded as BER length, long or short form, and shall indicate the size of the Audio Metadata Payload Value field (the Audio Metadata Payload content).

9.4 Audio Metadata Payload Value

The Payload Value of the Audio Metadata Payload shall contain the specific Audio Metadata content as identified by the Audio Metadata Payload Key. For Audio Metadata Payloads identified as public, the registration shall point to a publicly available specification on the payload contents. For Audio Metadata Payloads identified as private, the specification and disclosure of the contents is at the discretion of the registering entity. Specific Audio Metadata Payloads may place requirements on the Audio Metadata Pack beyond those specified in this standard. Such requirements are beyond the scope of this standard.

Audio Metadata Payload values are defined as integer numbers of bytes. When an Audio Metadata Payload specification defines a payload that is not an integer number of bytes, unused bits in the last payload value byte shall be set to ‘0’.

10 Payload Static Tag Assignments

10.1 ST 2109 Administrative Register

This clause defines the ST 2109 Administrative Register which is an Administrative Register as specified in the SMPTE Standards Operations Manual (SMPTE OM). Each entry in the ST 2109 Administrative Register shall consist of the following information:

- a Payload Tag (as specified in clause 7.5)
- a unique SMPTE UL (as specified in ST 298)

No two entries shall have identical Payload Tag values.

No two entries shall have identical SMPTE ULs.

Payload Tag values shall be allocated consecutively, starting with value 0x01.

The SMPTE UL shall not be designated as experimental by its defining register.

The ST 2109 Administrative Register shall consist of at least the following entries:

Table 7 – ST 2109 Administrative Register Initial Entries

Payload Tag	SMPTE UL
0x01	06.0E.2B.34.04.01.01.0D.04.04.02.01.00.00.00.00
0x02	06.0E.2B.34.04.01.01.0D.04.04.02.02.00.00.00.00
0x03	06.0E.2B.34.04.01.01.0D.04.04.02.03.00.00.00.00

A new entry may be added if it meets the criteria in this clause.

An entry shall not be modified or removed once added.

The ST 2109 Administrative Register shall be contained in an online register at: https://smpte-ra.org/smpte-st-2109-ar.

10.2 Private Payloads

As new payloads, and payloads identified by Private ULs, may be added to this specification in the future, conformant decoders shall be able to parse any payload encoded according to this standard and to extract the recognized data while ignoring payloads whose format is not recognized.

Decoders shall always follow the payload syntax described in this standard whether or not the format of the payload is recognized. For an unrecognized payload, whose Payload Tag corresponds to an unrecognized SMPTE UL, the decoder shall use the Payload Length field to skip over the Payload Value of the item and continue processing at the start of the next Payload Tag.

Decoders are not required to react to a payload identified by a Private UL. They should preserve any item identified by a Private UL in the Audio Metadata Pack to enable later processes to use it.

11 Mapping of Audio Metadata Packs to AES3

11.1 Introduction

Audio Metadata Packs shall be mapped to the AES3 interface according to SMPTE ST 355, which specifies carriage of SMPTE ST 336 (KLV) data bursts within an AES3 interface according to SMPTE ST 337. This clause contains additional constraints beyond the SMPTE ST 355 specification when the SMPTE ST 336 data bursts contain Audio Metadata Packs.

11.2 burst_preamble

11.2.1 burst_preamble Overview

The Pc word (burst_info value) of the burst_preamble carries the data_type identifier, the data_type_dependent and the data_stream_number information, as defined by SMPTE ST 337.

11.2.2 data_type identifier

As per SMPTE ST 355, the data_type identifier shall be set to 27.

NOTE: ST 338 data type 27 is not reserved exclusively for ST 2109 Audio Metadata transport. Receiving devices should not rely on data type 27 as indication that ST 2109 formatted Audio Metadata is present.

11.2.3 data_type_dependent

As per SMPTE ST 355, the burst_preamble shall include a data_type_dependent field as shown in ST 355 Table 1, with bit 4 defined as the key_flag. As per ST 355, when this flag is set to 1, a full SMPTE Universal Label (UL) key for the Audio Metadata Pack shall be present at the beginning of the burst_payload indicating that the data burst contains the start of a KLV packet containing audio metadata. When this flag is set to 0, no key shall be present in the burst_payload indicating that the data burst is a continuation of an Audio Metadata Pack KLV packet.

11.2.4 data_stream_number

The data_stream_number shall be set to any number from 0 to 6. See SMPTE ST 337.

11.3 Reference Point

The Reference Point of an Audio Metadata Pack data burst shall be bit 0 of the Pa sync word of the data burst in which the Key of the Audio Metadata Pack is present.

11.4 Reference Position

The Reference Position of an Audio Metadata Pack burst_payload is defined in relation to other AES3 streams that contain the PCM (or other coded) audio essence that the metadata is associated with. An Audio Metadata Pack data burst is defined as being in the default Reference Position when the Reference Point of the burst_payload is aligned with the first sample of the segment of audio essence the Audio Metadata Pack is associated with. The time period of the associated audio essence that the Audio Metadata Pack burst_payload applies to shall not exceed 250 msec. Specific Audio Metadata Payloads contained within the Audio Metadata Pack can indicate alignment with the audio that differs from the segment in both location and duration. Differences may be indicated as sample offsets relative to the audio segment aligned with the Reference Position by using of the Sync payload, or as other timing offsets included in the Audio Metadata Payload itself.

The requirements stipulate that the SMPTE Universal Label Key for the Audio Metadata Pack shall always be present at the beginning of a SMPTE ST 337 burst_payload. An ST 337 encoded Audio Metadata Pack may span multiple ST 337 data bursts using the key_flag described in clause 11.2.3 and in SMPTE ST 355, however a single ST 337 data burst shall contain no more than one Audio Metadata Pack as required by SMPTE ST 355.

11.5 Repetition Rate

To meet synchronization and latency requirements for real time applications, Audio Metadata Pack burst_payloads shall be transmitted in the AES3 interface at a time period not to exceed 250 msec. This requires Audio Metadata Packs to be restricted in size to no more than can be conveyed in the AES3 interface in a time period of 250 msec. When the AES3 sample rate is 48 kHz, this means the Audio Metadata Pack shall not exceed 12,000 AES3 frames, and the time period between Audio Metadata Pack burst_payloads shall not exceed 12,000 AES3 frames

This standard does not place requirements on the repetition rate of specific payloads included in Audio Metadata Packs; however, payloads shall not exceed the maximum size available in a single Audio Metadata Pack. Other standards or recommended practices may specify repetition rates and Audio Metadata Pack size restrictions less that those specified in this standard for specific audio metadata types or for specific applications.

NOTE: Many applications require low latency for real-time operation and will require repetition rates less than the maximum 250 msec. Other applications with higher tolerances for latency might benefit from metadata corresponding to larger audio block sizes and/or increased data space for audio metadata within the AES3 transport interface. ST 2109 is not intended for file transfer applications.

Bibliography (informative)

SMPTE Standards Operations Manual (SMPTE OM) v.3.1, April 4, 2015

SMPTE Elements Register

SMPTE Groups Register

SMPTE Labels Register

SMPTE Types Register

SMPTE ST 298:2009, Universal Labels for Unique Identification of Digital Data

SMPTE ST 338:2016, Format for Non-PCM Audio and Data in AES3 - Data Types