

SMPTE REGISTERED DISCLOSURE DOCUMENT

SMPTE RDD 9:2013

Revision of RDD 9-2009

MXF Interoperability Specification of Sony MPEG Long GOP Products



Page 1 of 33 pages

The attached document is a Registered Disclosure Document prepared by the proponent identified below. It has been examined by the appropriate SMPTE Technology Committee and is believed to contain adequate information to satisfy the objectives defined in the Scope, and to be technically consistent.

This document is NOT a Standard, Recommended Practice or Engineering Guideline, and does NOT imply a finding or representation of the Society.

Errors in this document should be reported to the proponent identified below, with a copy to eng@smpte.org. All other inquiries in respect of this document, including inquiries as to intellectual property requirements that may be attached to use of the disclosed technology, should be addressed to the proponent identified below.

Proponent contact information:

Hiroshi Nakano
Sony Corporation
4-14-1 Asahi-cho, Atsugi
Kanagawa, 243-0014
Japan

Email: hiroshi.nakano@jp.sony.com

Table of Contents	Page
1 Scope.....	3
2 Reference Documents.....	3
3 Introduction	3
4 Outline of MXF File Structure for this Mapping	4
5 MPEG Picture Data and AES3 Data Mapping	5
5.1 Frame Wrappings	5
5.2 System Item Mapping	6
5.2.1 Overview of System Item	6
5.2.2 System Metadata Pack	6
5.2.3 Package Metadata Set	7
5.2.4 Picture Metadata Set	7
5.3 Picture Item Mapping	7
5.3.1 MPEG Picture Element Key	8
5.3.2 MPEG Picture Element Length	8
5.3.3 MPEG Picture Element Value	8
5.4 AES3 Sound Item Mapping	9
5.4.1 AES3 Sound Element Key	9
5.4.2 AES3 Sound Element Length	10
5.4.3 AES3 Sound Element Value	10
5.5 Data Item Mapping (Optional)	10
5.6 Temporal Reordering	10
6 SMPTE Labels for Essence Container Identification	11
7 SMPTE Labels for Essence Coding Identification	11
8 Application Issues	12
8.1 Application of the KLV Fill Item	12
8.2 Application of MXF Structure and Indexing Style	12
8.2.1 Segmented Body Partition Style	12
8.2.2 Single Body Partition Style	14
8.3 Application of Index Table for Frame Wrapped MPEG Picture and AES Sound Essence	15
8.3.1 Essence Container and Index Table	15
8.3.2 Index Table Items	15
8.3.3 Delta Entry Array	16
8.3.4 Index Entry Array	18
8.3.5 Setting the Properties Specific for Long GOP MPEG	20
8.4 Application of Random Index Pack	22
Annex A UL Code List	24
Annex B Constraints of the Conformant Implementation	26
B.1 Structure	26
B.2 Header and Body Partition Pack Values	26
B.3 Essence Descriptors	26
B.4 Identification Set Value	27
B.5 Timecode Representation in MXF Header and an Essence Container	27
B.6 Index Table Segments	27
B.7 Random Index Pack	28
B.8 Essence	28
B.8.1 System Item	28
B.8.2 Picture Item	28
B.8.3 Sound Item	28
B.8.4 Data Item	28
Annex C Property Values of the Essence Descriptors	29

1 Scope

This document specifies the mapping of MPEG-2 Picture (ES), AES3 audio and ANC packets into the MXF Generic Container. The MXF files created according to the details of this specification comply with the MXF specifications defined in the normative references, with a variation with respect to one new rule introduced in SMPTE ST 377-1. In conjunction with the referenced Standards, this RDD is intended to provide sufficient information to enable a developer to construct MXF files that will be compatible with Sony MPEG-2 Long GOP products.

2 Reference Documents

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 326:2000) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 326-2000). Documents with the same root number (e.g. 326) and publication year (e.g. 2000) are functionally identical.

SMPTE ST 326:2000, Television — SDTI Content Package Format (SDTI-CP)

SMPTE ST 331:2011, Element and Metadata Definitions for the SDTI-CP

SMPTE ST 377-1:2011, Material Exchange Format (MXF) — File Format Specification

Amendment 2:2012 to SMPTE ST 377-1:2011

SMPTE ST 378:2004, Television — Material Exchange Format (MXF) — Operational Pattern 1a (Single Item, Single Package)

SMPTE ST 379-1:2009, Material Exchange Format (MXF) — MXF Generic Container

SMPTE ST 381-1:2005, Television — Material Exchange Format (MXF) — Mapping MPEG Streams into the MXF Generic Container

SMPTE ST 382:2007, Material Exchange Format — Mapping AES3 and Broadcast Wave Audio into the MXF Generic Container

SMPTE ST 385:2012, Material Exchange Format (MXF) — Mapping SDTI-CP Essence and Metadata into the MXF Generic Container

SMPTE ST 400:2012, Television — SMPTE Labels Structure

SMPTE ST 436:2006, Television — MXF Mappings for VBI Lines and Ancillary Data Packets

SMPTE RP 210, Metadata Element Descriptions

SMPTE RP 224, SMPTE Labels Register

SMPTE RDD 18:2012, Acquisition Metadata Sets for Video Camera Parameters

Recommendation ITU-R BT.709-5 (04/02), Parameter Values for the HDTV Standards for Production and International Programme Exchange

3 Introduction

The MXF Generic Container is a streamable Essence Container that can be placed on any suitable transport and stored. SMPTE ST 379-1 defines the MXF Generic Container as the native Essence Container in MXF

files. SMPTE ST 381-1 defines how MPEG streams, as identified by an ISO 13818-1 stream_id value, can be mapped in the MXF Constrained Generic Container. SMPTE ST 382 defines how AES3 and Broadcast Wave Audio can be mapped in the MXF Generic Container. SMPTE ST 385 defines the System Item that is compatible with SMPTE ST 326 (SDTI-CP) and also defines how SDTI-CP essence and metadata can be used in the MXF Generic Container.

This document specifies the mapping of MPEG Picture (ES) compatible with MPEG stream and AES3 audio into the MXF Generic Container. This document also specifies the MXF file format which includes unique identifiers, Operation Pattern, Partitions, Index Table Segments and RIP. The common basic structure is described in this document

4 Outline of MXF File Structure for this Mapping

Figure 1 shows the outline of the MXF file structure. The file consists of a Header Partition, segmented Body partition(s), a Footer Partition and a Random Index pack. Picture, Sound and System Items are mapped into the Essence Container and placed in each Body Partitions. The Data Item is optional. Because of the MPEG Long GOP structure of Picture Item, segmented Index Table is used together with Random Index pack. More detailed explanation can be found in Section 8 (Application Issues).

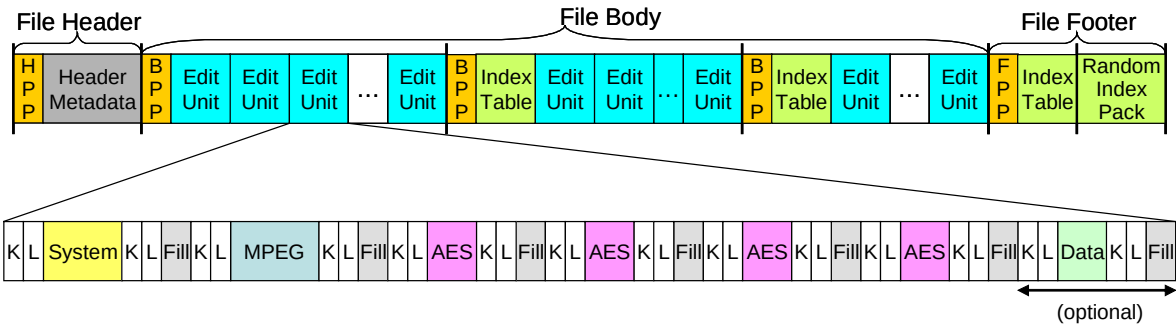


Figure 1 – Outline of MXF File

- **HPP**: Header Partition Pack
- **BPP**: Body Partition Pack
- **FPP**: Footer Partition Pack
- **Fill**²: The length of the KLV Fill Item should be required to align to a KAG boundary.

Some of the aspects of this structure are shown below.

- It is only necessary to include one Index Table Segment for each Body Partition period on the sender side.
- It is easy to perform the function “Play while receiving file” on the receiver side.
- It is easy to pick extract a “Partial file”.
- A list of major constraints for this file structure is given in Table 1.
- Detailed constraints are listed in Annex B.

Note: In Figure 1, the order of Items in the Content Package is System, Picture, Sound and Data Items, as described in other documents (e.g. SMPTE ST 379-1, SMPTE ST 386, SMPTE ST 387, and so on), while SMPTE ST 436 defines Data Item precedes Picture and Sound Items. This Element ordering issue is being discussed in the ST 436:2006 revision group and it is expected to be resolved in its revision.

Table 1 – Constraints for RDD 9 Stream Products

Item	Constraints
Operational Pattern	1a - <i>Origin</i> and <i>Duration</i> ¹ are used to express GOP Pre-Charge and Roll-Out.
Wrapping (Interleaving)	Frame by Frame (coded order)
KAG ² size	512
System Item	Compliant to ST 326 and ST 385, includes the Frame by Frame Timecode and UMID
Video packetization	Compliant to ST 381-1, MPEG-2 ES
GOP structure	Max 15 frames/GOP (N<=15, M<=3, open GOP / closed GOP) // 1920x1080i Max 12 frames/GOP (N<=12, M<=3, open GOP / closed GOP) // 1280x720p Variable length GOPs are permitted.
Audio sampling	48 kHz locked
Audio packetization	Compliant to ST 382, AES3, 1ch/Element (min 2 to max 8 channels)
Data Item (optional)	Compliant to ST 436, Optionally used for Ancillary packet
Timecode	System Item and Header Metadata

5 MPEG Picture Data and AES3 Data Mapping

The mapping of MPEG Picture (ES) data is as defined in SMPTE ST 381-1. The mapping of AES3 digital audio data is defined by SMPTE ST 382. This specification uses Frame Wrapping as defined by SMPTE ST 379-1. The System Item is defined by SMPTE ST 326, and mapped into the MXF by SMPTE ST 385. The order of Items in each Edit Unit is System, Picture, Sound and Data (where the Data Item is optional).

5.1 Frame Wrappings

This document requires the use of Frame Wrapping as defined by SMPTE ST 379-1, Section 5.4.1.

In the case of audio locked to video at 25 (or 50) content packages per second, each Element will contain the same number of samples, for example 1920 (or 960).

In the case of audio locked to video at 30*1000/1001 content packages per second, the number of samples in each Element will vary to maintain a correct aggregate rate. Typically the number of samples varies according to a 5-frame sequence, 1602, 1601, 1602, 1601, and 1602.

In the case of audio locked to video at 60*1000/1001 content packages per second, the number of samples in each Element will vary to maintain a correct aggregate rate. Typically the number of samples varies according to a 5-frame sequence, 801, 801, 800, 801, and 801.

The number of samples in each content package is calculated from the Length field of the surrounding KLV packet, divided by the value of the BlockAlign Property of the AES3 Audio Essence Descriptor.

In the case of audio locked to video at 24*1000/1001 content packages per second, each Element will contain the same number of samples, for example 2002.

An arrangement of System, Picture, and Sound Items in a Frame Wrapping is shown in Figure 2. It shows the case of 4 channels AES3 audio.

¹ In legacy RDD 9 files (version 1.2), the duration of the top level file package does not contain the Pre-Charge or Roll-Out as defined in SMPTE ST 378. In RDD 9 files (version 1.3), the duration of the top level file package contains the Pre-Charge but doesn't contain Roll-Out as defined in SMPTE ST 377-1.

² Refer to Section 8.1 Application of the KLV Fill Item.

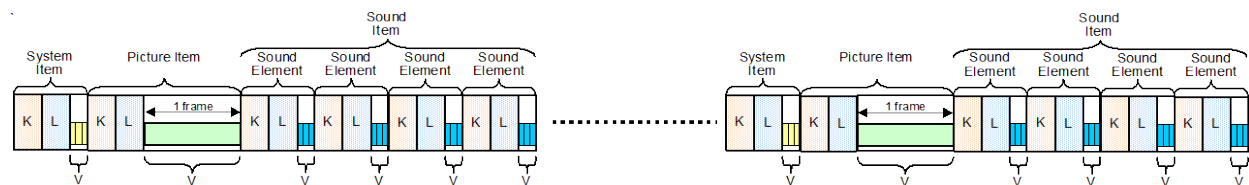


Figure 2 – Frame Wrapping of System, Picture and Sound Elements

5.2 System Item Mapping

The System Item in each Edit Unit consists of System Metadata Pack, a Package Metadata set and Picture Metadata Set.

5.2.1 Overview of System Item

System Item is placed at the beginning of every Edit Unit and contains information on the essence item and the metadata attached to the frames, and it shall comply with SMPTE ST 385 (Mapping SDTI-CP Essence and Metadata into the MXF Generic Container).

Typical System Item consists of the following four Elements and its size is the same as one KAG size (200_h).

- System Metadata Pack contains Package Rate, Multiple EC UL, LTC
- Package Metadata Set contains UMID, Essence Mark (optional)
- Picture Metadata Set contains Acquisition Metadata Sets (optional)
- Fill Item

Figure 3 shows the outline of System Item.

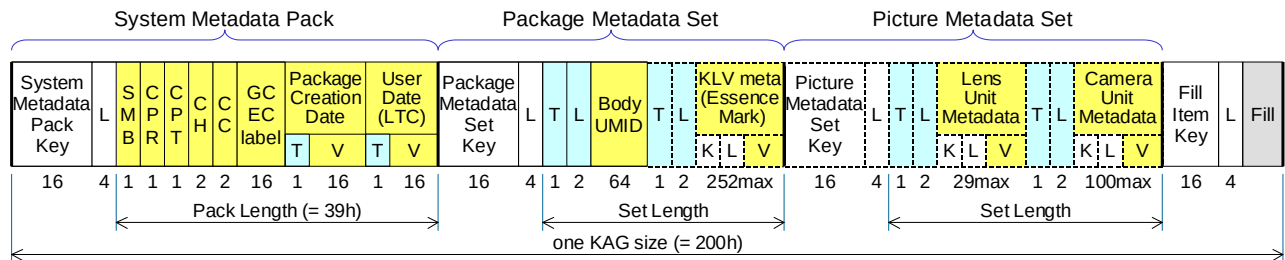


Figure 3 – Typical System Item structure

5.2.2 System Metadata Pack

The Pack Key is 06.0E.2B.34.02.05.01.01.0D.01.03.01.04.01.01.00 as defined in Table 1 of SMPTE ST 385. Length of this pack shall be fixed, i.e. 57-byte payload. Also, each Property shall be described in the provided field without tag and length. The sequence and values shall comply with SMPTE ST 326.

- System Metadata Bitmap ("SMB" in the figure) indicates the presence of metadata in the Pack, and of essence data within the Edit Unit, should be set to 0101_1100_b, when Data Item is not recorded or 0101_1110_b when Data Item is recorded.

- The value of Continuity Count ("CC" in the figure) shall be monotonically increasing within a file. It does not have to start from 0, and is back to 0000_h following full count FFFF_h.
- SMPTE Universal Label ("GC EC label" in the figure) shall be set to a suited MXF Generic Container label which is the same with the Essence Container Property of Multiple Descriptor Set.
- Package Creation Date should be blank. Tag ("T" in the figure) and the remains are filled with 00_h.
- LTC shall be described in the User Date column. Since it complies with SMPTE ST 331, it starts with CP-Tag 81_h and digits of Frame, Second, Minute, and Hour are placed with flags such as DF, and then Binary Group data (4 bytes) is placed, and remaining 8 bytes are filled with 0.

5.2.3 Package Metadata Set

The Set Key is 06.0E.2B.34.02.43.01.01.0D.01.03.01.04.01.02.nn as defined in Table 2 of SMPTE ST 385. The parameter nn indicates the number of Metadata Blocks within the Package Metadata Set, and the value is set to 2 when there are UMID and KLV Metadata, 1 when there is just UMID, and 0 when there is no UMID nor KLV Metadata.

Each metadata block is described with 1-byte CP-Tag and 2-byte Length field. Typical kind of metadata in this specification, shown in Figure 3, is defined as follows:

- Frame-by frame UMID should be described as the first block.
 - Extended UMID (64 bytes) should be described with CP-Tag 83_h.
 - When omitting the block, Tag and later are not described.
 - Decoders should support the cases where only Basic UMID is described or the item is blank.
- EssenceMark™ is normally described as KLV Metadata as the second block.
 - Value is less than 32 bytes, recorded with CP-Tag 88_h.
 - When omitting the item, Tag and later are not described.
- Other kind of KLV Metadata may be present with CP-Tag 88_h instead of EssenceMark™.

If there are no items in Package Metadata Set, the length field shall be zero.

5.2.4 Picture Metadata Set

The Set Key is 06.0E.2B.34.02.43.01.01.0D.01.03.01.04.01.03.nn as defined in Table 2 of SMPTE ST 385. The parameter nn indicates the number of Metadata Blocks within the Picture Metadata Set.

This Set contains Acquisition Metadata Sets specified in SMPTE RDD 18. Those Sets are described as KLV Metadata with 1-byte CP-Tag 88_h and 2-byte Length field.

When the clip does not contain Acquisition Metadata at all, the Picture Metadata Set should be omitted as a whole. When there is nothing to record as Picture Metadata item for the Edit Unit, those later than Set Key may be omitted.

5.3 Picture Item Mapping

As shown in Figure 4, the Picture Item for this mapping specification is Frame Wrapped and the data length of each Picture Element in an Edit Unit has a variable value.

This Element contains the MPEG-2 Video elementary stream, and shall comply with SMPTE ST 381-1 (Mapping MPEG Streams into the MXF Generic Container).

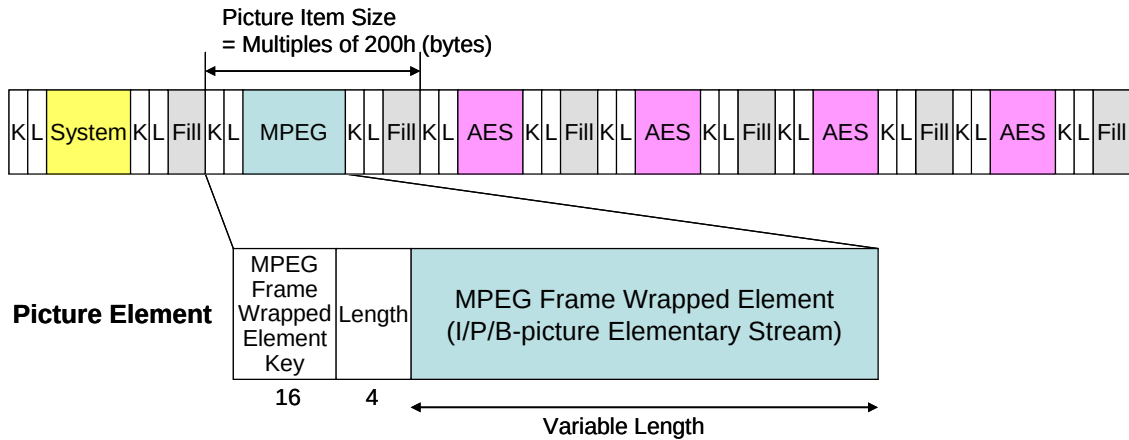


Figure 4 – Mapping of Picture Item in Edit Unit

5.3.1 MPEG Picture Element Key

The values of the first 12 bytes of the Essence Element Key are defined in SMPTE ST 379-1, MXF Generic Container Format. The Picture Element is defined by SMPTE ST 381-1, Mapping MPEG streams into the MXF Generic Container Format. This defines byte 13 with the value of 15h. The values of the last four bytes of the Essence Element Key are given in Table 2.

Table 2 – Key Value for the MPEG Picture Element

Byte No.	Description	Value (hex)	Meaning
1-12	Specified by the MXF Generic Container Specification SMPTE ST 379-1		
13	Item Type Identifier	15h	Picture Item
14	Essence Element Count	01h	Count of Elements in this Item
15	Essence Element Type	05h	Frame Wrapped Picture Element
16	Essence Element Number	00h	Element Unique Number

5.3.1.1 MPEG Picture Element Number — Byte 16

This is a number used as an index to identify this instance of the Element within the Item. It does not change within any instance in the Generic Container.

5.3.2 MPEG Picture Element Length

The length field of the KLV coded Element is 4 bytes BER long-form encoded (i.e. 83h.xx.yy.zz) for Frame wrapping.

5.3.3 MPEG Picture Element Value

The MPEG-2 Picture Element complies with SMPTE ST 381-1. The maximum frame number per GOP is fifteen frames for 1080-line systems and twelve frames for 720-line systems, and both open and closed GOP modes are supported.³

³ Refer to Table 1.

5.4 AES3 Sound Item Mapping

The AES3 Sound Element complies with SMPTE ST 382, Mapping AES3 and Broadcast Wave Audio into the MXF Generic Container, as shown in Figure 5. Sound Item for this mapping specification is Frame Wrapped.

This Element contains Linear-PCM Audio data stream, and shall comply with SMPTE ST 382 (Mapping AES3 and Broadcast Wave Audio into the MXF Generic Container).

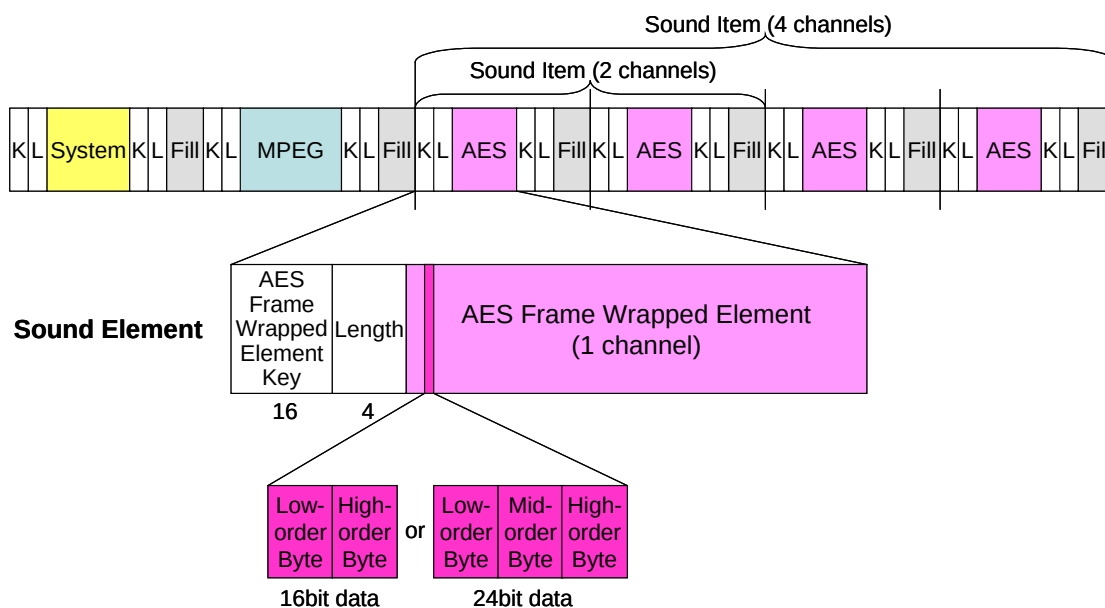


Figure 5 – Mapping of Sound Item

5.4.1 AES3 Sound Element Key

The values of the first 12 bytes of the Essence Element Key are defined in SMPTE ST 379-1, MXF Generic Container Format. The values of the last four bytes of the Essence Element Key are given in Table 3.

Table 3 – Key Value for the AES3 Sound Element

Byte No.	Description	Value (hex)	Meaning
1-12	Specified by the MXF Generic Container Specification SMPTE ST 379-1		
13	Item Type Identifier	16h	GC Sound Item
14	Essence Element Count	02, 04 or 08h	Count of Sound Elements in this Generic Container
15	Essence Element Type	03h	AES Frame Wrapped Element as listed in SMPTE RP 224
16	Essence Element Number	00, 01,02,03, 04,05,06 or 07h	Element Unique Number

5.4.1.1 AES3 Sound Element Count — Byte 14

This is a count of the number of Elements in the Sound Items in the Generic Container. The count of the Sound Elements for this mapping specification may be 2, 4 or 8 according to the count of available audio channels.

5.4.1.2 AES3 Sound Element Number — Byte 16

This is a number used as an index to identify this instance of the Element Type within the Item. Each Element within an Item has a unique value in sequence from 00h upwards as defined by SMPTE ST 379-1, which remain constant throughout any instance of a Generic Container.

5.4.2 AES3 Sound Element Length

The length field of the KLV coded Element is 4 bytes BER long-form encoded (i.e. 83h.xx.yy.zz) for Frame wrapping.

5.4.3 AES3 Sound Element Value

In a Frame Wrapped file, the Start Position of the first sample of AES3 audio data should be the same as the Start Position of the synchronized Picture Element. Note that the order of Frame Wrapped Picture Item of MPEG Picture Element is not display order, but coding order (i.e. I2 B0 B1 P5 B3 B4 P8 B6 B7...). Therefore, Sound and Picture Elements are not necessarily synchronized within each Edit Unit.

5.5 Data Item Mapping (Optional)

The Data Item contains data stream, e.g. closed captions or subtitles, and shall comply with SMPTE ST 436 (MXF Mappings for VBI Lines and Ancillary Data Packets). Figure 6 shows the ANC packets wrapped in a Data Item.

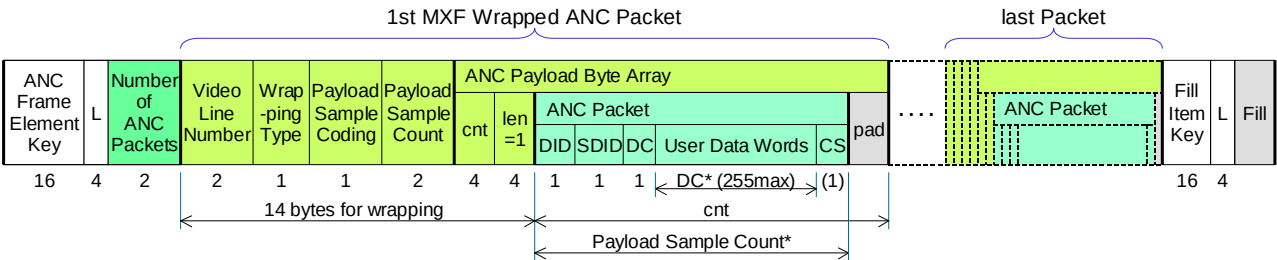


Figure 6 – Mapping of ANC Data in an Edit Unit

The Element Key is 06.0E.2B.34.01.02.01.01.0D.01.03.01.17.01.02.01 as defined in Table 6 of SMPTE ST 436. Essence Element Type is limited to Frame Wrapped ANC Data and Essence Element Count is limited to one.

Note: In SMPTE ST 436, Byte 16 of the Element Key is defined as 01h and indicates the Number (used as an Index) of this Element. RDD-9 specifies the constraint value of 01h irrespective of any revision of SMPTE ST 436.

5.6 Temporal Reordering

In the case of MPEG long GOP video, the compressed video pictures are reordered from their display order according to the MPEG specification. This reordering is applied only to the video Elements⁴.

⁴ Refer to Figure 15 in Section 8.3.5.