
SMPTE STABLE DOCUMENT



The attached SMPTE Engineering Document has been declared “Stable” by the controlling Technology Committee.

The SMPTE Operations Manual for Standards states:

A document should be stabilized if it is believed to be substantially correct, does not contain harmful or misleading recommendations, may still be relevant to equipment or practices in use, is stable, but does not represent current technology, and need not be subject to future reviews.

A Stable document shall still be made available and offered for sale by the Society, but it shall be prefaced by a cover page explaining its current status.

At any time, a Technology Committee may revise, amend, or otherwise initiate a new Project on a Stable document.

A Stable document is “In Force”, and not deprecated or withdrawn.

*** * * * ***

Note:

SMPTE “Stable” documents were previously described as “Archived” and the attached document may be marked as “Archived”. The status of a SMPTE document described as “Archived” is exactly as described above for a “Stable” document.

Stable documents may not adhere to the latest style and format of SMPTE documents, or to current usage of normative language. Suitable care should be taken in interpretation.

SMPTE STANDARD

for Television —

Material Exchange Format (MXF) — MXF Constrained Generic Container



Table of Contents	Page
Foreword	2
Intellectual Property	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	3
4 Glossary of Acronyms, Terms and Data Types	4
5 Overview	4
6 Legacy Considerations (Informative)	5
7 MXF Container Constraints Signaling	5
8 MXF Generic Container Format	6
9 System Item Coding	14
10 Picture, Sound, Data and Compound Item Coding	17
11 SMPTE Essence Container Label	19
Annex A Bibliography (Informative)	20

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE ST 379-2 was prepared by Technology Committee 31FS.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. 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.

1 Scope

This standard specifies numerous refinements based on SMPTE 379-1. This standard specifies the format of the MXF Generic Container (GC). The MXF GC is the native essence container of the material exchange format (MXF) file body. The MXF GC is defined for the interchange of streamable audio-visual material.

This standard defines the data structure at the signal interfaces of networks or storage media. This standard does not define internal storage formats for MXF compliant devices.

Appropriate essence and metadata payloads that can be mapped into the MXF GC are defined in associated documents

The MXF specification includes operation pattern specifications that may define restrictions on the way in which this essence container type should be implemented. The reader is advised to carefully study the appropriate operational pattern document for compliance to a defined implementation.

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; followed by 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 standard. At the time of publication, the editions indicated were valid. All standards are subject to revision and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent of the standards indicated below.

SMPTE 331M-2004, Television — Element and Metadata Definitions for the SDTI-CP

SMPTE 336M-2007, Data Coding Protocol Using Key-Length-Value

SMPTE 377-1-2009, Material Exchange Format (MXF) — File Format Specification

4 Glossary of Acronyms, Terms and Data Types

The general glossary of acronyms, terms and data types used in the MXF specification is given in SMPTE 377-1. Most definitions are not repeated here to minimize any divergence of meaning.

4.1 Acronyms Used in this Standard

CI: Content Item – a unit that consists of one or more Content Elements.

CE: Content Element – a single KLV element of any kind belonging to a Content Item.

CP: Content Package – a generic term for a grouping of some combination of System, Picture, Sound, Data and / or Compound Items.

GC: Generic container.

GOP: Group of pictures, as used in ISO/IEC 13818-2.

KLV: Key-Length-Value, defined in SMPTE 336M.

4.2 Terms Used in this Standard

Picture Essence: A general term for all types of Picture Essence including video, still images, graphics, etc.

Sound Essence: A general term for all types of Sound Essence including audio, MIDI, sampled data, etc.

Data Essence: A general term for all types of Data Essence including teletext, closed caption data, etc.

Compound Essence: A general term for essence that contains an indivisible mixture of different essence types.

Essence Element: The entire essence stream of a single track wrapped by one or more Content Elements, depending on the wrapping type.

5 Overview

The MXF Generic Container (GC) is a streamable data container that can be placed on any suitable transport and potentially stored. The concept of this container was based on the work done by the EBU/SMPTE Task Force in the Wrappers and Metadata sub-group. The MXF GC defined in this standard is fully compatible with the work of the EBU/SMPTE Task Force Report.

The MXF GC format is intended for inclusion into a MXF (Material eXchange Format) file as an essence container. This standard defines the MXF GC for use in an MXF file body.

Notes:

1 A streamable data container is designed to allow the audio-visual material to be continuously decoded through mechanisms such as interleaving essence components with stream-based metadata.

2 The Task Force report defines: "Content is composed of Content Packages, which in turn are composed of Content Items, which are further composed of Content Elements". These Content Packages are convenient groupings of the various Items where each Item is a group of similar element types. Although the term "Content Package" is also used in the SDTI-CP specification, the GC Content Package is a more generalized arrangement that retains backwards compatibility with the SDTI-CP Content Package.

The MXF GC comprises a contiguous sequence of Content Packages, each of which has up to five basic components known generically as Content Items, and more specifically known as one of the following:

- A System Item is a group of up to 127 metadata or Control Data Elements. The metadata contained in the System Item can include local links which associate any metadata item uniquely with its corresponding Content Element. In many cases, metadata can be embedded into each Content

Element. In the case of MPEG-2, the metadata can be embedded in the various headers of the MPEG-2 essence bitstream). The metadata link from the System Item to a Content Element provides metadata in addition to the essence bitstream. The system metadata can be a partial or whole extraction of embedded metadata extracted at the data packing process to provide quick access to key metadata without a requirement to re-parse the essence bitstream. The metadata can also be temporally sensitive metadata such as time-code information or camera coordinates.

- A Picture Item is a group of up to 127 Picture Content Elements. Each Content Element in a Picture Item contains a predominance of picture essence although the element could contain metadata and other ancillary essence.
- A Sound Item is a group of up to 127 Sound Content Elements. Each Content Element in a Sound Item contains a predominance of sound essence although the element could contain metadata and other ancillary essence.
- A Data Item is a group of up to 127 Data Content Elements. Each Content Element in a Data Item contains a predominance of data essence although the element could contain metadata and other ancillary essence.
- A Compound Item is a group of up to 127 Compound Content Elements. Each Content Element in a Compound Item contains a mixture of essentially indivisible essence and metadata components that, as a group, do not match the intent of the Picture, Sound or Data Items.

Picture, Sound and Compound Items are essentially carrying the primary Video and Audio Elements that are often routed to specialist storage or processing equipment. The Data Item is used to carry data-centric elements such as sub-titles and teletext data and is frequently created, processed and stored on computer media. The System Item provides services for each Content Package (CP) through Metadata Elements such as time stamps, metadata for Content Elements in the other Items and, optionally, downstream Control Data Elements.

The premise for the MXF GC format is that of a general purpose Essence and Metadata container for the containment of many different kinds of Essence and Metadata Elements into a single entity by interleaving the data streams in a defined and time-synchronous manner (typically over a 1-frame duration). Associated SMPTE GC mapping documents define the Essence and Metadata Elements that can be placed in the container. Some SMPTE GC mapping documents could define complete mappings for an entire Content Package while others could simply define mapping of Metadata or Essence into an element.

The MXF GC defined by this document complies with the requirements for essence containers defined in the MXF file format specification.

6 Legacy Considerations (Informative)

SMPTE ST 379-2 is intended to be a proper subset of SMPTE 379-1. In some cases, the scope of SMPTE 379-1 allows greater latitude by leaving constraints unspecified. This version is generally consistent with SMPTE 379-1, but includes additional constraints with the objective of permitting simpler implementation of compliant decoders.

7 MXF Container Constraints Signaling

Encoders that conform to this specification shall add a `ContainerConstraintSubDescriptor` to the `GenericDescriptor::SubDescriptors` property of the top-most File Descriptor that describes the essence container.

The GenericDescriptor::SubDescriptors is defined by SMPTE 377-1, Annex B.2:

Element Name	Type	Len	Local Tag	UL Designator	Req ?	Element Description
All elements from the Generic Descriptor defined in SMPTE 377M-1, Annex B.2						
Sub Descriptors	StrongRefArray (Sub Descriptors)	8+16 n	Dyn	06.0e.2b.34 01.01.01.09 06.01.01.04 06.10.00.00	Opt	Ordered array of strong references to sub descriptor sets

The set Key of the ContainerConstraintsSubDescriptor shall be as defined in the table below.

Byte No.	Description	Value	Meaning
1~13	As defined in SMPTE 377-1, Table 18		Values for all MXF structural metadata sets
14~15	Set Kind	01.67h	Defines the Key value for the Container Constraints Sub Descriptor
16	Reserved	00h	Reserved value

The Container ConstraintsSubDescriptor does not define any required properties in addition to those inherited from the Generic SubDescriptor.

Item Name	Type	Len	Local Tag	Item UL	Req ?	Meaning
Container Constraints SubDescriptor	Group UL	16		As defined in SMPTE 377-1, Annex B.3	Req	Defines the Container Constraints SubDescriptor Class
All items in SMPTE 377-1, Annex B.3 except the Key or Group UL and the Length, if present	See SMPTE 377-1, Annex B.3					

8 MXF Generic Container Format

The MXF Generic Container (GC) comprises a contiguous sequence of one or more Content Packages (CP) as illustrated in Figure 1. The CPs may be of constant or variable length depending on the application. The example in Figure 1 shows a GC with Content Packages of variable length.

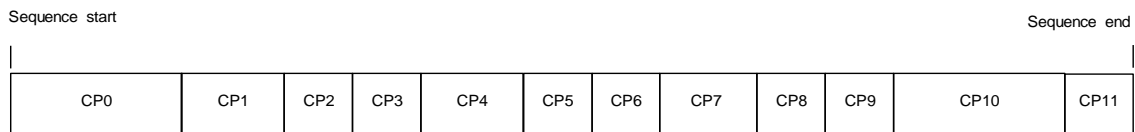


Figure 1 – GC as a contiguous sequence of CPs

8.1 Content Package Definition

A Content Package (CP) shall consist of a collection of one or more Content Items (CI). A CI shall consist of one or more Content Elements (CE). A CE shall consist of zero or more pieces of the Essence Element, one Metadata or System Element or one of any other kind of element. A CP may include Metadata directly related to the CP itself, or used to associate the component parts (CIs and CEs). A CP may be a complex audiovisual sequence video, audio, subtitles and description; or it could be as simple as a single frame of video with or without the associated audio.

CP instances support mutual synchronization of multiple CIs. In the context of this document, CP instances shall be encoded as an ordered sequence of KLV packets of CEs and, optionally, a KLV Fill item to form the byte stream of the essence container. Depending on the duration of the CP instances and the number of CP instances in the essence container, this document distinguishes three different encoded CP representations: frame wrapped, clip wrapped and custom wrapped.

In frame wrapped CP representations, the duration of a CP shall equal the duration of a frame or field. If no Picture Item is present in the CP, and no Compound Item containing a picture is present in the CP, then a constant duration shall be assumed. In clip wrapped CP representations, the duration of the CP shall equal the total duration of the material in the essence container. If the total duration of the material equals one picture, the CP representation shall be identified as frame-wrapped.

All other wrappings shall be custom wrappings. These are further distinguished as sub-frame custom wrapping and super-frame custom wrapping. In sub-frame custom wrapping, a frame shall consist of multiple CP instances; for example, one KLV packet per MPEG slice, or one per MPEG-coded video field. In super-frame custom wrappings, a CP instance shall contain multiple frames and the essence container shall consist of one or more CPs. An example is an essence container which has one KLV packet per GOP in long GOP MPEG encodings.

8.2 Content Items and Elements

Each Content Package (CP) shall hold the Content Items (CI), which in turn holds the Content Elements (CE) which, in turn holds the Essence Elements and Metadata Elements. The CEs are interleaved over a defined duration (typically 1 picture in frame wrapped containers), which shall be the primary timebase. Logically, the CP shall be constructed of up to five Content Items (CI) which are the System, Picture, Sound, Data and Compound Items. All Items in the MXF generic package are optional and their presence shall depend on the requirements of the associated SMPTE GC mapping document.

Each CI shall consist of up to 127 CE, each of which may be followed by a KLV Fill item. The System Item in any CP may carry up to 127 metadata or control data CE. These system CEs are also referred to as System Elements. The Picture, Sound, Data and Compound Items in any CP may each comprise up to 127 CE. These CEs contain zero or more pieces of the Essence Element, depending on the essence type and the CP duration. It is a key feature that each CP shall contain the associated contents of one defined duration. Figure 2 illustrates the logical structure of a CP.

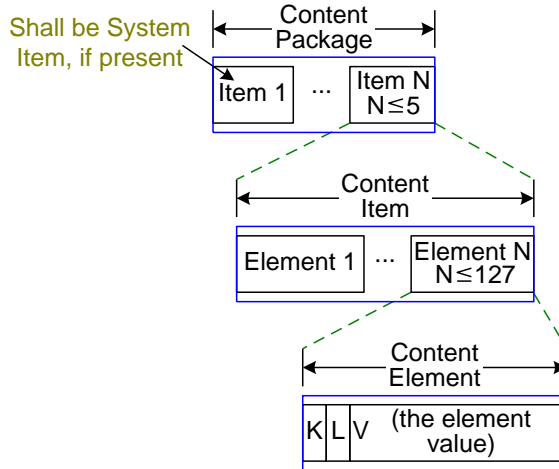


Figure 2 – Logical structure of Items and Elements in a CP

Note: There exists no marker that indicates the start of a CP or CI. The start of a GC and the first CP and the first CI and the first CE is inferred by the first occurrence of an Essence Element Key or a System Element Key. These keys are the only physical markers in the GC.

8.3 Content Package Structure

There shall be zero or one instance of each type of Content Item (CI) in any one Content Package (CP); e.g., you may not have two Sound Items in one CP. If a System Item is present, then it shall be the first CI of every CP. The package may be completed with any other CI as needed, in any order. If a System Item is not present, then the CP may start with any CI.

Within each CI, all CEs that belong to the same CI shall occupy adjacent KLV packets, between which there may only be optional KLV Fill item but shall be no other KLV items.

CEs are comprised of a key, a length and a value. For a System Element or Metadata Element, the key shall be constructed as defined in Section 9.2.1 and the value shall be one System Element or Metadata Element as defined in Section 9.2.2. For an Essence Element, the key shall be constructed as defined in Section 10.110.1 and the value shall be zero or more pieces if the Essence Element as defined by the Content Package wrapping in Section 8.4.

Encoders shall ensure that the order and presence of CI, CE and KLV Fill item is identical for all CP of the essence container.

Note: Given the guarantee that a System Item, if present, must be the first item, this can be used to identify the start of any CP. If no System Item is present, the guarantee of unchanging order and presence of every element and item provides a similar mechanism. In this case, it is possible to achieve synchronization by discovering the Essence Element key of the first CE at the beginning of the essence container. This key can be used to identify the start of all other CPs.

8.4 Content Package Wrappings

The MXF Generic Container (GC) offers three forms of essence wrapping; Frame Wrapping, Clip Wrapping and Custom Wrapping. Each SMPTE GC mapping document shall identify which wrapping is enabled. Each enabled wrapping type shall be specified by providing two things:

- a) A unique Essence Element Key for each enabled wrapping type. This is done with a unique value for byte 15 in the key, as described in Section 10.1.

- b) A unique Essence Container Label for each enabled wrapping type. This is done with a unique label, as described in Section 11.

This absence of an enabled wrapping type in one SMPTE GC mapping document shall not prevent other SMPTE GC mapping documents from enabling it.

8.4.1 Frame Wrapping

With Frame Wrapping, there shall be one or more Content Packages (CP) in the essence container. If there is only one CP, it shall represent the contents of a single frame or field. The duration of every CP shall be constant and shall match the Edit Unit duration (1/Edit Rate) of the related File Package tracks as defined in 377-1. This duration shall be the primary timebase.

This duration may be a video frame, a video field, a motion picture frame or any other value that represents the basic unit of duration of the primary essence component.

If no Picture Item is present in the CP, and no Compound item containing a picture is present in the CP, then a suitable nominally constant duration shall be used to define the CP duration. This CP duration shall match the Edit Unit duration as above, and shall consist of an integer multiple (zero or more) of indivisible pieces of the Essence Element. This grouping may be, for example, an audio frame (as a 192-sample block), a compressed audio packet, or a set of audio samples covering a duration equivalent to some picture duration.

An example Arrangement of Frame Wrapped system, picture, sound and Data Items is shown in Figure 3.

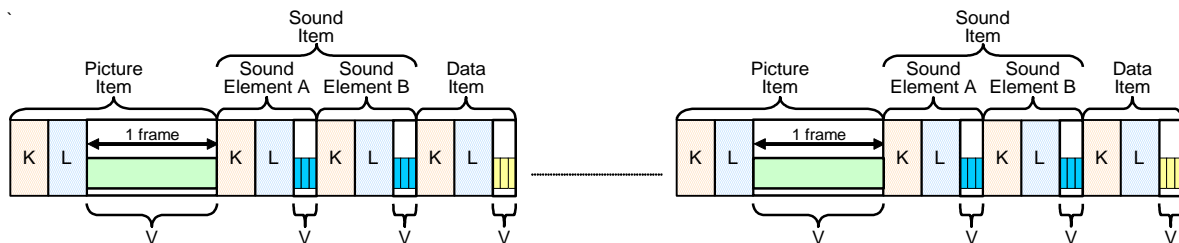


Figure 3 – Example of Frame Wrapped items and elements in the GC

Note 1: The CP and Content Item (CI) durations are logically constant and equal, but it is possible that the Content Element (CE) duration varies slightly, such as 48kHz audio wrapped in a container where the CP duration is 30000/1001. Every CP and CI duration is equal, but the duration of the audio CE might vary in a 5-frame cycle: 1602, 1601, 1602, 1601, 1602 samples.

Note 2: If the rate of indivisible sound packets, such as compressed audio, is greater than the rate of Picture packets, a possible solution to handle this would periodically place multiple CE with the same key in one CI. This standard does not support this approach and prescribes that multiple CE from the same track are combined into a single CE in each CP. Likewise, if the rate of sound packets is less than the rate of Picture packets, a possible solution to handle this would periodically omit one CE from the CI. This standard does not support this approach. This standard requires that empty CE be used to replace missing elements. Provisions for handling multiple or missing elements are described in Section 8.5.1.

Note 3: An example arrangement of the System Item followed by one CE in the Picture Item, two CE in the Sound Item and one Data Element in the Data Item is shown in Figure 4. Another valid arrangement is to replace the Picture Item and Video Essence Element in Figure 4 with a Compound Item and Compound Element.

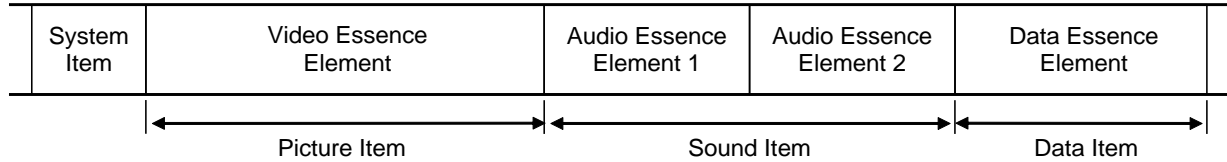


Figure 4 – Example arrangement of System Item, Video, Audio and Data Essence Elements in a Content Package

8.4.2 Clip Wrapping

A Clip Wrapping is where the entire duration of the essence container shall be contained within a single Content Package (CP). The CP duration shall be equal to the duration of the related File Package tracks as defined in 377-1, and shall be an integer multiple of indivisible Essence Elements. The duration of the CP shall comply with Section 8.1. That is, if the CP contains Picture Elements or Compound Elements containing Picture Essence, the value of duration shall be two or more. If the CP does not contain Picture Elements or Compound Elements containing Picture Essence, the value of duration shall be one or more.

Encoders producing Clip Wrapped essence shall produce only mono-essence streams. The structure of the container shall be a single CP, containing a single CI, containing a single CE, comprised of a single KLV, with the value (V) comprised of an integer multiple of indivisible Essence Elements. The duration of this CP is as defined above. For definitions elsewhere in this document, the primary timebase for the single CP shall be defined as the Edit Unit duration (1/Edit Rate). An example arrangement of edit units in a mono-essence, Clip Wrapping is shown in Figure 5.

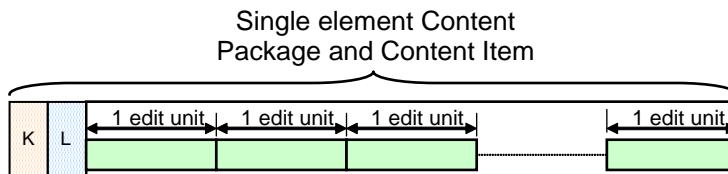


Figure 5 – Example of Clip Wrapping mono-essence CE in the GC

8.4.3 Custom Wrappings

All wrapping methods that do not meet the requirements for either Frame or Clip Wrapping as specified in Sections 8.4.1 and 8.4.2 shall be identified as Custom Wrapping. Specifically, this includes sub-frame and super-frame wrapping as specified in Section 8.1. SMPTE Generic Container (GC) mapping documents that enable Custom Wrapping shall specify the customization details. This shall include, but is not limited to:

- The primary timebase for the duration of every CP.
- Whether the duration of each CE is constant for the entire GC.
- The structure of each CE and how the essence maps into it.
- Whether indexing is possible under SMPTE 377-1 rules and how those rules apply to the CE described in the SMPTE GC mapping document.

Custom wrappings shall follow all other normative provisions of this standard including those provisions regarding the CP structure, the KLV coding structure, the element key values and the essence container label.

Note: When custom wrappings are used, it might be difficult or impossible to construct Index Tables that comply with SMPTE 377-1, Section 11. Index Tables have a fixed value of their index edit rate. If the custom wrapping is super- or sub-frame-wrapping there might be a simple relationship between index edit rate and CE edit rate, in which case Index Tables could be easily constructed and interpreted. Custom wrappings that use a varying CE edit rate are not easily mapped to a fixed index edit rate.

8.4.4 Index Semantics

SMPTE 377-1 Index Table Segments can specify a stream offset for every CP. Within the CP, it can specify the offset of some or all CE. For the wrapping types specified in this document, the meaning of stream offsets shall be as show in Table 1.

Table 1 – Wrapping Types

Frame Wrapped	The stream shall begin at the first byte of the first KLV in the Container. For every indexed CE, the Stream offset specified by the SMPTE 377-1 Index Table Segment shall correspond with the first byte of the CE's KLV.
Clip Wrapped	The stream shall begin at the first byte of the V within the KLV of the first KLV in the Container. If the Clip Wrapping consists of multiple tracks, the essence within the second and following KLV elements should not be indexed.
Custom Wrapped	The stream may begin at the first byte of the first KLV in the Container or the first byte of the V within the KLV. That is, the byte indexed by the SMPTE 377-1 index table shall be defined by the SMPTE document that defines the Custom Wrapping. Note: If the duration of each CE varies, it can be impossible to construct an index table segment.

8.5 MXF Encoder Rules

The Generic Container (GC) is intended to wrap a wide variety of essence types. Individual SMPTE GC mapping documents can constrain MXF encoder behavior so that consistent MXF files are created. This section provides overall rules which shall be followed. Other GC documents can further constrain encoder behaviour. The rules are intended to aid interoperability by ensuring MXF encoder implementations create consistent MXF files.

1. Each Content Package (CP) shall have a constant number of Content Element (CE).
2. The order of CE in the CP shall not change.
3. Logically, every Content Item (CI) in every CP shall have identical duration.
4. Every CE in the CP should have the same nominal duration, although, for example, some complex audio codings may have indivisible units whose actual duration is not the same as the actual picture duration.
5. Each CP should have one CE which is the primary timebase (usually the video).
6. Each CP should have a duration which is an integer multiple of the atomic size of the underlying essence of the primary timebase (video frame, audio frame, audio block etc.).
7. Synchronized CE should be grouped in the same CP. This guideline has two parts: one for intra-coded essence, and one for reordered inter-coded essence, such as long-GOP MPEG.

- 7.a *Intra-coded essence*: if the presentation order of a CE is the same as storage order, such essence shall be considered as intra coded and not reordered. All synchronized intra-coded element shall be grouped in the same CP.
- 7.b *Inter-coded essence*: If presentation order is not the same as storage order, such essence shall be considered to be reordered. CE Elements that are reordered due to inter-coding, such as Long GOP MPEG, may result in synchronised CE Elements being physically located in different CPs. With this in mind, the rule for CE grouping applies to CE as if they were stored in presentation order. The presentation order can be discovered through the reordering mechanism within the index table specification of SMPTE 377-1. Essence such as MPEG can also be directly examined to discover the presentation order.

8.5.1 Multiple or Missing Elements

The GC may be used for some interleaved streams where the primary timebase element duration (e.g., a video frame) is not the same as the other elements (e.g., compressed audio frames). The MXF encoder behavior above may lead to a situation where there are multiple indivisible elements from the same track in a Content Package (CP). For example, the time duration of a 29.97-Hz video frame is 33 ms, whereas the duration of an indivisible, compressed audio frame can be 32 ms. Most CPs would contain one compressed audio frame and one video frame, and on average, one CP in every 32 would contain two compressed audio frames and one video frame in order to maintain audio/video synchronization.

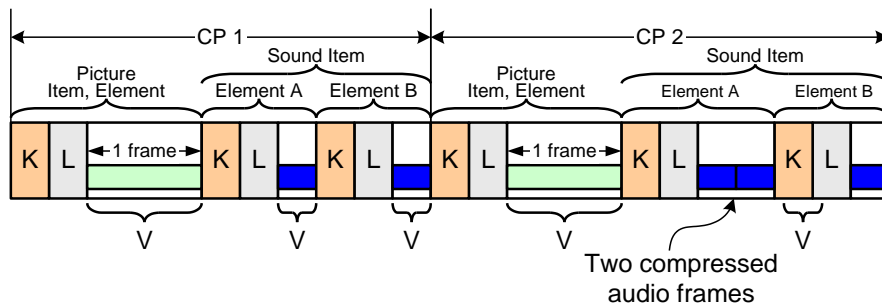


Figure 6 – An element with two, large, indivisible, compressed audio frames

Where more than one indivisible element from the same track is needed in a single CP, MXF encoder implementations shall combine multiple elements into a single KLV element. Figure 6 shows an example where the second occurrence of CE (A) has included two compressed audio frames in order to meet this interleaving constraint.

Note: Any Index Table for this audio CE is required by SMPTE 377-1 to reflect the fact that the number of bytes per CP is not constant.

A similar situation arises when the picture rate is greater than the sound rate thus leaving some CPs without a Sound Element for one of the tracks. In this case, MXF encoder implementations shall replace the missing element with a KLV element with a size of zero and same key as the missing element. This is illustrated in Figure 7 with two CPs, the first with a non-empty Sound Element, and the second with an empty Sound Element. These two rules preserve the number of elements, as specified in Sections 8.3 and 8.5.

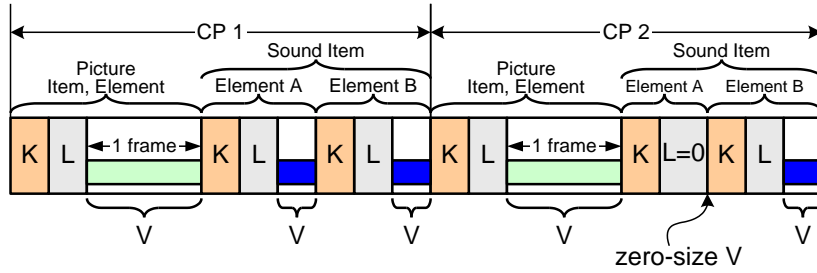


Figure 7 – A zero-size element when essence is periodically missing

The picture in Figure 8 shows three Sound Elements in a CP. Sound Elements A1 and A2 are from the same sound track, and Sound Element B is from a different sound track. Sound Elements A1 and A2 have identical keys, and therefore this arrangement shall not be allowed.

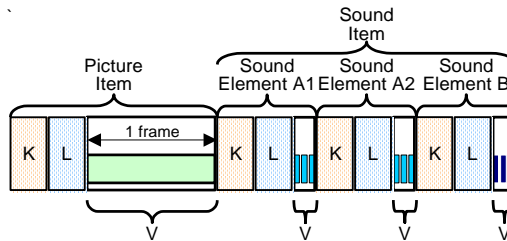


Figure 8 – Illegal use of multiple elements from the same track

Note: The rules exclude this behavior because containers that oscillate between one and two identical elements introduce potential problems:

- Changing element order violates the rules in Sections 8.3 and 8.5 for unchanging element order and presence.
- Index tables entries cannot be constructed to index a position in a CP where element presence changes, or any point past that.
- Decoders might inadvertently miss the second and following elements with the same key.

8.5.2 Behavior for Adding or Removing Content Elements After a Recording Begins

When MXF is used to record multiple essence streams, it is possible to construct a system that can accept new essence streams after a recording begins. Similarly, it is possible that one of the essence streams terminates before others. Some of the operational patterns have constraints that prevent streams from being added or terminated at will. For example, only higher operational patterns allow tracks to begin and end on different boundaries. Also, the construction of index table segments with slice offsets is only possible when the number and order of indexed elements does not change.

Any time a new essence stream begins or ends independent of other streams, encoders shall restructure the file to match the constraints in Section 8.3 and 8.5, and this may dictate a selection of a different operational pattern.

8.6 KLV Coding Structure

The System elements, together with all picture, sound, data and Compound Essence Elements, shall each be coded using the key-length-value (KLV) coding protocol according to SMPTE 336M.

KLV coding allows a decoder to identify each component by its 16-byte Universal label 'key' and skip any component it cannot recognize using the 'length' value to continue decoding data types with recognized 'key' values.

The general data structure of each KLV packet is shown diagrammatically in Figure 9.



Figure 9 – Data structure of each KLV packet

SMPTE 336M defines the structure of the Key value and the options for the format of the length field. Section 9 defines the KLV coding of the Metadata Elements in the System Item, while section 10 defines the KLV coding of Essence Elements in the Picture, Sound, Data and Compound Items.

8.6.1 KLV length field

The KLV length field and its application shall comply with SMPTE 377-1. Essence container specifications may define further constraints.

9 System Item Coding

The System Item may contain Metadata Elements which describe the operation of the Content Package in various modes and may provide Metadata Items related to the whole package. It may include Metadata Elements linked to Essence Elements in the Picture, Sound, Data and Compound Items. The System Item may include optional downstream control elements. This section defines the details of the System Item coding and format.

9.1 System Item Components

The System Item shall comprise a contiguous sequence of up to 127 KLV packets where each packet comprises Metadata Element or Control Data Elements for support of different aspects of the Content Package. Each packet may be coded as a fixed-length (FL) pack, a variable-length (VL) pack or a local set according to SMPTE 336M.

Figure 10 illustrates the System Item data structure.

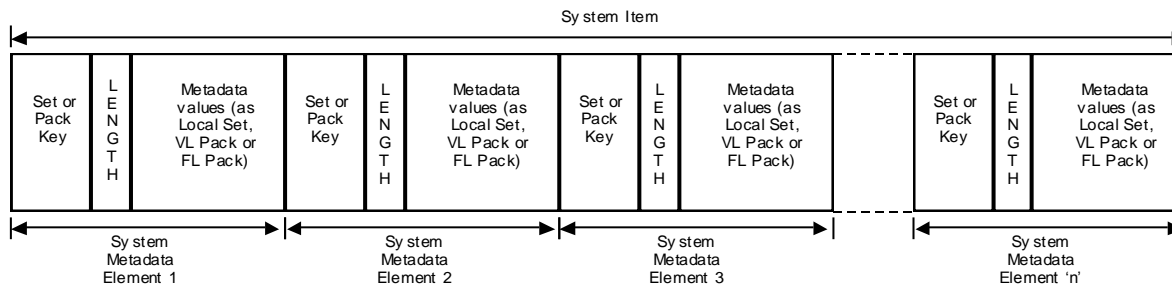


Figure 10 – System Item as a sequence of Metadata Elements

9.2 System Item Metadata Element Definitions

9.2.1 Pack and set keys

The key of a System Item Metadata Element shall be as defined in Table 2.

Table 2 – Specification of the set or pack key for a System Item Metadata Element

Byte No.	Description	Value (hex)	Meaning
1	Object Identifier	06h	
2	Label size	0Eh	
3	Designator	2Bh	ISO, ORG
4	Designator	34h	SMPTE
5	Registry Category Designator	02h	Sets and packs
6	Registry Designator	xxh (See SMPTE 336M)	Fixed-length Pack, Variable-length Pack or Local Set as required
7	Structure Designator	01h	Sets and Packs Registry
8	Version Number	01h	Fixed Version
9	Item Designator	0Dh	Organisationally Registered
10	Organisation	01h	AAF Association
11	Application	03h	MXF Generic Container Keys
12	Structure Version	01h	MXF-GC Version 1
13	Item Type Identifier	04h 14h	SDTI-CP Compatible System Item (see SMPTE 326M) GC-Compatible System Item
14	System Scheme Identifier	Xxh	Defined in other SMPTE documents
15	Metadata or Control Element Identifier	Yyh	Defined in other SMPTE documents
16	Reserved for use by Metadata Element	Zzh	

The choice of fixed-length pack, variable-length pack or local set coding shall be defined by an associated Metadata Element or Control Data Element specification.

Where a System Item is present in the Content Package, the first Metadata Element shall set the Metadata Element identifier (byte 15) to '01h'. No other System Item element shall precede this element.

The System Element key shall be unique within the scope of the Generic Container (GC) in which it is used and shall be constant for the duration of the GC.

Note: This allows decoders to identify an unambiguous starting point of a Content Package.

9.2.2 System Element value

The value of a System Item Metadata Element or Control Data Element shall be a sequence of metadata properties coded as a local set, a variable-length pack or a fixed-length pack as defined by byte 6 of the key value. The definition of the properties in a Metadata Element or Control Data Element shall be found in other SMPTE documents.

The value of a System Element may use a track number as a link to another System Element (see Section 9.2.3, bullet 4) or an Essence Element (see Section 10.3), both of which shall be in the same content package. If a System Element specifies a linkage to an Essence Element, this four-byte value shall be used to specify the linkage.

9.2.3 Requirements for the definition of a System Element

If the System Item is required for a SMPTE GC mapping document, then the document shall specify the following:

- Bytes 13 to 16 of the System Element key value (see Table 2);
- Either the definition of the System Element value or the reference document where the definition of the element value can be found;
- If the System Element specifies an explicit linkage to an Essence Element in the Content Package, then that linkage shall use the four-byte value described in Section 10.3.
- If the System Element specifies an explicit linkage to another System Element in the Content Package, then that linkage shall use the four-byte value described in Section 9.3.

9.3 Relationship between an MXF Header Metadata Track and a System Element

SMPTE 377-1 defines a mechanism that links Header Metadata tracks to Content Elements (CE). Each System Element may be described by a metadata track in a MXF Header Metadata package. The track number shall be a UInt32 value comprising bytes: 'A.B.C.D' (most significant byte first). These bytes shall have the same values as bytes 13-16 of the key used to KLV wrap the System Element in the GC. These byte values shall be assigned to the System Element key value as follows:

A = Byte 13 of the element key value;

B = Byte 14 of the element key value;

C = Byte 15 of the element key value;

D = Byte 16 of the element key value.

This technique ensures that each track referenced by the Header Metadata package has a unique number for each System Element key value.

The element key shall be unique and shall be constant for the duration of the Generic Container (GC). To identify the correct GC in which to resolve the track number item, the BodySID mechanism shall be used as detailed in the SMPTE 377-1-2009. This mechanism shall be used for all MXF GC mappings.

This 4-byte GC-unique value should be used to convey linkage to a System Element. It is possible to use fewer than 4-bytes to create a link to a System Element; however, the document that defines such a link shall also define a way to disambiguate the missing bytes. For example, if the document describing the link specifies that the System Element is SDTI-CP compatible, then byte 13 shall have a known value defined above, and the link value could use three bytes: 14, 15 and 16.

Note 1: The System Item is frequently composed of Metadata Elements associated with Essence Element tracks in the same GC. But the same mechanism can be used to link Header Metadata Descriptors to a System Element in the GC. But examples do exist such as the need for a Header Metadata Descriptor to define the kind

of timecode used in a System Timecode Element, or whether an extended System UMID Element contains geo-spatial information. An example use can be found in SMPTE 385M, Section 5.3.1.

Note 2: The method of linking a System Element to an associated Essence Track within a given MXF GC is not defined in this standard. Other MXF standards define the specific linking method used.

10 Picture, Sound, Data and Compound Item Coding

A valid Generic Container (GC) shall be comprised of one or more Content Packages (CP) as defined by Section 8.1. A valid CP is comprised of one or more Content Items (CI) as defined by Section 8.2. A valid CI is comprised of one or more Content Elements (CE) as defined by Sections 8.3 and 8.4. A valid CE payload is one that is standardized and normatively referenced by the respective MXF essence container specification. Where picture, sound, data and Compound Items are present in a CP, each CE shall be coded as a single KLV item according to SMPTE 336M. Each KLV coded element starts with a 16-byte element key value to identify the type of element and the Item to which the element belongs, followed by a length field and completed by the element value itself.

10.1 Essence Element Key

The Essence Element key is defined in Table 3.

Table 3 – Key values for Picture, Sound, Data and Compound Elements

Byte No.	Description	Value (hex)	Meaning
1	Object Identifier	06h	
2	Label size	0Eh	
3	Designator	2Bh	ISO, ORG
4	Designator	34h	SMPTE
5	Registry Category Designator	01h	Dictionaries
6	Registry Designator	02h	Essence Dictionary
7	Structure Designator	01h	Dictionary Standard
8	Version Number	01h	Fixed Version
9	Item Designator	0Dh	Organisationally Registered
10	Organisation	01h	AAF Association
11	Application	03h	MXF Generic Container Keys
12	Structure Version	01h	MXF-GC Version 1
13	Item Type Identifier	05h 06h 07h 15h 16h 17h 18h	SDTI-CP Picture (SMPTE 326M) SDTI-CP Sound (SMPTE 326M) SDTI-CP Data (SMPTE 326M) GC Picture GC Sound GC Data GC Compound
14	Essence Element Count	Xxh	See below
15	Essence Element Type	yyh	See below
16	Essence Element Number	Zzh	See below

Byte 13 of the key value identifies the Item to which the element belongs and the correct item value shall be entered.

Values for byte 13 of '05h', '06h' and '07h' shall be reserved for Essence Elements defined in SMPTE 331M.

Note: Item type '07h' is known in SMPTE 326M as an auxiliary item, but is identical to the Data Item of the MXF Generic Container.

Values of '15h', '16h', '17h' and '18h' shall be reserved for Essence Elements defined in other SMPTE MXF documents which specifically define essence mappings for the MXF Generic Container (GC).

Byte 14 of the key value shall be used to define the number of Essence Elements in this Item of the Content Package. A single Essence Element within an Item shall result in an Essence Element count value of '01'h. For a given Essence Element this value shall be constant within the entire GC (even when new elements are added). This is to maintain track linking as detailed in Section 10.3.

Note: This byte is the same value as the Item Header (which is the Essence Element count limited to the range 01h~7Fh) in SMPTE 326M.

Byte 15 shall be the value of the element type as defined by either SMPTE 331M or an associated SMPTE GC mapping document. Element type values shall be constrained to the range 01h~7Fh, and for each value of byte 13, each value of byte 15 shall be assigned by no more than one SMPTE GC mapping document. For a given Essence Element, this value shall be constant within the entire GC (even when new elements are added). This is to maintain track linking as detailed in Section 10.3.

Byte 16 shall be used to define the value of the element number in the range 00h~7Fh. It shall be set by the encoder to be unique amongst the elements in any one Item. The element number shall be different for each new Essence Element. For example, the difference may be achieved by incrementing each element number by one for each new Essence Element in sequence within an item. For a given Essence Element, this value shall be constant within the entire GC (even when new elements are added). This is to maintain track linking as detailed in Section 10.3.

The Essence Element key shall be unique within the scope of the GC in which it is used and shall be constant for the duration of the GC.

10.2 Essence Element value

The picture, sound, data or Compound Element value is as defined in an associated SMPTE GC mapping document.

10.2.1 Picture, Sound, Data and Compound Item status

If an Essence Element is required for a SMPTE GC mapping document, then the document shall specify the following:

- Bytes 13 to 16 of the Essence Element key value (see Table 3);
- Either the definition of Essence Element value or the reference document where the definition of the Essence Element value can be found.

10.3 Relationship between an MXF Header Metadata Essence Track and an Essence Element

SMPTE 377-1 defines a mechanism that links Header Metadata tracks to Content Elements (CE). Each Essence Element shall be described by an essence track in a MXF Header Metadata package. The track number shall be a UInt32 value comprising bytes: 'A.B.C.D' (most significant byte first). These bytes shall have the same values as bytes 13-16 of the key used to KLV wrap the Essence Element in the GC. These byte values shall be assigned to the Essence Element key value as follows:

- A = Byte 13 of the element key value;
- B = Byte 14 of the element key value;
- C = Byte 15 of the element key value;
- D = Byte 16 of the element key value.

This technique ensures that each track referenced by the Header Metadata package has a unique number for each Essence Element key value.

This 4-byte GC-unique value should be used to convey linkage to an Essence Element.

It is possible to use fewer than 4-bytes to create a link to an Essence Element; however, the document that defines such a link should also define a way to disambiguate the missing bytes.

Note 1: SMPTE 331M, Section 8.1, takes a unique approach where it records three of the four bytes directly in header metadata, but leaves the fourth byte implied. It defines a link that specifies metadata type, element type and element number, which corresponds to bytes 13, 15, and 16. Since the element count, byte 14 is always one in SMPTE 331M, bytes 13, 15 and 16 plus the implied value of byte 14 results in a 4-byte value that meets the requirements above.

Note 2: If there is one video element in the Picture Item and there are two Sound Elements in the Sound Item, then the Header Metadata package contains one picture track and two sound tracks. The value of the track number item in each track can be linked to the Essence Element keys in the Generic Container essence using the mechanism described above.

11 SMPTE Essence Container Label

Encoders shall identify essence container payload as specified in Table 4.

Table 4 – Specification of the Essence Container Label

Byte No.	Description	Value (hex)	Meaning
1	Object Identifier	06h	
2	Label size	0Eh	
3	Designator	2Bh	ISO, ORG
4	Designator	34h	SMPTE
5	Registry Category Designator	04h	Labels
6	Registry Designator	01h	Labels Registry
7	Structure Designator	01h	Labels Structure
8	Version Number	vvh	Version of the Registry
9	Item Designator	0Dh	Organizationally Registered
10	Organization	01h	AAF Association
11	Application	03h	Essence containers
12	Structure Version	01h	Version 1
13	Essence container Kind	02h 01h	MXF Generic Container Experimental MXF Generic Container Shall be used for prototyping only
14	Mapping Kind	xxh	Defines the kind of mapping
15~16	Locally defined	yyh	Defined by the application specification

Byte 14 shall be defined by the appropriate SMPTE GC mapping document and shall be a value in the range '01'h - '7F'h.

This SMPTE label is the individual 'essence container' property used in the partition pack, in the preface set and in the appropriate file descriptor.

This SMPTE label may also be added to the System Item where the definition of the System Item allows.

A value of 01h in byte 13 is provided to document Generic Container mappings which were experimental. This value shall not be used.

Annex A (Informative)
Bibliography

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

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

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

SMPTE 385M-2004, Television — Material Exchange Format (MXF) — Mapping SDTI-CP Essence and Metadata into the MXF Generic Container

SMPTE RP 210, Metadata Dictionary Registry of Metadata Element Descriptions

SMPTE RP 224, SMPTE Labels Registry

SMPTE EG 41-2004, Television — Material Exchange Format (MXF) — Engineering Guideline

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

ISO/IEC 13818-2:2000, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Video

EBU/SMPTE Task Force for Harmonized Standards for the Exchange of Programme Material as Bitstreams, Final Report: Analyses and Results, Sept 1998

SMPTE Journal, Vol 109, No 3, March 2000, pp 205..210, "A Tutorial on SDTI-CP"