

SMPTE STANDARD

Interoperable Master Format — Core Constraints



Page 1 of 43

Table of Contents	Page
Foreword	3
Intellectual Property	3
Introduction	3
1 Scope	4
2 Conformance Notation	4
3 Normative References	4
4 Overall	6
4.1 XML Schema and Namespace	6
4.2 UUID Generation	6
4.3 XML Character Encoding	6
5 Track Files	7
5.1 Overall	7
5.2 Image Track Files	8
5.3 Audio Track Files	9
5.4 Timed Text Track Files	15
5.5 Ancillary Data Track Files	17
5.6 Caption Description Packet (CDP) Track File	17
6 Composition	18
6.1 Application Identification	18
6.2 Homogeneous Essence	18
6.3 Virtual Tracks	19
6.4 Composition Edit Rate	22
6.5 Track File Identification	22
6.6 Segment Duration	22
6.7 Content Version	22
6.8 EssenceDescriptor	22
6.9 Digital Signature	23
6.10 Digital Certificates	23
6.11 Definition	23

7	Packaging	23
7.1	Overview (Informative).....	23
7.2	Interoperable Master Package (IMP).....	24
7.3	IMF Packing List Constraints	24
8	Delivery	25
8.1	Introduction	25
8.2	Definition	25
8.3	Ingest	25
8.4	Delivery Media	26
Annex A	Consolidated Schema (Informative).....	27
Annex B	Timed Text Definitions (Normative).....	28
B.1	Subtitles	28
B.2	Captions for the Hearing Impaired	28
B.3	Text for the Visually Impaired	28
B.4	Commentary.....	28
B.5	Karaoke.....	28
B.6	Forced Narrative	28
Annex C	StereolmageTrackFileResourceType (Normative).....	29
Annex D	Sparse Encoding of Timed Text Essence (Normative)	30
Annex E	Additional Generic Sound Essence Descriptor Items (Normative)	31
E.1	General	31
E.2	Reference Image Edit Rate.....	31
E.3	Reference Audio Alignment Level	31
Annex F	Alternative Center Cuts (Normative)	32
Annex G	Active Area Rectangle (Normative)	34
Annex H	RFC 5646 Language Tag List (Normative).....	37
Annex I	Packaging and Delivery Examples (Informative).....	38
I.1	Replacement Asset.....	38
I.2	Original Version (OV) and Version File (VF)	38
Annex J	IMF Packing List (Normative).....	40
Annex K	IMF HTTP Map Profile (Normative)	41
K.1	General (Informative).....	41
K.2	Asset Map Location and Encoding	41
K.3	Asset Location and Path Constraints.....	41
K.4	Retrieving Resources.....	42
Bibliography	(Informative).....	43

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 35PM.

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.

A single TV or movie title is transformed into multiple content versions (airline edits, special edition, languages...) These versions, which share common assets sourced from high-quality source masters, are ultimately made available to multiple distribution channels (Internet, optical media, broadcast...) across multiple territories and over the span of many months to over a year.

The IMF is a file-based framework that allows these high-quality versions, called Compositions, to be efficiently represented, managed, played back, processed and transformed on file-based systems. For example, it facilitates the generation of multiple outputs of the same Composition (through instructions contained in an Output Profile List defined in other documents) to accommodate the specific needs of distribution channels. Since management and processing of Compositions are performed across multiple devices and service providers, interoperability is desirable.

1 Scope

This document specifies provisions used by multiple IMF Applications. These provisions include essence and metadata constraints, but exclude image essence constraints, which are specified along with other parameters by each IMF Application.

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 330:2011, Unique Material Identifier (UMID)

SMPTE ST 334-1:2015, Vertical Ancillary Data Mapping of Caption Data and Other Related Data

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

SMPTE ST 377-4:2012, MXF Multichannel Audio Labeling Framework

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

SMPTE ST 400:2012, SMPTE Labels Structure

SMPTE ST 429-5:2017, D-Cinema Packaging — Timed Text Track File

SMPTE ST 429-6:2006, D-Cinema Packaging — MXF Track File Essence Encryption

SMPTE ST 429-8:2007, D-Cinema Packaging — Packing List

SMPTE ST 429-9:2014, D-Cinema Packaging — Asset Mapping and File Segmentation

SMPTE ST 430-2:2017, D-Cinema Operations — Digital Certificate

SMPTE ST 436-1:2013, MXF Mappings for VI Lines and Ancillary Data Packets

SMPTE ST 2001-1:2015, XML Representation of SMPTE Registered Data — Mapping Rules

SMPTE ST 2029:2009, Uniform Resource Names for SMPTE Resources

SMPTE ST 2067-3:2020, Interoperable Master Format — Composition Playlist

SMPTE ST 2067-5:2020, Interoperable Master Format — Essence Component

SMPTE ST 2067-8:2013, Interoperable Master Format — Common Audio Labels

SMPTE ST 2067-9:2018, Interoperable Master Format — Sidecar Composition Map

SMPTE ST 2067-100:2014, Interoperable Master Format — Output Profile List

ISO/IEC 14496-22:2019, Open Font Format

IEC 61966-2-1:1999, Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB

Internet Engineering Task Force (IETF) (December 1997). RFC 2253 — Lightweight Directory Access Protocol (v3): UTF-8 String Representation of Distinguished Names

Internet Engineering Task Force (IETF) (April 2005). RFC 4051 — Additional XML Security Uniform Resource Identifiers (URIs).

Internet Engineering Task Force (IETF) (July 2005). RFC 4122 — A Universally Unique Identifier (UUID) URN Namespace.

Internet Engineering Task Force (IETF) (May 2006). RFC 4539 — Media Type Registration for the Society of Motion Picture and Television Engineers (SMPTE) Material Exchange Format (MXF).

Internet Engineering Task Force (IETF) (September 2009). RFC 5646 — Tags for Identifying Languages.

World Wide Web Consortium (W3C) (26 November 2008). Extensible Markup Language (XML) 1.0 (Fifth Edition)

World Wide Web Consortium (W3C) (08 November 2018), TTML Profiles for Internet Media Subtitles and Captions 1.1 (IMSC1.1)

World Wide Web Consortium (W3C) (28 October 2004). XML Schema Part 1: Structures (Second Edition)

World Wide Web Consortium (W3C) (11 April 2013). XML Signature Syntax and Processing Version 1.1

4 Overall

4.1 XML Schema and Namespace

XML elements defined by this specification shall conform to the XML schema definitions (see W3C XML Schema Part 1: Structures) found in this specification. In the event of a conflict between schema definitions and the prose, the prose shall take precedence.

The XML schema root element shall be as defined in Table 1.

Table 1. XML Schema root element definition

```
<xs:schema targetNamespace="http://www.smpte-ra.org/ns/2067-2/2020"
  xmlns:cpl="http://www.smpte-ra.org/schemas/2067-3/2016"
  xmlns:cc="http://www.smpte-ra.org/ns/2067-2/2020"
  xmlns:dcml="http://www.smpte-ra.org/schemas/433/2008/dcmlTypes/"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified" attributeFormDefault="unqualified">
  <xs:import namespace="http://www.smpte-ra.org/schemas/2067-3/2016"/>
  <xs:import namespace="http://www.smpte-ra.org/schemas/433/2008/dcmlTypes/" />
  <!-- schema definitions found in this document -->
</xs:schema>
```

4.2 UUID Generation

UUID values used by this specification shall be generated as specified in IETF RFC 4122.

Unless specified otherwise, e.g. in Section 7.3.1, UUID values that identify assets or cryptographic information shall be generated using a truly-random or pseudo-random number source, and shall have a Version field value of '4' (or 0100b).

Note: The 'b' suffix on this value indicates a binary encoding, most significant bit (MSB) first.

4.3 XML Character Encoding

XML documents shall use UTF-8 encoding, as specified in W3C Extensible Markup Language (XML).

5 Track Files

5.1 Overall

5.1.1 Format

Track Files shall conform to SMPTE ST 2067-5.

5.1.2 Shim Parameters

Track Files shall be associated with the shim parameter values (see SMPTE ST 2067-5) specified in Table 2.

Table 2. Shim Parameter Values Definitions

<i>Shim Parameter</i>	<i>Value</i>	<i>Notes</i>
audio_family	ST 2067-2	Section 5.3.2
audio_file_arrangement	ST 382	Section 5.3.3
data_family	CDP (ST334-1), IMSC1.1	Sections 5.5 and 5.6
data_file_arrangement	ST 436-1, ST429-5	Section 5.4 and 5.5
sys_item_tc	true	Metadata in the System Item can be copied or moved to header metadata.
partition_spacing	60s	Section 5.1.3
index_strategy_frame	Follow	Section 5.1.4
Index_strategy_clip	Follow	Section 5.1.4
track_tag_policy	none	Section 5.3.6
ingest_TC	none	No dominant source of timecode during ingest is specified.
lead_TC	CPL	The Composition Playlist (CPL) specified in SMPTE ST 2067-3 defines a synthetic timecode Composition timeline.
dropframe_tc	none	No constraints are specified on timecode tracks in Track File header metadata.

Note: The Composition Playlist specified in SMPTE ST 2067-3 provides synchronization across the Composition timeline and timecode is not used.

The shim parameters listed in Table 3 shall be defined by individual Applications.

Table 3. Shim Parameter defined by Applications

<i>Shim Parameter</i>
shim_id
gc_type
picture_family

picture_bitrate
 picture_format
 picture_custom_ANC
 picture_render_ANC

5.1.3 Body Partition

With the exception of the last body partition, the body partitions of a frame-wrapped Track File shall occur every 60 ± 6 s.

5.1.4 Index tables

At each partition point in a Track File, the index partition shall follow the essence partition that it indexes.

5.1.5 Identification

The Package UID of a given Track File shall be a basic UMID per SMPTE ST 330, having a UUID value in the material number part and a material number generation method of UUID/UL. The Package UID value shall be further constrained as follows:

- Byte 11 of the UL portion of the UMID shall be 0Fh (unidentified material type).
- Byte 12 of the UL portion of the UMID shall be 20h (UUID/UL material number generation method and undefined instance number generation method).
- The three bytes of the instance number shall be 0 (zero).

Package UID values generated in accordance with the normative provisions of this subsection will thus have the following contents in the first 16 bytes: 060a2b34h 01010105h 01010f20h 13000000h.

5.1.6 MIME Type

The media type of a Track File shall be application/mxf as specified in IETF RFC 4539.

5.1.7 Origin Property

The Origin Property of all Track items in the top-level File Package shall be zero.

Note: When the Origin Property of a Track items in the top-level File Package is zero, the Zero Point, as defined in SMPTE ST 377-1, refers to the first Content Element in the Essence Element.

5.1.8 Empty Essence Elements

The Structural Components item of the Top-Level File Package shall not reference any Filler component, i.e., the Essence Element of the Essence Container is not empty.

5.2 Image Track Files

5.2.1 General

An Image Track File shall conform to Sections 5.1 and 5.2.

An Image Track File primarily contains image essence. This specification does not specify the nature of the image essence.

5.2.2 Alternative Center Cut

The item defined in Annex F may be present in the Generic Picture Essence Descriptor (see SMPTE ST 377-1).

5.2.3 Active Area Rectangle

The items defined in Annex G may be present in the Generic Picture Essence Descriptor (see SMPTE ST 377-1).

5.2.4 Essence Container Integrity

The Composition Playlist (SMPTE ST 2067-3) and IMF Packing List (Section 7) allow the integrity of an entire Track File to be checked.

To enable the integrity of individual image frames to be checked, Image Track Files may conform to SMPTE ST 429-6 with the following constraints:

- the value of the Cipher Algorithm ID shall be 00000000.00000000.00000000.00000000 ("No cipher algorithm used").
- the value of the MIC Algorithm ID shall be 060e2b34.04010107.02090202.01000000 ("HMAC-SHA1 with 128-bit key").
- the MIC item shall be present
- Plaintext Offset shall be equal to Source Length
- Cryptographic Key ID shall be equal to 42b5ea14-233d-415c-b7d8-8b2dbd95ad39
- Cipher Key shall be equal to 0xcdea9b5a9bf00a0d564dd5ff0c19757d

5.3 Audio Track Files

5.3.1 General

An Audio Track File shall conform to Sections 5.1 and 5.3.

An Audio Track File primarily contains audio essence.

Each Audio Track File is intended to carry a single soundfield group. If the audio that is to be carried consists of more than one soundfield group, then each is carried in its own Audio Track File(s).

Example: 3 Audio Track Files carry a 6+2 Music and Effects content: one containing the "Music and Effects Main" 5.1 soundfield group and two others containing each a monaural soundfield group with a single center channel, one for "Music and Effects Optional" and one for "Dialog Guide". The GroupOfSoundfieldGroupsLabelSubDescriptor (5.3.6.4) is used to associate the soundfield groups with each other.

5.3.2 Essence

5.3.2.1 General

The audio essence shall consist of an audio waveform digitally sampled at regular interval into a sequence of audio samples.

5.3.2.2 Audio Sampling Rate

The sampling rate of the audio samples shall be either 48000 Hz or 96000 Hz.

Note: The sampling rates specified above are absolute and specified independently of a video signal. The Composition Playlist (SMPTE ST 2067-3) provides the mechanism to synchronize audio and image essence.

5.3.2.3 Quantization

Each audio sample shall be uniformly quantized as a 24-bit signed integer in the range $[-(2^{23}) .. 2^{23} - 1]$.

Source audio uniformly quantized using signed integers with lower bit depth, e.g. 16- or 20-bit, shall be padded in the least significant bits to create 24-bit audio samples.

5.3.3 Wrapping

The audio essence within an Audio Track File shall be:

- clip-wrapped as a Wave Clip-Wrapped Element, as specified in SMPTE ST 382; and
- packed bit-by-bit as specified by AES31-2, as specified in Section 7.2.1 of SMPTE ST 382:2007.

5.3.4 Wave Audio Essence Descriptor

5.3.4.1 General

The Top-Level File Package shall reference a Wave Audio Essence Descriptor as specified in SMPTE ST 382.

The constraints specified in Table 4 shall apply. Items labeled “shall be present” shall be present in the Wave Audio Essence Descriptor. Items labeled “shall be ignored” shall be ignored if present.

Table 4. Wave Audio Essence Descriptor Constraints

<i>Item</i>	<i>Constraints</i>
Sample Rate	Same as "Audio sampling rate"
Audio sampling rate	See Section 5.3.2.2
Channel Count	See Section 5.3.6.6
Quantization bits	See Section 5.3.2.3
Sound Essence Coding	See Section 7.2.1 of SMPTE ST 382:2007
Locked/Unlocked	Shall be ignored
Audio Ref Level	Shall be ignored
Electro-Spatial Formulation	Shall be ignored
Dial Norm	Shall be ignored
Channel Assignment	See Section 5.3.4.2

5.3.4.2 Channel Assignment

The Channel Assignment item shall be present and shall be equal to the UL specified in Table 5.

Table 5. Specification of the Channel Assignment Label

Byte No.	Description	Value (hex)	Meaning
1-7	Registry Designator	See SMPTE ST 400	
8	Registry Version Number	0Dh	Version of SMPTE register (see SMPTE ST 400) in which this label first appears
9	Parametric	04h	Node used to define parametric data
10	Sound Essence	02h	Identifies sound essence coding
11	Sound Coding Characteristics	02h	Identifies sound coding characteristics
12	Sound Channel Labeling	10h	Identifies sound channel labeling
13	Sound Channel Labeling SMPTE ST2067-2	04h	Identifies sound channel labeling as defined in SMPTE ST2067-2
14	Application of the MXF Multichannel Audio Framework	01h	Indicates that the MXF Multichannel Audio Framework is used
15	Reserved	00h	Reserved
16	Reserved	00h	Reserved

5.3.5 Additional Generic Sound Essence Descriptor Items

The items specified in Annex E shall be subject to the constraints specified in Table 6.

Table 6. Additional Generic Sound Essence Descriptor Items Constraints

Item	Constraints
Reference Audio Alignment Level	Should be present.
Reference Image Edit Rate	Should be present. See Section 6.4. Note: This is the edit rate of the picture essence to which the sound essence was conformed initially. Audio Track Files can be used in Compositions that have different Composition Edit Rates than the value of the Reference Image Edit Rate item.

5.3.6 Multichannel Audio Labeling

5.3.6.1 General

An Audio Track File shall conform to SMPTE ST 377-4.

5.3.6.2 AudioChannelLabelSubDescriptor Instances

Each audio channel in an Audio Track File shall be associated with exactly one AudioChannelLabelSubDescriptor instance, and each AudioChannelLabelSubDescriptor instance shall be associated with exactly one audio channel.

Note: The MCA Channel ID item is used to make this association.

5.3.6.3 SoundfieldGroupLabelSubDescriptor Instances

An Audio Track File shall contain exactly one SoundfieldGroupLabelSubDescriptor instance, which all AudioChannelLabelSubDescriptor instances shall reference.

5.3.6.4 GroupOfSoundfieldGroupsLabelSubDescriptor Instances

An Audio Track File shall contain zero or one GroupOfSoundfieldGroupsLabelSubDescriptor.

5.3.6.5 Item Constraints

Within a given Audio Track File, the constraints of Table 7 shall apply in addition to those specified in SMPTE ST 377-4.