

SMPTE STANDARD

Interoperable Master Format —
Composition Playlist



Table of Contents	Page
Foreword	3
Intellectual Property	3
Introduction	3
1 Scope	6
2 Conformance Notation	6
3 Normative References	6
4 Glossary	7
5 Instance	8
5.1 Schema	8
5.2 Character Encoding	8
5.3 MIME Type	8
5.4 Structure Versioning	9
6 Structure	9
6.1 CompositionPlaylistType	9
6.2 CompositionTimecodeType	14
6.3 TimecodeType	15
6.4 ContentKindType	15
6.5 LocaleType	15
6.6 ContentMaturityRatingType	17
6.7 EssenceDescriptorBaseType	18
6.8 ContentVersionType	18
6.9 SegmentType	19
6.10 SequenceType	20
6.11 BaseResourceType	20
6.12 TrackFileResourceType	22
6.13 MarkerResourceType	23
6.14 MarkerType	23
7 Timeline and Synchronization Model	27
7.1 Composition	27

7.2	Segment	27
7.3	Sequence	28
7.4	Resource	28
8	Canonical Synthetic Timecode	29
Annex A	Consolidated Schema (Informative)	30
Annex B	Consideration for Track File Specifications (Informative)	31
Annex C	Considerations for Applications (Informative)	32
C.1	Content Version	32
C.2	EditRate	32
C.3	Timeline	32
C.4	Splice Points	32
C.5	Sequences	32
C.6	Entry Points and Source Durations	32
C.7	Track File characteristics	32
C.8	Essence characteristics	32
C.9	Mapping of TrackFileId values to actual Track File locations	32
C.10	Mapping of KeyId values to cryptographic keys	32
C.11	Markers	33
C.12	Digital Signature	33
C.13	Minimum and Maximum Track File Duration	33
C.14	ExtensionProperties	33
C.15	EssenceDescriptors	33
C.16	Content Maturity Rating Agencies and Audience	33
C.17	Sequences	33
C.18	Resource	33
Annex D	Sample Instance (Informative)	34
Bibliography	(Informative)	35

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.

The IMF is an interoperable file-based framework designed to facilitate the management and processing of multiple content versions (airline edits, special edition, languages...) of the same high-quality finished work (feature, episode, trailer, advertisement, etc) destined for distribution channels worldwide. Each content version is embodied in a composition, which combines metadata and essence.

The Composition Playlist defines the playback timeline for the Composition and includes metadata applicable to the Composition as a whole. It is a human-readable structure expressed using XML and specified using XML Schema. It includes multiple extension points supporting both backward- and forward-compatibility.

The Composition Playlist is not designed to contain essence but rather reference external Track Files that contain the actual essence. This allows multiple compositions to be managed and processed without duplicating the essence in common. For convenience, metadata contained in the Track Files is exposed in the Composition Playlist (Section 6.1.10.)

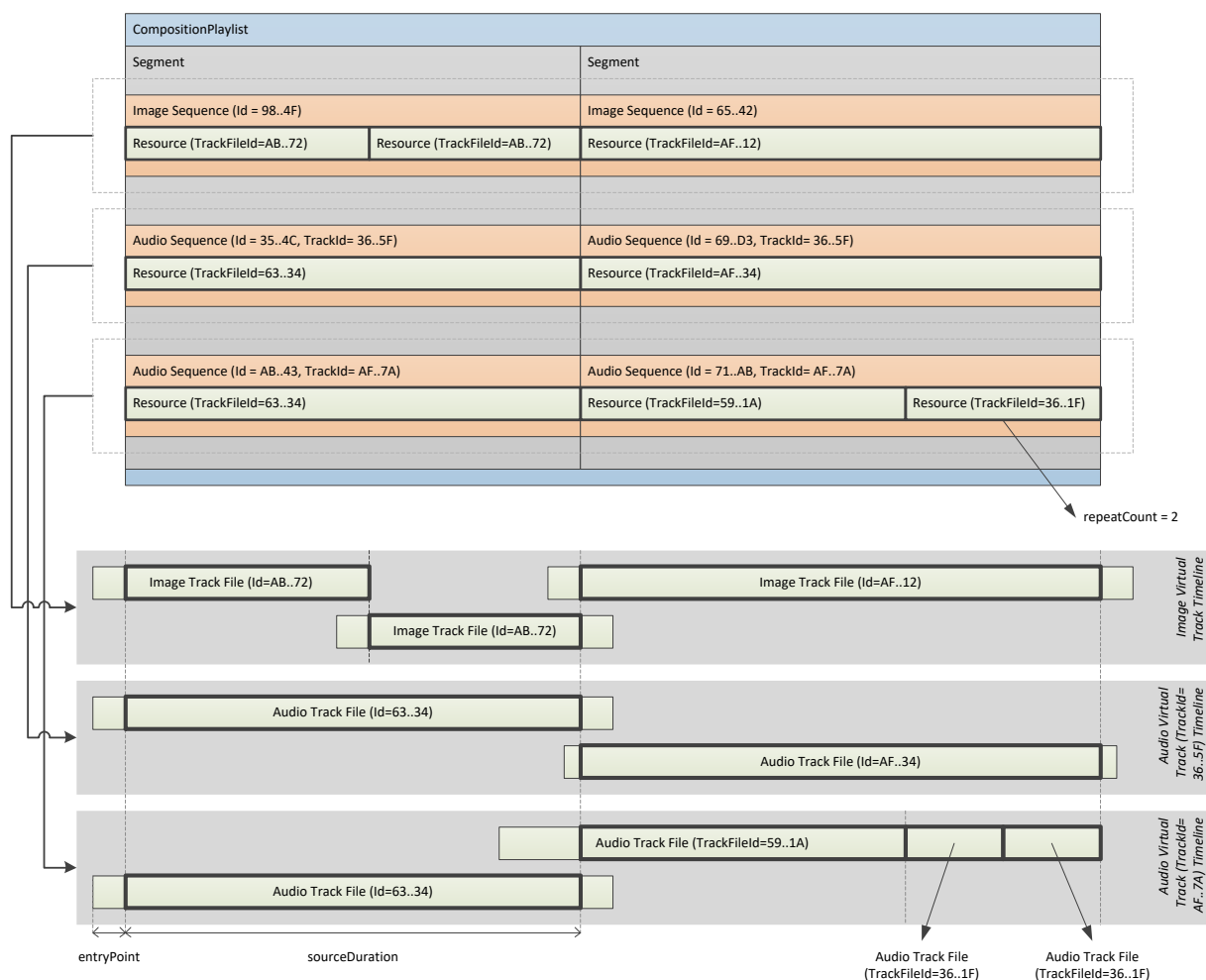


Figure 1. Sample Monoscopic Composition
Only the first and last bytes of UUIDs are represented

The Composition¹ is illustrated in Figure 1. It consists of a list of Segments that are reproduced sequentially and without gaps. The number of Segments and the duration of each Segment are determined by the content provider to fit both the particular workflow and content type. For example, a Segment can be the same duration as a reel or it can be the running time of a television program between commercials. Each Segment consists of a list of Sequences that are played in parallel and have the same duration. Each Sequence is associated with a single kind of essence or metadata and consists of a list of Resources that are reproduced sequentially.

Each Resource references a playable region of essence (audio, image, data, etc.) or metadata contained in an underlying Asset. The Composition timeline makes no assumption on the format of the asset, other than its ability to reproduce essence or metadata in the playable region.

¹ The Composition Playlist does not use timecode to synchronize Assets but defines a canonical synthetic timecode, which implementations can use.

Assets will typically be external Track Files, which can be referenced by multiple Resources, within and outside a particular Virtual Track. In Figure 1, Audio Track File Id=63..34 is referenced by Audio Sequences Id= 35..4C and Id=AB..43. Image Track File Id=AB..72 is referenced twice within the first Image Sequence. A Track File can be repeated multiple times within a single Resource. This allows, for instance, gaps in Audio Tracks to be filled by repeating a Track File containing a few samples of silence. If an Asset contains a very small amount of XML information and is referenced by a single Resource, it can be included directly within the Composition Playlist structure, e.g. MarkerResourceType in Section 6.13.

The Composition Playlist does not define any form of transition or overlap between resources. In other words, the essence contained in Assets is prepared for artifact-free reproduction across these boundaries.

Temporal offsets and durations on the Composition timeline are expressed as integer multiples of Edit Units, the smallest temporal increment accessible on the timeline and the inverse of the Edit Rate. Edit Rates and Edit Units can be defined independently for the Composition and each Resource. The Composition Edit Rate will generally correspond to the image frame rate.

To support a range of applications within its stated scope, the Composition Playlist allows flexibility in the nature and structure of Sequences, Resources and Track Files. Applications are therefore expected to constrain these characteristics to fit their respective requirements. Similarly, the method by which essence is wrapped in an Asset is left to the specification defining the Asset.

1 Scope

This standard specifies the Composition Playlist (CPL) for the Interoperable Master Format (IMF). The Composition Playlist is a representation of a single version of a finished IMF composition (feature, episode, trailer, advertisement, etc). It contains the information necessary to describe the composition and synchronize its underlying essence; e.g., for playout or transcoding. It is an extensible, human-readable structure designed for file-based operations.

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 standard. 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 12-1:2014, Time and Control Code

SMPTE ST 433:2008, D-Cinema — XML Data Types

Amendment 1:2011 to SMPTE ST 433:2008

Internet Engineering Task Force (IETF) (November 1996). RFC 2045 — Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies

Internet Engineering Task Force (IETF) RFC 2046 (November 1996) Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types

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) (28 October 2004). XML Schema Part 1: Structures (Second Edition)

World Wide Web Consortium (W3C) (28 October 2004). XML Schema Part 2: Datatypes (Second Edition)

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

4 Glossary

4.1 Asset: Metadata or essence underlying a Resource. A Track File is one kind of Asset.

4.2 Composition: Embodies a version of a complete work, combining metadata and essence.

4.3 Edit Rate: A number of Edit Units to be reproduced per second.

4.4 Edit Unit: The smallest temporal increment of access to a timeline.

4.5 Essence: Content, e.g. image, audio or data essence, meant for presentation.

4.6 IMF: Acronym for Interoperable Master Format.

4.7 Main Titles: A credit sequence generally shown near the beginning of a motion picture.

4.8 Metadata: Data about data or data describing other data. Information that is considered ancillary to or otherwise directly complementary to essence. Information that is useful or of value when associated with the essence being provided.

4.9 MXF: Material Exchange Format.

4.10 Native Duration: The total number of Resource Edit Units in an Asset.

4.11 Native Start Point: The first Edit Unit of an Asset.

4.12 Playable Region: A contiguous sequence of Resource Edit Units within an Asset that is intended to be reproduced. An Asset can contain Edit Units before and/or after the Playable Region.

4.13 Resource: A portion of an Asset selected for reproduction on the Composition timeline.

4.14 Sequence: An ordered collection of Resources to be reproduced sequentially.

4.15 Segment: A collection of Sequences intended to be reproduced in parallel.

4.16 Track File: An Asset consisting of a file containing essence of a single kind, e.g. as audio, image or data essence.

4.17 UUID: Acronym for Universal Unique Identifier.

4.18 XML: Acronym for extensible Markup Language.

5 Instance

A Composition Playlist instance shall be an XML document, as specified in W3C XML 1.0, that consists of a single CompositionPlaylist element (see Section 6.1).

5.1 Schema

Each Composition Playlist instance shall conform to the XML schema definitions (see W3C XML Schema Part 1: Structures and W3C XML Schema Part 2: Datatypes) found in this specification. In the event of a conflict between schema definitions and the prose, the prose shall take precedence.

Table 1 specifies the XML schema root element.

Table 1. XML Schema root element definition

```
<xs:schema targetNamespace="http://www.smpte-ra.org/schemas/2067-3/2016"
  xmlns:cpl="http://www.smpte-ra.org/schemas/2067-3/2016"
  xmlns:dcml="http://www.smpte-ra.org/schemas/433/2008/dcmlTypes/"
  xmlns:ds="http://www.w3.org/2000/09/xmldsig#"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified" attributeFormDefault="unqualified">
  <xs:import namespace="http://www.smpte-ra.org/schemas/433/2008/dcmlTypes/" />
  <xs:import namespace="http://www.w3.org/2000/09/xmldsig#" />
  <xs:element name="CompositionPlaylist" type="cpl:CompositionPlaylistType"/>
  <!-- schema definitions found in this document excluding this one -->
</xs:schema>
```

The namespace prefixes used in XML Schema definitions herein are not normative values and implementations shall perform correctly with any XML compliant prefix values. The XML Schema definitions found in this specification include elements specified in W3C XML Signature Syntax and Processing Version 1.1 and SMPTE ST 433.

A schema instance that consolidates all inline schema definitions made in this document is provided for convenience in Annex A. In case of conflict, the schema definitions included herein shall take precedence.

5.2 Character Encoding

Composition Playlist instances shall be encoded using the UTF-8 character encoding.

5.3 MIME Type

The MIME type, as defined in IETF RFC 2046, of a Composition Playlist instance shall be text/xml.

5.4 Structure Versioning

The target namespace specified in Section 5.1, i.e. the value of the targetNamespace attribute of the schema element, shall only be used by Composition Playlist instances that conform to this specification as expressed by the combination of its prose and schema definitions.

Instances using specifications that modify the schema definitions or the semantics of the elements defined herein, including future versions of this specification, shall use a different namespace and no two distinct schemas shall have the same target namespace. As such, the target namespace allows implementations to unambiguously identify the defining specification of a Composition Playlist instance.

6 Structure

In order to avoid duplication between text and schema, the cardinality and default values of elements are specified in the schema only.

6.1 CompositionPlaylistType

Table 2. CompositionPlaylistType schema definition

```
<xs:complexType name="CompositionPlaylistType">
  <xs:sequence>
    <xs:element name="Id" type="dcml:UUIDType" />
    <xs:element name="Annotation" type="dcml:UserTextType" minOccurs="0" />
    <xs:element name="IssueDate" type="xs:dateTime" />
    <xs:element name="Issuer" type="dcml:UserTextType" minOccurs="0" />
    <xs:element name="Creator" type="dcml:UserTextType" minOccurs="0" />
    <xs:element name="ContentOriginator" type="dcml:UserTextType" minOccurs="0" />
    <xs:element name="ContentTitle" type="dcml:UserTextType" />
    <xs:element name="ContentKind" type="cpl:ContentKindType" minOccurs="0" />
    <xs:element name="ContentVersionList" minOccurs="0">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="ContentVersion" type="cpl:ContentVersionType"
            maxOccurs="unbounded" />
        </xs:sequence>
      </xs:complexType>
    </xs:element>
    <xs:element name="EssenceDescriptorList" minOccurs="0">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="EssenceDescriptor" type="cpl:EssenceDescriptorBaseType"
            maxOccurs="unbounded" />
        </xs:sequence>
      </xs:complexType>
    </xs:element>
    <xs:element name="CompositionTimecode" type="cpl:CompositionTimecodeType"
      minOccurs="0" />
    <xs:element name="EditRate" type="dcml:RationalType" />
    <xs:element name="TotalRunningTime" minOccurs="0">
      <xs:simpleType>
        <xs:restriction base="xs:string">
          <xs:pattern value="[0-9][0-9]:[0-5][0-9]:[0-5][0-9]" />
        </xs:restriction>
      </xs:simpleType>
    </xs:element>
    <xs:element name="LocaleList" minOccurs="0">
      <xs:complexType>
        <xs:sequence>
```

```

        <xs:element name="Locale" type="cpl:LocaleType" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="ExtensionProperties" minOccurs="0">
    <xs:complexType>
      <xs:sequence>
        <xs:any namespace="##other" processContents="lax" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="SegmentList">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="Segment" type="cpl:SegmentType" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="Signer" type="ds:KeyInfoType" minOccurs="0" />
  <xs:element ref="ds:Signature" minOccurs="0" />
</xs:sequence>
</xs:complexType>

```

6.1.1 Id

The Id element shall identify the Composition Playlist instance. No two Composition Playlists shall have the same Id values unless they are identical.

6.1.2 Annotation

The Annotation element shall be a free-form, human-readable annotation describing the composition. It is meant strictly as a display hint to the user.

6.1.3 IssueDate

The IssueDate element shall indicate the time and date at which the Composition Playlist was issued.

6.1.4 Issuer

The Issuer element shall be a free-form, human-readable annotation that identifies the entity that created the Composition Playlist. It is meant strictly for display to the user.

Note: The Signer element defined in Section 6.1.17 is used to identify the entity that digitally signed the Composition Playlist.

6.1.5 Creator

The Creator element shall be a free-form, human-readable annotation that identifies the device or software program used to create the Composition Playlist, the facility that created the Composition Playlist and the operator that created the Composition Playlist. It is meant strictly for display to the user.

6.1.6 ContentOriginator

The ContentOriginator element shall be a free-form, human-readable annotation that identifies the originator of the content underlying the composition, as opposed to the Issuer, which identifies the entity that created the Composition Playlist. It is meant strictly for display to the user.

6.1.7 ContentTitle

The ContentTitle element shall contain a human-readable title for the composition, e.g. *The Jazz Singer*. It is strictly meant as a display hint to the user.

6.1.8 ContentKind

The ContentKind element shall be human-readable and indicate the kind of work represented by the Composition. ContentKindType is defined in Section 6.4.

The scope attribute of the ContentKind element shall determine the permissible values of the ContentKind element:

- if the scope attribute is absent from the ContentKind element, or set to its default value, the value of the ContentKind element shall match one of the values listed in Table 3;
- if the scope attribute is present and equal to "http://www.smpte-ra.org/schemas/2067-3/2016#content-kind", the value of the ContentKind element shall match one of the values listed in Table 4.
- otherwise, the value of the ContentKind element is outside the scope of this specification. In any case, this element is intended to be displayed to the user.

Table 3. Permitted Content Kind values under the default scope

<i>Kind</i>	<i>Description</i>
advertisement	Content promoting a product or service other than an upcoming feature.
feature	A theatrical feature.
psa	Public service announcement.
rating	Slate/still image indicating the recommended age group permitted to view the content to follow. This rating is generally unique per country.
short	Non advertising/promotional content (3 to 15 minutes) typically before a theatrical feature.
teaser	Very short (typically less than 1 minute) content promoting an upcoming theatrical feature.
test	Content used to test, calibrate or setup equipment.
trailer	Short (2 to 3 minutes) content promoting an upcoming theatrical feature.
transitional	Extremely short content (1 to 15 seconds) separating unrelated compositions.
episode	Part of a dramatic work such as a serial television program.
highlights	Edited sequence of one or more memorable parts of an event, feature, episode or period of time.
event	Activity that happens at a given place and time, especially one of importance, including a music performance, a contest making up a sports competition.