SMPTE ST 2067-201:2021
Revision of SMPTE ST 2067-201:2019

SMPTE STANDARD

Immersive Audio Bitstream Level 0 Plug-in

Table of Contents

Approved November 9, 2020

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.

SMPTE ST 2098-2 has proven (i) effective in representing cinematic immersive sound essence, and (ii) straightforward to wrap in MXF. The objective of this standard is to define a baseline method for the carriage of such sound essence for use with feature and episodic content in IMF compositions (Level 0).

1 Scope

This document specifies requirements for a plug-in mechanism to add Immersive Audio Bitstream (IAB) essence, as specified in SMPTE ST 2098-2, to IMF compositions.

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 377-1:2019, Material Exchange Format (MXF) — File Format Specification

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

SMPTE ST 429-18:2019, D-Cinema Packaging — Immersive Audio Track File

SMPTE ST 2067-2:2020, Interoperable Master Format — Core Constraints

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

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

SMPTE ST 2098-2:2019, Immersive Audio Bitstream Specification

4 Terms and Definitions

No terms and definitions are listed in this document.

5 IAB Track File

5.1 General

An IAB Track File is a Track File, as specified in Section 5.1 of SMPTE ST 2067-2:2020, which contains SMPTE ST 2098-2 Immersive Audio Bitstream essence, and is further constrained by the provisions of this Section.

5.2 Identification

The ConformsToSpecifications element specified in Annex B shall be present and contain a single instance of the Universal Label (UL) specified in Table 4.1.

Table 4.1. IMF IAB Track File Level 0 Label.

Kind	Leaf
Name	IMF IAB Track File Level 0
Symbol	IMF_IABTrackFileLevel0
UL	urn:smpte:ul:060E2B34.0401010D.01010201.02000000

5.3 Single Immersive Audio Bitstream

The Essence Track of the IAB Track File shall reference an Essence Container, as defined in SMPTE ST 377-1, that contains a single Immersive Audio Bitstream specified in SMPTE ST 2098-2.

5.4 Single Sound Essence Track

The following shall apply to the Essence Track referencing the Immersive Audio Bitstream essence:

- It shall be a Sound Essence Track.
- Its Edit Rate, as specified at Annex B.12 of SMPTE ST 377-1:2019, shall be equal to the Sample Rate Property of the IAB Essence Descriptor, as defined in Section 5.9.
- In accordance with Section 4.1 of SMPTE ST 377-1:2019, an Edit Unit shall correspond to a single complete frame of the Immersive Audio Bitstream, as defined in Section 9 of SMPTE ST 2098-2:2019.

NOTE The first requirement above determines the value of the Data Definition item of the Sequence referenced by the Sound Essence Track.

5.5 Single Sound Element

The Immersive Audio Bitstream shall be contained in a single Sound Element, as defined in SMPTE ST 377-1, whose Essence Element Key shall be as specified in Table 4.2.

Table 4.2. Essence Element Key.

Kind	Leaf
Name	IMF IAB Essence Clip-Wrapped Element
Symbol	IMF_IABEssenceClipWrappedElement
UL	urn:smpte:ul:060E2B34.01020101.0D010301.16cc0Dnn

NOTE As audio essence, the Immersive Audio Bitstream is clip-wrapped as specified in SMPTE ST 2067-5.

5.6 Bitstream Characteristics

The following characteristics of the Immersive Audio Bitstream shall remain constant in the Essence Container:

- BitDepth
- FrameRate
- SampleRate

5.6.1 Bitstream Constraints

- The Bit Depth of the Immersive Audio Bitstream shall be 24 bits.
- All AudioData elements contained within the Immersive Audio Bitstream shall be AudioDataPCM elements.
- AudioDataDLC elements shall not be present in the Immersive Audio Bitstream.

5.6.2 Bitstream Element Constraints

The following sections describe IAB bed related elements that shall remain fixed in configuration, and fields within these elements that shall remain fixed.

5.6.2.1 BedDefinition

All IAFrames in a given Immersive Audio Bitstream shall have the same number of BedDefinition elements. For each BedDefinition element in a given IAFrame, there shall be exactly one BedDefinition element in each additional IAFrame that has the same MetaID value. BedDefinition elements with the same MetaID shall have the same value(s) for ConditionalBed, BedUseCase (if present), ChannelCount, and ChannelID. No element which is a child of an ObjectDefinition or BedDefinition shall be used.

5.6.2.2 BedRemap

BedRemap elements shall not be present in the Immersive Audio Bitstream.

5.6.2.3 Conditional Elements

No conditional elements (i.e. BedDefinition elements that have ConditionalBed set to 1 or ObjectDefinition elements that have ConditionalObject set to 1) shall be used unless their UseCase is set to 0xFF (Always Use).

5.7 Index Table

As defined in Section 11.2.3 of SMPTE ST 377-1:2019, the Index Edit Rate shall be equal to the Edit Rate of the MXF Essence Track, as defined in Section 5.4.

5.8 IAB Essence Descriptor Definition

5.8.1 General

The IAB Essence Descriptor is a subclass of Generic Sound Essence Descriptor, as defined in SMPTE ST 377-1 and extended in SMPTE ST 2067-2, and contains the items specified in Table 4.3 with the Set Key specified in Table 4.4.

Table 4.3. IAB Essence Descriptor.

Name	Type	Req	Definition
IAB Essence Descriptor	Set Key	Req	Identifies an Immersive Audio Bitstream Essence Descriptor
Length	BER length	Req	Set length
All items in Generic Sound Essence Descriptor, as specified in SMPTE ST 377-1 and extended in SMPTE ST 2067-2, except Group UL and Length			

Table 4.4. IAB Essence Descriptor Key.

Kind	Leaf
Name	IAB Essence Descriptor
Symbol	IABEssenceDescriptor
UL	urn:smpte:ul:060E2B34.027F0101.0D010101.01017B00

NOTE The value of byte 6 (0x7F) in the IAB Essence Descriptor Key is a placeholder that is used in the registers, but is not used in actual implementations. Please consult SMPTE ST 377-1:2019, Section 9 for the proper value(s).

5.9 IAB Essence Descriptor Constraints

A single IAB Essence Descriptor, as specified in Section 5.8, shall be associated with the Essence Track.

The Sample Rate item of the IAB Essence Descriptor, as inherited from File Descriptor, shall be set to a value corresponding to the frame rate represented by the code value of the first instance of FrameRate in the Immersive Audio Bitstream, as defined in Section 10.2.4 of SMPTE ST 2098-2:2019.

NOTE 1 FrameRate may occur multiple times within an Immersive Audio Bitstream, but is constrained by Section 5.6, above, to be remain constant throughout the bitstream.

The Essence Container Label item of the IAB Essence Descriptor, as inherited from File Descriptor, shall be as specified in Table 4.5.

Table 4.5. Essence Container Label.

Kind	Leaf
Name	IMF Clip-Wrapped IAB Essence Container
Symbol	IMF_IABEssenceClipWrappedContainer
UL	urn:smpte:ul:060E2B34.0401010D.0D010301.021D0101

The IMF Clip-Wrapped IAB Essence Container label is a child of the MXF-GC IAB Audio node label, which is a child of the MXF-GC Immersive Audio node label. The MXF-GC IAB Audio and MXF-GC Immersive Audio node labels are defined in Annex D of this specification.

The Codec item of the IAB Essence Descriptor, as inherited from File Descriptor, shall not be present.

The Sound Essence Coding item of the IAB Essence Descriptor, as inherited from Generic Sound Essence Descriptor, shall be set to the UL value specified in Table 4.6, which is defined in SMPTE ST 429-18 and has the meaning: Immersive Audio Coding per SMPTE ST 2098-2.

Table 4.6. Coding Label.

urn:smpte:ul:060E2B34.04010105.0E090604.00000000

The Audio Sampling Rate item of the IAB Essence Descriptor shall be present and shall be set to a value corresponding to the sampling frequency of the audio sample data represented by the code value of the first instance of SampleRate in the Immersive Audio Bitstream, as defined in Section 10.2.2 of SMPTE ST 2098-2:2019.

NOTE 2 SampleRate may occur multiple times within an Immersive Audio Bitstream, but is constrained by Section 5.6, above, to be remain constant throughout the bitstream.

The Locked/Unlocked item of the IAB Essence Descriptor shall be ignored.

The Audio Ref Level item of the IAB Essence Descriptor shall be ignored.

If present, the Electro-Spatial Formulation item of the IAB Essence Descriptor shall be set to a value of 15 (multi-channel mode default).

The ChannelCount item of the IAB Essence Descriptor shall be ignored.

The Quantization Bits item in the IAB Essence Descriptor shall be set to a value corresponding to the bit depth of the audio sample data represented by the code value of the first instance of BitDepth in the Immersive Audio Bitstream, as defined in Section 10.2.3 of SMPTE ST 2098-2:2019.

NOTE 3 BitDepth may occur multiple times within an Immersive Audio Bitstream, but is constrained by Section 5.6, above, to be remain constant throughout the bitstream.

The Dial Norm item in the IAB Essence Descriptor shall be ignored.

The Reference Image Edit Rate item in the IAB Essence Descriptor, as defined in section E.2 of SMPTE ST 2067-2:2020 to extend Generic Sound Essence Descriptor, should be present.

The Reference Audio Alignment Level item in the IAB Essence Descriptor, as defined in section E.3 of SMPTE ST 2067-2:2020 to extend Generic Sound Essence Descriptor, should be present.

5.10 Immersive Audio Labeling

5.10.1 General

An IAB Track File shall utilize the MCA Label framework as defined in SMPTE ST 377-4 and further extended in Annex C.

5.10.2 SubDescriptor Instances

An IAB Track File shall contain exactly one instance of an IAB Soundfield Label SubDescriptor, as defined in Annex C.

An IAB Track File shall not contain instances of AudioChannelLabelSubDescriptor, SoundfieldGroupLabelSubDescriptor, or GroupOfSoundfieldGroupsLabelSubDescriptor.

5.10.3 Item Constraints

Within a given IAB Track File, the constraints of Table 4.7 shall apply.

Table 4.7. IAB Soundfield Label SubDescriptor Constraints.

Item	IAB Soundfield Label SubDescriptor Constraints
MCA Tag Name	Shall be present.
RFC 5646 Spoken Language	Shall be equal to the primary spoken language associated with the IAB soundfield. It shall be absent if and only if the IAB soundfield is not associated with a primary spoken language.
MCA Audio Content Kind	Should be present.
MCA Audio Element Kind	
MCA Title	
MCA Title Version	

Other items defined for IAB Soundfield Label SubDescriptor but not required by this specification may be present and may be safely ignored by implementations.

NOTE 1 MCA Tag Symbol and MCA Tag Name contain a human-readable text intended for display to the user. The MCA Label Dictionary ID is used to unambiguously determine the nature of the underlying soundfield.

NOTE 2 MCA Audio Content Kind, MCA Element Kind, MCA Title and MCA Title Version contain human-readable descriptive text intended for display to the user.

5.10.4 Common IAB Soundfield Label SubDescriptor Values

The MCA Label Dictionary ID, MCA Tag Symbol and MCA Tag Name item values shall be set according to Table 4.8.

Table 4.8. Common IAB Soundfield Label SubDescriptor Values.

Item	Value
MCA Label Dictionary ID	UL of IAB Soundfield defined in Table 4.9.
MCA Tag Symbol	'IAB'
MCA Tag Name	'IAB'

Table 4.9. IAB Soundfield Label.

Kind	Leaf
Name	IAB Soundfield
Symbol	IABSoundfield
UL	urn:smpte:ul:060E2B34.0401010D.03020221.00000000
Description	Identifies an IAB soundfield.

5.11 IAB Track File Descriptor Object Model

The object model for the descriptors in an IAB Track File is shown in Figure 1.

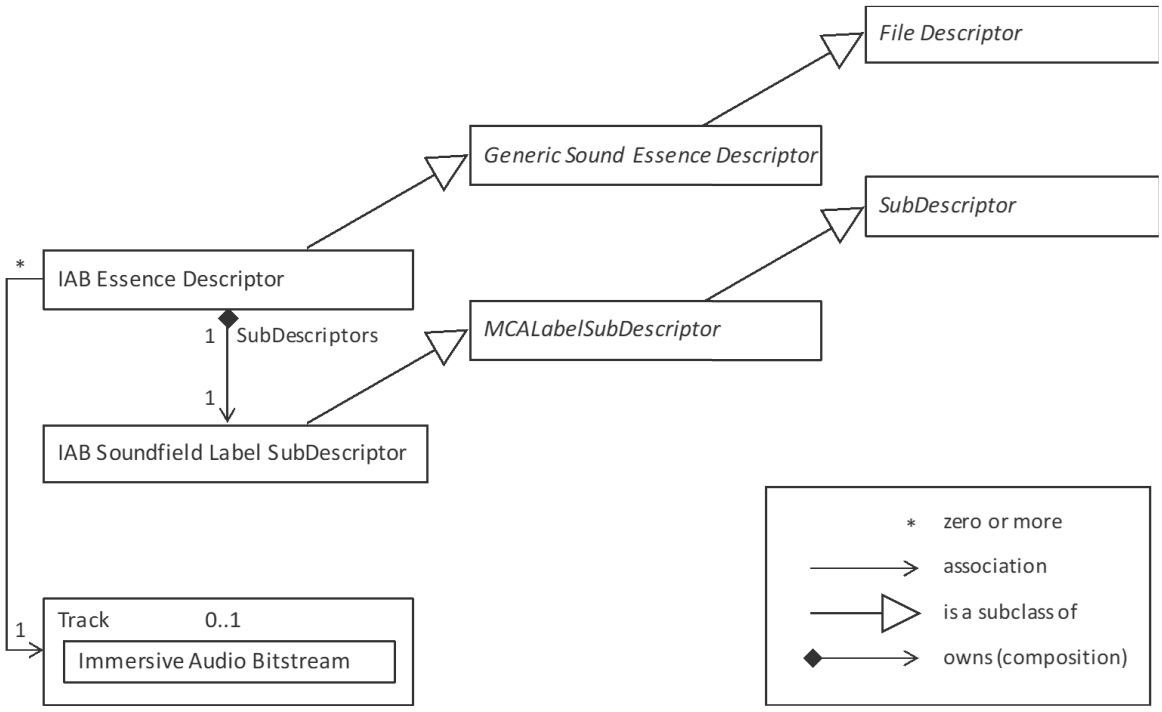

Figure 1 – IAB Track File Descriptor Object Model.

6 Composition

6.1 General

A Composition, as defined in SMPTE ST 2067-3, shall contain zero or more IAB Virtual Tracks.

6.2 IAB Virtual Track

An IAB Virtual Track consists of one or more instances of IABSequence elements as specified in Table 5.1.

Table 5.1. IABSequence element schema definition.

```
<xs:schema
  targetNamespace="http://www.smpte-ra.org/ns/2067-201/2019"
  xmlns:cpl="http://www.smpte-ra.org/schemas/2067-3/2016"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified"
  attributeFormDefault="unqualified">

<xs:import namespace="http://www.smpte-ra.org/schemas/2067-3/2016" />

<!-- schema definitions found in this document -->

<xs:element name="IABSequence" type="cpl:SequenceType"/>
</xs:schema>
```

Each IABSequence element shall contain Resource elements of type TrackFileResourceType, as defined in Section 6.12 of SMPTE ST 2067-3:2020, with each Resource element referencing an IAB Track File.

The IAB Virtual Track shall reference an IAB Track File whose Edit Rate, as defined in Section 5.4, is an integer multiple of the Edit Rate of the Main Image Virtual Track.

The EditRate of the Resource element, as defined in Section 6.11.3 of SMPTE ST 2067-3:2020, shall be set to a value equal to the Edit Rate of the Essence Track in the referenced MXF file, as defined in Section 5.4.

6.2.1 Homogeneous Essence

The following characteristics of the Immersive Audio Bitstream shall remain constant for all IAB Track Files referenced by a given IAB Virtual Track:

- BitDepth
- FrameRate
- SampleRate

Annex A Supplemental Informative Metadata (informative)

A.1 General

The following applies to informative metadata that does not alter rendering of the Immersive Audio Bitstream essence.

A.2 Internal Storage

XML-based informative metadata associated with the Immersive Audio Bitstream essence can be stored within the IAB Track File as specified in SMPTE RP 2057.

A.3 External Storage

Informative metadata associated with the Immersive Audio Bitstream essence can be stored separately from the IAB Track File by introducing additional Virtual Tracks to the Composition, as specified in SMPTE ST 2067-3.

Annex B ConformsToSpecifications Element Definition (Normative)

The ConformsToSpecifications element identifies the specification(s) to which an MXF file conforms.

NOTE 1 Other MXF Header Metadata items identify the defining specifications of specific components of the file. For instance, the Codec element of the File Descriptor identifies a codec compatible with an Essence Container. Similarly, the DM Schemes element of the Preface identifies the Descriptive Metadata Schemes used in the file.

The Preface Set specified in SMPTE ST 377-1 shall contain 0 or 1 instances of the ConformsToSpecifications element specified in Table B.1.

Table B.1. ConformsToSpecifications Element.

Item Name	Item Symbol	Type	Len	Local Tag	Item UL	Req?	Meaning	Default
Conforms To Specifications	ConformsToSpecifications	Batch of UL	var	dyn	06.0E.2B.34.01.01.01.0E.01.02.02.10.02.04.00.00	Opt	Identifies the specification(s) to which an MXF file conforms.	

The value of the ConformsToSpecifications element is a batch of Universal Labels (UL), where each UL element shall identify a specification to which the MXF file conforms in its entirety.

NOTE 2 A single file can conform to multiple specifications simultaneously.

NOTE 3 The absence of a particular UL in the ConformsToSpecifications element, or the absence of the element as a whole, is not itself an indication that the file does not conform with the specification associated with the UL.

Annex C IAB Soundfield Label SubDescriptor Definition (Normative)

C.1 General

The Immersive Audio Bitstream (IAB) defined in SMPTE ST 2098-2 carries all information (bed channels, objects and metadata) necessary to generate a Soundfield, as defined in SMPTE ST 377-4, and thus functions in a similar manner as a Soundfield Group, which carries all Audio Channels necessary to create a Soundfield. The IAB Soundfield Label SubDescriptor extends the MCA Label framework in accordance with Section 5.3 of SMPTE ST 377-4:2012 to provide labeling for immersive audio, and is comparable to the Soundfield Group Label SubDescriptor defined in that specification for channel-based audio.

C.2 Definition

The IAB Soundfield Label SubDescriptor is a concrete subclass of MCALabelSubDescriptor defined in Section 6.3 of SMPTE ST 377-4:2012, and contains the items specified in Table C.1 with the Set Key specified in Table C.2.

MCA Channel ID shall not be present in the IAB Soundfield Label SubDescriptor.

Table C.1. IAB Soundfield Label SubDescriptor.

Name	Type	Req	Definition
IAB Soundfield Label SubDescriptor Key	Set Key	Req	Identifies an IAB Soundfield Label SubDescriptor
Length	BER length	Req	Set length
All items in MCALabelSubDescriptor, as specified in SMPTE ST 377-4, except Set Key, Length, and MCA Channel ID.			

Table C.2. IAB Soundfield Label SubDescriptor Key.

Kind	Leaf
Name	IAB Soundfield Label SubDescriptor
Symbol	IABSoundfieldLabelSubDescriptor
UL	urn:smpte:ul:060E2B34.027F0101.0D010101.01017C00

NOTE The value of byte 6 (0x7F) in the IAB Soundfield Label SubDescriptor Key is a placeholder that is used in the registers, but is not used in actual implementations. Please consult SMPTE ST 377-1:2019, Section 9 for the proper value(s).

Annex D MXF-GC Immersive Audio and MXF-GC IAB Audio Label Definitions (Normative)

D.1 General

The MXF-GC Immersive Audio and MXF-GC IAB Audio labels are newly defined node labels under the MXF Generic Container node that provide for reasonable grouping of labels identifying essence container types related to immersive, object-based audio, generally, and SMPTE ST 2098-2 Immersive Audio Bitstream essence, specifically.

D.2 MXF-GC Immersive Audio Label Definition

The MXF-GC Immersive Audio label is a node-type label that is a child of the MXF Generic Container label (symbol: MXFGenericContainer). The MXF-GC Immersive Audio label shall be defined as specified in Table D.1.

Table D.1. MXF-GC Immersive Audio Label.

Kind	Node
Name	MXF-GC Immersive Audio
Symbol	MXFGCImmersiveAudio
UL	urn:smpte:ul:060E2B34.0401010D.0D010301.021D0000
Definition	Identifiers for MXF-GC Mappings of immersive audio data

D.3 MXF-GC IAB Audio Label Definition

The MXF-GC IAB Audio label is a node-type label that is a child of the MXF-GC Immersive Audio label (symbol: MXFGCImmersiveAudio). The MXF-GC IAB Audio label shall be defined as specified in Table D.2.

Table D.2. MXF-GC IAB Audio Label.

Kind	Node
Name	MXF-GC IAB Audio
Symbol	MXFGCIABAudio
UL	urn:smpte:ul:060E2B34.0401010D.0D010301.021D0100
Definition	Identifiers for MXF-GC Mappings of Immersive Audio Bitstream (SMPTE ST 2098-2) audio data