

IMMERSIVE AUDIO BITSTREAM (IAB) : THE PARENT OF THE SPATIALIZED SOUND

PREFACE

Immersive Audio Bitstream (known as IAB) is a standard for **spatial audio technologies**.

It aims to be the common basis of all spatial audio technologies. This standard avoids several distinct and incompatible standards that each manufacturer will have to maintain over time. It aims to be "future-proof".

The spatial audio technologies relying on the IAB standard :

- [Dolby Atmos](#) : the leader: almost 80% of the market.
- [Barco Auro](#) : approximately 15%.
- [DTS:X](#) : approximately 5%.

For example, the **Atmos** standard is based on the IAB standard and adds its own rules on top to create the **IAB-compliant Dolby Atmos**.

- For more information on **Dolby Atmos**, refer to the chapter [Dolby Atmos](#).
- For more information on **Barco Auro** and **DTS-X**, refer to the chapter [Barco Auro & DTS-X](#) (a bit empty for now)

The main purpose of the IAB standard is to ensure that all IAB assets can be played on any device that supports the IAB specifications. Thus, a Dolby Atmos asset could theoretically be played on Barco Auro or DTS-X devices (and vice versa).

Also, it prevents to distribute different versions of the same movie. Imagine creating different versions of MXF for Atmos, Auro, and DTS-X (or even much more if more manufacturers join the party) in addition to all DCP features. Here, we only have a single IAB version. The hardware in booth will adapt the playback process according to its own capabilities.

This standard is relatively new (initiated in 2013, published in 2018). Here is the first documents :

- **SMPTE 2098-1:2018** - Immersive Audio Metadata
- **SMPTE 2098-2:2018** - Immersive Audio Bitstream Specification
- **SMPTE 2098-5:2018** - Immersive Audio Channels and Soundfield Groups.

Not all devices may support the **Immersive Audio Bitstream**; However, the entire industry is pushing for widespread adoption and a large migration toward this new audio format.

TECHNICAL STANDARDS AND SPECIFICATIONS

Immersive Audio (IAB) has a common data structure defined in the following standards :

- The basis :
 - [SMPTE 429-18 - Immersive Audio Track File](#)
 - [SMPTE 429-19 - DCP Operational Constraints for Immersive Audio](#)
 - [SMPTE 2098-1 - Immersive Audio Metadata](#)
 - [SMPTE 2098-2 - Immersive Audio Bitstream Specification](#) : la plus importante, elle décrit le format IAB.

- [SMPTE 2098-5 - Immersive Audio Channels and Soundfield Groups](#)
- [SMPTE RDD 57 - Immersive Audio Bitstream & Packaging Constraints - IAB Application Profile 1](#)
- Addons :
 - [SMPTE 430-1 Addendum 2019 - Immersive Audio KDM](#) : Addition for integrating IAB into KDMs.
 - SMPTE RDD-52 - DCP Bv2.1 Application Profile
 - SMPTE 430-12 - FSK Synchronization Signal
 - SMPTE 2067-201 - Immersive Audio Bitstream Level 0 Plug-in
 - SMPTE EG-2098-3 - Immersive Audio Renderer Expectations & Testing Recommendations
 - [ISDCF #15: SMPTE 2098-2 Immersive Audio Bitstream Constraints - IAB Application Profile 1](#)

You can also use a search engine with the keyword [SMPTE Immersive Audio](#) to find any new publications.

TECHNICAL SPECIFICATIONS OF IAB

The IAB Constraint standard ^[SMPTE:2098-2] gives the following technical constraints :

- SampleRate at 48 kHz ¹
Note that que **SMPTE 2098-2:2018 - IAB** standard also provides 96 kHz; However, the **SMPTE RDD-57:2021 - IAB Constraints** standard requires using only 48 kHz for now ².
Assume support for both SampleRates :)
- Sample Bit Depth : 24 bits
- Sample Rate available: 24, 25, 30, 48, 50, 60 fps ³
(96, 100, 120 fps in the future or according to the IAB-DolbyAtmos standard).
SampleRate is linked to the global FrameRate (Picture)
- Maximum BedChannels number: 10 ⁴
(up to 64 but only on a Dolby Atmos master ⁵)
- Maximum Objects number: 118 ⁶
- Maxium Rendered Field (Bed + Objects): 128 ⁷
- At least 1 BedChannel per frame
- AudioData Elements: only AudioDataDLC elements (AudioDataPCM element is forbidden)

MXF

For the **Dolby Atmos** MXF following the Immersive Audio Bitstream standard, you can find an analysis of **IAB** in the chapter [Dolby Atmos](#) and the chapter [MXF Dolby Atmos](#).

Concerning the main audio MXF file ([MainSound](#)), if your DCP contains a **Immersive Audio** MXF :

- You can find a **FSK Sync Signal** ⁸ (synchronization signal) on the [track #14](#), it used for the sound sync with the external audio device (such as Atmos hardware or DTS-X device)
- [Other sound channels](#) are not used (except for additional data such as hard-of-hearing/deaf tracks (HI, VIN and SLV), the **Motion Data** track (used by DBOX for example)
- All reels containing an Immersive Audio track must include an FSK Sync signal in one of the sound tracks within a [MainSound](#) MXF.
- The [MainSound](#) MXF must not contain any **Immersive Audio Bitstream** data. The main sound channels are empty and must not contain PCM audio data, except the data for the channels HI, VIN and Motion Data. There is therefore no audio fallback system if the Immersive Audio don't work correctly. The MainSound will use only to store additional data (eg. hard-of-hearing/deaf tracks, DBOX, Sync Signal).

MXF ANALYSIS

You can find a complete analysis of a IAB implementation using the Dolby Atmos specifications in the chapter [MXF Dolby Atmos](#)

You have an Immersive Audio Bistream parser [available here](#) and [IAB assets here](#).

To use it, just execute `iab-reader` on a IAB frame (such as [this frame](#)) or directly on a [directory](#) containing several IAB frames :

```
+++ atmos-stresstest-64objects.iab/*-ImmersiveAudioDataElement.value.bin
|----- IABitstream Frame -----| [1]
|----- Preamble -----| [1]
| PreambleTag : 0x01
| PreambleLength : 1603 bytes
| PreambleValue : 11010effffffffffffffffffff(...1571 bytes...)ffffffffffffffffffffffffffff
|----- IAFrame -----| [1]
| IAFrameTag : 0x02
| IAFrameLength : 44501 bytes
| IAFrameValue : 08ffadd10110529bfff020e640(...44469 bytes...)ffc1bfff03000000200000000500
|----- IAElement -----| [1]
.
.
.
(....)
```

CPL

In **CPL**, **ContentTitle** must include the `IAB` code in its **AudioType** field. See the chapter [Digital Cinema Naming Convention](#).

IAB reuses the concept and the syntax of the AssetType [AuxData](#) with a specific **DataType** for Immersive Audio (**Immersive Audio Data Essence Coding**):

```
<!-- Universal Label for Immersive Audio Data Essence Coding -->
<DataType>urn:smp:ul:060e2b34.04010105.0e090604.00000000</DataType>
```

And add (not mandatory) an IAB **ExtensionMetadata** in the [CompositionMetadataAsset](#) section within the **ExtensionMetadataList** section :

```
<CompositionMetadataAsset>
  <!-- (...) -->
  <ExtensionMetadataList>

    <ExtensionMetadata scope="http://isdcf.com/ns/cplmd/app">
      <Name>Application</Name>
      <PropertyList>
        <Property>
          <Name>IAB Profile</Name>
          <Value>SMPTE-ST-2098-2:2019-P1</Value>
        </Property>
      </PropertyList>
    </ExtensionMetadata>

  </ExtensionMetadataList>
  <!-- (...) -->
</CompositionMetadataAsset>
```

By default, if we have an IAB without an **IAB ExtensionMetadata** section, then it is considered an **IAB Profile 1**.

KDM

For the needs of the IAB, [SMPTE 430-1 KDM](#) standard was supplemented with an [Addendum in 2019](#) to integrate the Immersive Audio specifications.

The [Dolby Documentation](#) states that IAB must be encrypted with a key of type **MDEK**, originally used to

designate an **AuxData Essence Key** and now also used to designate an **Immersive Audio Essence Key**.

For the older players, you may probably use the [ForensicMarkFlag](#) to disable all watermark on the channel audio track #12 and higher (DBOX channel 13, Atmos channel 14) :

```
<ForensicMarkFlag>http://www.dcmovies.com/430-1/2006/KDM#mrkflg-audio-disable-above-channel-12</ForensicMarkFlag>
```

Or even disable all watermarks on the sound tracks (not recommended) :

```
<ForensicMarkFlag>http://www.smp-te-ra.org/430-1/2006/KDM#mrkflg-audio-disable</ForensicMarkFlag>
```

You can find further informations in the chapter [KDM - Authenticated Public](#), paragraphs **KeyType** and **ForensicMarkFlag**

KDM: MULTI-HARDWARE MANAGEMENT

It may be necessary to generate multiple KDMs to support heterogeneous hardware management:

- one KDM for the DCP player to handle its data assets (Picture, Sound, Subtitle),
- and another for the IAB-compatible audio hardware to handle its data assets (AuxData-IAB).

For example, in a case where the DCP player is a Dolby/Doremi device and the audio hardware is a DTS:X device, and they are not in the same technical "block" but can communicate with each other.

The DCP player will read keys of type MDIK (Image), MDAK (Audio), and MDSK (Subtitle), while ignoring MDEK keys (AuxData/IAB). Conversely, the audio hardware will ignore all keys except for MDEK keys (AuxData/IAB).

REFERENCES

- [Excellent article about Immersive Audio Bitstream, its history and its technical aspect](#) written by C.J. Flynn for [CelluloidJunkie](#)
- [Brief technical documentation on the Dolby Cinema](#) (in japanese)
- [Excellent SMPTE Webcast of Steve Llamb about IAB](#)

NOTES

1. « *The value of all instances of Sample Rate, as defined by SMPTE ST 2098-2, shall be 48 kHz* » -- SMPTE.RDD.57-2020 - Immersive Audio Bitstream and Packaging Constraints - IAB Application Profile 1 [↩](#)
2. Supports and Constraints for 48 kHz / 96 kHz : [↩](#)
 - « *The compression system supports both 48 kHz and 96 kHz sample rates.* » -- SMPTE 2098-2:2018 - Immersive Audio Bitstream Specification
 - « *AudioDataDLC Element : The AudioDataDLC element supports sample rates of 48 kHz or 96 kHz with 24-bit resolution.* » -- SMPTE RDD-29:2019 - Dolby Atmos Bitstream.
 - « *ST 2098-2 Bitstream Constraints : Sample Rate : The value of all instances of Sample Rate, as defined by SMPTE ST 2098-2, shall be 48 kHz.* » -- SMPTE RDD-57:2021 - Immersive Audio Bitstream and Packaging Constraints - IAB Application Profile 1

These constraints are likely due to the hardware, which must support 48 kHz and could potentially support 96 kHz.

3. « *The only values that shall be used for IAFramerate, as defined in SMPTE ST 2098-2, are: 24/1, 25/1, 30/1, 48/1, 50/1, or 60/1.* » -- SMPTE RDD.57-2021 - Immersive Audio Bitstream and Packaging Constraints - IAB Application Profile 1 [↩](#)

With a Maximum Frame Size for each IAB frame according to the EditRate :

EditRate	Maxium size
24	781.250 octets
25	750.000 octets
30	625.000 octets
48	390.625 octets
50	375.000 octets
60	312.500 octets
96	195.313 octets
100	187.500 octets
120	156.250 octets

4. « *The maximum number of bed channels, as defined in SMPTE ST 2098-2, shall be 10* » -- SMPTE.RDD.57-2020 - Immersive Audio Bitstream and Packaging Constraints - IAB Application Profile 1 [↩](#)
5. « *64: Maximum number of audio beds* » -- Dolby Atmos Master ADM Profile [↩](#)
6. « *The maximum number of objects, as defined in SMPTE ST 2098-2, shall be 118* » -- SMPTE.RDD.57-2020 - Immersive Audio Bitstream and Packaging Constraints - IAB Application Profile 1 [↩](#)
7. « *The MaxRendered field (which is the sum of objects and bed channels) of the IAFFrame, both as defined in SMPTE ST 2098-2, shall have a value of 128 or less.* » -- SMPTE.RDD.57-2020 - Immersive Audio Bitstream and Packaging Constraints - IAB Application Profile 1 [↩](#)
8. [SMPTE 430-12](#) - D-Cinema Operations - FSK Synchronization Signal [↩](#)