

SMPTE Education
SMPTE - Enabling Global Education



SMPTE ST 2098-2 Immersive Audio Bitstream – An overview

Welcome!

2018-08-16

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

SMPTE Standards Update Webcasts



SMPTE - Enabling Global Education

- Series of quarterly 1-hour online (this is 90 minutes), interactive webcasts covering select SMPTE standards
- Free to everyone
- Sessions are recorded for on-demand viewing convenience SMPTE.ORG and YouTube
- Please tell others about SMPTE's webcasts on social media:

@smpteconnect

#SMPTEWebcast

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Your Host

Joel E. Welch

*Director of Education
SMPTE*



SMPTE - Enabling Global Education

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

3

SMPTE Education Webcast Series
SMPTE - Enabling Global Education



Today's Guest Speaker

Stan Cossette

*Senior Staff Engineer
Dolby Laboratories*



© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

SMPTE Education
SMPTE - Enabling Global Education



SMPTE ST 2098-2 Immersive Audio Bitstream – An overview

Presented By Stan Cossette

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Presentation Overview



SMPTE - Enabling Global Education

- Background and Terms
- High-level bitstream frame and element structure
- Hierarchy of the bitstream frame elements
- Audio Object coordinate system
- Extensibility, backward compatibility, and efficiency mechanisms designed into the standard
- Examples of bitstream usage
- Bitstream packaging in Cinema and IMF

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Background and terms



SMPTE - Enabling Global Education

SMPTE 25CSS Immersive Audio Suite

ST 2098-1 Immersive Audio Metadata

ST 2098-2 Immersive Audio Bitstream

EG 2098-3 Immersive Audio Renderer Expectations (in WD development)

RP 2098-4 Immersive Audio Renderer Interoperability Testing Procedure (in WD development)

ST 2098-5 D-Cinema Immersive Audio Channels and Soundfield Groups

Scope of ST 2098-2 is generic ‘...defines a coded representation (bitstream) of an audio program’, but the intended application is Digital Cinema. This affected:

- Choice of supported channels
- Choice of supported speaker configurations
- Assumptions regarding playback environment

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Background and terms



SMPTE - Enabling Global Education

Immersive Sound vs Surround Sound

From ST 2098-5:

Immersive Sound

Includes sound that emanates from sources at and beyond the Base Layer.

Base Layer

Refers to the nominally horizontal layer of Loudspeakers used in 5.1 and 7.1 Soundfield Configurations (*arrangement of speakers intended to reproduce a specific soundfield*).

So **Surround Sound** uses the Base Layer only, while **Immersive Sound** uses the Base Layer and additional (typically height) layers or speakers. Also, note that **Immersive Audio** does not have to use **Audio Objects**

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)



SMPTE - Enabling Global Education

Background and terms

Audio Channel

Distinct collection of sequenced audio samples that are intended for delivery to a single loudspeaker, loudspeaker array or other reproduction device

Audio Object

Segment of audio essence with associated metadata describing positional and other properties which may vary with time

Bed

Soundfield Group (*group of channels associated with a Soundfield*), such as a 5.1, 7.1 or 9.1, that serves as the foundation of the immersive soundtrack mix

Target Environment

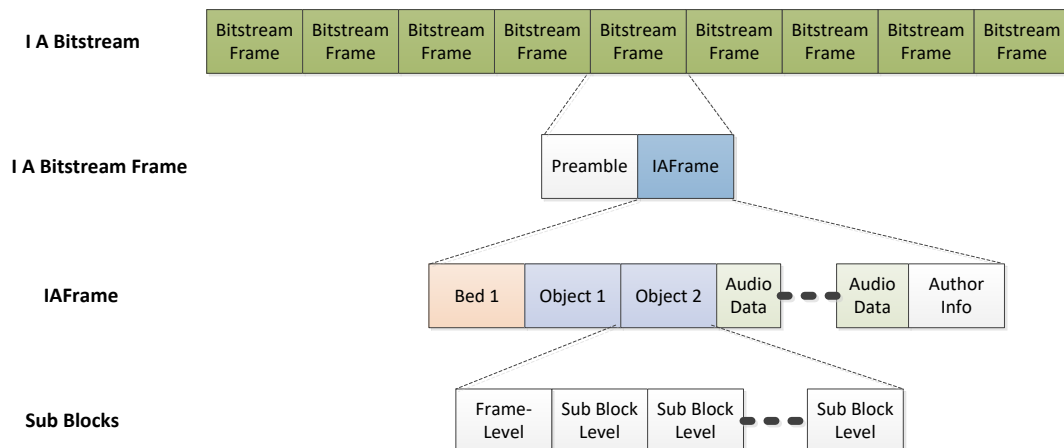
Specific set of conditions that is present in the playback environment

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Bitstream Structure



SMPTE - Enabling Global Education



© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Basic Bitstream Element



SMPTE - Enabling Global Education

| IAElement |
|---------------------------------|
| ElementID |
| ElementSize |
| (Contents – based on ElementID) |

- ElementID – Identifies the element and its syntax. Decoders may skip unrecognized IDs
- ElementSize – Can be used to skip an unknown element
- The combination of these two items allows extensibility
- All bitstream elements are based on this basic structure

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

IAFrame



SMPTE - Enabling Global Education

| IAFrame |
|-----------------|
| ElementID |
| ElementSize |
| BitDepth |
| FrameRate |
| MaxRendered |
| SubElementCount |
| SubElement 1 |

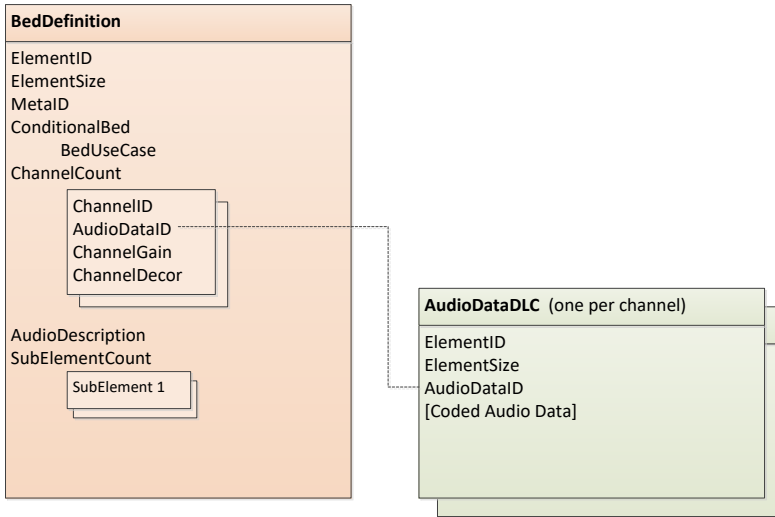
- ST 2098-2 specifies a single IAFrame without constraint on adjacent bitstream frame content
- Each IAFrame is independently decodable
- Applications can constrain bitstreams to contain IAFrames with consistent parameters

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Bed Structure (simplified)



SMPTE - Enabling Global Education



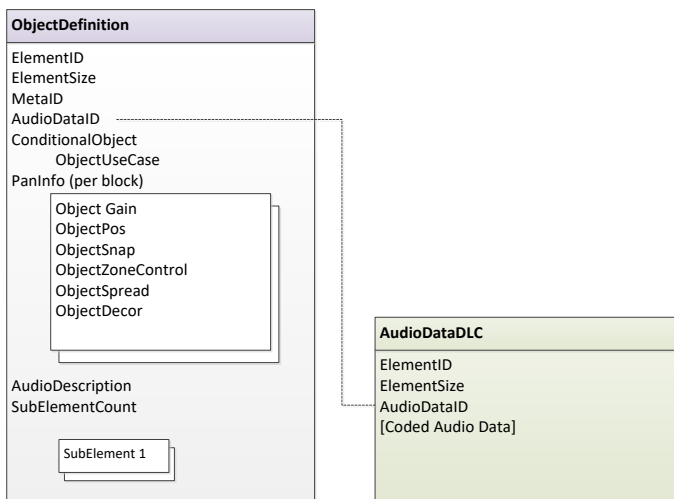
- Forms the basis of the mix
- ChannelID determines routing destination
- ChannelCount and ChannelID not constrained
- Each Channel References an AudioData element

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Object Structure (simplified)



SMPTE - Enabling Global Education



- Is not routed, but 'rendered'
- Each ObjectDefinition element references audio essence carried as PCM or DLC using an AudioDataID
- Supports rendering modifications: Spread (size), Snap, Zone exclusion and Decorrelation
- Metadata can be updated on a sub block basis (typically 8 per frame or around every 5ms)

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)



SMPTE - Enabling Global Education

Object Dynamic Metadata (PanInfo)

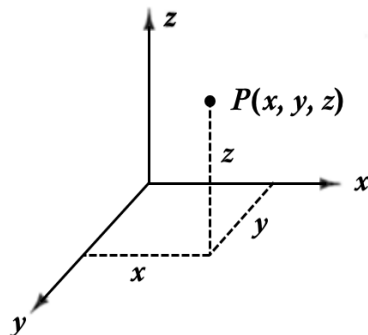
- Object Gain – Defines gain for a sub block
 - ObjectPos – Defines position intent for the sub-block
 - ObjectSnap – Defines importance of Timbre vs Position
 - ObjectZoneControl – Defines extent of use of zones
 - ObjectSpread – Defines object 'size' intent
 - ObjectDecor – Defines decorrelation effect intent
-
- At least one sub block per frame
 - Decoder uses previous PanInfo if current block has none

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Object Coordinate System and Frame of Reference



SMPTE - Enabling Global Education



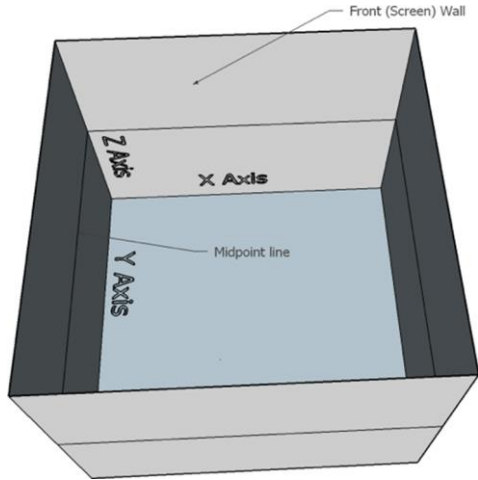
- Positional metadata is Allocentric and uses a Cartesian coordinate system
- Positional reference is to walls of a unit cube
- The origin is at the intersection of the front (screen) wall and the left wall at the height of the Base Layer

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Cube and reference points



SMPTE - Enabling Global Education



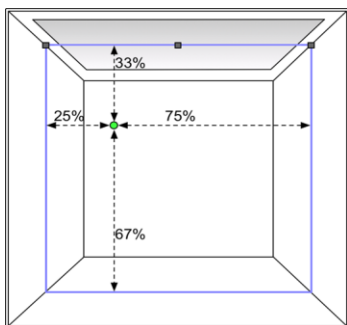
- 'Midpoint line' is at Z=0 and is height of Base Layer speakers
- Left and Right are relative to a viewer facing the screen
- All walls (except floor) are located at 0 or 1

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

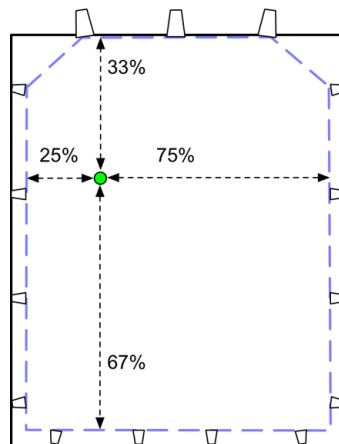
Mapping location to a cinema



SMPTE - Enabling Global Education



Position $[x,y,z] = [0.25, 0.33, 0]$



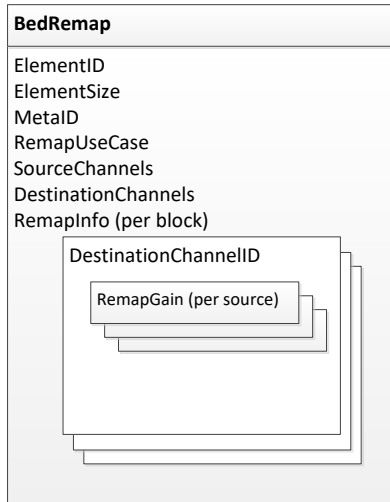
- Figure on left shows idealized location
- Figure on right shows mapping example
- Cinema layout affects precise location, but mapping preserves intent

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Bed Remap Structure (simplified)



SMPTE - Enabling Global Education



- Allows user to create optional re-mix of existing bed to be used under a certain playback condition (Target Environment)
- Remix is created using a matrix with each new output channel created by summing the existing bed channels with a gain parameter for each.

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

New Elements



SMPTE - Enabling Global Education

| ObjectZoneDefinition19 |
|------------------------|
| ElementID |
| ElementSize |
| ZoneInfo (per block) |
| ZoneGain (per zone) |

| AuthoringToolInfo |
|-------------------|
| ElementID |
| ElementSize |
| AuthoringToolURI |

| UserData |
|---------------|
| ElementID |
| ElementSize |
| UserID |
| UserDataBytes |

ObjectZoneDefinition19 – Defines zones for use with the Auro speaker configs. If used, it only replaces the existing ObjectZoneControl metadata

AuthoringToolInfo – Carries a URI to identify the authoring tool

UserData – Carries unspecified data whose syntax is indicated by the UserID (not used in Cinema)

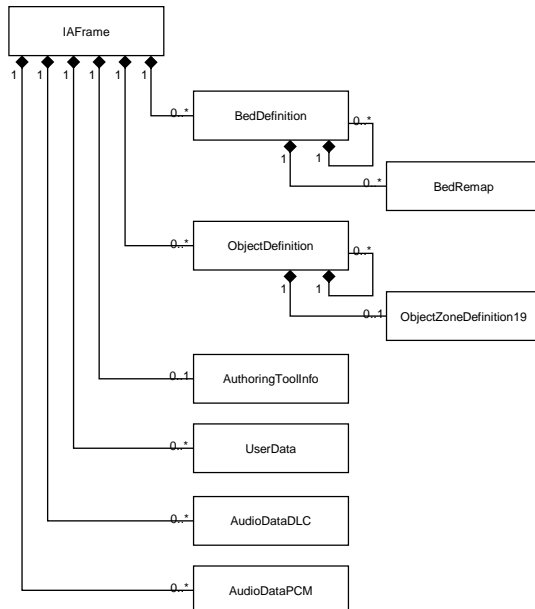
© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Backward Compatibility

- Many elements based on existing bitstreams (Dolby Atmos®)
- No new required elements
- No change to syntax for existing elements (although some new codes were added)
- Decoders can skip unrecognized elements
- Result: Existing bitstreams are compliant, but new bitstreams may not work properly on existing decoders

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Element Hierarchy



- IAFRames are composed of the shown elements
- All elements are optional
- Some elements can contain sub-elements
- Upper-level elements have Parent-Child relationship with the lower-level element
- The Child will replace all or a portion of the Parent, if used

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Conditional Bitstream Elements



SMPTE - Enabling Global Education

- Allows bitstream creator to supplement the bitstream for alternate Target Environments
- Always associated with a UseCase code which represents a specific playback condition
- A conditional element can be a Child or a Top-Level element
- A conditional Child element replaces the Parent (or a part)
- The renderer will determine the UseCase based on its configuration, which is a one-time setup
- Only rendered if renderer configuration matches UseCase

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Conditional Bitstream Elements



SMPTE - Enabling Global Education

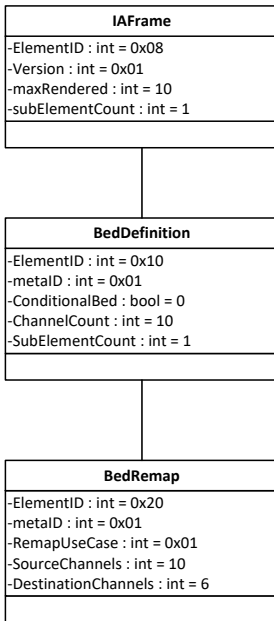
- **Conditional Bed**
 - Replaces parent Bed with a new mix (new audio essence)
- **Conditional Object**
 - Replaces parent Object with a substitute Object (may use new audio essence)
- **BedRemap**
 - Remixes parent bed channels to create alternate mix

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)



SMPTE - Enabling Global Education

Examples - BedRemap



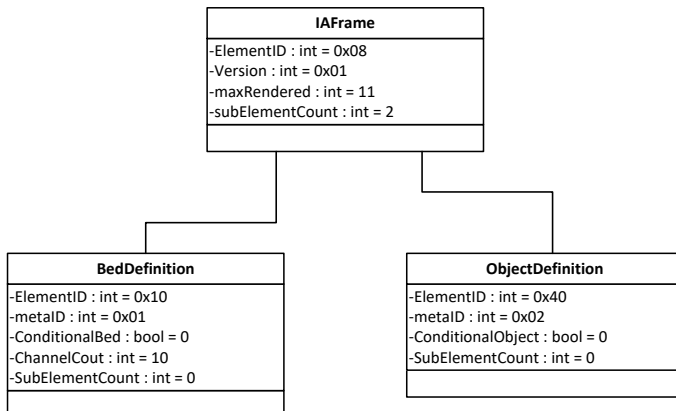
- Diagram shows element relationships and hierarchy: Lower elements are contained in upper (parent) elements
- RemapUseCase = 1 (5.1). Note SourceChannel count and DestinationChannel count.
- BedRemap is used IFF the renderer configuration indicates a 5.1 playback environment (not an immersive audio system)

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Examples – Typical Stream



SMPTE - Enabling Global Education



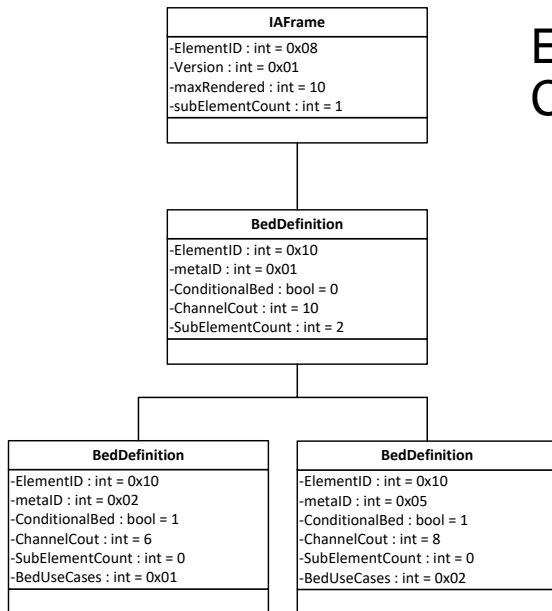
- Typical element structure of Bed plus Object(s)
- Active elements are rendered simultaneously
- Neither element is conditional, so both will always be rendered
- MaxRendered is the total of bed channels plus objects

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Examples – Conditional Bed



SMPTE - Enabling Global Education



- Conditional Child Beds will replace the parent under a specific UseCase
- Child Beds carry additional essence i.e., a different mix intended for the Target Environment
- The bitstream standard constrains the number of child elements
 - Does not allow Children of Children
 - MaxRendered cannot exceed 128 (at 48kHz) or 64 (at 96kHz)

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Coding Efficiency



SMPTE - Enabling Global Education

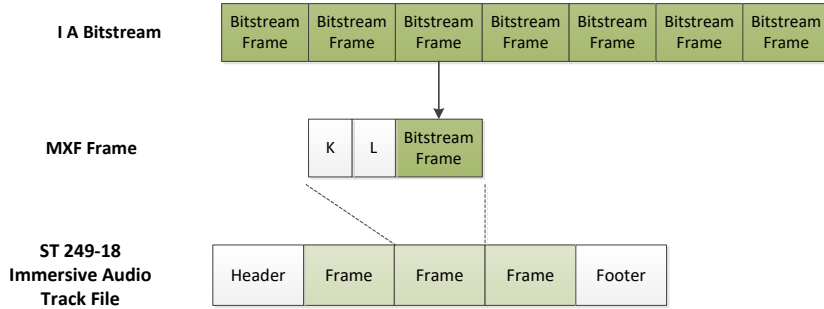
- Metadata and essence only present in a frame if needed in that frame. Objects need not persist if they are not used
- Sub Blocks are only used when needed to carry metadata that changes over a frame
- Audio Essence can be losslessly coded (2:1 typical)
- In each frame, for each channel or object, silence can be carried as metadata (AudioDataID = 0)
- Anecdotally: IAB track file approx. equal to 14-chan Main Audio

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Cinema Packaging (in progress)



SMPTE - Enabling Global Education

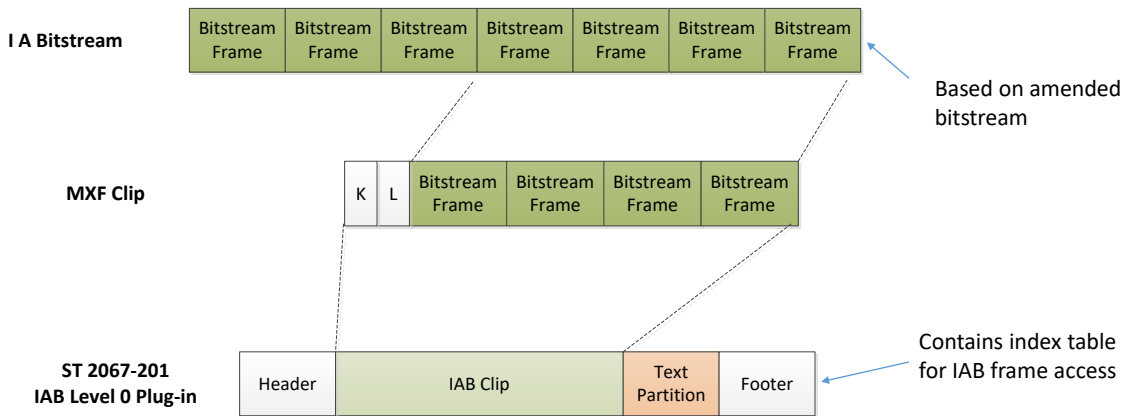


© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

IMF application (in progress)



SMPTE - Enabling Global Education



© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Bitstream Standard Summary



SMPTE - Enabling Global Education

- Efficient, flexible, extensible standard that supports the features needed to convey immersive audio essence and rendering intent
- Backward compatible with existing immersive audio content (Dolby Atmos®)
- File transport will soon be supported by SMPTE standards for D-Cinema Packaging and IMF

© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)

Questions Please! (Verbal Questions Take Precedence)



SMPTE - Enabling Global Education

Stan Cossette

*Senior Staff Engineer
Dolby Laboratories*



Joel E. Welch



© 2018 by the Society of Motion Picture & Television Engineers®, Inc. (SMPTE®)