

**SONY
PICTURES**

Sony Pictures Entertainment

Master and Archive Specifications

Version	Date	Name	Action	File Location
1.04	Nov, 2015	David Deelo	Release: version	local
1.05	July, 2018	David Deelo and Brian Vessa	Release: version	local
1.06	May, 2020	David Deelo and Brian Vessa	Release: version	local
1.07	July, 2021	Asset Stewardship and Data Preservation Council	Release: version	local
1.07.1	October, 2021	Asset Stewardship and Data Preservation Council	Release: version	local
1.07.1	July, 2024	David Deelo	Release: version	local

TABLE OF CONTENTS

OVERVIEW6

TELEVISION ELEMENT SPECIFICATIONS7

TELEVISION PICTURE MASTER SPECIFICATIONS.....7

 Color Timed Master (CTM).....7

 UHD Standard Dynamic Range7

 HD Standard Dynamic Range.....7

 High Dynamic Range Distribution Master (HDRDM)8

 Academy Color Encoding System (ACES)8

 Video Assembly Master (VAM)9

TELEVISION AUDIO MASTER SPECIFICATIONS.....10

 Broadcast-Level Audio Definition.....10

 Full-Range Audio Definition.....10

 General Television Audio Specifications10

 Television Audio Element Specifications10

 Printmasters (PM):.....10

 Stems (SM) (a.k.a. printmaster stems) (SM):10

 Music and Effects (ME) and M&E Optionals (OP) Supersession (MESP):10

 Music and Effects Stems (MESM):11

 Fully-Filled Effects Stem (FFXSM):11

 DME “split track” (DME):11

 Predubs (PD):11

 Note On Two-Track Audio.....11

THEATRICAL ELEMENT SPECIFICATIONS12

THEATRICAL DCP, DCDM AND DKDM12

 Digital Cinema Package (DCP)12

 Digital Cinema Distribution Master (DCDM)12

 Distribution Key Delivery Message (DKDM).....12

 Textless, Text Files, Title and Spotting List Materials12

THEATRICAL PICTURE SPECIFICATIONS14

 Digital Intermediate (DI)14

 Digital Source Master (DSM)14

 Academy Color Encoding System (ACES)15

 YCM Color Separation.....16

 Flash-cut Negative.....16

THEATRICAL-LEVEL AUDIO MASTER SPECIFICATIONS.....17

 Theatrical-Level Audio Definition.....17

 General Theatrical-Level Audio Specifications.....17

 Theatrical-Level Audio Element Specifications17

 Printmasters (PM):.....17

 Stems (SM) (a.k.a. Printmaster Stems) (SM):17

 Music and Effects (ME) and M&E Optionals (OP) Supersession (MESP):18

 Music and Effects Stems (MESM):18

 Fully-Filled Effects Stem (FFXSM):18

 DME “split track” (DME):18

 Predubs (PD):18

 Note On Two-Track Audio19

FEATURE HOME MASTERING SPECIFICATIONS20

FEATURE HOME MASTERING PICTURE SPECIFICATIONS.....20

 High Dynamic Range Distribution Master (HDRDM)20

 Color Timed Master (CTM).....20

 UHD Standard Dynamic Range20

 HD Standard Dynamic Range.....21

HOME THEATER AUDIO MASTER SPECIFICATIONS.....22

 Home Theater Audio Definition.....22

 General Home Theater Audio Specifications22

Home Theater Audio Element Specifications.....	22
Printmasters (PM):.....	22
Stems (SM) (a.k.a. Printmaster Stems) (SM):.....	22
Music and Effects (ME) and M&E Optionals (OP) Supersession (MESP):	22
Music and Effects Stems (MESM):.....	23
Fully-Filled Effects Stem (FFXSM):.....	23
DME “split track” (DME):.....	23
Home Theater Audio Specification Notes.....	23
Note On Two-Track Audio	23
RESTORATION AND REMASTERING SPECIFICATIONS.....	24
Raw Scan Clone (RSC).....	24
Intermediate Picture Data Clone (IPDC)	24
Source Files (SRC)	24
Final Picture Data Clone (FPDC).....	24
REFERENCE PICTURE SPECIFICATION.....	25
Video / Audio Formatting	25
PRODUCTION ARCHIVE SPECIFICATIONS.....	26
Camera files, dailies and visual effects (VFX) materials.....	26
Editorial drives and related process metadata.....	26
All script material.....	26
Wrap Drive.....	26
Picture Editorial Delivery Methodology.....	26
Audio archive:.....	26
PHYSICAL ARCHIVE SPECIFICATIONS.....	28
LTFS Archive	28
Standard Metadata Block.....	28
APPENDIX A –CONTENT FORMATTING GUIDELINES.....	32
Overall Content Formatting (For All Content Types)	32
Scaling Requirements.....	32
Directory & File Naming	32
General:.....	32
Directory structure:.....	33
Audio File naming.....	33
Picture File naming.....	33
Formatting For Long-Form Content.....	33
Head Format:.....	33
Tail Format:.....	34
Textless Formatting:.....	34
Formatting For Reel-Based Audio Content.....	34
Reel-Based Audio File Formatting.....	35
APPENDIX B–EPISODIC / MOW /MFT MASTER CONTENT FORMATTING.....	36
EPISODIC SPECIFIC FORMATTING.....	36
APPENDIX C –AUDIO TRACK LAYOUTS.....	37
Track Layouts by Audio Configuration.....	37
Commonly Used Audio Configurations.....	37
Legacy Audio Configurations for ProTools Sessions	38
Legacy Audio Configurations for standard DCP	38
Stems:.....	39
5.1:.....	39
7.1:.....	39
Atmos:	39
Music and Effects Supersessions (MESP):.....	40
5.1 MESP:.....	40
7.1 MESP:.....	40
LtRt MESP.....	40
Atmos MESP:.....	41
M&E Stems (MESMSP):.....	42
5.1:.....	42
7.1:.....	42
Atmos:	42
Audio Layouts for MOV Containers.....	44

Audio Channel Layout for MOV container for Feature content	44
Audio Channel Layouts for MOV container for Episodic content only:	44
APPENDIX D –MUSIC AND EFFECTS STEMS SPECIFICATIONS.....	45
<i>PREFACE:</i>	45
<i>DESCRIPTION:</i>	45
<i>SPECIFICATIONS:</i>	45
<i>FORMATTING:</i>	46
<i>DELIVERY:</i>	46
<i>FILE NAMING AND GOLD VALUES:</i>	46
APPENDIX E –GLOSSARY OF TERMS.....	48
<i>PICTURE-RELATED TERMS</i>	48
3DDCDM: 3D Digital Cinema Distribution Master	48
3DDCP: 3D Digital Cinema Package	48
CTM: Color Timed Master.....	48
DCDM: Digital Cinema Distribution Master.....	48
DCP: Digital Cinema Package.....	48
DI: Digital Intermediate.....	48
DSM: Digital Source Master.....	48
FPDC: Final Picture Data Clone	49
HDRDM: High Dynamic Range Digital Master	49
IPDC: Intermediate Picture Data Clone.....	49
Production Archive -	49
RSC: Raw Scans Clone.....	49
Source Files (SRC).....	50
VAM: Video Assembled Master.....	50
<i>AUDIO ELEMENTS</i>	50
DME (Dialogue Music and Effects) (a.k.a. split track)	50
Fully-Filled Effects Stem (FFX)	50
Music and Effects (ME) (a.k.a. Main M&E)	50
Music and Effects Stems (MESM).....	50
Music and Effects Supersession (MESP).....	50
Optional (OP) (a.k.a. M&E Optional).....	50
Predub (PD).....	51
Printmaster (PM).....	51
Stem (SM) (a.k.a. Printmaster Stem)	51
Wide Stems.....	51
<i>AUDIO-TERMINOLOGY AND FORMATS</i>	51
Audio Channel (a.k.a. “Channel”).....	51
Audio Configuration (a.k.a. Soundfield Group).....	51
Dual Mono (DM)	51
IMAX 5.0.....	51
IMAX 12.0	51
IMAX Enhanced.....	51
LtRt (a.k.a. Stereo Surround).....	52
Monaural (M) (a.k.a. “Mono”).....	52
ProTools.....	52
Sony Dynamic Digital Sound (SDDS).....	52
Soundfield.....	52
Soundfield Configuration.....	52
Standard Stereo (ST) (a.k.a. LoRo)	52
5.1.....	52
7.1 (a.k.a. 7.1DS)	52
.wav (a.k.a. “Wave file”)	52
<i>AUDIO-IMMERSIVE AUDIO TERMINOLOGY</i>	52
ATMOS®.....	52
Audio Object (a.k.a. “Object”)	53
Immersive Audio.....	53
Immersive Audio Bed (a.k.a. “Bed”).....	53
Immersive Audio Bitstream (IAB).....	53
Immersive Audio Configuration.....	53
Immersive Audio Renderer	53
Immersive Sound.....	53
Immersive Soundfield Configuration.....	53

Immersive Sound System53
CONTACT INFO54

OVERVIEW

This document is divided into specifications and formatting requirements. The specifications are divided into sections depending on the content type being delivered. Formatting guidelines follow and are located in the Appendix section at the tail of the document.

TELEVISION ELEMENT SPECIFICATIONS

TELEVISION PICTURE MASTER SPECIFICATIONS

Color Timed Master (CTM)

UHD Standard Dynamic Range

DPX Encoding:

File Format: DPX (.dpx)
Bit Depth: 10-bit
Compression: none
Color Space: Rec. 709 full-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:4:4 (RGB)
Image Width: 3840 (UHD)
Image Height: 2160 (UHD)

ITU-R. 2035, specifies the viewing environment. Studio mastering reference display should be in compliance with ITU-R BT 1886 as follows:

Max Luminance: 100 cd/m² (nits)
Black Luminance: <0.1 cd/m² (nits)
EOTF: Gamma 2.4
Color Primaries (CIE 1931 x, y):
Red: 0.64, 0.33
Green: 0.30, 0.60
Blue: 0.15, 0.06
White Point (D65): 0.3127, 0.329

HD Standard Dynamic Range

DPX Encoding:

File Format: DPX (.dpx)
Bit Depth: 10-bit
Compression: none
Color Space: Rec. 709 full-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:4:4 (RGB)
Image Width: 1920 (HD)
Image Height: 1080 (HD)

ITU-R. 2035, specifies the viewing environment. Studio mastering reference display should be in compliance with ITU-R BT 1886 as follows:

Max Luminance: 100 cd/m² (nits)
Black Luminance: <0.1 cd/m² (nits)
EOTF: Gamma 2.4
Color Primaries (CIE 1931 x, y):
Red: 0.64, 0.33
Green: 0.30, 0.60
Blue: 0.15, 0.06
White Point (D65): 0.3127, 0.329

ProRes Encoding:

Codec: Apple ProRes 422 (HQ)
Image Width: 1920 (HD)
Image Height: 1080 (HD)
Frame Rate: Native
Data Rate: 220 Mbps VBR
Color Space: Rec. 709 legal-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:2:2 (YCbCr)
Color Depth: 10 Bit
Container: .MOV

Codec: Apple ProRes 4444 (no alpha)
Image Width: 1920 (HD)
Image Height: 1080 (HD)
Frame Rate: Native
Data Rate: 330 Mbps VBR
Color Space: Rec. 709 legal-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:2:2 (YCbCr) or 4:4:4 (YCbCr)
Color Depth: 10 Bit
Container: .MOV

High Dynamic Range Distribution Master (HDRDM)**TIFF Encoding:**

File Format: TIFF (.tiff)
Bit Depth: 16-bit
Compression: none
Color Space: Rec. 2020 full-range (ITU-R BT.2020-2)
EOTF: ST-2084 (PQ)
Chroma Sample: 4:4:4 (RGB)
Max Luminance: uncapped
Image Width: 3840 (UHD)
Image Height: 2160 (UHD)

ITU-R. 2035, specifies the viewing environment. Studio mastering reference display specifications are as follows:

Max Luminance: 4000 cd/m2 (nits)
Black Luminance: <0.005 cd/m2 (nits)
EOTF: ST-2084 (PQ)
Color Primaries (CIE 1931 x, y):
Red: 0.708, 0.292
Green: 0.170, 0.797
Blue: 0.131, 0.046
White Point (D65): 0.3127, 0.329

Academy Color Encoding System (ACES)

ACES Encoding: ACES2065-1, ACEScg & ACESProxy – see: <https://acescentral.com/aces-documentation/>

Deliverables must include all metadata, IDTs, LMTs, CLFs and any supporting content or files necessary for a complete ACES package. Please consult SPE technical representative for details if delivering in this format.

Video Assembly Master (VAM)

EXR Encoding:

File format: OpenEXR (.exr)

Bit depth: 16-bit half float

Compression: none

Color space: Mastering or As-Source

Image Width: 1920 (HD), 3840 (UHD), 2048 (2K DCI), 4096 (4K DCI)

Image Height: 1080 (HD), 2160 (UHD), unspecified (2K DCI), unspecified (4K DCI)

DPX Encoding:

File format: DPX (.dpx)

Bit depth: 16-bit

Compression: none

Color space: Mastering or As-Source

Chroma Sample: 4:4:4 (RGB)

Image Width: 1920 (HD), 3840 (UHD), 2048 (2K DCI), 4096 (4K DCI)

Image Height: 1080 (HD), 2160 (UHD), unspecified (2K DCI), unspecified (4K DCI)

TIFF Encoding:

File Format: TIFF (.tiff)

Bit Depth: 16-bit

Compression: none

Color Space: Mastering or As-Source

Chroma Sample: 4:4:4 (RGB)

Image Width: 1920 (HD), 3840 (UHD), 2048 (2K DCI), 4096 (4K DCI)

Image Height: 1080 (HD), 2160 (UHD), unspecified (2K DCI), unspecified (4K DCI)

TELEVISION AUDIO MASTER SPECIFICATIONS

Broadcast-Level Audio Definition

Broadcast-Level Audio conforms to the loudness specifications set by the first window broadcaster. This may be A/85 for domestic broadcast, R128 for European broadcast, or a specific loudness specification that the broadcaster requires.

Full-Range Audio Definition

Full-Range audio is mixed using nearfield loudspeakers. The dynamic range is appropriate for listening at 75-79 dB SPL through a consumer home theater system and is not constrained to a broadcast specification. This serves to future-proof the content for later delivery to any specification.

General Television Audio Specifications

- 48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample. (48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample is not accepted unless previously agreed between the production and SPE.)
- **Discrete audio must be delivered in ProTools sessions** per Television Audio Element Specifications (below) and APPENDIX C –AUDIO TRACK LAYOUTS.
 - **Raw .wav files are not accepted.**
 - **All audio utilizing plug-ins, processing or automation must be rendered into flat .wav audio files.**
 - Session Type: ProTools Version 11 or later, v12 or later preferred.
- Audio within a ProRes or Quick Time .mov file does not substitute for discrete audio delivery
- Television content is always long form per the instructions in APPENDIX B–EPISODIC / MOW /MFT MASTER CONTENT FORMATTING.
- Any frame rates other than 23.976, (for example, 25.00 fps), must be discussed and agreed to by Sony Pictures in writing prior to delivery.
- For approved multiple frame rate titles, a full set of audio deliverables for each frame rate must be delivered.
- ALL NON-NATIVE FRAME RATE AUDIO MUST BE PITCH CORRECTED TO THE NATIVE PITCH. Note, no pitch correction is necessary between 23.976 and 24.00 fps.

Television Audio Element Specifications

Printmasters (PM):

- Each printmaster is to be in its own ProTools session. Mixing audio configurations in the same session is not permitted (e.g. cannot put 5.1 printmaster and LtRt printmaster in the same session)

Stems (SM) (a.k.a. printmaster stems) (SM):

- Stems are to be delivered as wide as possible, it is encouraged to separate the stems wider than just dialog, music and effects. A wide stem example layout is Dialog (DX), Crowd (CD), Walla (WL), Music A (MX_A), Music B (MX_B), Backgrounds (BG), Foley (FL), Effects (FX)
- The sum of all stems must exactly equal the printmaster
- All stem types are to be in the same ProTools session, with tracks clearly labeled.

Music and Effects (ME) and M&E Optionals (OP) Supersession (MESP):

- Music and effects must be fully-filled, covering all production effects lost when muting the dialog
- M&E optionals: Any material that might not be dubbed into a foreign language should be put into one or more “Optionals”. The optional audio configuration generally matches the audio configuration of the main M&E, smaller audio configurations can be used if desired. **Any content that is in an Optional must NOT be in the main M&E.** Examples of optional material are: actor efforts, dialog in a language native to a location, vocals for songs.

- “M&E Supersession” (MESP): The main M&E and all Optionals are to be in the same ProTools session, with tracks clearly labeled.

Music and Effects Stems (MESM):

M&E Stems are made from the printmaster stems, and are modified as needed to create the main M&E.

- The M&E stems are constructed as follows:
 - The “Fill M&E stem’ (FILL_MESM) is the dialog stem with all dialog muted, leaving only production effects that were in the clear from the original recording. Additional effects and foley are added to “fill” and cover the production effects that were lost by muting the dialog.
 - The other M&E stems are the same layout and content as the printmaster stems-any changes specific to the M&E process are reflected in these M&E stems. For example, if vocals were removed from a song and put to an optional, the music stem is modified to not contain those vocals.
 - See APPENDIX D –MUSIC AND EFFECTS STEMS SPECIFICATIONS for further details.
- M&E Stems are to be delivered as wide as possible, matching the layout of the printmaster stems.
- The sum of all M&E stems must exactly equal the main M&E.
- All M&E stem types are to be in the same ProTools session, with tracks clearly labeled.

Fully-Filled Effects Stem (FFXSM):

- The fully-filled effects stem (FFX) is a combined effects stem **consisting of** Backgrounds, Foley and Effects, plus the production effects and fill from the **Fill M&E stem. Adding music to the FFX stem equals the fully-filled M&E**
- The fully-filled effects stem is in its own ProTools session. It is not in the same session with the M&E and Optionals (MESP)

DME “split track” (DME):

- The DME “split track” is a reduction of the stems that is used for downstream markets. It consists of Dialog (dialog+crowd+walla), Music (all music) and Effects (BG, Foley and Effects stems combined). Each of these is a LtRt pair.
- The DME is in its own ProTools session, with each track clearly labeled

Predubs (PD):

Predubs are to be created by rendering flat .wav audio files from the source ProTools mix sessions. Typically, the source ProTools sessions contain plug-ins, panning, mix moves and other processing, all of which must be rendered into the flat audio files.

- Predub organization is determined by each production and must be sufficiently granular to represent the individual components of the soundtrack. For example, Dialog is typically organized into four or more predubs, ADR is similarly organized. Backgrounds, Foley and hard Effects are organized into twenty or more individual predubs.
- Music is generally not predubbed on a dubbing stage and is handled separately.
- The predubs are split into ProTools sessions that represent each major component, the grouping is determined by the production. A typical split might be a Dialog/ADR predub session, BG predubs, Foley predubs, and FX predubs. Each track is to be clearly labeled.
- Predubs are generally delivered long form, See Formatting For Long-Form Content in Appendix A.

Note On Two-Track Audio

Two-Track audio must be delivered as surround encoded LtRt (Left total - Right total) elements that contains the surround information from the 5.1 or higher format. Standard stereo LoRo, which is simple left-right stereo that does not contain any surround information, is not accepted. By delivering LtRt two-track elements, the surround information is preserved and can be perceived by the listener regardless of the playback environment-it does not have to be decoded to be perceived. Note that processors in today’s consumer devices can make use of the surround encoding to enhance the listening experience, whereas LoRo audio can only be perceived as simple stereo.

The LtRt surround encoding must be compatible with Pro Logic II, Pro Logic I compatibility is accepted. The surround encoder **does not have to be a Dolby encoder**, it can be any brand that provides quality acceptable to the production-for example, Neyrinck Soundcode.

THEATRICAL ELEMENT SPECIFICATIONS

THEATRICAL DCP, DCDM AND DKDM

Digital Cinema Package (DCP)

- 2D 2K (4K if available) DCI-Compatible DCP for each studio-approved sound format being released in Digital Cinema (e.g. 5.1, 7.1, Atmos), manufactured in accordance with DCI and SMPTE Digital Cinema Package standards in effect at the time of manufacture [Compressed (JPEG 2000), Encrypted (AES-128), Wrapped (MXF) file] on a CRU, 500GB USB 3.0 HDD, 500GB SSD or other Sony approved hard drive, compiled with all applicable Image, uncompressed Audio, and Subtitle files
- 3D 2K (4K if available) DCI-Compatible DCP for each studio-approved sound format being released in Digital Cinema (e.g. 5.1, 7.1, Atmos), manufactured in accordance with DCI and SMPTE Digital Cinema Package standards in effect at the time of manufacture [Compressed (JPEG 2000), Encrypted (AES-128), Wrapped (MXF) file] on a CRU, 500GB USB 3.0 hard drive or other Sony approved hard drive, compiled with all applicable Image, uncompressed Audio, and Subtitle files

Digital Cinema Distribution Master (DCDM)

- 2D 2K (4K if available) Image DCDM in TIFF (TIFF) 16 bit, XYZ file format on LTO-7 or LTO-8 (LTFS 2.2 or later), in reel lengths. Long form accepted only by agreement with Sony Pictures
- 2D Audio DCDM for each studio-approved sound format being released in Digital Cinema (e.g., 5.1, 7.1, Atmos), consisting of .wav files, 48k, 24bit, 24fps, in reel lengths, tested for “butt splice click-free reel changeovers”, on LTO-7 or LTO-8 (LTFS 2.2 or later). Long form accepted only by agreement with Sony Pictures
- 3D 2K (4K if available) Left-Eye Image DCDM in TIFF (TIFF) 16 bit, XYZ file format on LTO-7 or LTO-8 (LTFS 2.2 or later), in reel lengths. Long form accepted only by agreement with Sony Pictures
- 3D 2K (4K if available) Right-Eye Image DCDM in TIFF (TIFF) 16 bit, XYZ file format on LTO-7 or LTO-8 (LTFS 2.2 or later), in reel lengths. Long form accepted only by agreement with Sony Pictures
- 3D Audio DCDM for each studio-approved sound format being released in Digital Cinema (e.g., 5.1, 7.1, Atmos), consisting of .wav files, 48k, 24bit, 24fps, in reel lengths, tested for “butt splice click-free reel changeovers”, on LTO-7 or LTO-8 (LTFS 2.2 or later). Long form accepted only by agreement with Sony Pictures
- Include frame count for each reel, identify first frame and last frame of each reel.

Distribution Key Delivery Message (DKDM)

- 2D DCP DKDM Files
- 3D DCP DKDM Files

Textless, Text Files, Title and Spotting List Materials

- Textless Background Digital DCDM Data Files: The original 2D (and 3D if applicable) Data Files (format to be approved by Columbia Pictures) on LTO-7 or LTO-8 (LTFS 2.2 or later), of ALL background material (textless, i.e., without any superimposed lettering) for the main and end title credits of the Pictures and any inserts along with all photographic effects present in the titles or inserts such as fades, dissolves, blowups, freeze frames, multiple exposures, etc.
- All title and text elements, including subtitles, free from picture background, to be delivered as Files on LTO-7 or LTO-8 (LTFS 2.2 or later), exactly matching the Original Version (O.V.) picture
- Master Dialogue List - One (1) electronic copy (.doc, .rtf, or .pdf) of a complete industry standard theatrical language annotated dialogue list, including footage notations of all scene ends, all verbatim dialogue (including all grunts, groans, efforts, and the like), lyrics (if any), translations and phonetic transcriptions of all spoken dialogue spoken in other languages than the main spoken language, and annotations of all colloquial slang, historical events, technical terms, and the like. Footages for dialogue lists should be calculated on an AB-reel basis (2,000-foot reels) and referenced to 35mm film running at 24 frames per second. The Master Dialogue List must be created by a Columbia Pictures approved vendor.

- Master Spotting List - One (1) electronic copy (.doc, .rtf, or .pdf) of a complete industry standard theatrical language annotated spotting list, including subtitle-by-subtitle in, out, and length footages, lyrics (if any), speaker and addressee identification, annotations of all colloquial slang, historical events, technical terms, and the like, and laboratory and translator instructions. Footages for spotting lists should be calculated on an AB-reel basis (2,000-foot reels) and referenced to 35mm film running at 24 frames per second. The Master Spotting List must be created by a Columbia Pictures approved vendor.
- Combined Continuity & Spotting List - One (1) electronic copy (.doc, .rtf, or .pdf) of a complete industry standard language Combined Continuity and Spotting List (CCSL), including all dialogue and spotting as referenced above (note that only spotting need have annotations in CCSL), cut-by-cut frame and footage counts of all shots including location and camera angle, meticulous scene description, soundtrack music starts and stops, and including complete main and end credits. Footages for CCSLs should be calculated on an AB-reel basis (2,000-foot reels) and referenced to 35mm film running at 24 frames per second. The Combined Continuity & Spotting List must be created by a Columbia Pictures approved vendor.

THEATRICAL PICTURE SPECIFICATIONS

Digital Intermediate (DI)

The primary deliverable is a fully finished, full show version of the feature and all textless background components, uncompressed and unprocessed, at the highest resolution and bit depth available. The “full show” includes slate, head and tail leader, original theatrical main title and end credits, all alternate main title elements, textless units, alternate versions (director’s or extended cut) shots or scenes, and Standard Metadata Block. For example, a fully delivered DI might include the full theatrical show and textless as one set of LTOs, and a second set of LTOs for an extended version. This data is generally P3 D65 or Rec709 color delivered in the form of DPX files or EXR files. This component is a color-timed master. The DI deliverable should also contain the conformed audio (LtRt, 5.1, DME, in .wav) as a reference.

- DI Project shall be of the same Color Grading project, software and version used by colorist to generate the DSM.
- DI Project timeline shall include all color decisions and Viewing LUTs (where applicable) in the timeline.
- DI Project shall include all nodes/layers used for comps, resizes, shakes, stabilization, or convergence adjustments performed during the grade to reflect all creative intent of the image framing and composition.
- DI Project shall include an export of EDLs and XMLs which reflects each timeline graded in the DI.

Metadata: Standard per Documentation and Labeling Requirements v.1.9. In the case of material produced in ACES, all transforms, LUTs or other devices must be included in text files.

Requirement: 1 copy (check-summed clone, LTFS 2.2 or later) on LTO-7 or LTO-8. The format for the LTO must conform to the PHYSICAL ARCHIVE SPECIFICATIONS.

Project Naming: The DI Project shall be named with the following naming scheme:

[showname]_[REEL##]_[ISODate]_[SoftwareName][SoftwareVersion#].[EXTENSION]

- **Show Name** (in CamelCaseFormat)
- **Reel ##** (reel01, reel02, etc.)
- **ISO DATE of last file save date** (YYYYMMDD)
- **SOFTWARE_NAME** (Resolve, Baselight)
- **SOFTWARE_VERSION NUMBER**

spiderman_reel04_20210112_resolve_16.1.drp

Digital Source Master (DSM)

The DSM or Digital Source Master is a conformed, cleaned master created at the highest bit depth and largest raster dimension (i.e., sometimes 4K+). This element is colored, but with color not baked in but provided in the form of a LUT. Typically, the element may be in EXR, TIFF or DPX, and in P3D65, Log-C or S-log or other non-standard color space. This element is generally composed of camera exports, often in Log-C or S-log space, and uncorrected units such as VFX and opticals. It is a large file (2.8K+ or 4K, 16bit) in EXR or DPX, depending on the production workflow. The element is not sized or formatted for distribution and may include files of different sizes (for example, camera files might be 4096x2304, while VFX might be 3840x2160). Since most DI sources are mixed resolution, it is common to choose a “hero” resolution based on the largest raster in the DI. This element is a rough conform based on the most original or raw data available. It is the parent of elements such as the DI and DCDM, as well as 4K HV masters. and all textless background material. It can be delivered as a fully textless version of the show, but if delivered as texted, it must also include textless for the texted sections.

DSM shall be rendered in the native resolution of the DI timeline to minimize scaling of archival images. The exceptions to this requirement are the following:

- **Mixed Resolutions:** If DI timeline contains source images of mixed resolutions, all images shall be rendered to a single resolution for delivery. This rendered resolution shall be based on the native “hero” camera resolution, or the most common

resolution on the DI timeline, or the highest resolution found in the DI timeline as determined with and approved and must adhere to SPE approved scaling methodology.

- **Upscaled Resolution DSM:** On some productions, it may be requested that the DSM be delivered in a higher resolution than the native DI timeline (e.g., scaled output from 2k to 4K). The target resolution shall be determined and approved and must adhere to SPE approved scaling methodology.
- For shows whose theatrical projection aspect ratio is 1.85, the final DSM archive shall be rendered to a minimum aspect ratio of 1.78.
- For shows whose theatrical projection aspect ratio is 2.39, the final DSM archive shall be rendered to a minimum aspect ratio of 2.0.
- If these delivery aspect ratios cannot be achieved for any reason (limited resolution, anamorphic unsqueezed, etc.), please contact SPE.

EXR Encoding:

File format: OpenEXR (.exr)

Bit depth: 16-bit half float

Compression: none

Color space: Mastering or As-Source

Image Width: 1920 (HD), 3840 (UHD), 2048 (2K DCI), 4096 (4K DCI)

Image Height: 1080 (HD), 2160 (UHD), unspecified (2K DCI), unspecified (4K DCI)

DPX Encoding:

File format: DPX (.dpx)

Bit depth: 16-bit

Compression: none

Color space: Mastering or As-Source

Chroma Sample: 4:4:4 (RGB)

Image Width: 1920 (HD), 3840 (UHD), 2048 (2K DCI), 4096 (4K DCI)

Image Height: 1080 (HD), 2160 (UHD), unspecified (2K DCI), unspecified (4K DCI)

TIFF Encoding:

File Format: TIFF (.tiff)

Bit Depth: 16-bit

Compression: none

Color Space: Mastering or As-Source

Chroma Sample: 4:4:4 (RGB)

Image Width: 1920 (HD), 3840 (UHD), 2048 (2K DCI), 4096 (4K DCI)

Image Height: 1080 (HD), 2160 (UHD), unspecified (2K DCI), unspecified (4K DCI)

Metadata: Standard per Documentation and Labeling Requirements v.1.9, but also includes all transforms, LUTs, EDLs, CDLs, HDR documentation (where available) as well as Baselight scenes and Resolve sessions that were used to correct the data to create downstream products.

Requirement: 1 copy (check-summed clone, LTFS 2.2 or later) on LTO-7 or LTO-8. The format for the LTO must conform to the PHYSICAL ARCHIVE SPECIFICATIONS.

Academy Color Encoding System (ACES)

ACES Encoding: ACES2065-1, ACEScG & ACESProxy – see: <https://acescentral.com/aces-documentation/>

Deliverables must include all metadata, IDTs, LMTs, CLFs and any supporting content or files necessary for a complete ACES package. Please consult SPE technical representative for details if delivering in this format.

YCM Color Separation

Some Sony-owned features are written to film as YCM color separations. This is a secondary operation, using the picture data of the conformed and completed DI data (including the textless component of the show), to manufacture a set of three-color stable monochrome negatives per reel representing the primary colors of the digital data.

Flash-cut Negative

Sony-owned features shot on film are submitted to a process whereby the film negative is inventoried and collated with editorial drives. The shots used in the film are assembled and documented as a series of flash-to-flash camera takes in conform order and organized into 1000' reels, which allows the film negative to be accessed effectively for scanning and preservation. This is a secondary activity and involves organization of material in the production archive, not creation of new physical material.

THEATRICAL-LEVEL AUDIO MASTER SPECIFICATIONS

Theatrical-Level Audio Definition

Theatrical-level audio is typically mixed in a dubbing theater with large loudspeakers behind a projection screen and surround speakers on the walls at a good distance from the listening position. The reference sound level is 85dBc per SMPTE RP 200, which allows for a very wide dynamic range, typically 50dB to 70dB. It is designed to be heard in a cinema auditorium environment.

General Theatrical-Level Audio Specifications

- The frame rate and sample rate for each discrete audio element is specified below in Theatrical-Level Audio Element Specifications.
- **Discrete audio must be delivered in ProTools sessions** per Theatrical-Level Audio Element Specifications (below) and APPENDIX C –AUDIO TRACK LAYOUTS.
 - **Raw .wav files are not accepted**
 - **All audio utilizing plug-ins, processing or automation must be rendered into flat .wav audio files**
 - Session Type: ProTools Version 11 or later, v12 or later preferred
- Audio within a ProRes or Quick Time .mov file does not substitute for discrete audio delivery
- Theatrical-level audio can be formatted in reels or in long form. It is preferred that the audio be formatted the way it was mixed.
 - If in reels, all reels are to be in the same session. Reels are placed such that reel number = hour number. For example, reel 5 would be at 05:00:00:00. See **Formatting For Reel-Based Audio Content** in Appendix A.
 - If in long form, see **Formatting For Long-Form Content** in Appendix A.
- Any frame rates other than 24.00 or 23.976, (for example, 25.00 fps), must be discussed and agreed to by Sony Pictures in writing prior to delivery.
- For approved multiple frame rate titles, a full set of audio deliverables for each frame rate must be delivered
- ALL NON-NATIVE FRAME RATE AUDIO MUST BE PITCH CORRECTED TO THE NATIVE PITCH. Note, no pitch correction is necessary between 23.976 and 24.00 fps

Theatrical-Level Audio Element Specifications

Printmasters (PM):

- 48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample. (48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample is not accepted for theatrical printmasters.)
- Each printmaster is to be in its own ProTools session. Mixing audio configurations in the same session is not permitted (e.g., cannot put 5.1 printmaster and LtRt printmaster in the same session)
- DCP Printmaster Audio: Printmaster audio designated for DCP must be delivered in reels unless previously agreed between the production and SPE. Reel changeovers must be sample accurate and be able to be “butt spliced” with no audible clicks or other artifacts.

Stems (SM) (a.k.a. Printmaster Stems) (SM):

- 48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample. It is permitted to deliver stems at 48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample if the feature was natively mixed at 23.976.
- Stems are to be delivered as wide as possible. It is encouraged to separate the stems wider than just dialog, music and effects. A wide stem example layout is Dialog (DX), Crowd (CD), Walla (WL), Music A (MX_A), Music B (MX_B), Backgrounds (BG), Foley (FL), Effects (FX).
- The sum of all theatrical-level stems must exactly equal the theatrical-level printmaster.
- All stem types are to be in the same ProTools session, with tracks clearly labeled.

Music and Effects (ME) and M&E Optionals (OP) Supersession (MESP):

- 48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample. (48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample is not accepted for theatrical Music and Effects.)
- Music and effects must be fully-filled, covering all production effects lost when muting the dialog
- M&E optionals: Any material that might not be dubbed into a foreign language should be put into one or more “Optionals”. The Optional audio configuration generally matches the audio configuration of the main M&E, smaller audio configurations can be used if desired. **Any content that is in an Optional must NOT be in the main M&E.** Examples of optional material are: Actor efforts, dialog in a language native to a location, vocals for songs.
- “M&E Supersession” (MESP): The main M&E and all Optionals are to be in the same ProTools session, with tracks clearly labeled.

Music and Effects Stems (MESM):

M&E Stems are made from the printmaster stems, and are modified as needed to create the main M&E

- The M&E stems are constructed as follows:
 - The “Fill M&E stem’ (FILL_MESM) is the dialog stem with all dialog muted, leaving only production effects that were in the clear from the original recording. Additional effects and foley are added to “fill” and cover the production effects that were lost by muting the dialog.
 - The other M&E stems are the same layout and content as the printmaster stems-any changes specific to the M&E process are reflected in these M&E stems. For example, if vocals were removed from a song and put to an optional, the music stem is modified to not contain those vocals.
 - See APPENDIX D –MUSIC AND EFFECTS STEMS SPECIFICATIONS for further details.
- 48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample. (48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample is not accepted for theatrical M&E stems.)
- M&E Stems are to be delivered as wide as possible, matching the layout of the printmaster stems.
- The sum of all theatrical-level M&E stems must exactly equal the main theatrical-level M&E.
- All M&E stem types are to be in the same ProTools session, with tracks clearly labeled.

Fully-Filled Effects Stem (FFXSM):

The fully-filled effects stem (FFX) is a combined effects stem consisting of Backgrounds, Foley and Effects, plus the production effects and fill from the Fill M&E stem. Adding music to the FFX stem equals the fully-filled M&E

- 48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample. (48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample is not accepted for theatrical-level FFXSM)
- The fully-filled effects stem is in its own ProTools session. It is not in the same session with the M&E and Optionals (MESP).

DME “split track” (DME):

The DME “split track” is a reduction of the stems that is used for downstream markets. It consists of Dialog (dialog+crowd+walla), Music (all music) and Effects (BG, Foley and Effects stems combined). Each of these is a LtRt pair.

- 48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample. It is permitted to deliver the DME at 48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample if the feature was natively mixed at 23.976.
- The DME is in its own ProTools session, with each track clearly labeled.

Predubs (PD):

Predubs are to be created by rendering flat .wav audio files from the source ProTools mix sessions. Typically, the source ProTools sessions contain plug-ins, panning, mix moves and other processing, all of which must be rendered into the flat audio files.

- Predub organization is determined by each production and must be sufficiently granular to represent the individual components of the soundtrack. For example, Dialog is typically organized into four or more predubs, ADR is similarly organized. Backgrounds, Foley and hard Effects are organized into twenty or more individual predubs.
- Music is generally not predubbed on a dubbing stage and is handled separately.
- 48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample. It is permitted to deliver the predubs at 48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample if the feature was natively mixed at 23.976.

- The predubs are split into ProTools sessions that represent each major component, the grouping is determined by the production. A typical split might be a Dialog/ADR predub session, BG predubs, Foley predubs, and FX predubs. Each track is to be clearly labeled.

Note On Two-Track Audio

Two-Track audio must be delivered as surround encoded LtRt (Left total - Right total) elements that contains the surround information from the 5.1 or higher format. Standard stereo LoRo, which is simple left-right stereo that does not contain any surround information, is not accepted. By delivering LtRt two-track elements, the surround information is preserved and can be perceived by the listener regardless of the playback environment-it does not have to be decoded to be perceived. Note that processors in today's consumer devices can make use of the surround encoding to enhance the listening experience, whereas LoRo audio can only be perceived as simple stereo.

The LtRt surround encoding must be compatible with Pro Logic II, Pro Logic I compatibility is accepted. The surround encoder **does not have to be a Dolby encoder**, it can be any brand that provides quality acceptable to the production-for example, Neyrinck Soundcode.

FEATURE HOME MASTERING SPECIFICATIONS

FEATURE HOME MASTERING PICTURE SPECIFICATIONS

High Dynamic Range Distribution Master (HDRDM)

TIFF Encoding:

File Format: TIFF (TIFF)
Bit Depth: 16-bit
Compression: none
Color Space: Rec. 2020 full-range (ITU-R BT.2020-2)
EOTF: ST-2084 (PQ)
Chroma Sample: 4:4:4 (RGB)
Max Luminance: uncapped
Image Width: 3840 (UHD)
Image Height: 2160 (UHD)

ITU-R. 2035, specifies the viewing environment. Studio mastering reference display specifications are as follows:

Max Luminance: 4000 cd/m² (nits)
Black Luminance: <0.005 cd/m² (nits)
EOTF: ST-2084 (PQ)
Color Primaries (CIE 1931 x, y):
Red: 0.708, 0.292
Green: 0.170, 0.797
Blue: 0.131, 0.046
White Point (D65): 0.3127, 0.329

Color Timed Master (CTM)

UHD Standard Dynamic Range

DPX Encoding:

File Format: DPX (.dpx)
Bit Depth: 10-bit
Compression: none
Color Space: Rec. 709 full-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:4:4 (RGB)
Image Width: 3840 (UHD)
Image Height: 2160 (UHD)

ITU-R. 2035, specifies the viewing environment. Studio mastering reference display should be in compliance with ITU-R BT 1886 as follows:

Max Luminance: 100 cd/m² (nits)
Black Luminance: <0.1 cd/m² (nits)
EOTF: Gamma 2.4
Color Primaries (CIE 1931 x, y):
Red: 0.64, 0.33
Green: 0.30, 0.60
Blue: 0.15, 0.06
White Point (D65): 0.3127, 0.329

HD Standard Dynamic Range

DPX Encoding:

File Format: DPX (.dpx)
Bit Depth: 10-bit
Compression: none
Color Space: Rec. 709 full-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:4:4 (RGB)
Image Width: 1920 (HD)
Image Height: 1080 (HD)

ITU-R. 2035, specifies the viewing environment. Studio mastering reference display should be in compliance with ITU-R BT 1886 as follows:

Max Luminance: 100 cd/m² (nits)
Black Luminance: <0.1 cd/m² (nits)
EOTF: Gamma 2.4
Color Primaries (CIE 1931 x, y):
Red: 0.64, 0.33
Green: 0.30, 0.60
Blue: 0.15, 0.06
White Point (D65): 0.3127, 0.329

ProRes Encoding:

Codec: Apple ProRes 422 (HQ)
Image Width: 1920 (HD)
Image Height: 1080 (HD)
Frame Rate: Native
Data Rate: 220 Mbps VBR
Color Space: Rec. 709 legal-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:2:2 (YCbCr)
Color Depth: 10 Bit
Container: .MOV

Codec: Apple ProRes 4444 (no alpha)
Image Width: 1920 (HD)
Image Height: 1080 (HD)
Frame Rate: Native
Data Rate: 330 Mbps VBR
Color Space: Rec. 709 legal-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:2:2 (YCbCr) or 4:4:4 (YCbCr)
Color Depth: 10 Bit
Container: .MOV

HOME THEATER AUDIO MASTER SPECIFICATIONS

Home Theater Audio Definition

Home Theater audio is full-range audio mixed using nearfield loudspeakers. The dynamic range is appropriate for listening at 75dB -79dB SPL through a consumer home theater system and is not constrained to a broadcast specification.

General Home Theater Audio Specifications

- 48.00 kHz or 96.00 kHz, 23.976 fps, 24 bits per sample. (48.00 kHz or 96.00 kHz, 24.00 fps, 24 bits per sample is not accepted unless previously agreed between the production and SPE.)
- **Discrete audio must be delivered in ProTools sessions** per Home Theater Audio Element Specifications (below) and APPENDIX C –AUDIO TRACK LAYOUTS.
 - **Raw .wav files are not accepted.**
 - **All audio utilizing plug-ins, processing or automation must be rendered into flat .wav audio files.**
 - Session Type: ProTools Version 11 or later, v12 or later preferred.
- Audio within a ProRes or Quick Time .mov file does not substitute for discrete audio delivery.
- Home Theater audio can be formatted in reels or in long form. It is preferred that the audio be formatted the way it was mixed.
 - If in reels, all reels are to be in the same session. Reels are placed such that reel number = hour number. For example, reel 5 would be at 05:00:00:00. See **Formatting For Reel-Based Audio Content** in Appendix A.
 - If in long form, see **Formatting For Long-Form Content** in Appendix A.
- Any frame rates other than 23.976, (for example, 24.00 or 25.00 fps), must be discussed and agreed to by Sony Pictures in writing prior to delivery.
- For approved multiple frame rate titles, a full set of audio deliverables for each frame rate must be delivered
- ALL NON-NATIVE FRAME RATE AUDIO MUST BE PITCH CORRECTED TO THE NATIVE PITCH. Note, no pitch correction is necessary between 23.976 and 24.00 fps.

Home Theater Audio Element Specifications

Printmasters (PM):

- Each printmaster is to be in its own ProTools session.
- Combining multiple printmaster audio configurations in the same session is not permitted (e.g., cannot put 5.1 printmaster and LtRt printmaster in the same session).

Stems (SM) (a.k.a. Printmaster Stems) (SM):

- Stems are to be delivered as wide as possible. It is encouraged to separate the stems wider than just dialog, music and effects. A wide stem example layout is Dialog (DX), Crowd (CD), Walla (WL), Music A (MX_A), Music B (MX_B), Backgrounds (BG), Foley (FL), Effects (FX).
- The sum of all home theater stems must exactly equal the home theater printmaster.
- All stem types are to be in the same session, with tracks clearly labeled.

Music and Effects (ME) and M&E Optionals (OP) Supersession (MESP):

- Music and effects must be fully-filled, covering all production effects lost when muting the dialog.
- M&E Optionals: Any material that might not be dubbed into a foreign language should be put into one or more “Optionals”. The Optional audio configuration generally matches the audio configuration of the main M&E, smaller audio configurations can be used if desired. **Any content that is in an Optional must NOT be in the main M&E.** Examples of optional material are: actor efforts, dialog in a language native to a location, vocals for songs.
- “M&E Supersession” (MESP): The main M&E and all Optionals are to be in the same ProTools session, with tracks clearly labeled.

Music and Effects Stems (MESM):

- M&E Stems are made from the Printmaster Stems, and are modified as needed to create the main M&E.
- The M&E stems are constructed as follows:
 - The “Fill M&E stem” (FILL_MESM) is the dialog stem with all dialog muted, leaving only production effects that were in the clear from the original recording. Additional effects and foley are added to “fill” and cover the production effects that were lost by muting the dialog.
 - The other M&E stems are the same layout and content as the printmaster stems-any changes specific to the M&E process are reflected in these M&E stems. For example, if vocals were removed from a song and put to an optional, the music stem is modified to not contain those vocals.
 - See APPENDIX D –MUSIC AND EFFECTS STEMS SPECIFICATIONS for further details.
- M&E Stems are to be delivered as wide as possible, matching the layout of the printmaster stems.
- The sum of all home theater M&E stems must exactly equal the main home theater M&E.
- All M&E stem types are to be in the same session, with tracks clearly labeled.

Fully-Filled Effects Stem (FFXSM):

- The fully-filled effects stem (FFX) is a combined effects stem consisting of Backgrounds, Foley and Effects, plus the production effects and fill from the Fill M&E stem. Adding music to the FFX stem equals the fully-filled M&E
- The fully-filled effects stem is in its own ProTools session. It is not in the same session with the M&E and Optionals (MESP).

DME “split track” (DME):

- The DME “split track” is a reduction of the stems that is used for downstream markets. It consists of Dialog (dialog+crowd+walla), Music (all music) and Effects (BG, Foley and Effects stems combined). Each of these is a LtRt pair.
- The DME is in its own ProTools session, with each track clearly labeled.

Home Theater Audio Specification Notes

The audio for feature home masters should ideally be home theater audio, which is full-range audio mixed while monitoring with nearfield loudspeakers to ensure the audio plays well at home. Note that feature home theater audio is NOT broadcast audio and is NOT compressed to meet a broadcast loudness spec such as A/85 or R128.

Though home theater audio is best for feature home mastering, theatrical-level audio can be used for feature home theater masters if home theater audio will not be created.

Broadcast-level mixes (e.g., A/85 or R128) may be optionally delivered for the archive. Broadcast-level mixes will NOT be used for feature home mastering unless a specific broadcast master is requested that uses this audio.

Note On Two-Track Audio

Two-Track audio must be delivered as surround encoded LtRt (Left total - Right total) elements that contains the surround information from the 5.1 or higher format. Standard stereo LoRo, which is simple left-right stereo that does not contain any surround information, is not accepted. By delivering LtRt two-track elements, the surround information is preserved and can be perceived by the listener regardless of the playback environment-it does not have to be decoded to be perceived. Note that processors in today’s consumer devices can make use of the surround encoding to enhance the listening experience, whereas LoRo audio can only be reproduced as simple stereo.

The LtRt surround encoding must be compatible with Pro Logic II, Pro Logic I compatibility is accepted. The surround encoder **does not have to be a Dolby encoder**, it can be any brand that provides quality acceptable to the production-for example, Neyrinck Soundcode.

RESTORATION AND REMASTERING SPECIFICATIONS

Raw Scan Clone (RSC)

In cases where the original or source element for the restoration is film, the first deliverable are the raw scans. These should be completely raw and free of processing of any kind (stabilization, sharpening, dust-busting, color correction, grain adjustment). Normally, these scans are in the form of DPX files, but EXR, TIFF or ADX files are equally acceptable. In cases where the scan implements ACES, the reference render transform and any additional metadata must be included in the delivery. This material must be original file clones or rendered material.

Requirement: 1 copy (check-summed clone, LTFS 2.2 or later) on LTO-7 or LTO-8, of the complete RSC, including textless components. The format for the LTO must conform to PHYSICAL ARCHIVE SPECIFICATIONS. The data cannot be in proprietary formats such as Retrospect or BRU. All archive deliverables must be on new LTOs, not re-used stock.

Intermediate Picture Data Clone (IPDC)

This element is a fully conformed version of the data; it should be cleaned and color corrected, but without the color baked-in. Normally this is a 16bit data resource and represents the highest resolution and widest dynamic range extracted from the raw scans. It is not necessarily sized for final delivery – it may have a resolution characteristic that reflects the sensor rather than a 4K or UHD size. This element is sometimes referred to as a “master render.” It may be in P3 or in a space such as S-Log, depending on workflow. IPDCs are generally in DPX format unless derived from an ACES workflow, in which case they may be in EXR format. The IPDC deliverable is not constrained by the EXR and DPX templates outlined in SPE HD/UHD Master Archive Delivery Specifications document. The IPDC must include the textless component. This material must be original file clones or rendered material.

Requirement: 1 copy (check-summed clone, LTFS 2.2 or later) on LTO-7 or LTO-8, of the complete IPDC. The format for the LTO must conform to PHYSICAL ARCHIVE SPECIFICATIONS. The data cannot be in proprietary formats such as Retrospect or BRU. All archive deliverables must be on new LTOs, not re-used stock.

Source Files (SRC) – The SRC is a specific configuration of IPDC based on raw scans that have been restored – that is, a conformed element that has been cleaned and repaired to the point where it can serve as the data matrix for the remainder of workflow processes on the way to a final data set. Restoration work and QC fixes are integrated into the SRC. If titles or graphics are created in restoration, these are included in the SRC archive LTO. The SRC is not color corrected. The SRC must be accompanied by all available Baselight scenes (or documentation of a similar working environment such as Resolve sessions) and EDLs. This material must be original file clones or rendered material.

Requirement: 1 copy (check-summed clone, LTFS 2.2 or later) on LTO-7 or LTO-8, of the complete CTM, including textless components. The format for the LTO must conform to PHYSICAL ARCHIVE SPECIFICATIONS. The data cannot be in proprietary formats such as Retrospect or BRU. All archive deliverables must be on new LTOs, not re-used stock.

Final Picture Data Clone (FPDC)

The FPDC is the primary deliverable; it is the final, talent-approved, broadcast-ready full show with textless component and any required inserts. This element includes all color correction and complete picture in a fully integrated, audio-conformed state. The FPDC may be delivered in DPX or EXR files, and it may have been finished at HD, UHD or 4K resolution. Conformed audio for the CTM must also be included on the LTO. This material must be original file clones or rendered material.

Requirement: 1 copy (check-summed clone, LTFS 2.2 or later) on LTO-7 or LTO-8, of the complete production archive. The format for the LTO must conform to PHYSICAL ARCHIVE SPECIFICATIONS. The data cannot be in proprietary formats such as Retrospect or BRU. All archive deliverables must be on new LTOs, not re-used stock.

REFERENCE PICTURE SPECIFICATION

Video / Audio Formatting

ProRes Encoding:

Codec: Apple ProRes 422 (Proxy)
Image width: 960
Image height: 540
Frame Rate: Native
Data Rate: approx. 45 Mbps VBR
Color Space: Rec. 709 legal-range (ITU-R BT.709-6)
EOTF: Gamma 2.4
Chroma Sample: 4:2:2 (YCbCr)
Color Depth: 8 or 10 Bit
Container: .MOV

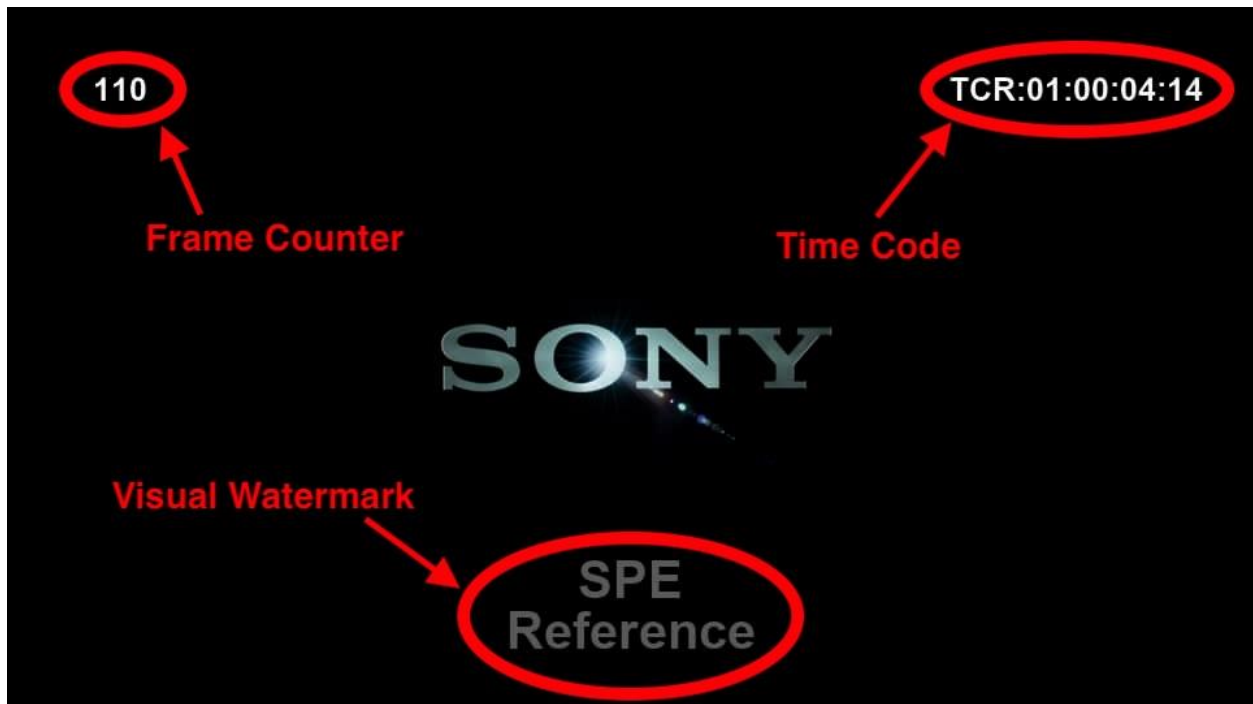
Audio Encoding:

Audio Codec: PCM
Sampling rate: 48 kHz
Sample size: 24 bits per sample
Channels: Original Language Stereo Comp

Formatting:

Picture is to contain the following visual overlays (with example below):

- **Frame Counter** – starting with frame 0 at file start, placed in the upper left of frame
- **Time Code** – visual time code and should match embedded file time code, placed in the upper right of frame
- **Visual Watermark** – “SPE Reference” placed in the lower center of frame, text should be white at 50% opacity
- **Layout** - See Formatting For Long-Form Content in Appendix A.



PRODUCTION ARCHIVE SPECIFICATIONS

The production archive is not formally specified insofar as it is comprised of diverse and not entirely standardized data forms. The production archive consists of all the product of the workflow from camera tests to final delivery, with the exception of the conformed deliverables described above. This is an all-encompassing and variable deliverable, which will be different for every workflow and production. The data in the production archive is not constrained by the specifications for the conformed deliverables.

Camera files, dailies and visual effects (VFX) materials.

- A. Typically, the largest part of the production archive would include any raw camera files, camera exports (deBayered camera file exports, in camera color space such as s-log and Arri log-c), dailies and daily proxies. This body of material represents the original camera data and the transcodes used to view that data. A production might use several cameras - a principle cinematic camera, a DSLR, a drone camera or even a phone camera – all of which will have their own file types which need to be captured in the camera file section of the archive. Reference implements such as framing, focus and color charts should be included.
- B. Image materials acquired or produced other than principle photography. The most significant of these are visual effects units. These must be captured as delivered, uncolored and accompanied by all plates, layers, implements and metadata used to composite the effects. All stock footage, titles, text, opticals, graphics, logos and other pictorial material created for the production must also be included.

Editorial drives and related process metadata.

These are the drives that contain editorial decisions, various assemblies and versions of the work, including the finished work, lifts and alternate scenes, temporary edits, annotations, etc. Typically, these will be Avid drives. All Baselight scenes (or Resolve sessions or process records produced by other editorial software environments), all EDLs, CDLs, LUTs, transforms, HDR information and any other process metadata used to create the production must also be included. Color science and LUT stack with all related items must be included.

All script material

Production tracking and related information.

Wrap Drive

The omnibus collection of all textual components of the production history, correspondence, logs, reports, etc. This comprehensive resource is normally delivered on a USB drive or as a part of the editorial drive.

Picture Editorial Delivery Methodology

The production archive should be delivered on LTO-7 or LTO-8 in LTFS 2.2 (or later). The editorial drives may be delivered as HDD or RAID, or on LTO. The material must not be formatted in a proprietary container (i.e., Retrospect, BRU, Cache-A, etc.). The data must not be encrypted or compressed (except where it is natively compressed, such as the camera files or editorial drives).

Audio archive:

This includes the Sound Editorial Turnover with production audio and post-production audio and the Collated Archive of all master audio. The Sound Editorial Turnover is generally delivered on one or more HDD's formatted for macOS compatibility. It may also be delivered on LTO-7 or LTO-8 in LTFS 2.2 (or later). It is to be delivered no later than two weeks after the end of the final mix. The Collated Archive is delivered on LTO-7 or LTO-8 in LTFS 2.2 (or later) and is to be delivered within six months of the first release.

1. Sound Editorial Turnover

- Production sound recordings (“sound rolls”), recorded during principle photography (“sync sound”)
- Production sound recordings (“sound rolls”), recorded during any pickups, retakes, reshoots, B-rolls and additional photography (“sync sound”)
- Any audio recorded “wild” (not in sync with action) on location or set. Wild sounds may include such items as nature backgrounds, room tone, effects performed and recorded off camera at location.
- Any other audio recorded specifically for the show
- All production sound logs with circled master takes for the above
- Lined script with production sound notes
- All ADR recorded for the show. This includes Principle ADR and TV/Airline “coverage” ADR

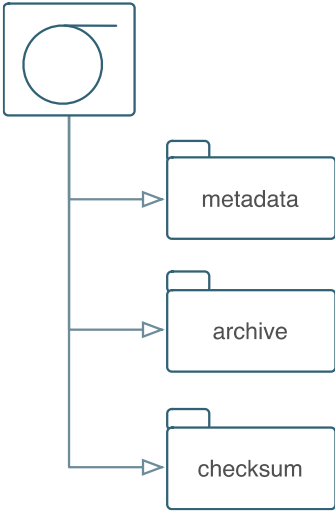
Master and Archive	SPE	Specs	Ver 1.07.2	RELEASE	Page 26
--------------------	-----	-------	------------	---------	---------

- ADR logs with circled master takes with special callouts for TV/Airline ADR
 - OMF files (generally cut production dialog from picture editorial as given to sound editorial)
 - LFOA (Last Frame of Action) list
 - All Sound Editorial sessions: Dialog, ADR, Effects, Backgrounds, Foley. The session type and version must be clearly noted e.g., ProTools v12.6
 - “Skeleton Sessions”-sessions with automation and plug-ins and no media. These are often created from the dialog mix sessions to aid in dubbed dialog mixes.
 - Mixing stage equipment: Long-form console used, list of all plug-ins (name and version) and outboard processors used for the mix
 - Mixing stage data. This may include such items as voice processing settings, reverb settings and automation files
2. Collated Archive of master audio (conformed to final picture)
- Original mixing stage outputs of all audio elements in all audio formats in the native language (OV). This includes everything that the mix stage created. The mix stage deliverables are detailed in SPE’s delivery specifications for each content type.
 - Conformed OV master audio to final home master (IMF or Video Master)
 - All dubbed language mixing stage outputs and deliveries from territory. For each language, this is typically the dubbed dialog stem and all printmaster(s) in each audio format
 - Conformed dubbed language audio to final home master (IMF or Video Master)

PHYSICAL ARCHIVE SPECIFICATIONS

LTFS Archive

- **Physical Archive Medium:** LTO-7 or LTO-8
- **Tape Format:** LTFS Version 2.2.0 or later
- **File compression:** None
- **Data Structure:** only 3 folders at the root level of each tape
 - 'metadata' – used for storing Archive Metadata
 - 'archive' – used to store the Archive Contents
 - 'checksum' – used to store the MD5 Hash Manifests



Standard Metadata Block

The standard metadata block is descriptive metadata which applies to the structure and content of the LTO. The standard metadata block has two parts, (1) the (singular) container part (refers to the LTO) and (2) the (plural) content part (refers to the essence data). There will be only one “container” data block per LTO, but there may be several essences on the LTO. The “standard metadata block” contains standard information that is also found on the LTO label, the archival delivery statement when the data is turned over to Sony, the LTO itself (in the metadata root), the GOLD record (Sony’s inventory database), file paths and wrappers. This fictitious example, illustrates metadata for an LTO containing three episodes of THE SHOW:

LABEL: For each **LTO label**, please include:

Show: THE SHOW S3 Ep 305, 306, 307 (Must include name of production, season and show number)

Media Class: CTM or VAM

Tape: 1 of 1 (as per the tape configuration)

Reel Number: per reel configuration or n/a

Blocking Factor: variable

Data Org: LTFS 2.2

Read Method: n/a

Total Files: 167,481

Total Size: 5.253 TB

Post House: The Post House

Date of LTO Manufacture: 8/21/2020

File type: DPX

Resolution: 3840x2160

If there is space (in most cases, there may not be), please also include the following data:

SHOW_s3_ep305/show_rnd/show_305_4K_20131015_co3_grd03_exr/3840x2160/show_305_4K_20131015_co3_grd03.[0085680-0110852].dpx
SHOW_s3_ep306/show_rnd/show_306_4K_20131015_co3_grd03_exr/3840x2160/show_306_4K_20131015_co3_grd03.[0085680-0110852].dpx
SHOW_s3_ep307/show_rnd/show_307_4K_20131015_co3_grd03_exr/3840x2160/show_307_4K_20131015_co3_grd03.[0085680-0110852].dpx

LTO On-board Metadata

For each LTO, please provide the following standard metadata block information on the LTO in a folder and in an electronic form (e-mail, ASCII text) accompanying your pre-invoice:

Part 1 - Media metadata: this set of data points describes the physical object and method of data storage. [television example]

Show: THE SHOW S3 Ep 305, Ep 306, Ep307

Barcode: PJ1594

Tape: 1 of 1

Date (production date): 20180315

Tape Type: LTO-7 or LTO-8

Blocking Factor: variable

Data Org: tar per shot

Read Method: LTFS 2.2

Written by: datalinux1: /dev/lto7-1nst

Software used to create: Linux

Total Files: 215,174

Total Size: 5.253 TB

Tar Files: n/a

Post House: The Post House

Date of LTO Manufacture: 20200815

Part 2 - Content Metadata: this section describes each essence data on the LTO – so there should be one for Ep 305, one for Ep306 and one for Ep307. Also note that our standard delivery includes the conformed audio (LtRt stereo, 5.1 and DME).

Show Title: THE SHOW S3 Ep305

Version Info: Broadcast, Uncensored, other as appropriate

Media Class: FPDC (CTM) or IPDC (VAM)

Reel Numbers: n/a

Content Type: full show, graded, with textless (CTM) or full show with textless (VAM).

Source material: raw scans, Baselight conforms (as appropriate)

File type: DPX

Resolution: 3840x2160

Bit Depth: 16bit (or 12bit or 10bit as appropriate)

Color space: Rec709 (or as appropriate)

EOTF: [applicable for HDR deliveries only]

Reference Luminance: [applicable for HDR deliveries only]

Reference White: [as applicable]

Frame range: [0085680-0110852]

Processing: IR channel automated restoration; conformed, titled, color timed (this is used to describe the various processes to which the data has been submitted)

Note: HDR deliveries require metadata augmentation corresponding to SMPTE 2084 (in the form of a CSV file): MAXCLL (CSV) and MAXFALL (CSV) expressing maximum luminance values. This metadata should be included in the metadata root file along with the standard metadata blocks.

Show Title: THE SHOW S3 Ep305

Version Info: Broadcast, Uncensored, other as appropriate

Media Class: conformed audio

Reel Numbers: n/a

Content Type: full show audio as LtRt stereo, 5.1 and DME

Source material: audio source as appropriate

File type: .wav

Resolution: n/a

Bit Depth: n/a

Color space: n/a

EOTF: n/a Reference Luminance: n/a

Reference White: n/a

Frame range: n/a

Processing: n/a

STANDARD METADATA DATA BLOCK on LTFS 2.2 LTO-7 or LTO-8:

(container data)

(SHOW CTM Episode 305 picture)

(SHOW CTM Episode 306 picture)

(SHOW CTM Episode 307 picture)

(SHOW CTM Episode 305 audio)

(SHOW CTM Episode 306 audio)

(SHOW CTM Episode 307 audio)

On the LTO, this would translate to 20 files in LTFS:

Root file: metadata (ASCII):

Container metadata

SHOW CTM Ep 305 metadata (picture)

SHOW CTM Ep 306 metadata (picture)

SHOW CTM Ep 307 metadata (picture)

SHOW CTM Ep 305 metadata (audio)

SHOW CTM Ep 306 metadata (audio)

SHOW CTM Ep 307 metadata (audio)

Root file: Essence

SHOW CTM EP 305 essence (picture)

SHOW CTM EP 306 essence (picture)

SHOW CTM EP 307 essence (picture)

SHOW CTM EP 305 essence (audio)

SHOW CTM EP 306 essence (audio)

SHOW CTM EP 307 essence (audio)

Root file: Checksum:

Container checksum

SHOW CTM EP 305 checksum (picture)

SHOW CTM EP 306 checksum (picture)

SHOW CTM EP 307 checksum (picture)

SHOW CTM EP 305 checksum (audio)

SHOW CTM EP 306 checksum (audio)

SHOW CTM EP 307 checksum (audio)

Normally our television is delivered as longplay, but if data comes by reel, then each reel becomes an essence, and the metadata and checksum files would be created as per the data configuration.

Note that the title used for labels and file paths must be the correct formal title of the show. i.e., THE SHOW or the authorized abbreviation, SHOW. The GPMS name (authorized abbreviation) is preferred for file paths because it is shorter than the natural language name. Please consult the Sony Digital Archive personnel to get the abbreviated name for any show.

APPENDIX A –CONTENT FORMATTING GUIDELINES

Overall Content Formatting (For All Content Types)

- Master image elements must be accompanied by audio that matches their frame rate and will sync with the image when clocked at 48.00 kHz.
- Video Frame Rate to be 23.976 fps. Any other frame rates such as 25.00 fps (for example), must be discussed and agreed to by Sony Pictures in writing prior to delivery.
- Aspect Ratio must remain fixed and consistent for the length of the program.
- SDR Video content to fall within Rec.709 legal video signal specifications, including the horizontal and vertical blanking.
- HDR Video content to adhere to Rec.2020 / ST-2084 (PQ) video signal specifications.
- Video content shall contain no crushed black levels, clipped white levels, dropouts, glitches and/or other technical flaws.
- No Time Compression, Time Expansion, Enhancement, Noise Reduction or Electronic Dirt Concealment Processing shall be applied.
- Audio reference level is -20dBFS and the program content shall fall within AES/EBU specifications. No clipping or distortion of the audio content shall be present. A 1kHz tone at -20dBFS is to be placed over the bars in the head leader on all channels.
- Audio files shall be contiguous from the beginning of the head leader to the end of the tail leader. Files must be consolidated this way prior to delivery and must conform to the proper spec.
- The beginning of the picture file and the beginning of the audio files must be aligned at 00:59:30:00 and play in perfect sync.
- First Frame Of Picture (FFOP) and First Frame Of Audio (FFOA) must align to the Start of Program, program audio must not precede FFOP.
- Last Frame Of Picture (LFOP) and Last Frame Of Audio (LFOA) must take into consideration any audio ring-out. The LFOP/LFOA is the last frame of visible picture (including logos) OR the last frame of audio ring-out, whichever comes later. Make sure that the 20 seconds of tail leader black begins AFTER the LFOP/LFOA as described here.

Scaling Requirements

SPE considers their UHD content a premium experience to clients and customers, therefore SPE doesn't allow upscaling of content to meet higher resolution targets. In instances where upscaling may be necessary, it must be approved in writing by SPE prior to implementation.

In instances where a production records principle photography with a camera that is sub-4K resolution (such as Arri Mini) but with expanded color space, upscaling may be permitted if the scaling occurs from the original camera resolution, original color space (or better) and using a high-quality scaler (available in most professional color grading systems). In comparison, upscaling from constrained SDR Rec.709 color space falls outside of the Sony allowed parameters as it would result in a suboptimal experience.

Directory & File Naming

General:

- File and directory names shall be case-sensitive.
- File and directory names shall consist only of the following 8-bit ASCII characters:
 - Letters A-Z and a-z
 - Numerals 0-9
 - Underscore “_”
 - Period “.”
 - Dash “-”
- Filenames will not contain whitespace and will begin with alphanumeric characters (a-z, A-Z, 0-9)
- All frames shall be regular files and directories (no symbolic references to other files, directories or devices).
- All file and directory paths will be relative paths (no absolutes)

Directory structure:

The directory structure delivered should meet the following specifications:

- 1) The project GMDM title should be included in the directory naming.
- 2) The GMDM title abbreviation should be included as part of the file and/or sequence directory name.
- 3) The aspect ratio should be included in the file or sequence directory name to differentiate between versions.
- 4) The directory that contains the image files should be labeled by the pixel resolution (X x Y) of the images.

Audio File naming

Audio sessions and audio files must be named and include their channel assignment as a suffix to the name per the SPE Audio File Naming Convention, available at <https://tekzone.spe.sony.com/tekzone>. (Click “new user” at the log in screen and fill in the form to gain access.) The GMDM title abbreviation for file naming is also available on Tekzone.

Picture File naming

ProRes file naming shall include a basename prefix that matches the sequence directory name with additional audio and video configuration of the content separated by underscores “_” examples:

/InterviewThe/InterviewThe_DomTxt_IMF9755002_EngLtRt_MELtRt_Eng51_OPME_16x9_240_2398_ProResHD_Archival.mov
/CrossingLines/CrossingLines_EP301_DomTxt_EngLtRt_MELtRt_16x9_178_2398.mov

Image file/sequence should meet the following specifications:

- 1) Image sequences should include a basename prefix that matches the sequence directory name
- 2) Image sequences should have 7-digit number padding

See specification example here:

GMDMname/GMDMabbrev_uhd_185_master/3840x2160/GMDMabbrev_uhd_185_master.[0085680-0216745].dpx
GMDMname/GMDMabbrev_hd_178_master/1920x1080/GMDMabbrev_hd_178_master.[0085680-0216745].dpx
GMDMname/GMDMabbrev_hdr_185_master/3840x2160/GMDMabbrev_hdr_185_master.[0085680-0216745].tiff
GMDMname/ep_301/GMDMabbrev_ep_301_uhd_178_master/3840x2160/GMDMabbrev_ep_301_uhd_178_master.[0085680-0149822].dpx

Actual example for the GMDM feature title “Chappie” and GMDM abbreviation “CHAPP0”:

chappie/chapp0_uhd_185_master/3840x2160/chapp0_uhd_185_master.[0085680-0216745].dpx
chappie/chapp0_hd_178_master/1920x1080/chapp0_hd_178_master.[0085680-0216745].dpx
chappie/chapp0_hdr_185_master/3840x2160/chapp0_hdr_185_master.[0085680-0216745].tiff

Actual example for the GMDM television title “The Blacklist” and GMDM abbreviation “BLACKL”:

blacklist_the/ep_301/blackl_ep_301_uhd_178_master/3840x2160/blackl_ep_301_uhd_178_master.[0085680-0149822].dpx

Formatting For Long-Form Content

Head Format:

All files should have embedded record TC into the file container (where supported). Image sequence files should be numbered according to TC frame numbering (included in example below). Formatting for all audio/video files should be as follows. Please note the first sequenced frame number is 85680 which is equivalent to 00:59:30:00 in 24 frame-based timecode.

Head Leader (TC)

Head Leader (Frames)

00:59:30:00 - 00:59:33:23

85680 – 85775

Slate

00:59:34:00 - 00:59:36:23

85776 – 85847

Black

00:59:37:00 - 00:59:46:23

85848 – 86087

Test Pattern (Bars) and 1kHz tone at -20dBFS

00:59:47:00 - 00:59:49:23	86088 – 86159	Black
00:59:50:00 - 00:59:50:01	86160	10-pop (picture and audio)
00:59:50:01 - 00:59:59:23	86161 – 86399	Black
01:00:00:00 - 01:00:00:01	86400 -	First Frame of Picture (FFOP)

Tail Format:

All files should contain 20 seconds of black after the last frame of picture (LFOP), followed by a tail pop and an additional 5 seconds of black:

- The LFOP/LFOA must take into consideration any audio ring-out. The LFOP/LFOA is the last frame of visible picture (including logos) OR the last frame of audio ring-out, whichever comes later. **Make sure that the 20 seconds of tail leader black begins AFTER the LFOP/LFOA as described here.**

Tail Leader

LFOP + 00:00:00:01 - LFOP + 00:00:19:23	Black
LFOP + 00:00:20:00 - LFOP + 00:00:20:01	Tail 20 Pop (picture and audio)
LFOP + 00:00:20:02 - LFOP + 00:00:25:01	Black

Textless Formatting:

Each texted main picture element requires a corresponding textless replacement. A textless slate (4 seconds), followed by black (3 seconds), should precede all the textless inserts.

In the case where image sequences are delivered, these shall be formatted as non-continuous sequences with accurate TC numbering that matches the appropriate insert TC. Textless image sequences are to only contain the exact frames matching the main program, textless image sequences shall not contain handles.

An example is provided below, which includes textless slate and black (continuous sequence 85680-85847), and three textless picture inserts (non-continuous sequences):

chapp0_uhd_185_master_txtls.[0085680-0085775]DPX	Textless Slate
chapp0_uhd_185_master_txtls.[0085776-0085847]DPX	Black
chapp0_uhd_185_master_txtls.[0087378-0087431]DPX	Textless Insert 1
chapp0_uhd_185_master_txtls.[0087846-0088245]DPX	Textless Insert 2
chapp0_uhd_185_master_txtls.[0088517-0089128]DPX	Textless Insert 3
chapp0_uhd_185_master_txtls.[0089139-0089258]DPX	Black

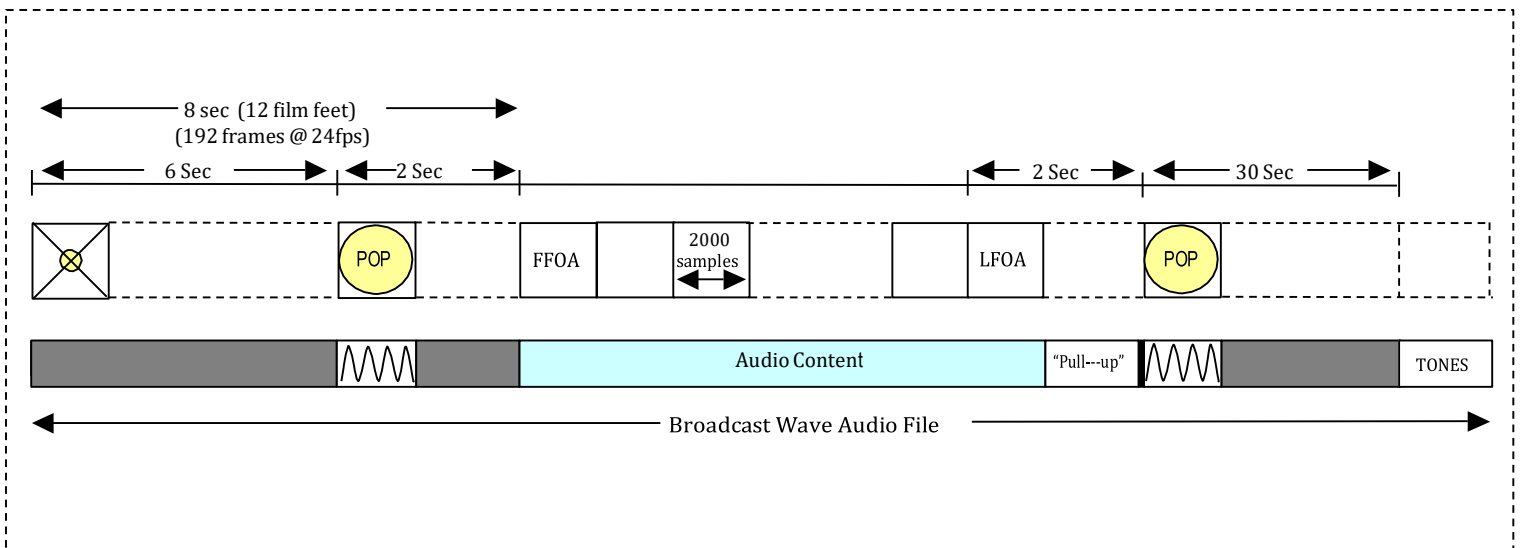
In the case where ProRes is delivered, textless shall be continuous without breaks in timecode and be delivered as part of the main program file also known as “textless at tail”. An EDL must be provided that accurately replaces the program content with the textless.

Formatting For Reel-Based Audio Content

- Audio files must start at exactly at zero feet (i.e., aligned exactly to the left edge of the picture start frame). Typically, this is at an even hour time code location (e.g., 03:00:00:00 for reel 3), but the spec is referred to the picture start frame independent of time code.
- There must be exactly 8 seconds (which is 12 film feet or 192 frames at 24 fps) from the start of the audio file to the FFOA.
- There must be a head pop at exactly 6 seconds (9 film feet) from the start of the audio file, one frame in duration. This pop therefore begins exactly 2 seconds before FFOA. All master tracks must contain this pop.
- There is to be a tail pop exactly 2 seconds from the beginning of the LFOA, one frame in duration. All master tracks must contain this pop.
- Tones are to be included at the end of the reel ONLY, starting at approximately 30 seconds (45 film feet) after the tail pop.
- Tones for each reel are to be consolidated into the audio files for that reel and are not to be in a separate session (i.e., “Tone Reels” are not permitted).
- Audio files should be contiguous from 0’ to end of tones for each reel.

- In order to attain this if not recorded this way on the stage, select this duration and then consolidate. The resultant file will have the proper spec. This must be done prior to delivery.
- Reel changeovers must be sample accurate between the LFOA of the outgoing reel and the FFOA of the incoming reel. This must be tested prior to delivery.
- Track audio files must include their channel assignment as a suffix to the name.
- Each stem type is in its own session, one per reel. No combined stem sessions.
- Audio file start = 0'
- Alignment tones at tail, 30 seconds after tail pop.

Reel-Based Audio File Formatting



APPENDIX B–EPISODIC / MOW /MFT MASTER CONTENT FORMATTING

CONTACTS:

Primary Contact: DCS_TITLESTEWARDSHIP@spe.sony.com

Episodic Mastering Contact: TV_Mastering@spe.sony.com

EPISODIC SPECIFIC FORMATTING

- Production logos must be at the tail of the episode, followed by the appropriate Sony Pictures Entertainment logo (see contract for correct logo).
- All commercial blacks are to be pulled between :01 and :02 seconds in length.
- No in-show bumpers (i.e., commercial in or out bumpers).
- Rapid detailed motion credits should be kept to a minimum and within 16x9 (4x3 preferred) center picture safe area. Static credit cards are preferable to crawls for reasons of standards conversion.
- Consolidated episodes (i.e., special 1 hour of Seinfeld) must be delivered in original length format (i.e., Two 30 minute episodes).
- No Network TV Ratings, Closed Captioning or In-Stereo logos are permitted.
- No voiceovers for bumpers (i.e., "We'll Be Right Back") in program.
- No address, telephone number or URL references in program.
- Audio sessions and audio files must be named and include their channel assignment as a suffix to the name per the SPE Audio File Naming Convention, available at <https://tekzone.spe.sony.com/tekzone>.
- Home Theater content should be 23.976 fps at 48.00 OR 96.00 kHz. 24.00 fps is accepted. (25.00 fps home theater content accepted only with prior arrangement with SPE.)
- For Episodic, Movie of the Week (MOW) and Made For Television (MFT) content; the audio should be monitored through standard television speakers or soundbar to ensure it plays well in a home environment, and as such is considered home theater audio for the purposes of this document.
- Labeling & Slating:
 - Episode labeling must comply with SPE labeling procedures: the first digit equates to the season of production, followed by the episode number production has assigned. For example, episode #101 is the first episode of the first season, episode #210 is the tenth episode of the second season, etc.
 - Material must be labeled and slated as follows:

Show Title (English/Native Language)
Episode #/Version Production #
Episode Title
Ch1- audio / Ch2- audio / Ch3- audio / Ch4- audio
Standard, Version - Runtime: XX:XX
First / Last Frame of Picture (FFOP / LFOP)

APPENDIX C –AUDIO TRACK LAYOUTS

Track Layouts by Audio Configuration

Below are the required ProTools session track layouts for each audio configuration. Alternate layouts permitted only with permission of Sony Pictures.

The layouts depict track 1 first and then subsequent tracks follow e.g., track 1, track 2, track 3.....

Commonly Used Audio Configurations

Monaural (M) for home use: dual mono pair.

M, M

Standard Stereo (ST) a.k.a. LoRo **Note, do NOT use the moniker 2.0--this audio configuration name is not permitted**

L, R

LtRt surround (DS) **Note, do NOT use the moniker 2.0--this audio configuration name is not permitted**

Lt, Rt

5.1 (51)

L, R, C, LFE, Ls, Rs

(L, C, R, Ls, Rs, LFE permitted only with prior permission)

5.0 (50) Note, this configuration does not have a LFE channel, silence (MOS) must be used in its stead

L, R, C, MOS, Ls, Rs

(L, C, R, Ls, Rs, MOS permitted only with prior permission)

7.1 (71) This is the 7.1 Discrete Surround (71DS) SMPTE MCA configuration in common use

L, R, C, LFE, Lss, Rss, Lrs, Rrs

(L, C, R, Lss, Rss, Lrs, Rrs, LFE permitted only with prior permission)

7.0 (70) Note, this configuration does not have a LFE channel, silence (MOS) must be used in its stead. SMPTE MCA label is 70DS

L, R, C, MOS, Lss, Rss, Lrs, Rrs

(L, C, R, Lss, Rss, Lrs, Rrs, MOS permitted only with prior permission)

ATMOS (ATM) (DTSX has the same layout) Note that Dolby and ProTools may have different channel abbreviations for rear surrounds and top surrounds. The SMPTE MCA abbreviations are used here:

Bed(s) L, R, C, LFE, Lss, Rss, Lrs, Rrs, Lts, Rts. If Lts and Rts are not used, MOS is used instead for each. Beds are in 7.1.2 tracks
Objects OBJ11-128. Objects are numbered starting from the first slot open after the beds up to 128

IMAX (IX) This is the standard theatrical IMAX configuration, often called IMAX 5.0 or IMAX 5.

Ls, Rs, C, L, R,

IMAX (IX60) This is the original IMAX configuration, often called IMAX 6.0 or IMAX 6, which includes a center height for very tall screens. The SMPTE MCA label is 60CH

Ls, Rs, C, L, R, Ch

IMAX 12.0 (IX12) This is the theatrical IMAX immersive audio configuration. The SMPTE MCA label is 120CH

L, R, C, Ch, Lrs, Rrs, Lss, Rss, Ltfs, Rtfs, Ltrs, Rtrs

IMAX-E 5.1 (IX51) This is the IMAX 5.1 enhanced configuration for the home, which includes an LFE channel

Ls, Rs, C, L, R, LFE

IMAX-E 12.1 (IX121) This is the IMAX 12.1 enhanced configuration for the home, which includes an LFE channel

L, R, C, Ch, Lrs, Rrs, Lss, Rss, Ltfs, Rtfs, Ltrs, Rtrs, LFE

Legacy Audio Configurations for ProTools Sessions

LCR (3.0) (Legacy content only)

L, R, C

LCRS (4.0) (Legacy content only)

L, R, C, S

6.1 (6D) (This is discrete Dolby EX and is legacy content only)

L, R, C, LFE, Ls, Rs, Cs

(L, C, R, Ls, Rs, Cs, LFE permitted only with prior permission)

“SDDS” (SDS) This is the 7.1 SDDS (71SDS) SMPTE MCA configuration that is legacy and not in common use today

L, R, C, LFE, Ls, Rs, Lc, Rc

(L, Lc, C, Rc, R, Ls, Rs, LFE is permitted)

Legacy Audio Configurations for standard DCP

The only legal audio configurations for standard DCP in SMPTE ST 429-2 that are in practical use are Configuration 1 (5.1 with HI and VI) and Configuration 5 (7.1 with HI and VI). There are others for SDDS and 6.1 but these are not widely implemented.

To convey legacy audio configurations in DCP, 5.1 or 7.1 is used, and MOS channels (silence channels with all digital zeros) are used for channels that are not part of the legacy audio configuration

Monaural (M) for DCP: single monaural channel to play from center speaker. Use 5.1 as follows:

MOS, MOS, C, MOS, MOS, MOS

Standard Stereo for DCP: Use 5.1 as follows:

L, R, MOS, MOS, MOS, MOS

LCR for DCP: Use 5,1 as follows:

L, R, C, MOS, MOS, MOS

Dolby Surround (Dolby A or SR) for DCP: There is no surround decoding in D-Cinema, so this is conveyed as LCRS. Use 5.1 as follows:

L, R, C, MOS, S, S The “S” mono surround should be reduced by 3dB

Dolby EX (6.1) for DCP: Decode the Lst and Rst channels into Ls, Rs, and Cs. Then use 7.1 as follows (Do not use Configuration 2, it is not widely implemented):

L, R, C, MOS, Ls, Rs, Cs, Cs (The rear center surround (Cs) should be reduced by 3dB

SDDS in DCP: (for special screenings only in known SDDS-compatible auditoriums), use Configuration 3

L, R, C, LFE, Ls, Rs, Lc, Rc

Session Layouts for Stems, M&E Stems and M&E with Optional Supersessions:

The below are suggested ProTools session layouts for Stems and M&E with Optional (MESP) supersessions. Track numbers can vary if individual items have more or less tracks

Stems:

5.1:

- 1-6 Dialog stem
- 7-12 Group (GP) or Crowd (CD) stem
- 13-18 Walla (WL) or Group (GP) stem
- 19-24 Music A (MX_A) stem
- 25-30 Music B (MX_B) stem
- 31-36 Effects (FX) stem
- 37-42 Foley (FL) stem
- 43-48 Backgrounds (BG) stem

7.1:

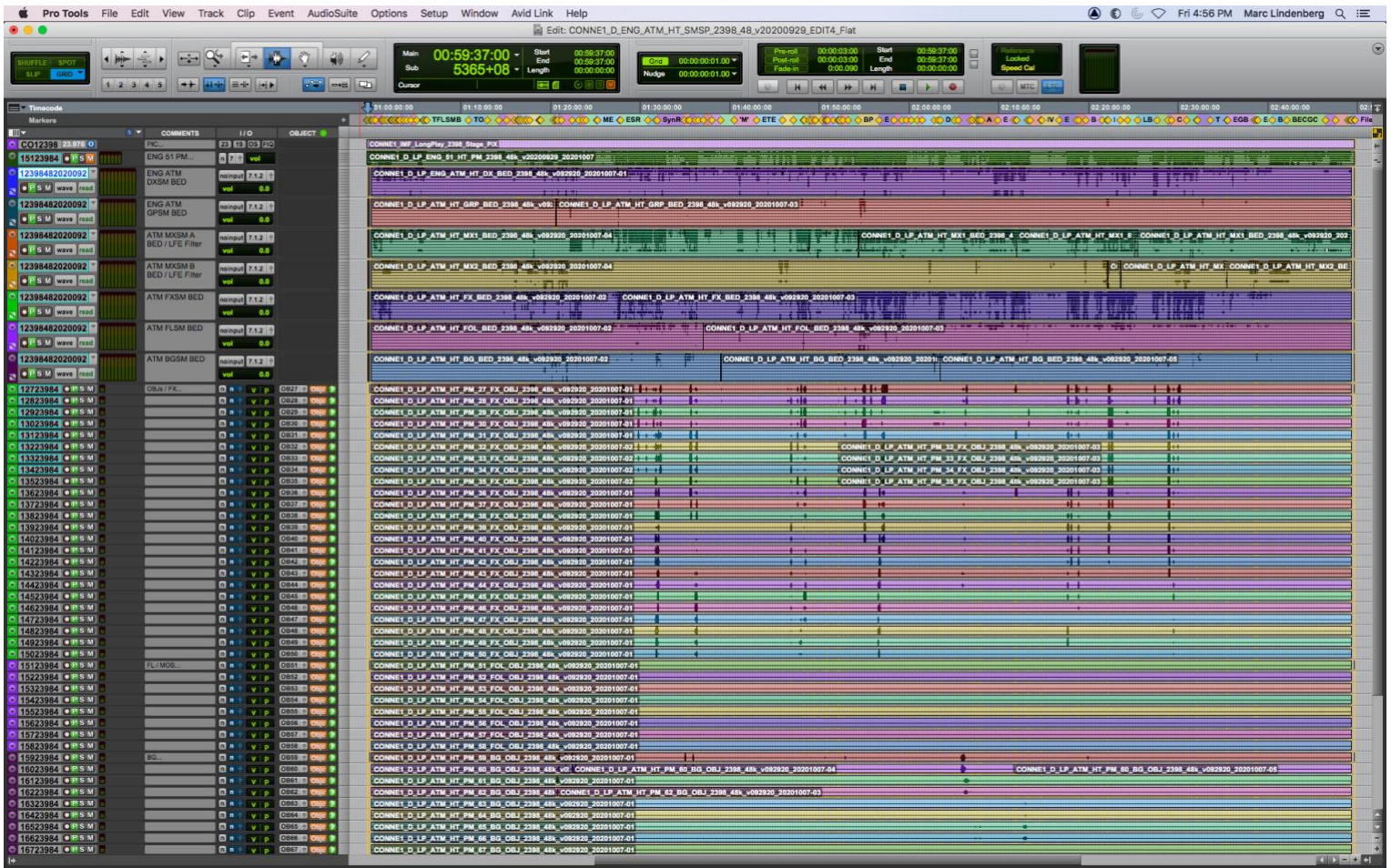
- 1-8 Dialog
- 9-16 Group (GP) or Crowd (CD) stem
- 17-24 Walla (WL) or Group (GP) stem
- 25-32 Music A (MX_A) stem
- 33-40 Music B (MX_B) stem
- 41-48 Effects (FX) stem
- 49-56 Foley (FL) stem
- 57-64 Backgrounds (BG) stem

Atmos:

For Atmos, put the stem beds (in 7.1.2 tracks) first, and put the objects after the beds in the same order. See figure.

- Dialog stem bed
- Group (GP) or Crowd (CD) stem bed
- Walla (WL) or Group (GP) stem bed
- Music A (MX_A) stem bed
- Music B (MX_B) stem bed
- Effects (FX) stem bed
- Foley (FL) stem bed
- Backgrounds (BG) stem bed

- Dialog (DX) objects
- Group (GP) or Crowd (CD) objects
- Walla (WL) or Group (GP) objects
- Music A (MX_A) objects
- Music B (MX_B) objects
- Effects (FX) objects
- Foley (FL) objects
- Backgrounds (BG) objects



Music and Effects Supersessions (MESP):

5.1 MESP:

- 1-7 6+1 Main M&E (ME) (This is 5.1 plus mono dialog guide (DXG), track 7 has Dialog Guide (DXG) (crash down of dialog stem)
- 8-14 Optional A (OPME_A)
- 15-21 Optional B (OPME_B)
- 22-27 Optional C (OPME_C)
- 28-34 Optional D (OPME_D)

7.1 MESP:

- 1-8 8+1 Main M&E (ME) (This is 7.1 plus mono dialog guide (DXG), track 9 has Dialog Guide (DXG) (crash down of dialog stem)
- 10-18 Optional A (OPME_A)
- 19-27 Optional B (OPME_B)
- 28-36 Optional C (OPME_C)
- 37-45 Optional D (OPME_D)

LtRt MESP

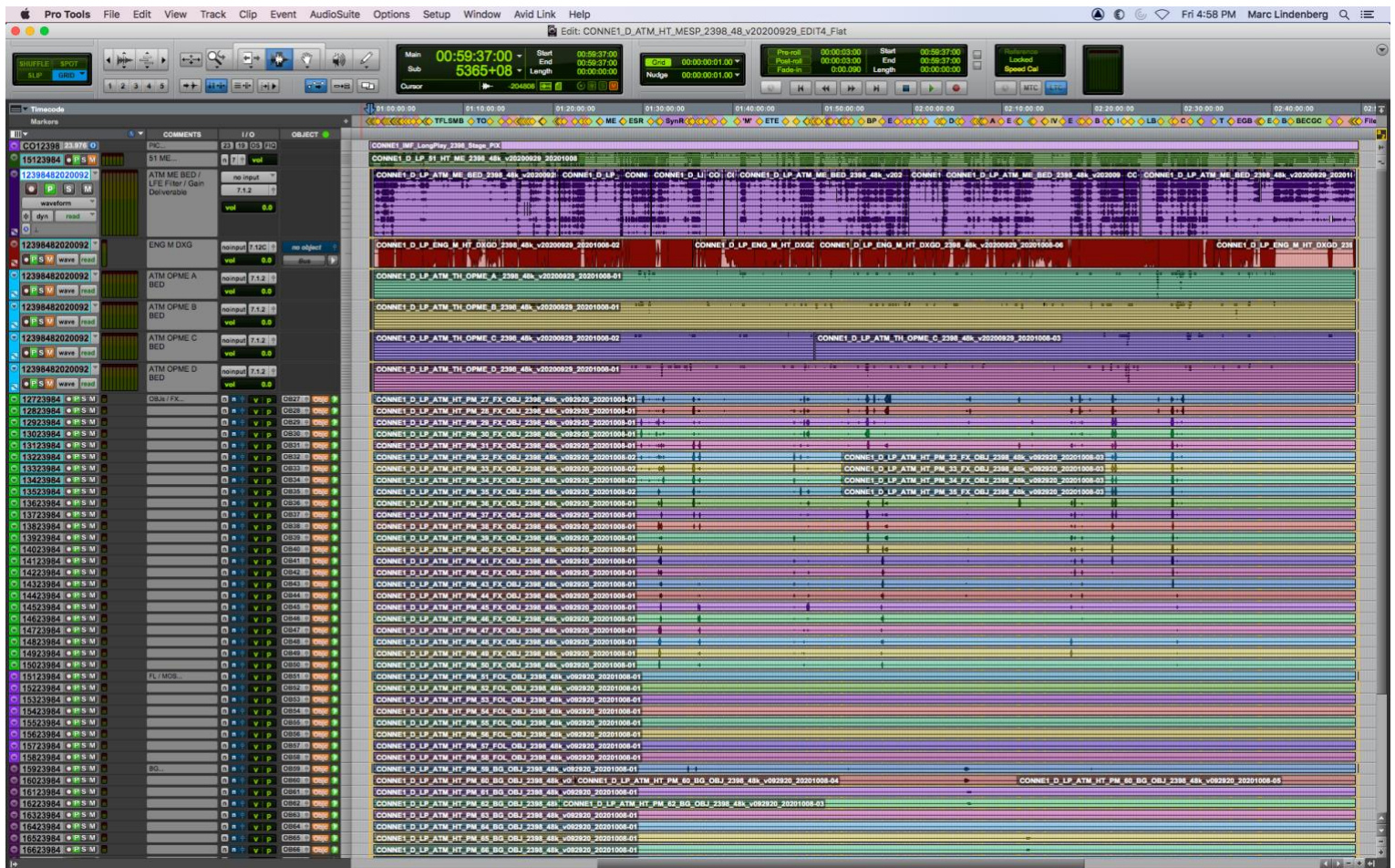
- 1-2 LtRt Main M&E (ME) (fold-down through surround encoder of 5.1 Main M&E)
- 3 Mono Optional 1 (simple crash down of chosen multichannel optionals)
- 4 Mono Optional 2 (simple crash down of chosen multichannel optionals)
- 5 Mono Dialog Guide (DXG) (simple crash down of dialog stem)

Atmos MESP:

For Atmos, put the M&E and Optional beds (in 7.1.2 tracks) first, and put the stem objects after the beds in the same order as the stems supersession. See figure.

- Main M&E (ME)
- Dialog Guide (DXG)
- Optional A (OPME_A)
- Optional B (OPME_B)
- Optional C (OPME_C)
- Optional D (OPME_D)

- Dialog (DX) objects (muted except for ones intended as optionals)
- Group (GP) or Crowd (CD) objects
- Walla (WL) or Group (GP) objects
- Music A (MX_A) objects
- Music B (MX_B) objects
- Effects (FX) objects
- Foley (FL) objects
- Backgrounds (BG) objects



M&E Stems (MESMSP):

5.1:

1-6 Fill M&E Stem (Dialog stem with dialog muted and fill added)
7-12 Group (GP) or Crowd (CD) M&E Stem (Subtract anything that is put to an optional)
13-18 Walla (WL) or Group (GP) M&E Stem (Subtract anything that is put to an optional)
19-24 Music A (MX_A) M&E Stem (Subtract anything that is put to an optional)
25-30 Music B (MX_B) M&E Stem (Subtract anything that is put to an optional)
31-36 Effects (FX) M&E Stem (usually same as regular FX stem)
37-42 Foley (FL) M&E Stem (usually same as regular FL stem)
43-48 Backgrounds (BG) M&E Stem (usually same as regular BG stem)

7.1:

1-8 Fill M&E Stem (Dialog stem with dialog muted and fill added)
9-16 Group (GP) or Crowd (CD) M&E Stem (Subtract anything that is put to an optional)
17-24 Walla (WL) or Group (GP) M&E Stem (Subtract anything that is put to an optional)
25-32 Music A (MX_A) M&E Stem (Subtract anything that is put to an optional)
33-40 Music B (MX_B) M&E Stem (Subtract anything that is put to an optional)
41-48 Effects (FX) (usually same as regular FX stem)
49-56 Foley (FL) (usually same as regular FL stem)
57-64 Backgrounds (BG) (usually same as regular BG stem)

Atmos:

For Atmos, put the stem beds (in 7.1.2 tracks) first, and put the objects after the beds in the same order. See figure.

Fill M&E Stem bed (Dialog stem with dialog muted and fill added)
Group (GP) or Crowd (CD) M&E Stem bed (Subtract anything that is put to an optional)
Walla (WL) or Group (GP) M&E Stem bed (Subtract anything that is put to an optional)
Music A (MX_A) M&E Stem bed (Subtract anything that is put to an optional)
Music B (MX_B) M&E Stem bed (Subtract anything that is put to an optional)
Effects (FX) stem bed (usually same as regular FX stem bed)
Foley (FL) stem bed (usually same as regular FL stem bed)
Backgrounds (BG) stem bed (usually same as regular FX stem bed)

Dialog (DX) objects
Group (GP) or Crowd (CD) objects
Walla (WL) or Group (GP) objects
Music A (MX_A) objects
Music B (MX_B) objects
Effects (FX) objects
Foley (FL) objects
Backgrounds (BG) objects

Pro Tools | File | Edit | View | Track | Clip | Event | AudioSuite | Options | Setup | Window | Avid Link | Help

00:59:37.00 | 5365+08 | 00:59:37.00 | 00:59:37.00 | 00:59:37.00 | 00:59:37.00

CONNET1_D_ATM_HT_MESMSP_2398_48_v20200929_EDIT4_Flat

Timecode: 01:50:00.00 | 01:50:00.00 | 01:50:00.00 | 01:50:00.00 | 01:50:00.00 | 01:50:00.00 | 01:50:00.00 | 01:50:00.00 | 01:50:00.00 | 01:50:00.00

COMMENTS | I/O | OBJECT

1239842020092 | ATM FLL | val | 0.0

1239842020092 | MESM BED / Includes Elements | val | 0.0

1239842020092 | ATM MX A MESM BED / LFE Filter | val | 0.0

1239842020092 | ATM MX B MESM BED / LFE Filter | val | 0.0

1239842020092 | ATM FX MESM BED | val | 0.0

1239842020092 | ATM FL MESM BED | val | 0.0

1239842020092 | ATM BG MESM BED | val | 0.0

12723984 | OBV / FX | 0:27 | 0:00

12823984 | 0:28 | 0:00

12923984 | 0:29 | 0:00

13023984 | 0:30 | 0:00

13123984 | 0:31 | 0:00

13223984 | 0:32 | 0:00

13323984 | 0:33 | 0:00

13423984 | 0:34 | 0:00

13523984 | 0:35 | 0:00

13623984 | 0:36 | 0:00

13723984 | 0:37 | 0:00

13823984 | 0:38 | 0:00

13923984 | 0:39 | 0:00

14023984 | 0:40 | 0:00

14123984 | 0:41 | 0:00

14223984 | 0:42 | 0:00

14323984 | 0:43 | 0:00

14423984 | 0:44 | 0:00

14523984 | 0:45 | 0:00

14623984 | 0:46 | 0:00

14723984 | 0:47 | 0:00

14823984 | 0:48 | 0:00

14923984 | 0:49 | 0:00

15023984 | 0:50 | 0:00

15123984 | FIL / MOD | 0:51 | 0:00

15223984 | 0:52 | 0:00

15323984 | 0:53 | 0:00

15423984 | 0:54 | 0:00

15523984 | 0:55 | 0:00

15623984 | 0:56 | 0:00

15723984 | 0:57 | 0:00

15823984 | 0:58 | 0:00

15923984 | BG | 0:59 | 0:00

16023984 | 1:00 | 0:00

16123984 | 1:01 | 0:00

16223984 | 1:02 | 0:00

16323984 | 1:03 | 0:00

16423984 | 1:04 | 0:00

16523984 | 1:05 | 0:00

16623984 | 1:06 | 0:00

16723984 | 1:07 | 0:00

16823984 | 1:08 | 0:00

16923984 | 1:09 | 0:00

17023984 | 1:10 | 0:00

CONNET1_LP_HT_ME_2398_48_v20200929_20201008-01

CONNET1_LP_ATM_HT_ME_2398_48_v20200929_20201008-02

CONNET1_LP_ATM_HT_ME_MX1_BED_2398_48_v20200929_20201008-04

CONNET1_LP_ATM_HT_ME_MX2_BED_2398_48_v20200929_20201008-04

CONNET1_LP_ATM_HT_ME_MX3_BED_2398_48_v20200929_20201008-04

CONNET1_LP_ATM_HT_ME_MX4_BED_2398_48_v20200929_20201008-04

CONNET1_LP_ATM_HT_ME_MX5_BED_2398_48_v20200929_20201008-04

CONNET1_LP_ATM_HT_ME_FOL_BED_2398_48_v20200929_20201008-02

CONNET1_LP_ATM_HT_ME_FOL_BED_2398_48_v20200929_20201008-03

CONNET1_LP_ATM_HT_ME_BG_BED_2398_48_v20200929_20201008-05

CONNET1_LP_ATM_HT_PM_27_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_28_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_29_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_30_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_31_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_32_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_33_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_34_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_35_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_36_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_37_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_38_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_39_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_40_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_41_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_42_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_43_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_44_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_45_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_46_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_47_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_48_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_49_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_50_FX_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_51_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_52_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_53_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_54_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_55_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_56_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_57_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_58_FOL_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_59_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_60_BG_ORJ_2398_48_v092920_20201008-04

CONNET1_LP_ATM_HT_PM_60_BG_ORJ_2398_48_v092920_20201008-05

CONNET1_LP_ATM_HT_PM_61_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_62_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_63_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_64_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_65_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_66_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_67_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_68_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_69_BG_ORJ_2398_48_v092920_20201008-01

CONNET1_LP_ATM_HT_PM_70_BG_ORJ_2398_48_v092920_20201008-01

Audio Layouts for MOV Containers

Discrete audio must always be delivered. Audio delivered within a MOV file is for reference only and does not substitute for discrete audio delivery.

- Formatting for audio delivered with a ProRes video: .wav in .mov container
- The Frame Rate of audio in a .mov file must match frame rate of the .mov picture.
- If a non-native frame rate .mov picture has been agreed for delivery the audio in the .mov container should be pitch corrected. Note, no pitch correction is necessary between 23.976 and 24.00 fps.

Audio Channel Layout for MOV container for Feature content

- Channel 1 = LtRt Original Language Left Total (Surround encoded)
- Channel 2 = LtRt Original Language Right Total (Surround encoded)

Audio Channel Layouts for MOV container for Episodic content only:

Configuration 1 (Preferred):

- Channel 1 = LtRt Original Language Left Total (Surround encoded)
- Channel 2 = LtRt Original Language Right Total (Surround encoded)
- Channel 3 = LtRt Music & Effects Left Total (Surround encoded)
- Channel 4 = LtRt Music & Effects Right Total (Surround encoded)
- Channel 5 = 5.1 Original Language Left
- Channel 6 = 5.1 Original Language Right
- Channel 7 = 5.1 Original Language Center
- Channel 8 = 5.1 Original Language Subwoofer
- Channel 9 = 5.1 Original Language Left Surround
- Channel 10 = 5.1 Original Language Right Surround
- Channel 11 = 5.1 Music & Effects Left
- Channel 12 = 5.1 Music & Effects Right
- Channel 13 = 5.1 Music & Effects Center
- Channel 14 = 5.1 Music & Effects Subwoofer
- Channel 15 = 5.1 Music & Effects Left Surround
- Channel 16 = 5.1 Music & Effects Right Surround

Configuration 2:

- Channel 1 = LtRt Original Language Left Total (Surround encoded)
- Channel 2 = LtRt Original Language Right Total (Surround encoded)
- Channel 3 = LtRt Music & Effects Left Total (Surround encoded)
- Channel 4 = LtRt Music & Effects Right Total (Surround encoded)

Configuration 3:

- Channel 1 = LtRt Original Language Left Total (Surround encoded)
- Channel 2 = LtRt Original Language Right Total (Surround encoded)

APPENDIX D – MUSIC AND EFFECTS STEMS SPECIFICATIONS

PREFACE:

M&E stems are to be delivered **in addition to** the mixed fully filled M&E, multichannel optionals and fully filled Effects Stem.

M&E Stem configuration should match the main Stem configuration of the Domestic mix, with the Fill M&E stem taking the place of the Dialog stem

DESCRIPTION:

M&E Stems are the individual stems that make up the fully-filled music and effects element. The M&E stems are defined as follows:

Fill M&E Stem: This is all of the fill that needs to be added to the music and effects stems to create the filled M&E. It is typically the dialog stem with all dialog muted (leaving only production effects), plus any fill that was created to cover areas where production effects were lost when the dialog was muted.

Group M&E Stem: This is the Group stem with any specifics removed (i.e., discernable callouts, etc.). This is delivered only if a Group stem was created for the title.

Crowd M&E Stem: This is the Crowd stem with any specifics removed (i.e., discernable callouts, etc.) This is delivered only if a Crowd stem was created for the title.

Walla M&E Stem: This is the Walla stem with any specifics removed (i.e., discernable dialog, etc.) This is delivered only if a Walla stem was created for the title.

Foley M&E Stem: This is generally the same as the regular Foley Stem, but may have been updated depending on decisions made during the M&E process. It is renamed and delivered with the M&E stems

Background M&E Stem: This is generally the same as the regular BG Stem, renamed and delivered with the M&E stems

Effects M&E stem: This is generally the same as the regular FX Stem, renamed and delivered with the M&E stems

Music M&E stem: This is generally the same as the regular MX Stem, renamed and delivered with the M&E stems. However, in some cases this may be modified to remove vocals or an entire piece of music or song that is then put into an optional. This allows the vocal to be dubbed or the song replaced by music from a particular territory.

Any other stem that was created for the title that is part of the M&E would be included similarly to the above.

SPECIFICATIONS:

Theatrical M&E stems: ProTools 12 or above, 48.00 kHz or 96.00 kHz at 24.00 fps (47.952 kHz or 95.904 kHz at 23.976 fps is equivalent). 48.00 kHz or 96.00 kHz at 23.976 fps is accepted.

Home Theater M&E stems: ProTools 12 or above, 48.00 kHz or 96.00 kHz at 23.976 fps

FORMATTING:

M&E stems are delivered in their own ProTools supersession that contains all elements, placed at hour=reel on the timeline. If the project is long form, the M&E stems supersession is formatted to start at the beginning of the picture file (e.g., 00:59:30:00).

M&E Stem configuration should match the main Stem configuration of the Domestic mix, with the Fill M&E stem taking the place of the Dialog stem

NOTE, **optionals are not delivered in the M&E stem supersession.** Optionals are delivered in the M&E supersession.

DELIVERY:

All titles must deliver a set of 5.1 M&E stems. A typical example is below, other stems are possible on a per-title basis:

- 5.1 Fill M&E Stem
- 5.1 Crowd or Group M&E Stem (if created for the title). Depending on the content, this may have specific dialog removed for the M&E
- 5.1 Walla M&E Stem (if created for the title)
- 5.1 Background M&E Stem
- 5.1 Foley M&E Stem
- 5.1 Effects M&E Stem
- 5.1 Music M&E Stem

If other formats are created for the title, deliver to the same pattern for:

- 7.1
- Atmos
- IMAX 5.0
- IMAX 12.0
- DTS:X
- Auro 11.1
- Auromax

FILE NAMING AND GOLD VALUES:

Theatrical (TH) and home theater (HT) are in the “usage” field

In order to clearly differentiate M&E stems supersession from the stems that make up the printmaster, the “Audio Element Type” abbreviation is MESM. Do not use ME_SM.

The “Audio Content” abbreviation precedes the element abbreviation and is separated by an underscore. The abbreviations are the same, with the addition of “FILL”

The individual M&E stems within the supersession are named in this style:

TH_FILL_MESM
TH_FX_MESM
TH_MX_MESM

HT_FILL_MESM
HT_FX_MESM
HT_MX_MESM

<u>Item</u>	<u>File Name</u>	<u>Gold Element Audio Type Value</u>
M&E Stems Supersession	MESMSP	M&E Stems Supersession
<u>Item</u>	<u>File Name</u>	<u>Gold Track Audio Type Value</u>
M&E Background stem	BG_MESM	M&E BG Stem
M&E Background and Foley stem	BG_FOL_MESM	M&E BG/Foley Stem
M&E Crowd Stem	CRD_MESM	M&E Crowd Stem
M&E Effects Stem	FX_MESM	M&E Effects Stem
M&E Extra Stem	EX_MESM	M&E Extra Stem
M&E Fill Stem	FILL_MESM	M&E Fill Stem
M&E Foley Stem	FOL_MESM	M&E Foley Stem
M&E Group Stem	GRP_MESM	M&E Group Stem
M&E Laughs Stem	LF_MESM	M&E Laughs Stem
M&E Music Stem	MX_MESM	M&E Music Stem
M&E Sound Design Stem	DZN_MESM	M&E SND Design Stem
M&E Walla Stem	WL_MESM	M&E Walla Stem

APPENDIX E – GLOSSARY OF TERMS

PICTURE-RELATED TERMS

3DDCDM: 3D Digital Cinema Distribution Master - The 3DDCDM is a DCI-compliant element that is the parent to a 3D version of the show. Productions with a 3D version must be delivered to Sony’s designated distribution partner (Deluxe). After the feature has been exploited, this element must be delivered to Sony Digital Archive. If there are alternate versions, one version of each must be delivered. This data is delivered as X’Y’Z’ TIFFs, and must include all textless background material.

Metadata: Standard per **Documentation and Labeling Requirements** document. Identification of the type of 3D and any special technical characteristics must be included.

3DDCP: 3D Digital Cinema Package - The 3DDCP is a DCI-compliant element that is the theatrical display element for digital cinema. This element is normally created by Sony’s designated distribution partner (Deluxe). One full show unencrypted SMPTE DCP of the original theatrical version must be delivered to the Sony Digital Archive. If there are alternate versions, one version of each must be delivered, and must include all textless background material.

Metadata: Standard per **Documentation and Labeling Requirements** document.

CTM: Color Timed Master - is the primary conformed deliverable for television. The CTM is the final, color-corrected, legal and talent-approved, broadcast-ready full show with textless component and any required inserts and logos, at the highest resolution and bit depth available. The “full show” includes slate, head and tail leader, original approved main title and end credits, all alternate main title elements, textless units, alternate versions shots or scenes. When there are two versions of an episode, both versions must be delivered. All re-caps and previews must be included. This data must be delivered in P3 D65 or Rec709 color in the form of DPX files. If the show is derived from an ACES workflow, EXR files would be the optimal delivery. The CTM should also contain the conformed audio (LtRt, 5.1, DME) as a reference. This element includes all color correction and complete picture in a fully integrated, audio-conformed state. The CTM may be delivered in DPX (or EXR files when in ACES), and it is finished in HD, UHD or 4K resolution per Sony’s requirements. The CTM should be delivered to Sony Digital Archive within a month of the date of final show broadcast.

Metadata: Standard per **Documentation and Labeling Requirements** document.

DCDM: Digital Cinema Distribution Master - The DCDM is a DCI-compliant element that is the parent to digital cinema version of a feature. The DCDM must be delivered to Sony’s designated distribution partner (Deluxe). After the feature has been exploited, this element must be delivered to Sony Digital Archive. If there are alternate versions, each must be delivered. This data is delivered as X’Y’Z’ TIFFs and must include all textless background material.

Metadata: Standard per **Documentation and Labeling Requirements** document.

DCP: Digital Cinema Package - The DCP is a DCI-compliant element that is the theatrical display element for digital cinema. This element is normally created by Sony’s designated distribution partner (Deluxe). One full show unencrypted SMPTE DCP of the original theatrical version must be delivered to the Sony Digital Archive. If there are alternate versions, one version of each must be delivered, and must include all textless background material.

Metadata: Standard per **Documentation and Labeling Requirements** document.

DI: Digital Intermediate - the primary deliverable for features is a fully finished, full show version of the feature including all textless background components, uncompressed and unprocessed, at the highest resolution and bit depth available. The “full show” includes slate, head and tail leader, original theatrical main title and end credits, all alternate main title elements, textless units, alternate versions (director’s or extended cut) shots or scenes. For example, a fully delivered primary data delivery might include the full theatrical show and textless as one set of LTOs, and a second set of LTOs for an extended version. This data is generally P3 D65 or Rec709 color delivered in the form of DPX files. If the DI is derived from an ACES workflow, EXR files would be the optimal delivery. This component is a color-corrected master. The DI deliverable should also contain the conformed audio (LtRt, 5.1, DME, in .wav) as a reference.

Metadata: Standard per **Documentation and Labeling Requirements** document. In the case of material produced in ACES, all transforms, LUTs or other devices must be included in text files.

DSM: Digital Source Master - is a conformed, cleaned master of the feature created at the highest bit depth and fullest sensor dimensions (i.e., in some cases more than 4K). This element is colored, but with color not baked in but provided in the form of a LUT. Typically, the element may be in 16bit DPX or EXR, and in P3D65, camera log or other wide gamut color space. This element is generally composed of camera exports and uncorrected units such as VFX and opticals. The element is not sized or formatted for broadcast and may include files of different sizes (for example, camera files might be 4096x2304, while VFX might be 3820x2160).

This element is basically a rough conform based on the most original or raw data available. It is the parent of elements such as the DI, DCDM, and 4K HV masters. It can be delivered as a fully textless version of the show, but if delivered as texted, it must also include textless for the texted sections. The DSM should be delivered to Sony Digital Archive along with the DI within one month of the release of the feature.

Metadata: Standard per **Documentation and Labeling Requirements** document, but also includes all transforms, LUTs, EDLs, CDLs, HDR documentation (where applicable) as well as Baselight scenes (or Resolve sessions) that were used to modify the data to create downstream products.

FPDC: Final Picture Data Clone - The FPDC is the primary restoration deliverable; it is the final, talent-approved, distribution-ready full show with textless component and any required inserts. This element includes all color correction and complete picture in a fully integrated, audio-conformed state. The FPDC may be delivered in DPX or EXR files, and it should be finished at HD, UHD or 4K resolution. The FPDC should also contain the conformed audio (LtRt, 5.1, DME,) as a reference. The FPDC should be delivered to Sony Digital Archive within a week following the completion of the restoration project.

Metadata: Standard per **Documentation and Labeling Requirements** document.

HDRDM: High Dynamic Range Digital Master - The HDRDM is a deliverable which may be additionally specified by the Sony production contract. The HDRDM is the equivalent of the CTM, DI or FPDC, but graded for High Dynamic Range display. The HDRDM is only produced in cases where Sony specifically commissions its creation. This element includes all color correction and complete picture in a fully integrated, audio-conformed state, and must include the textless component. The HDRDM is Rec.2020 / ST-2084 (PQ) high dynamic range color space, formatted UHD, and delivered as TIFF image sequences. HDR data must be accompanied by specific additional metadata (EOTF, Reference Luminance, Reference White, MAXCLL, MAXFALL CSV files). Consult Sony Digital Archive for instructions on this requirement.

Metadata: Standard per **Documentation and Labeling Requirements** document.

IPDC: Intermediate Picture Data Clone - This element is a fully conformed version of the data; it should be cleaned and color-corrected, without the color baked in, but with accompanying color metadata. Normally this is a 16bit data resource and represents the highest resolution and widest dynamic range extracted from the raw scans. It is not necessarily sized for final delivery – it may have a resolution characteristic that reflects the sensor rather than a 4K or UHD size. This element is sometimes referred to as a “source file.” It may be in P3 or in a wide-gamut log space, depending on workflow. IPDCs are generally in DPX format unless derived from an ACES workflow, in which case they may be in the form of EXR files. The IPDC deliverable is not constrained by the EXR and DPX templates outlined in SPE HD/UHD Master Archive Delivery Specifications document. Preferably, the IPDC is delivered as a textless element, but in cases where the scanned element has text, the delivery requires the corresponding textless elements. The IPDC should be delivered with the FPDC to Sony Digital Archive within a week following the completion of the restoration project.

Production Archive - Includes all work product and applies to both theatrical and television productions. Delivery of the full production archive must include raw camera files, dailies, assemblies, effects, titles, opticals, lifts, editorial drives, EDLs, CDLs, all documentation, scripts, and other records of the production (including wrap drive) and must contain all metadata germane to the various data resources. It also must include the Sound Editorial Turnover, and Audio Collated Archive as specified in these technical specifications. This production material is not constrained by the specifications in SPE TV Master File Delivery Specifications, and this material does not require the Standard Metadata Block. This data must be delivered on LTO-7 or LTO-8 in LTFS 2.2 (or later). This data cannot be delivered in a proprietary format, such as Retrospect or BRU. When the production archive is ready to be delivered, the Digital Archive must be informed, and the LTOs that comprise the production archive will be moved to Sony Digital Archive for disposition. The data supplier must not purge any material until Sony Digital Archive has validated the delivery. The production archive must be delivered within one month after the release of the feature.

Metadata: All production documentation, LUTs, transforms and EDLs must be included and their location in the material identified.

RSC: Raw Scans Clone - In cases where the original or source element for the restoration is film, the first deliverable is the raw scans. These should be completely raw and free of processing of any kind (stabilization, sharpening, dust-busting, color correction, grain adjustment). Normally, these scans are in the form of 16bit DPX files, but TIFF, EXR or ADX files are equally acceptable. In cases where the scan implements ACES, the reference render transform and any additional metadata must be included in the delivery. The raw scans should be delivered to Sony Digital Archive at the completion of the restoration project.

Metadata: Standard per **Documentation and Labeling Requirements** document.

Source Files (SRC) – The SRC is a specific configuration of IPDC based on raw scans that have been restored – that is, a conformed element that has been cleaned and repaired to the point where it can serve as the data matrix for the remainder of workflow processes on the way to a final data set. Restoration work and QC fixes are integrated into the SRC. If titles or graphics are created in restoration, these are included in the SRC. The SRC is not color corrected. The SRC must be accompanied by all available Baselight scenes (or documentation of a similar working environment such as Resolve sessions) and EDLs.

Metadata: Standard per **Documentation and Labeling Requirements** document.

VAM: Video Assembled Master - is the second conformed deliverable. The VAM is an intermediate element that is created at the highest resolution and widest color range used for creation of the program. The VAM is a conformed, cleaned master of a television episode or MOW created at the highest bit depth and fullest sensor dimensions (i.e., in some cases more than 4K). This element is colored, but with color not baked in but provided in the form of a LUT. Typically, the element is 16bit DPX or EXR, and in P3D65, camera log or other wide gamut color space. This element is generally composed of camera exports and uncorrected units such as VFX and opticals. The element is not sized or formatted for broadcast and may include files of different sizes (for example, camera files might be 4096x2304, while VFX might be 3840x2160). This element is a rough conform based on the most original or data available. It is the parent of elements such as the 4K HV masters. It can be delivered as a fully textless version of the show, but if delivered as texted, it must also include textless for the texted sections. The VAM deliverable is not constrained by the EXR and DPX templates outlined in the SPE HD/UHD Master Archive Delivery Specifications document. For example, the VAM may have higher than 4K resolution. The VAM should be delivered to Sony Digital Archive along with the CTM within one month of the broadcast or streaming of the final episode of the series.

Metadata: Standard per **Documentation and Labeling Requirements** document.

AUDIO ELEMENTS

DME (Dialogue Music and Effects) (a.k.a. split track) – A reduction of the stems into a six-track element containing LtRt DX, LtRt MX and LtRt FX. The FX is a combination of BG, FL and FX. This is generally used for servicing needs such as advertising and menus and can also be used for certain edited versions.

Fully-Filled Effects Stem (FFX) – A stem that is the composite of all M&E Stems except the music. A FFX stem contains all of the fill and all of the effects (BG, FL, FX). Traditionally, the Effects stem was filled to create a FFX stem, and then this was mixed with the music to create the Main M&E. The advantage of the FFX stem today is the ability to create new M&E's easily if a title is later distributed to a medium or market where certain music is not cleared. This new M&E can then be used to create dubs for the new medium.

Music and Effects (ME) (a.k.a. Main M&E) – This is the soundtrack without the Dialog and is used to create dubbed language elements. It is the combination of the Music and Effects stems with additional material added to fill (substitute for) production effects in areas where dialog was removed. Depending on the soundtrack content, some content may be removed from the Main M&E and put into an “M&E optional”, which can be used as-is or dubbed. A Filled M&E can be created from M&E stems or by mixing the MX Stem with the Filled-Effects (FFX) stem.

Music and Effects Stems (MESM) – Stems that are used to create the Main M&E. These are created from the Printmaster Stems and mirror them in overall layout, but are modified as needed to create the Main M&E. The Dialog Stem is “flipped” to remove all of the dialog, and “fill” is substituted for the production effects lost when removing the dialog. This stem is renamed “M&E Fill Stem.” If any content is pulled out into an optional, it is also removed from the corresponding MESM. For example, if a vocal is put to an Optional, it is removed from the MX MESM. See APPENDIX D –MUSIC AND EFFECTS STEMS SPECIFICATIONS for details on creating and naming M&E stems.

Music and Effects Supersession (MESP) – This is a ProTools session containing the Main M&E, all of the Optionals and a Dialog Guide. By specific request it may also include the Dialog Stem. The MESP is the preferred delivery method rather than splitting up the Main M&E and the Optionals into separate sessions.

Optional (OP) (a.k.a. M&E Optional) – This is an element that contains content from the soundtrack that can either be dubbed or used as-is for localization. Examples of optional material are: actor efforts, dialog in a language native to a location, vocals for songs. More than one Optional may be created, each containing different optional material. Each Optional is labeled with a letter or number, and generally matches the audio configuration of the Main M&E. Any content that is in an Optional must not be in the Main M&E-the Main M&E and all Optionals must be able to play simultaneously with no overlap or common material.

Predub (PD) – Predubs are the product of the first round of mixing. All of the tracks that have been created by sound editorial are combined into manageable, fully-mixed elements in preparation for the final mix. For example, Effects Predubs are created by type, such as cars, helicopters, doors, room backgrounds, outdoor backgrounds, Foley footsteps, Foley props. Dialog is predubbed into principal and ancillary characters, and the ADR is predubbed to match the dialog it replaces or adds to. Music is generally not predubbed. The predubs are then fed to the final mix, where the Stems are created.

Printmaster (PM) – The combination of the stems into a single mixed element, which is used to create the final distribution master, such as a DCP, IMF or video master. The name derives from “printing master”, which was the name for the composite mix used as the source for the sound on a film print. Printmasters are traditionally in reel lengths but can be full length.

Stem (SM) (a.k.a. Printmaster Stem) – A stem is a fully-mixed audio element that is the output of the final mix automation and processing for a constituent part of the soundtrack. For example, the final mix of the dialog is the Dialog Stem (DXSM), the final mix of the music is the Music Stem (MXSM), and the final mix of the effects is the Effects Stem (FXSM). These stems are mixed in the intended native format (e.g., 5.1, Atmos). They are deliverables on their own, and are also combined to create Printmasters, Music & Effects and other Stems in the various audio configurations required for the final deliverables. The Stems are delivered together in one ProTools session.

Wide Stems – While the basic components of a soundtrack are Dialog, Music and Effects, it is best to split these up in finer components so that they are the most usable for the future. These are called “Wide Stems”. Typically, dialog is further split up into Dialog (DX), Crowd (CD) and Walla (WL). Music may have a Music A (MX_A) and Music B (MX_B), which differentiates source music and score music. Effects are further split into Backgrounds (BG), Foley (FL) and Effects (FX). Each of these is a fully-mixed Stem. SPE specifies that wide stems are delivered on each project, which are together in one ProTools session.

AUDIO-TERMINOLOGY AND FORMATS

Audio Channel (a.k.a. “Channel”) – a distinct collection of sequenced audio samples that are intended for delivery to a single loudspeaker or other reproduction device

Audio Configuration (a.k.a. Soundfield Group) – a collection (group) of Audio Channels meant to be played out simultaneously through a given Soundfield Configuration. In general, there is one audio channel for each loudspeaker, array of loudspeakers, or other sound reproducer. An example of an array of loudspeakers are the surround loudspeakers on the wall of a traditional movie theater.

Dual Mono (DM) – In some applications, mono can be conveyed as “Dual Mono”, with two identical channels of mono audio feeding two loudspeakers or other sound reproducers. While this might appear to be Standard Stereo, it is not because the two Audio Channels are identical.

IMAX 5.0 – The classic IMAX system. It has similar LCR front channels and two surrounds as 5.1 but does not have a dedicated LFE channel-the subwoofer information is derived via bass management (picking off frequencies below 70Hz from the main speakers and steering them to the subwoofer). The two surrounds are in the rear corners of the theater and are very large, essentially the same as the screen loudspeakers. The format actually offers eight channels, most notably a center height speaker above the tall screen, which is occasionally used, and the other two are rarely used today, hence the moniker “5.0.” IMAX is a cinema format. Recently, IMAX has introduced IMAX Enhanced, which allows the IMAX mix to be played on a home theater system.

IMAX 12.0 – The IMAX Immersive Sound system. This is the classic IMAX system with additional loudspeakers on the walls and ceiling to convey the height information, all of which are full range loudspeakers. IMAX 12.0 is a channel-based system and does not use Audio Objects or an Immersive Audio Renderer. If Audio Objects were used in the creation of the soundtrack, they are pre-rendered into channels for the IMAX 12.0 delivery.

IMAX Enhanced – The home IMAX delivery system. For audio, an LFE channel is added to the theatrical IMAX 5.0 or 12.0 track by bandpassing at 70Hz as a cinema IMAX system would do, making them 5.1 and 12.1 respectively. IMAX Enhanced audio is encoded with DTS:X and delivered to the home.

LtRt (a.k.a. Stereo Surround) – An Audio Configuration with two Audio Channels (Left total and Right total) that contain encoded surround information. When played through a modern surround decoder, it yields six Audio Channels that can be played through a 5.1 sound system. LtRt is the preferred two channel delivery for SPE because the encoded surround information can be perceived by the listener regardless of the playback environment-it does not have to be decoded to be perceived. LtRt plays well through a Standard Stereo system, and the surround information enhances the experience, whereas LoRo audio can only be reproduced as simple stereo.

Monaural (M) (a.k.a. “Mono”) – A Soundfield Configuration consisting of a single loudspeaker. In cinema this is a single loudspeaker behind the screen. A mono Audio Configuration has one Audio Channel.

ProTools – The Digital Audio Workstation (DAW) platform from Avid Technologies. The ProTools session format with .wav audio files is the Sony Pictures preferred audio delivery format.

Sony Dynamic Digital Sound (SDDS) – A legacy Soundfield Configuration that is similar to 5.1 and adds a Left Center (Lc) and Right Center (Rc) loudspeaker behind the screen. An SDDS Audio Configuration contains eight corresponding Audio Channels. Note that though there are eight channels and is sometimes mistakenly called 7.1, this is a distinctly different format than 7.1 and they are not interchangeable. There are no SDDS home theater systems, it is strictly a cinema system. SDDS was primarily active between 1993 and 2007, some SDDS releases continued into the next decade but were 5.1 only, the Lc and Rc channels were not used.

Soundfield – the acoustical space created by simultaneously reproducing one or more audio channels

Soundfield Configuration – a defined arrangement or configuration of loudspeakers or other sound reproducers that convey the intended Soundfield.

Standard Stereo (ST) (a.k.a. LoRo) – A Soundfield Configuration with left and right loudspeakers or other sound reproducers, which is intended to convey a front soundstage and has no surrounds. A Standard Stereo Audio Configuration has two unique Audio Channels (left and right) that contain no surround information.

5.1 – A Soundfield Configuration with Left (L), Center (C), Right (R), Left Surround (Ls), Right Surround (Rs), and Low Frequency Effects (LFE) loudspeakers. A 5.1 Audio Configuration contains six corresponding Audio Channels. In Cinema, the L, C, R loudspeakers are usually 2 to 4-way loudspeaker stacks behind the screen, and the Ls and Rs are generally arrays of smaller loudspeakers on the side and rear walls of the theater. The LFE channel is played through a number of subwoofers placed under the screen. In a typical home system, there are usually five individual loudspeaker boxes, each of which may be 2 or 3-way. The center speaker is below or above the video monitor, the L and R are offset 30 degrees from center, and the surrounds are 120 degrees from center. The subwoofer is usually a single cabinet placed near the front; larger systems may have multiple subwoofers.

7.1 (a.k.a. 7.1DS) – A Soundfield Configuration similar to 5.1 but adds two surrounds in order to split side and rear. It has L, C, R, Left Side Surround (Lss), Right Side Surround (Rss), Left Rear Surround (Lrs) and Right Rear Surround (Rrs). In Cinema, the Lss and Rss are on the walls of the theater, and the Lrs and Rrs are on the rear wall on either side of the projector. In a typical home system, the L, C, R, and subwoofer are the same as a 5.1 system, the side surrounds are 90 to 120 degrees offset, and the rear surrounds by offset 150-180 degrees depending on the room.

.wav (a.k.a. “Wave file”) – This stands for “Waveform Audio File Format”, which is a Pulse Code Modulation (PCM) audio format originally created by IBM and Microsoft. This was later modified to include metadata in the file header and called “Broadcast Wave Format” (BWF). This is a published specification, and .wav files are interoperable between many different platforms, making it the current industry choice for audio files. All audio files delivered to SPE must be .wav files.

AUDIO-IMMERSIVE AUDIO TERMINOLOGY

ATMOS® – The proprietary Immersive Audio system created by Dolby®. An Atmos sound system has a 7.1 Base Layer and two rows of Top Layer speakers overhead. It does not have Height Layer speakers. In a Cinema system, there are two rows of overhead loudspeakers on the ceiling that span the depth of the auditorium, which are driven by the Left Top Surround (Lts) and Right Top Surround (Rts) channels. All loudspeakers can be individually addressed via Audio Objects. The surround loudspeakers generally

have additional subwoofers to bass manage and extend the low frequency range of the surround speakers to more closely match the screen speakers. Dolby has its own proprietary Atmos renderer as well. A typical home system would be 7.1 with four overhead speakers, which is called 7.1.4. The four overhead speakers are in two rows and essentially mimic the role of the overhead speakers in a cinema auditorium. Other configurations using more top layer speakers and/or the addition of “wide” speakers are defined, such as 9.1.6. Atmos audio is carried in a proprietary bitstream that is similar to IAB.

Audio Object (a.k.a. “Object”) – A segment of audio with associated metadata describing positional and other properties which may vary with time that direct the audio within the Soundfield.

Immersive Audio – Audio that is designed to play through an Immersive Sound System. Immersive Audio consists of Audio Channels and/or Audio Objects, which can be utilized by the content creator to design a soundtrack with sounds above and around the listener. Immersive audio combines metadata with audio essence, which allows the Audio Objects and Audio Channels in the soundtrack to be rendered successfully into multiple Loudspeaker configurations. Note that Immersive Audio does not have to use Audio Objects to be considered Immersive Audio. Immersive Audio requires an Immersive Audio Renderer to play into a sound system.

Immersive Audio Bed (a.k.a. “Bed”) – An Audio Configuration, such as 5.1, 7.1 or 9.1, that serves as the foundation of an immersive soundtrack mix.

Immersive Audio Bitstream (IAB) – An interoperable bitstream designed to carry Immersive Audio as defined in SMPTE ST 2098-2. IAB can carry 128 items, which can be divided up between Immersive Audio Beds and Audio Objects. A typical arrangement has a single 10 channel Audio Bed and up to 118 Audio Objects.

Immersive Audio Configuration – The Audio Configuration of the Immersive Audio Bed e.g., 9.1, 7.1

Immersive Audio Renderer – A DSP device that has been programmed to receive and interpret channel and object metadata and directs both in real time into the set of loudspeakers or other reproduction device that is attached to the renderer.

Immersive Sound – Sound that has height in addition to 7.1 or 5.1 and pinpoint location of individual sounds.

Immersive Soundfield Configuration – Defined arrangement or configuration of Base Layer Loudspeakers, plus Height Layer and/or Top Layer Loudspeakers, that conveys the intended Immersive Soundfield. An example is 7.1.4, which is 7.1 with four Top Layer speakers.

Immersive Sound System – A sound system designed to play immersive audio. In an Immersive Sound System, there is the “Base Layer” of loudspeakers, which is 7.1 or 5.1, and the addition of a “Top Layer” (loudspeakers on the ceiling) and/or “Height Layer” (loudspeakers on the wall near the ceiling). Different brands of Immersive Sound Systems have their own approach to the use and layout of the Top and Height Layer loudspeakers, and the room also plays a part in the design. There may be subwoofers to extend the low frequency range of the surround speakers in some Cinema systems. Each brand of Immersive Sound System also has its own renderer design, which further distinguishes it.

Contact Info

Feature Picture QC & Mastering
Feature_Mastering@spe.sony.com

Television QC & Mastering
TV_Mastering@spe.sony.com