SMPTE REGISTERED DISCLOSURE DOCUMENT

D-Cinema Packaging — SMPTE DCP Bv2.1 Application Profile

Version 1.0.0 – May 18, 2020

The attached document is a Registered Disclosure Document prepared by the sponsor(s) identified below. It has been examined by the appropriate SMPTE Technology Committee and is believed to contain adequate information to satisfy the objectives defined in the Scope, and to be technically consistent.

This document is NOT a Standard, Recommended Practice or Engineering Guideline, and does not represent findings, representations, or recommendations by the Society..

Errors in this document should be reported to the sponsor(s) identified below, with a copy to eng@smpte.org.

All other inquiries in respect of this document, including inquiries as to intellectual property requirements that may be attached to use of the disclosed technology, should be addressed to the sponsor(s) identified below.

Sponsor(s) contact information:

Jack Watts
Trench Digital Ltd
Suite 23, 95 Mortimer Street
London, W1W 7GB

Email *jack.watts@trenchdigital.net*

Steve LLamb
Deluxe Entertainment Services Group
2233 N. Ontario St #300
Burbank, CA 91504

Email *steve.llamb@bydeluxe.com*

Matty Scoll
Eikon Group Limited
2777 N. Ontario St. Ste 180
Burbank, CA 91504

Email *matty.scoll@eikongroup.co.uk*

Andy Lampard
Motion Picture Solutions Ltd
Mission Hall, 9 -11 North End Road,
London , W14 8ST

Email *andyl@motionpicturesolutions.com*

Table of Contents Page

Introduction

SMPTE DCP Bv2.1 Application Profile is a defined feature set of and constraints on Digital Cinema (D-Cinema) Packages and DCP elements within current standards, applicable to theatrical distribution, which have been integrated, globally agreed upon, tested, and rolled out by various industry experts, vendors, manufacturers, content owners, and industry bodies such as Inter-Society Digital Cinema Forum (ISDCF) and European Digital Cinema Forum (EDCF). This profile is simply referred to in the industry as “Bv2.1”.

Background

D-Cinema Composition manufacturing and packaging has long been plagued by global interchange interoperability since its inception. SMPTE ratified a suite of standards within the 428, 429 and 430 families to constrain the implementation of the format. However, throughout this ratification process, ambiguities have been introduced at the manufacturing level as the format has gradually evolved and been adopted due to subjective interpretation of the standards by both manufacturing and servicing vendors alike.

These ambiguities have often been the root cause of financial loss to multiple content owners over the course of various releases throughout the years. Efforts have been made to remediate such ambiguities through ad hoc industry group recommended practices. Unfortunately, such practices, although based on real world faults, are not enforceable. Without their adoption, such ambiguities are still present and will continue to be so until such further constraints are included in the primary documents and are enforced.

By creating an RDD D-Cinema Application, it allows the sponsors and other DCP creators who work with the format daily and on a global scale, to agree and disseminate the constraints based upon their real-world experience. The aim; allow adoption to any market player and eventually achieving approximately 99%+ interoperability of the most commonly distributed formats within the D-Cinema distribution supply chain.

1 Scope

This document specifies requirements and constraints for an Application of D-Cinema Package (DCP) mastering and packaging for playback globally on the highest percentage of systems, and applies to standard general release packages used in current distribution. The document does not preclude the use of current or future formats and/or constraints for special releases. The document follows the Image and Audio Characteristics, Track File and Composition constraints specified in SMPTE ST 429-2 D-Cinema Packaging — DCP Operational Constraints whenever possible. Where applicable, other further constraints to ST 429-2 and references to other documents are included to further define current mastering and packaging practices.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:”

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

3 Normative References

The following documents contain provisions, which, through reference in this text, constitute provisions of this Application. At the time of publication, the editions indicated were valid. This document is subject to revision and parties to agreements based on this document are encouraged to investigate the possibility of applying the most recent edition of the documents indicated below.

SMPTE, ST 428-7:2014 “D-Cinema Distribution Master — Subtitle”

SMPTE, ST 428-12:2013 “D-Cinema Distribution Master — Common Audio Channels and Soundfield Groups”

SMPTE, ST 429-2:2019 “D-Cinema Packaging — DCP Operational Constraints”

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

SMPTE, ST 429-19:2019 “D-Cinema Packaging — DCP Operational Constraints for Immersive Audio”

SMPTE, ST “429-8:2007 “D-Cinema Packaging — Packing List”

SMPTE, ST 429-7:2006 “D-Cinema Packaging — Composition Playlist”

SMPTE, ST 429-16:2014 “D-Cinema Packaging — Additional Composition Metadata and Guidelines”

SMPTE, ST 2001-1:2015 “XML Representation of SMPTE Registered Data (Reg-XML) – Mapping Rules”

SMPTE, ST 429-4:2020 “D-Cinema Packaging — MXF JPEG 2000 Application”

SMPTE, ST 429-6:2006 Amd 1:2018 “D-Cinema Packaging — MXF Track File Essence Encryption”

SMPTE, ST 377:2004[1] “Material Exchange Format MXF – File Format Specification”

ISO/IEC Joint Technical Committee JTC 1 (ISO/IEC), 15444-1:2016, “Information Technology — JPEG 2000 Image Coding System: Core Coding System”

Digital Cinema Initiatives (DCI), “Digital Cinema System Specification, Version 1.3”, (all Addenda and Errata, as of May 18, 2020)

Internet Assigned Numbers Authority (IANA), "Language Subtag Registry, File-Date: 2019-09-16", Available: "http://www.iana.org/assignments/language-subtag-registry"

Internet Engineering Task Force (IETF), RFC 5646, "Tags for Identifying Languages", 2009-09

Internet Engineering Task Force (IETF), RFC 4122, "A Universally Unique Identifier (UUID) URN Namespace", 2005-07

Inter-Society Digital Cinema Forum (ISDCF), “Digital Cinema Naming Convention v.9.6.1”, Available: “http://isdcf.com/dcnc”

1 The reference to SMPTE ST 377:2004 is intentional. SMPTE ST 377-1:2011 or future versions are not appropriate for use with this document.

4 Glossary and Acronyms

In addition to the glossary terms and acronyms presented here, the reader should also be aware of terms defined in SMPTE ST 377, SMPTE ST 429-7, and SMPTE ST 428-7, and SMPTE ST 429-2.

CPL	Composition Playlist
DCP	Digital Cinema Package
IANA	Internet Assigned Numbers Authority
IETF	Internet Engineering Task Force
Immersive Audio	SMPTE ST 2098-2 Essence
MCA	Multichannel Audio
PCM	Pulse Code Modulation
PKL	Packing List
TMS	Theatre Management System

5 General Notes about this Document

This document and the specifications herein, like the industry they serve, are subject to change at any time without notice. Failure to provide 100% accurate data or information, and/or failure to deliver content conforming to this Application, may result in the following, including but not limited to:

- issue(s) and/or error(s) during ingestion and/or presentation of content in a theatrical environment

In some cases, ambiguous information within specification documents may have been clarified or refined in this document based on best industry practices and/or “known issues” with release versions of equipment in theatrical settings.

This is not a full list of the documents referenced herein. Where needed, further links to documents or websites may be noted in the text. Where a topic is unspecified, one shall defer to the above noted standards, specifications, and recommended practices.

6 DCP Constraints

6.1 Packaging Format

All compositions shall comply with the SMPTE packaging format as defined by SMPTE ST 429-2.

In addition, all compositions carrying an Immersive Audio essence track, as defined in SMPTE ST 429-18, shall comply with the SMPTE packaging constraints as defined by SMPTE ST 429-19.

6.2 XML Constraints

6.2.1 Language and Territory Codes

`Language` and `Territory` elements and attribute code values for XML documents shall conform to IETF RFC 5646.

> *Note: As of the publication date of this specification, the set of region subtags in the Language Subtag Registry consists of a subset of the geographical regions specified in ISO 3166-1 and U.N. M.49. The Language Subtag Registry is maintained by IANA at http://www.iana.org/assignments/language-subtag-registry and is the definitive list of region subtags.*

7 Composition Constraints

7.1 Picture Essence Encoding

For picture essence tracks, the pixel array size, frame rate, and 2D/3D shall be one of the formats listed in Table 1 below.

Source images having an aspect ratio not listed in Table 1 below shall be encoded so that the image fills either the horizontal or vertical dimension of the desired SCOPE or FLAT Container (2K or 4K). To fill the pixel array in the opposite dimension, the image should be padded with an equal number of black pixels on each side, *i.e.*, “letter-box” (top side, bottom side) or “pillar-box” (left side, right side).

Table 1 – Pixel Array Dimensions

Format	Horizontal Pixels	Vertical Pixels	Frame Rate	2D/3D
2K Scope (2.39:1)	2048	858	24/1, 25/1, or 48/1	2D/3D
2K Flat (1.85:1)	1998	1080	24/1, 25/1, or 48/1	2D/3D
4K Scope (2.39:1)	4096	1716	24/1	2D ONLY
4K Flat (1.85:1)	3996	2160	24/1	2D ONLY

> *NOTE: 4K 3D is only supported on a very limited number of systems in presentation environments such as museums, and special venues. Wide industry support does not yet exist, as it’s based on equipment capabilities. Use at your own risk.*

For purposes of this Application, 2K Full (1.90:1 – 2048x1080) and 4K Full (1.9:1 – 4096x2160) container size shall not be utilized.

7.1.1 Active Image Area

For full active image area resolutions SCOPE or FLAT, image shall fill the frame edge to edge in both vertical and horizontal directions. 3D “floating windows” are acceptable but must be within the edge boundaries.

All other aspect ratios (1.90, 1.40, 1.78, etc) and resolutions shall be sized and padded with black to fill the frame to the SCOPE or FLAT container resolutions in Table 1 above. The closest container to content’s frame size should be utilized to take advantage of as many pixels as possible to preserve quality. See Table 2 below for some examples of common padding of active image to FLAT and SCOPE containers.

Table 2 - Common Active Image Sizes and Padding

Active Image Aspect Ratio	Container Aspect Ratio	Active Image Size (2K/4K)	DCI Image Container Size (2K/4K)
SCOPE/2.39	FLAT/1.85	1998x836/3996x1672[2]	1998x1080/3996x2160
FLAT/1.85	SCOPE/2.39	1588x858/3176x1716[3]	2048x858/4096x1716
“Full Container”/1.90	FLAT/1.85	1998x1051/3996x2102	1998x1080/3996x2160
HD/1.78	FLAT/1.85	1922x1080/3844x2160	1998x1080/3996x2160
CinemaScope/2.35	SCOPE/2.39	2016x858/4032x1716	2048x858/4096x1716
Traditional TV/1.33 (4:3)	FLAT/1.85	1436x1080/2872x2160	1998x1080/3996x2160
Academy Ratio/1.375	FLAT/1.85	1485x1080/2970x2160	1998x1080/3996x2160
Paramount Format/1.66	FLAT/1.85	1792x1080/3585x2160	1998x1080/3996x2160

7.2 Timed Text

7.2.1 File Size Limitations

Timed Text DCDM XML files as defined in SMPTE ST 428-7, for usage as DisplayType:ClosedCaption files should not be larger than 256kB.

> *Note: In some rare cases Timed Text DCDM XML files might be used for both DisplayType:ClosedCaption and DisplayType:MainSubtitle. This size recommendation does not apply to their usage as DisplayType:MainSubtitle.*

2 For SCOPE trailers playing in front of FLAT features.

3 For FLAT trailers playing in front of SCOPE features.

The cumulative file size limit for a single reel Timed Text track, including DCDM XML and all ancillary assets such as PNGs or Fonts, shall not exceed 115 MB.

The font resource shall not be larger than 10MB.

7.2.2 Language Element

The `language` element as defined in SMPTE ST 428-7, shall be present, and the `language` attribute shall conform to IETF RFC 5646 as defined in section 6.2.1 above. Default behavior shall be as defined in *section 5.7* of SMPTE ST 428-7. The attribute shall be contiguous across all DCDM XML instances when the DisplayType:MainSubtitle is present.

7.2.3 `StartTime` Element

The `StartTime` element shall be present, and the value shall be `00:00:00:ZeroE` as defined In SMPTE ST 428-7.

7.2.4 First Timed Text Event

The composition's first Timed Text event's `TimeIn` attribute as defined In SMPTE ST 428-7 should be greater than or equal to the value of `00:00:04:ZeroE` (four [4] seconds).

> *Note: In a multi-reel composition, this only applies to the first reel. This is to accommodate the average buffer size of a wide range of rendering devices.*

7.2.5 Spot Duration

All `Subtitle` instances as defined In SMPTE ST 428-7 should have a minimum duration of 15 frames with a minimum duration of 2 frames between `Subtitle` instances.

> *Note: It is advised to avoid using persistent subtitles where possible, instead opting for a repetitive `Subtitle` instance within the `SubtitleList` as defined In SMPTE ST 428-7.*

7.2.6 Closed Captions Line Treatment

Each `Subtitle` instance as defined In SMPTE ST 428-7 intended for usage as DisplayType:ClosedCaption shall have a maximum of thirty-two [32] characters per line and a maximum of three [3] lines per `Subtitle` instance.

> *Note: Such a limitation is not applicable for all languages, as in some cases, the geometry of such languages' characters can require the character limit per line to be less (e.g. 14 or 16). For such instances, it is recommended to reference an appropriate style guide, which is outside the scope of this document.*

7.2.7 Main Subtitle Line Treatment (Informational)

Assuming common Font attributes are used, each `Subtitle` instance as defined In SMPTE ST 428-7, intended for usage as DisplayType:MainSubtitle should have a recommended limit of fifty-two [52] characters per line but should not exceed seventy-nine [79] characters per line. Each `Subtitle` instance should also have a maximum of three [3] lines.

7.3 Audio Sample Rates

The sample rate of audio essence in a Composition shall be 48 kHz.

8 Composition Playlist Constraints

8.1 Annotation Text

Some playback and theater management systems present AnnotationText from the Composition Playlist within a DCP to the user at the time of ingestion of content. To that end, the `AnnotationText` element as defined in SMPTE ST 429-7 in the Composition Playlist shall be present. The `AnnotationText` value should match exactly the CPL's ContentTitleText.

8.2 Reel Duration

The value of all `Duration` elements as defined in SMPTE ST 429-7 in a reel shall be equal, including for Timed Text and Aux Data tracks, if present.

8.3 Timed Text Tracks

8.3.1 CPL Elements

A timed text track is established by the presence of a timed text asset (e.g. `MainSubtitle`, or `ClosedCaption`, both as defined in SMPTE ST 429-7).

For purposes of this Application, `MainCaption` and `ClosedSubtitle` as defined in SMPTE ST 429-7 shall not be used.

If a `MainSubtitle` timed text track is present on any reel, a `MainSubtitle` timed text track shall be present on all reels. If (1) or more `ClosedCaption` timed text track(s) are present on any reel, the same number of `ClosedCaption` timed text tracks shall be present on all reels.

> *Note: For reels not needing Timed Text elements, the above may be accomplished by created a "blank" Timed Text asset track, which is outside the scope of this document.*

8.3.2 Entry Point

For all `MainSubtitle` or `ClosedCaption` timed text tracks, the Composition Playlist's `EntryPoint` element as defined in SMPTE ST 429-7 shall be present and have a value of "0".

8.4 `Hash` Element

To satisfy inter-SMS ingestion within an auditoria network without a TMS, the `Hash` element as defined in SMPTE ST 429-7 shall be present for every referenced track file.

8.5 Marker Tracks

When a Composition Playlist's `ContentKind` value is set to "feature" as specified in SMPTE ST 429-7, both marker labels `FFEC` (First Frame of End Credits) and `FFMC` (First Frame of Moving Credits) as described in SMPTE 429-7 shall be present.

Note: "End Credits" usually refer to director and/or actor credits done in card style or heavy styled end credits sequences. If neither of these types of credits exist, this can be the same frame number as FFMC. Moving Credits usually refer to the "rolling credits". If there are no rolling credits, can be the same frame number as FFEC.

Marker labels `FFOC` (First Frame of Composition) and `LFOC` (Last Frame of Compositon) as described in SMPTE ST 429-7, should be present for all Composition Playlists. If present, the value for `FFOC` should represent the first frame of the Composition and be equal to "1". If present, the value for `LFOC` should represent the last frame of the Compostion and be equal to (1) frame less than the `Duration` value of the Composition's last reel.

8.6 Composition Metadata

8.6.1 CompositionMetadataAsset

The `CompositionMetadataAsset` element defined in SMPTE ST 429-16 shall be present, and all fields shall be populated where values pertaining to such are available at the time the CPL is being created. Where applicable, the values used should match values for their respective fields used within the Composition Playlist's `ContentTitleText`. Absence of known and/or usable values shall be signaled by the absence of any optional element within the `CompositionMetadataAsset`.

Note: Informational examples of values and/or codes used within the `ContentTitleText` fields may be found at: http://isdcf.com/dcnc/. Mapping of those fields and values to specific elements within the `CompositionMetadataAsset` is outside the scope of this document.

8.6.2 Version Number

Where the `VersionNumber` value as defined in SMPTE ST 429-16 is not specified, the default value of "1" shall be present.

8.6.3 Application Extension

For CPLs conforming to this Application document, the following `ExtensionMetadata` shall be present:

```
<cpl-meta:ExtensionMetadata scope="http://isdcf.com/ns/cplmd/app">
  <cpl-meta:Name>Application</cpl-meta:Name>
  <cpl-meta:PropertyList>
    <cpl-meta:Property>
      <cpl-meta:Name>DCP Constraints Profile</cpl-meta:Name>
      <cpl-meta:Value>SMPTE-RDD-52:2020-Bv2.1</cpl-meta:Value>
    </cpl-meta:Property>
  </cpl-meta:PropertyList>
</cpl-meta:ExtensionMetadata>
```

8.6.4 MCA Sub Descriptors

The "xs:any" extension point at the top level of the `CompositionMetadataAsset` element as described in SMPTE ST 429-16 shall contain one `MCASubDescriptors` element conforming to the XML Schema definition below.

```
<?xml version="1.0" encoding="UTF-8"?>

<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
                   elementFormDefault="qualified"
 targetNamespace="http://isdcf.com/ns/cplmd/mca"
 xmlns:mxf="http://www.smpte-ra.org/reg/395/2014/13/1/aaf">

   <xs:import namespace="http://www.smpte-ra.org/reg/395/2014/13/1/aaf"/>
   <xs:element name="MCASubDescriptors">
      <xs:complexType>
         <xs:choice maxOccurs="unbounded" minOccurs="0">
            <xs:element ref="mxf:SoundfieldGroupLabelSubDescriptor"/>
            <xs:element ref="mxf:AudioChannelLabelSubDescriptor"/>
         </xs:choice>
      </xs:complexType>
   </xs:element>
</xs:schema>
```

The `SoundfieldGroupLabelSubDescriptor` and `AudioChannelLabelSubDescriptor` elements are the XML representation of all Audio Channel Label and Soundfield Group Label SubDescriptors, respectively, present in the Main Sound Track File. This XML representation is specified in SMPTE ST 2001-1.

8.7 Digital Signature

The Composition Playlist shall be digitally signed as described in SMPTE ST 429-7 when encrypted essence files are present.

9 Packing List Constraints

9.1 Annotation Text

Some playback and theater management systems present AnnotationText from the Packing List(s) within a DCP to the user at the time of ingestion of content. To that end, the `AnnotationText` element as defined in SMPTE ST 429-8 in the Packing List shall be present. In the case where a PKL contains a single (1) CPL, the PKL AnnotationText value shall match exactly the CPL's ContentTitleText. In the case where a PKL contains two or more CPLs, the primary CPL's ContentTitleText, or a string to best describe the collection of CPLs, should be used for the PKL AnnotationText value.

9.2 Digital Signature

The Packing List shall be digitally signed as described in SMPTE ST 429-8 when encrypted essence files are present.

10 Track File Constraints

10.1 Encryption

When any Track File in a composition is encrypted, all remaining Track Files (regardless of type) that exist as of the publication date of this document, in the composition shall also be encrypted.

The integrity of each frame of encrypted Sound and Picture essence shall be verifiable using the HMAC-SHA1 algorithm as described in SMPTE ST 429-6. Additionally, The optional Message Integrity Code (MIC) element shall be present as described in SMPTE ST 429-6.

10.2 Picture Track Files

10.2.1 JPEG2000 Compression

All codestreams shall fully conform with ISO/IEC 15444-1:2016 and SMPTE ST429-4:2020, with the additional constraints below as per DCI Digital Cinema System Specification.

All image frames shall be untiled. More precisely, the entire image shall be encoded as a single tile.

The number of guard bits to be specified in the QCD marker shall be a value of 1 for 2K content, and shall be a value of 2 for 4K content.

Codeblocks shall be of size 32x32.

The codeblock coding style shall be SPcod, SPcoc = 0b00000000.

The progression order for a 2K distribution shall be Component-Position-Resolution-Layer (CPRL). Progression Order Change (POC) marker segments are forbidden in 2K distributions.

For a 4K distribution, there shall be exactly one POC marker segment in the main header. Other POC marker segments are forbidden. The POC marker segment shall specify exactly two progressions having the following parameters:

- First progression: RSpoc = 0, CSpoc = 0, LYEpoc = L, REpoc = D, CEpoc = 3, Ppoc = 4
- Second progression: RSpoc = D, CSpoc = 0, LYEpoc = L, REpoc = D+1, CEpoc = 3, Ppoc = 4
- In the above, D is the number of wavelet transform levels and L is the number of quality layers. The constant 3 specifies the number of color components, and the constant 4 specifies CPRL progression.

  Note: This POC marker segment ensures that all 2K data precede all 4K data. Within each portion (2K, 4K), all data for color component 0 precede all data for color component 1, which in turn precede all data for color component 2.

Each compressed frame of a 2K distribution shall have exactly 3 tile parts. Each tile part shall contain all data from one color component.

Each compressed frame of a 4K distribution shall have exactly 6 tile parts. Each of the first 3 tile parts shall contain all data necessary to decompress one 2K color component. Each of the next 3 tile parts shall contain all additional data necessary to decompress one 4K color component.

Tile-part Lengths, Main header (TLM) marker segments shall be required in all frames of all distributions.

Note: This facilitates extraction of color components and resolutions (2K vs. 4K).

For a frame rate of 24 FPS, a 2K distribution shall have a maximum of 1,302,083 bytes per frame (aggregate of all three color components including headers). Additionally, it shall have a maximum of 1,041,666 bytes per color component per frame including all relevant tile-part headers.

For a frame rate of 48 FPS, a 2K distribution shall have a maximum of 651,041 bytes per frame (aggregate of all three color components including headers). Additionally, it shall have a maximum of 520,833 bytes per color component per frame including all relevant tile-part headers.
A 4K distribution shall have a maximum of 1,302,083 bytes per frame (aggregate of all three color components including headers). Additionally, the 2K portion of each frame shall satisfy the 24 FPS 2K distribution requirements as stated above.

Note: For information purposes only, this yields a maximum of 250 Mbits/sec total and a maximum of 200 Mbits/sec for the 2K portion of each color component.

10.3 Sound Track Files

10.3.1 Channel Assignments

Channel assignment defines what audio signal is carried in each channel of the distributed track. Sound Track File channel assignment shall be indicated by a UL value in the Channel Assignment property of the Wave Audio Essence Descriptor.

The Static Container Channel Configurations in A.1 of SMPTE ST 429-2 shall be used, and the Byte 15 Value shall be set to “04h” (Channel Configuration 4) as noted in Table A.2 of SMPTE ST 429-2.

Channel Configuration 4 is noted as an “open” configuration without inherent constraints on the soundfield it carries. Therefore, the Sound Track File’s channels shall be mapped to channel numbers as listed in Table 3 below for use within a common soundfield group. Channels 9,10, and 16 shall be unused and shall be recorded with silence (all samples have a value of 0). Further, any unused channels in a soundfield group (as denoted by “-“ in the below table) shall be recorded with silence (all samples have a value of 0.

Table 3 - Channel Layout

Container Channel	Mono	Stereo	5.1	7.1	Immersive Audio
1	-	Left	Left	Left	-
2	-	Right	Right	Right	-
3	Center	-	Center	Center	-
4	-	-	LFE	LFE	-
5	-	-	Left Surround	Left Side Surround	-
6	-	-	Right Surround	Right Side Surround	-
7	Hearing Impaired (HI)	Hearing Impaired (HI)	Hearing Impaired (HI)	Hearing Impaired (HI)	Hearing Impaired (HI)

8	Visually Impaired-Narrative (VIN)	Visually Impaired-Narrative (VIN)	Visually Impaired-Narrative (VIN)	Visually Impaired-Narrative (VIN)	Visually Impaired-Narrative (VIN)
9	-	-	-	-	-
10	-	-	-	-	-
11	-	-	-	Left Rear Surround	-
12	-	-	-	Right Rear Surround	-
13	Motion Data	Motion Data	Motion Data	Motion Data	Motion Data
14	-	-	-	-	Sync Signal
15	Sign Language Video	Sign Language Video	Sign Language Video	Sign Language Video	Sign Language Video
16	-	-	-	-	-

Note: The above "Immersive Audio" channel layout describes the layout noted in the MainSound constraints for Immersive Audio in SMPTE ST 429-19, and assumes the presence of an Immersive Audio essence Aux Data track, as defined in SMPTE ST 429-18, in the DCP.

Channel layout mappings and common soundfield groups used shall remain constant throughout the entire Composition.

10.3.2 Multichannel Audio

The Sound Track File shall conform to SMPTE ST 377-4 and include Audio Channel Label and Soundfield Group Label Sub-Descriptors ("MCA Sub Descriptors') as specified in Annex A.2.2 of SMPTE ST 429-2[4] that conform with the respective D-Cinema Soundfield Groups defined in SMPTE ST 428-12.

Sound Track Files shall not contain instances of `GroupOfSoundfieldGroupsLabelSubdescriptor`.

Note: In playback systems, the Static Container Channel Configuration "4" of the Sound Track file should take precedence, and the above described MCA Sub-Descriptors should only be utilized by playback systems that have the ability to ignore the Static Container Channel Configuration and support MCA.

4 This defines the combined usage of the Configuration 4 ChannelAssignment UL as defined in ST 429-2 Table A.1, in conjunction with ST 377-4 framework for AudioChannelLabelSubDescriptor as defined in ST 429-2 Section A.2.2 to label audio channels. This usage is not specified in the current version of ST 429-2.

10.3.3 Sign Language Video PCM

If present in the composition, the Sign Language Video PCM shall be placed on Sound Track channel 15 as noted in Table 3 above and shall conform to the encoding parameters and additional CPL Constraints as defined in Annex A.

> *Note: It is further recommended that the source VP9 be encoded as a single "longplay" file for the entire duration of the program. It can then, thereafter be divided into reels as needed for PCM conversion to ensure proper bitstream decode on current devices. Non-modulated PCM (silence) may be added at frames with an index less than the EntryPoint and more than the value of the EntryPoint+Duration frame index. This will allow each reel to conform to other channel lengths.*

The Audio Channel Label Subdescriptor, as described in section 10.3.2 above, associated with a Sign Language Video PCM channel shall be set according to Table 4 below.

Table 4 - Sign Language MCA Values

MCA Label Dictionary ID	MCA Tag Name	MCA Tag Symbol	RFC 5646 Spoken Language
06.0E.2B.34.04.01.01.0D.0D.0F.03.02.01.01.00.00	Sign Language Video Stream	SLVS	Same as the Soundfield Group Label Sub Descriptor

10.3.4 Motion Data PCM

If present in the composition, channel 13 is intended for Motion Data monoaural PCM as noted in Table 3 above.

10.4 Timed Text Track Files

10.4.1 Asset UUID

The MXF `AssetUUID` shall be identified by a unique `urn:uuid` of Type 4 value IETF RFC 4122

10.4.2 Resource ID

The MXF `ResourceID` shall be equal to the `Id` element `urn:uuid` of Type 4 value IETF RFC 4122 of the contained DCDM XML essence.

10.4.3 Container Duration

The `ContainerDuration` value of a Timed Text track file shall be equal to the `Duration` value of the reel element from which it is referenced.

Annex A Sign Language Video Encoding

A.1 Introduction

This annex describes a method for the encoding and CPL requirements of a sign language video track within a DCP for distribution to exhibition.

This scheme benefits from the following advantages:

- compatibility with all existing D-Cinema projection and playback systems
- supports random access playback
- synchronization with audio at the output of the Media Block

A.2 Video Format

A.2.1 Codec

The video codec shall be VP9.

Note: VP9 is a high-quality, open video format; see https://www.webmproject.org/vp9/

A.2.2 Resolution

The video shall be 480 pixels wide and 640 pixels high (*i.e.*, portrait orientation).

A.2.3 Framerate

The video frame rate shall be 24 FPS regardless of any other DCP frame rates in use.

A.2.4 Bitrate

The VP9 bitstream shall have a maximum bitrate of 1.0 Mbps.

A.2.5 Colorspace

The video shall be encoded as Y'UV.

A.2.6 Pixel Format

The video shall use Y'UV420p chroma sub-sampling.

A.3 Mapping Into the PCM

A.3.1 Background on D-Cinema Audio

DCPs carry audio as a sequence of uncompressed frames of (up to) 16 channels of 24-bit Pulse-Code Modulation (PCM) [See EBU Tech 3285]. The PCM has a sample rate of 48 kHz. Each audio frame has a duration:

```
duration = 1/e
```

Where *e* is Edit Rate of the composition.

During playback, projection systems sequentially decrypt each frame of audio and output each of its channels to their corresponding AES/EBU digital output to form 16 independent digital audio streams. Each of these 16 digital streams is therefore operating at the following fixed bitrate:

```
48000 samples/s * 24 bit/sample = 1.152 Mb/s
```

A.3.2 VP9 Chunking

VP9 is inherently a variable bitrate codec. To allow for carriage in fixed bitrate digital audio, the video must be encoded into discrete chunks that get distributed evenly throughout the digital audio program. Each chunk contains both an EBML Header and VP9 Segment.

The duration of each VP9 chunk shall be:

$$D_c = 2 \text{ seconds}$$

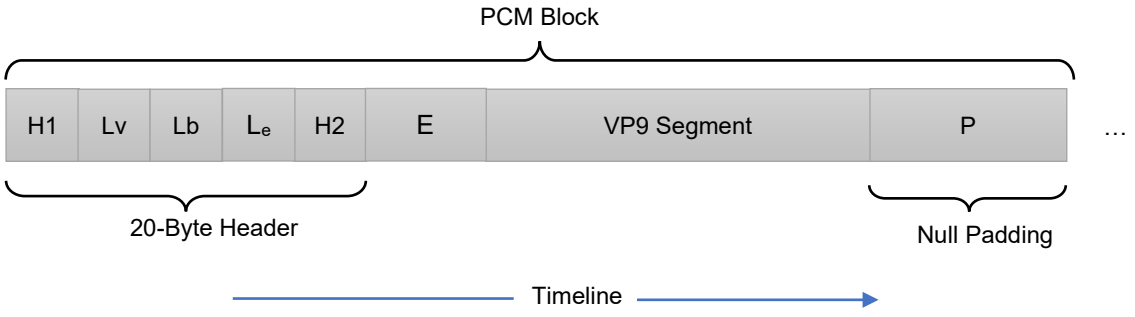
A.3.3 PCM Block Structure

Each segment of VP9 is carried in a block corresponding to PCM essence of equal duration. Thus the length of each PCM block is:

$$L_b = 48{,}000 \text{ samples/s} \cdot 3 \text{ bytes/sample} \cdot D_c = 288{,}000 \text{ bytes}$$

A PCM block is composed of a 20-byte header, followed by the VP9 EBML Header, followed by the VP9 segment, followed by zero or more null bytes. See Figure 1 below for details.

Figure 1 - PCM Block Structure



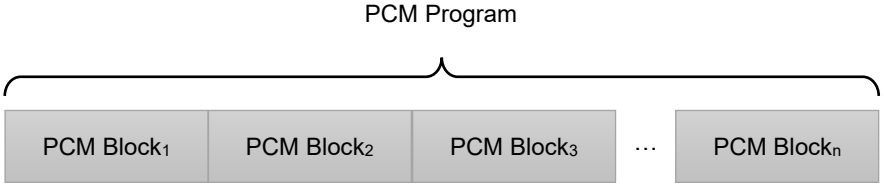
Where:

H_1 = 0xFFFFFFFF
L_v = Length of VP9 segment in bytes (32-bit unsigned integer, big-endian)
L_b = Length of PCM Block in bytes (32-bit unsigned integer, big-endian)
L_e = Length of VP9 EBML Header in bytes (32-bit unsigned integer, big-endian)
H_2 = 0xFFFFFFFF
E = VP9 EBML Header
P = A sequence of $L_b - L_v - L_e - 20$ null bytes

A.3.4 Complete PCM Program

Each PCM Block described above is combined to form the final PCM program as shown in Figure 2 below.

Figure 2 - PCM Program Structure



A.4 Composition Playlist Metadata

CPLs that carry a Sign Language Video track shall indicate the presence of this track; to support this, the following `ExtensionMetadata` shall be present:

```
<ExtensionMetadata scope="http://isdcf.com/2017/10/SignLanguageVideo">
  <Name>Sign Language Video</Name>
  <PropertyList>
    <Property>
      <Name>Language Tag</Name>
      <Value>DESCRIPTION</Value>
    </Property>
  </PropertyList>
</ExtensionMetadata>
```

Where *DESCRIPTION* is a Language-Tag, conforming to section 6.2.1 above, that identifies the sign language present in the Sign Language Video Track.

A.5 KDM Forensic Marking

To prevent the corruption of an encrypted sign language track by the audio forensic marking feature of the D-Cinema projection system, associated Key Delivery Messages (KDMs):

a) shall carry the selective audio FM mark flag as specified at section 9.4.6.2.3(d) of the DCI System Specification, set to a value smaller than or equal to 14 (see Note below).
b) may also carry the “http://www.smpte-ra.org/430-1/2006/KDM#mrkflg-audio-disable” flag for compatibility with legacy image media blocks (IMBs).

NOTE: The value of the selective audio FM mark flag is based on the specific sound channels that cannot be forensically-marked. In particular, when targeting a Composition containing Motion Data and External Sync Signal in addition to the Sign Language Video, the selective audio FM mark flag is set to a value smaller than or equal to 12.