SMPTE RECOMMENDED PRACTICE

D-Cinema Distribution Master — Audio File Format and Delivery Constraints

Table of Contents

Approved March 31, 2010

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE RP 428-4 was prepared by Technology Committee 21DC.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Recommended Practice. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This recommended practice is intended to constrain the audio that is delivered as a Digital Cinema Distribution Master (DCDM) in order to facilitate the process of creating a D-Cinema Package (DCP). By following the guidelines in this document, the audio DCDM will be interchangeable between facilities and will reliably sync with the intended picture essence.

This document is intended to work in conjunction with the existing SMPTE 428-2 DCDM Audio Characteristics and SMPTE 428-3 DCDM Audio Channel Mapping and Channel Labeling documents, utilizing the specifications already established therein and embellishing upon them with further detail and constraints. It brings into alignment traditional film terminology and techniques with today's digital technology and processes.

1 Scope

This document specifies constraints on the technical parameters, file structure, file naming and other relevant information pertaining to audio essence delivered from final post-production as a Digital Cinema Distribution Master.

2 Conformance Notation

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

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

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

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

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

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

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3 Normative References

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

SMPTE 428-2-2006, D-Cinema Distribution Master — Audio Characteristics

SMPTE 428-3-2006, D-Cinema Distribution Master — Audio Channel Mapping and Channel Labeling

SMPTE 429-2-2009, D-Cinema Packaging — DCP Operational Constraints

SMPTE 429-7-2006, D-Cinema Packaging — Composition Playlist

ITU-R BR.1352-1: 2002, Broadcast Wave Format (BWF), Annex 1

ISO/IEC 646:1991, Information Technology — ISO 7-Bit Coded Character Set for Information Interchange

ISO 639-2/T:1998, Codes for the Representation of Names of Languages — Part 2: Alpha-3 Code, First Edition

4 Acronyms and Terminology

4.1 Acronyms

DCDM: Digital Cinema Distribution Master

DCP: Digital Cinema Package

FFOA: First Frame Of Action (See definition in Section 4.2)

LFOA: Last Frame Of Action (See definition in Section 4.2)

4.2 Terminology

Audio Content and Type: A description of the audio that is contained in the essence. This is used to associate a given file with other files in the Soundfield Configuration group, or to differentiate a file containing other content that may be delivered at the same time to packaging. This may be comprised of an Audio Type descriptor and/or an Audio Content descriptor, which are defined below. In typical use, the two are combined to describe a unique audio element.

Examples include “Dialog Stem”, “Visually Impaired Narration”, “Hearing Impaired Time Code”, Director’s Commentary Narration”.

Audio Content: This is an optional modifier to “Audio Type” that can be included to further describe the audio if needed. Examples are “Dialog”, “Music”, “Effects”, “Narration”, “Director’s Commentary”, and “Time Code”.

Audio Type: This is the type of audio contained in the audio essence. Examples include “Printmaster”, “Stem”, “Music and Effects”, “Score”, “Composite Mix”, “Visually Impaired”, “Hearing Impaired”.

Bit Depth: The number of bits in a sample word.

Channel: The loudspeaker location intended for an audio essence. Loudspeaker locations are defined in SMPTE 428-3, Section 3.

Channel Label: A label to indicate the intended channel for audio reproduction. Channel labels may exist in multiple areas within the audio creation and delivery process including, but not limited to, audio essence, the audio reproduction chain, and individual speakers.

This document specifically refers to channel labels applied to the audio essence in an Audio DCDM for delivery to D-Cinema packaging.

Composition: (From SMPTE 429-7, DCP Composition Playlist, Section 3.) A composition is a self-contained representation of a single complete D-Cinema work such as a motion picture, or a trailer, or an advertisement, etc. It tangibly consists of a Composition Playlist file and one or more track files, which contain the actual essence.

Content Kind: A designation of what the content is (e.g., feature, trailer, advertisement, etc.) per Table 2 of SMPTE 429-7.

Content Title: The title of the composition. For example, the title of a feature film.

Content Version: The version of the composition, such as Domestic, International, or Director's Cut.

Duration: This is the duration of the actual program essence for a given post-production reel, expressed as an integer number of frame periods. This is counted from the left edge of the FFOA to the right edge (end) of the LFOA, and thus includes both the FFOA and LFOA in their entirety.

In practice, this may be calculated from an editor's "LFOA List" by taking the stated LFOA start, adding 1 frame and subtracting 192 frames (the leader). If the LFOA list is expressed in feet/frames, this must first be converted to frames and then the above formula is applied.

Example: A post-production reel has a stated LFOA of 1500'+4. Since one foot in film is 16 frames, 1500+4 is 24,000 + 4 = 24004 frames. Adding one and subtracting 192 gives a duration of 23,813 frames.

Editable Unit (Edit Unit): (From SMPTE 429-7, Section 4.) The smallest temporal increment of access to Essence, e.g., a *frame* or a *sample.*

In practice, the edit unit in D-Cinema is the duration of one picture frame for a monoscopic picture composition or the duration of the left eye/right eye pair of frames in a stereoscopic picture. Although audio itself is capable of being edited to the sample, in practice it is edited on the corresponding picture frame boundaries. In D-Cinema packaging, audio is "frame wrapped" such that a packet of audio essence is the duration of an edit unit and contains the number of audio samples corresponding to the audio sample rate and the composition edit rate.

Edit Rate: (From SMPTE 429-7, Section 4.) The number of Editable Units to be reproduced during a temporal interval having duration of exactly one (1.0) second. Because Edit Rate values are not always integer values and sometimes require many digits of precision, Edit Rate values are expressed as a rational number (the ratio of two integers).

For a composition containing monoscopic picture, the frame rate and edit rate are identical. For stereoscopic picture, the edit unit represents two picture frames, (left eye and right eye), which would always be edited as a pair, and thus the edit rate is half of the frame rate.

Essence: (From SMPTE 429-7, Section 4.) The sound, picture, and data resources that make up a Composition.

File Date: The date that an audio file was created or the most recent date it was modified. This is assigned by the creator of the file and is independent of the date that a file directory may indicate.

First Frame of Action (FFOA): The first frame of a post-production reel that contains image action pertinent to the program. For audio elements, it is the location of the sound that is intended to sync with the first frame of image action.

The location of this frame is depicted by the location of the left edge of the frame. For example, the FFOA is 8 seconds (192 frames at 24 fps) from the left edge of the picture start frame on the leader on a post-production reel of 35 mm film. The same concept applies to D-Cinema.

Frame Rate: (From SMPTE 429-7, Section 4.) The number of frames per second.

Language: The main spoken language of the audio essence. Language representation is depicted in Section 5.2.6.2.

Last Frame of Action (LFOA): The last frame of a post-production reel that contains image action pertinent to the program. For audio elements, it is the location of the sound that is intended to sync with the last frame of image action. It is common for this value to be expressed in feet and frames, but may be frames only.

The location of this frame is depicted by the location of the left edge of the frame, and is traditionally referenced as being the number of frames (or feet and frames) counting from the left edge of the picture start frame (a.k.a. "0 feet" in film terminology) to the left edge of the LFOA in a post-production reel. The picture start frame is therefore included in the count.

Note that the location of the LFOA is the start of the frame-the action actually finishes at the end of this frame. Therefore, the image action and sound actually ends at LFOA+1Frame.

Post-Production Reel: A partition of the essence, the duration of which is defined by the content provider's post-production process. For example, a theatrical presentation may be divided into 6 post-production reels, each of which is typically 22 minutes or less in length and corresponds to a 2100 foot or less reel of physical film. A post-production reel may typically be made into a D-Cinema track file.

Post-Production Reel Number: An identifier associated with a post-production reel to describe its place in the sequence of reels delivered by post-production, which are intended for a composition. This is generally a combination of letters and numbers. For example, "R5 A/B" would indicate that this is the 5th reel in the set of post-production reels delivered, and that it is a double length reel (a combination of an "A" reel and a "B" reel).

Prelap: An industry term that refers to audio that may be placed between the head "2 pop" and the FFOA, which is a copy of the outgoing audio of the previous post-production reel, in order to better facilitate the changeover between post-production reels.

Pullup: An industry term that refers to audio that may be placed after the LFOA and before the "tail pop", which is a copy of the incoming audio of the subsequent post-production reel, in order to better facilitate the changeover between post-production reels.

Note that pullups have been in use for many years in the film industry to facilitate film projection changeovers. Prelaps are more recent, and are often used when there is continuous music across the post-production reels.

Prelaps and Pullups are generally used in film projection and are not used in D-Cinema. If they are supplied as a DCDM, they shall be sample accurate with the previous and subsequent post-production reels.

Sample Rate: (From SMPTE 429-7, Section 4.) The number of essence samples per second.

Soundfield: The acoustical space within which the intended audio image is created.

Soundfield Configuration: A Soundfield Configuration is a defined arrangement or configuration of loudspeakers. A group of audio channels that are intended to be reproduced simultaneously through a defined *Soundfield Configuration* are said to belong to that *Soundfield Configuration* group and are labeled with a *Soundfield Configuration* name. For example, this may be represented by a *Soundfield Configuration* name such as "5.1", or by an abbreviation of the table headers in Section 4 of SMPTE 428-3, DCDM Audio Channel Labeling and Channel Mapping.

Watermark Type: The type of watermark that may have been placed on the audio at the creation stage and is carried in the DCDM audio. This is not to be confused with the Forensic watermark added at playout in the theater.

5 DCDM Audio Essence Constraints

5.1 Audio Characteristics Constraints

All Audio Characteristics shall conform to SMPTE 428-2 and are further constrained per the below sections.

In addition, DCDM Audio Essence intended for a given Composition is constrained to have identical audio parameter values as depicted below. Specifically, the combination of audio essences having differing audio parameters shall not be delivered for the same Composition.

5.1.1 Audio Reference Level and Alignment Tones

DCDM audio may contain alignment tones of any frequency at the SMPTE 428-2 reference level of –20dbfs. Alignment tones (if present) shall be placed in a location that will not be normally played out in a composition; for example, 30 seconds after the LFOA of a post-production reel.

5.1.2 Audio Frequency Response

The reader should be aware that the audio essence as delivered by the DCDM does not contain any frequency response constraints, nor are the audio tracks data reduced or compressed. Therefore, what is recorded on the mixing stage is what is delivered to the theater B chain.

5.1.3 Audio Speed Reference and Sample Rate Constraints

5.1.3.1 **Audio Speed, Sample Rate and Frame Rate Relationships** (Informative)

Audio speed is often expressed by the frame rate of the image to which it synchronizes. In reality, it is the action that determines the true audio speed, not the image format frame rate. If the image is being displayed at the same speed as the live action it captured, then the audio matching the image must be played at the same speed that it was recorded in order to synchronize. This is true regardless of how many frames per second the image format uses. The speed of audio essence is constant as long as the action depicted by the image is constant, regardless of the frame rate of the image format. Therefore, expressing audio speed by an absolute frame rate number can be misleading and does not truly depict the audio speed.

Since digital audio is captured at a given sample rate, such as 48.00K, it will match action when clocked at that same rate during playout. However, as above, it is the action that determines the true audio speed, not the sample rate. For example, if the image is projected at a different speed (such as a 24 fps image projected at 25 fps), the audio speed must increase to match. The audio could be clocked faster to attain this; however, D-Cinema does not support sample rates other than 48.00K and 96.00K, so a speed and sample rate conversion would be done so the audio will play at the faster speed while clocked at 48.00K. Thus, the sampling rate itself is also not a true measure of audio speed. In actual fact, it is the relationship between the audio sampling rate and the image speed that is truly relevant.

5.1.3.2 Audio Speed and Sample Rate Constraints

In D-Cinema the speed of DCDM audio essence is constrained to match the exact speed of the action in the picture essence for the intended composition. The edit rate of the intended composition is determined by the picture essence. As noted above, regardless of the picture edit rate, the intended speed of the action is the same, so the speed of the audio matching the action is the same. Therefore, the absolute speed of DCDM audio essence is a constant regardless of the composition edit rate.

When the audio DCDM is packaged into a DCP, it is wrapped into parcels that correspond to the edit unit(s) of the picture essence in the composition. Therefore, audio essence delivered as a DCDM is constrained as to the

number of samples that are contained in the duration of the edit unit for the intended composition. Since the audio speed is constant, the number of audio samples in the duration of an edit unit is constrained by the audio sample rate and the edit rate. SMPTE 428-2 defines the allowable audio sample rates to be 48.00 KHz and 96.00 KHz, so for an edit rate of 24/1, there must be exactly 2000 samples per edit unit at 48.00 kHz, and exactly 4000 samples per edit unit at 96.00 KHz.

Since Digital Cinema does not support fractional frame rates such as 23.976 fps or "pulled down" sampling rates such as 47.952K or 95.909K, the associated number of samples per edit unit are not allowed in an audio DCDM. For example, 2002 or 4004 samples per edit unit are not allowed.

5.1.4 Audio Bit Depth Constraints

In addition to the 24 bit justification constraint stated in Section 3.1 of SMPTE 428-2, audio sample words with fewer than 24 significant bits shall be padded to a full 24-bit word length by placing them in the most significant bits of a 24 bit word and setting the unused LSB's to zero to ensure a full 24 bit word length for delivery as an audio DCDM.

5.2 Audio File Constraints

5.2.1 Audio File Format

Audio essence shall be stored in single channel Broadcast Wave files, one Broadcast Wave file per audio channel. Broadcast Wave files are defined by ITU-R BR.1352-1, Broadcast Wave Format (BWF), Annex 1

Each post-production reel shall have its own separate set of Broadcast Wave files. For example, a 5.1 channel audio essence shall be conveyed as a set of six individual Broadcast Wave files, one set per post-production reel.

The file length shall conform to an integer value of editable units. All audio files for a given post-production reel shall have identical length, and shall be in sync with each other when played from the same sample point.

5.2.2 Audio File Date

Each audio file shall be specifically labeled as to its essence modification date. This date is assigned by the creator of the audio file and represents the most recent date that the audio essence was modified. This date may be used to determine the correct audio to use for the creation of a DCP.

The audio file date shall be conveyed per the audio information detailed in Section 6.3 and by the audio file name per Section 5.2.6. It may additionally be included in the Broadcast Chunk within the header of the broadcast wave file.

The date listed in a file directory (i.e., the file system date on a storage medium) is not an acceptable alternative to this constraint.

5.2.3 Audio File Channel Labeling

Each audio file shall be specifically labeled as to its intended reproduction channel. Audio file channel labels shall conform to SMPTE 428-3, Section 4, column 3 ("Label/Name"). For example, an audio file intended for the far left screen loudspeaker position would have the label "L".

The audio channel label shall be conveyed per the audio information detailed in Section 6.3 and also by the audio file name per Section 5.2.6. It may additionally be included in the Broadcast Chunk within the header of the Broadcast Wave file.

5.2.4 Audio File Framing and Synchronization

DCDM audio files shall start at the left edge (beginning of the frame period) of the "Picture Start" frame on the digital leader (a.k.a. "Zero Feet"), such that the left edge of the audio file is aligned to the left edge of the Picture Start frame. This establishes the sync reference for picture essence and audio essence. As a visual confirmation of synchronization, each audio file shall have a sync "two-pop" for one editable unit (frame), aligned exactly with frame boundaries, which starts exactly 2 seconds before the First Frame of Action (FFOA), and another sync "tail" pop for one editable unit (frame) which starts exactly 2 seconds after the beginning of the Last Frame of Action (LFOA).

Sync pops are to be a burst of sine wave. The frequency of this sine wave is typically 1 KHz, but other frequencies may be used as long as they are placed and framed according to the above constraints. The sync pops should be at –20 dbfs.

There shall be 8 seconds of duration (e.g., 192 frames at 24 fps) from the start of the audio file to the First Frame of Action. Other than the sync "2-pop" detailed above, this 8 seconds should be "audio black"; i.e., containing no audio modulation-unless a prelap is included.

The End of Program audio for a given post-production reel shall be inclusive of the Last Frame of Action ("LFOA"). The End of File for DCDM audio files shall be at least 2 seconds plus two editable units after the beginning of the Last Frame of Action, such that the tail pop is included plus at least one more editable unit (frame). Longer audio files are permitted and may include such items as alignment tones per Section 5.1.1. As noted above, all audio files for a given soundfield configuration and post-production reel shall be the same length.

The end of the audio file shall align to an editable unit (frame) boundary per Section 5.2.1.

The 2 seconds minus one editable unit area after the LFOA and before the tail pop may contain "audio black" or a pullup. The pullup must end a minimum of 2 editable units before the tail pop. At no time is the pullup to include or obscure the tail pop.

The Figure 1 illustrates correct and incorrect framing for 48 KHz, 24 fps audio essence.

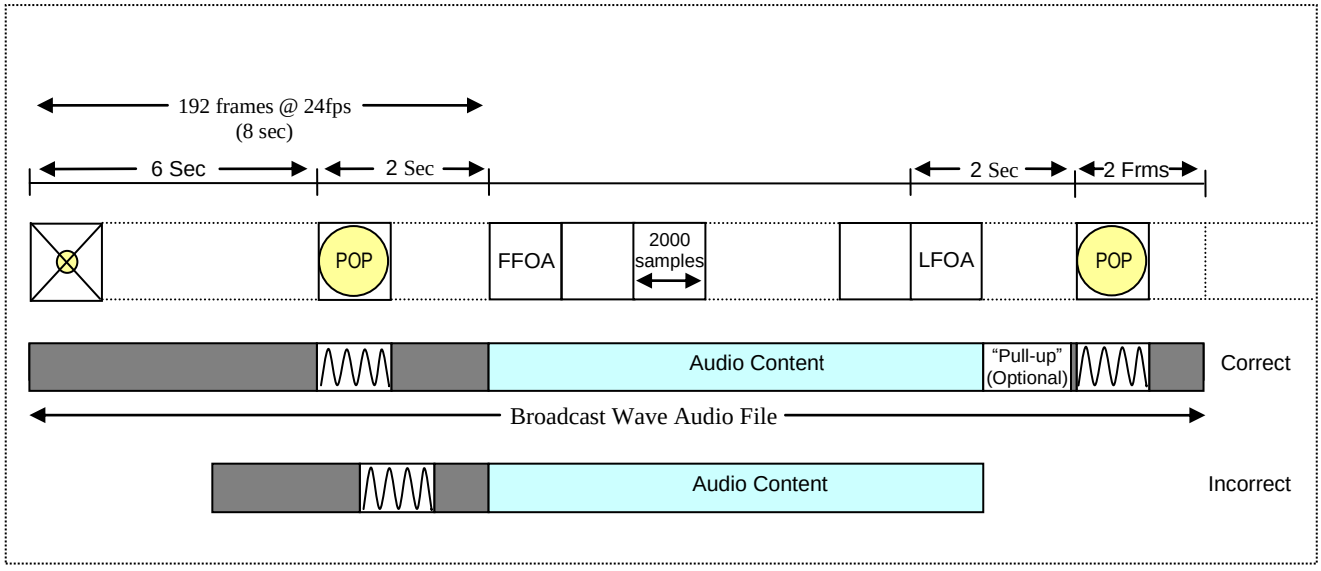

Figure 1 – Audio File Framing and Synchronization

5.2.5 Audio File Reel Changeover Constraints

The audio files must be constructed such that the last sample of the audio essence for the Last Frame of Action of an outgoing reel and the first sample of the First Frame of Action of the intended incoming reel can be played sequentially and result in a seamless transition when “butt spliced”. In order to achieve this, the changeovers must work with sample accuracy.

It is suggested that content providers test reel changeovers to insure that this requirement is met.

5.2.6 Audio File Naming Constraints

Audio file names shall be named according to the structure and constraints of this Section 5.2.6.

File names shall be unique for each file. The use of upper versus lower case letters shall not be used as the sole distinguishing attribute.

All audio files contained in the audio file set for a given soundfield configuration and post-production reel intended for a given Composition shall have the same name except for the individual audio file channel label.

Examples of file names are given in Annex A of this document.

5.2.6.1 File Naming Structure Constraints

The Audio File Naming Structure shall contain the following informational fields in the order defined below. Other informational fields may optionally be contained in the file name at the discretion of the content provider. Any optional fields shall be placed at the end of the required fields in the filename. An underscore “_” shall be used as a separator between each of the fields, with the exception of the separator between the date and channel label, which may be an underscore or a period (dot).

The content of these fields and the filename as a whole shall be restricted to the following ISO 646 characters: Upper and lower case letters (A-Z, a-z), decimal digits (0-9), periods (“dots”), underscores (_) and hyphens (-).

Note: None of the fields, nor the filename as a whole, may contain spaces, extraneous punctuation, or any character value other than the set defined above.

Field definitions are given above in Section 4; constraints on their use in the naming structure itself are detailed in Section 5.2.6.2.

The required fields and order are:

Content Title

Content Version

Content Kind

Post-Production Reel Number

Language

Soundfield Configuration

Audio Content and Type (may be two fields)

File Date

Channel Label

The file naming schema shall be as follows:

<ContentTitle>_<ContentVersion>_<ContentKind>_<PostProductionReelNumber>_<Language>_<SoundfieldConfiguration>_<AudioContent>_<AudioType>_<WatermarkType (*optional*)>_<FileDate>_<ChannelLabel>.wav

5.2.6.2 Audio File Name Information Field Constraints

The following constraints apply to the content in the information fields used in audio file naming. File names should be kept as short as possible while still conveying the required information. Character lengths for the information in each field are suggested but not mandated. The total number of characters shall not exceed 63.

Content Title: This can be the full title, an extraction of the title or an abbreviation. If the full title or an extraction is used, individual words are separated by underscores "_" or run together with no spaces. If a title abbreviation is used, it should be a single word that the content provider has determined will uniquely represent that title. The Content Title and its representation must be consistent for all audio essence intended for the same composition. This field should be no more than 15 characters including underscores.

Content Version: This can be the complete version description or an abbreviation. If the complete version description is used, individual words are separated by underscores "_" or run together with no spaces. If a version abbreviation is used, it should be a single word that the content provider has determined will uniquely represent that version. The Content Version and its representation must be consistent for all audio essence intended for the same composition. This field should be no more than 10 characters including underscores

The Content Version field shall always be included, even if there is only one known version of the content at the time of post-production, in order to clearly differentiate the essence in case another Content Version is later made.

Content Kind: This can be the complete content kind description as depicted in Table 2 of SMPTE 429-7, or an abbreviation. If an abbreviation is used, it should be a single word that the content provider has determined will uniquely represent that content kind. The Content Kind and its representation must be consistent for all audio essence intended for the same composition. This field should be no more than 15 characters including underscores.

Post-Production Reel Number: This can be a mixture of numbers and letters. The depiction used for the audio post-production reel number must be consistent for all audio post-production reels intended for the same composition. The post-production reel number and its representation must be consistent for all audio essence belonging to the same post-production reel in a given composition. This field should be no more than 6 characters.

Language: The language of the essence shall be specified using the 3-character codes found in "ISO 639-2/T, Codes for the Representation of Names of Languages – Part 2: Alpha-3 Code." Where conflicts arise in the language code, only the Terminological (T) codes shall be used. The code used for the audio language shall be consistent for all essence of that language intended for the same composition. This field shall contain only the 3 code characters.

Soundfield Configuration: This can be a mixture of numbers and letters. If individual words are used, they are to be separated by underscores "_" or run together with no spaces. The use of periods within the designation is heavily discouraged, as they may interfere with certain file structures. For example, a 5.1 audio configuration should be represented by "51" rather than "5.1". The soundfield configuration label or representation shall be consistent for all DCDM audio files to be used for a given composition. This field should contain no more than 5 characters.

Audio Content and Type: These are generally depicted by individual words or an abbreviation containing letters, and may be represented by one or two fields. If individual words are used, they shall be separated by underscores "_" or run together with no spaces. If an abbreviation is used, it should be one that the content provider has determined will uniquely represent that audio content and type. The Audio Content and Type and its representation must be consistent for all audio essence intended for the same composition. This field should be no more than 10 characters including underscores.

File Date: The File date shall be abbreviated YYYYMMDD.

Channel Label: This indicates the channel assignment that the post-production assigned it, the depiction of which is constrained in Section 5.2.3. It may be separated from the file date by an underscore "_" or a period ("dot").

6 DCDM Audio Information-Required

The following pertinent audio information shall be supplied with the audio DCDM.

Other audio information not pertaining to the Audio DCDM that may be of interest to the content providers may be provided separately.

Definitions of the below are in Section 4 of this document.

6.1 DCDM Composition Audio Information

The following information must be consistent for all audio in a given composition:

Content Title

Content Version

Content Kind

Sample Rate

Frame Rate

6.2 DCDM Post-Production Reel Information

This information pertains to a specific post-production reel:

Post-Production Reel Number

Last Frame of Action ("LFOA" or "LFOP") (expressed in feet/frames or frames, which includes the leader)

OR

Duration (expressed as the number of playable frames, which does not include the leader)

(For conversion of LFOA to duration, see Section 4.)

6.3 DCDM Audio Track Information

Language

Soundfield Configuration

Audio Content and Type

File Date

Channel Label

Alignment tone location (if present)

File name

7 DCDM Insert Audio Information-Optional

DCDM insert audio shall not include a leader, and shall comply with the changeover constraints in Section 5.2.5.

The following pertinent audio information may be supplied with the audio DCDM.

If the DCDM audio essence is insert or replacement audio which is intended to augment or replace existing audio to make a revised version, the following information shall be provided in addition to the requirements in Section 6:

Insert file (Y/N)

Insert File Type (Augment file, Replacement File). If no indication, the file is not an insert file.

Location of the first frame to be affected by the insert file (e.g. reel and frame to be replaced, displaced or augmented in the existing file.)

The number of frames in the existing file to be replaced or displaced by the insert file.

Reference information to existing essence to be revised

Existing CPL file to be revised (if known).

Annex A (Informative)
Audio File Naming Examples

Below is the overall audio file naming schema example:

<ContentTitle>_<ContentVersion>_<ContentKind>_<PostProductionReelNumber>_<Language>_<SoundfieldConfiguration>_<AudioContent>_<AudioType>_<WatermarkType (*optional*>_<FileDate>_<ChannelLabel>.wav

Below are examples of audio tracks and their corresponding file names:

"Your Movie", domestic version, feature, reel 3, English, 5.1 printmaster, created April 28th, 2008, center channel

yourmovie_domestic_feature_R3_ENG_51_PM_20080428_C.wav

"Your Movie", Director's Cut version, feature, reel 5, Italian, mono, Visually Impaired Narration, created April 30, 2008, center channel

yormov_DC_feat_R5_ITA_M_VI-N_20080430_C.wav

"Your Movie", International version, trailerB, reel 1, German, 5.1, created October 15, 2007, Left Surround channel

yormov_I_TRB_R1_DEU_51_20071015.Ls.wav