MPEG Interoperability Initiative: Sound and Picture Track File

Application Specification for Digital Cinema Packaging (AS-DCP)

NOTICE

This document is provided without warranty as to its fitness for a particular purpose.

Document type: Standard
Document subtype:
Document stage: Draft
Document language: English

Contents

MPEG Interoperability Initiative: Sound and Picture Track File
Application Specification for Digital Cinema Packaging (AS-DCP)

1   Scope

This document specifies the format of sound and picture tracks files, in the context of the overall Application Specification for Digital Cinema Packaging (AS-DCP). The specification follows the proposed SMPTE 377M MXF standard, but is further constrained to allow far fewer variations than those described in the latter. The constraints reflect specific requirements of Digital Cinema distribution, and simplify design, facilitate interoperability and ensure predictable processing.

2   Normative References

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

SMPTE, Proposed 377M, 2003. *For Television – MXF Format*

SMPTE, Proposed 390M, 2003. *For Television – MXF OP-ATOM*

SMPTE, Proposed 379M, 2003. *For Television – MXF Generic Container*

SMPTE, Proposed 381M, 2003. *For Television – MPEG stream in Generic Container*

SMPTE, Proposed 382M, 2003. *For Television – MXF Audio in GC*

SMPTE, Proposed Revision 330M, 2003. *For Television – UMID*

SMPTE, Proposed 000X, 2004. *For Cinema – AS-DCP Track File Essence Encryption*

Internet Engineering Task Force (IETF) (1996, November). *Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies* [WWW document]. URL http://www.ietf.org/rfc/rfc2045.txt

3   Overview

The MXF standards (proposed SMPTE 377M and related documents) describe a very flexible format with many options. To make effective use of MXF in a given application, it is necessary to choose which options to permit. These options fall in four broad categories:

- Pattern (in this case, a restricted version of SMPTE 390M OP-ATOM with internal essence)
- Essence (in this case, sound and picture essence)
- Header Metadata (in this case, strict compliance with the proposed SMPTE 377M)

▪ Descriptive Metadata (in this case, a small set of essential data)

These categories of options and choices for DCP are detailed in the sections below.

4   Pattern Constraints

4.1   General

The level of flexibility of the standard Operational Patterns (OPs) defined by SMPTE 378M and 390M are not required for d-cinema-specific use. The AS-DCP track file format specified is a constrained version of the MXF OP-ATOM with internal essence (SMPTE 390M).

However, it is useful to employ the same basic terminology as the standardized OPs, shown in Figure 1.
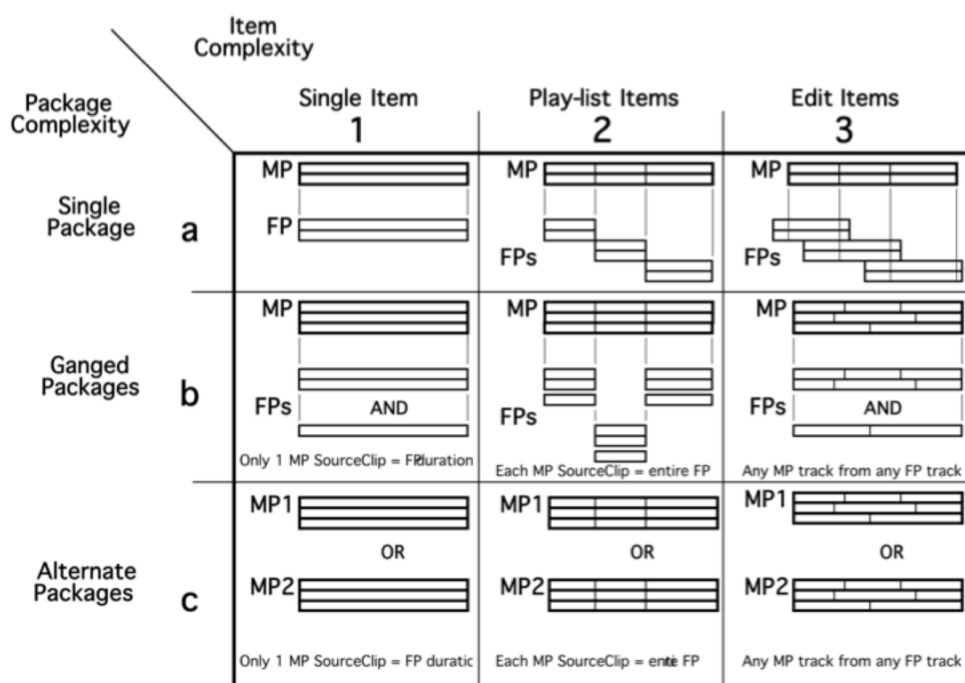
Figure 1 (Informative): Illustration of the Operational Pattern Axes

4.2   Baseline Operational Pattern

MXF defines Operational Patterns that place limits on the structure and complexity of content in MXF files. All standard OPs, including the most basic OP1A (SMPTE 378M), permit many variations within the essence stream that are unneeded or undesirable for d-cinema applications and optimized implementations. For instance, "Single Package" OPs (row a) assume the use of interleaved Essence Containers, which are undesirable in d-cinema operational scenarios. OP3c, which provide for multiple versions – column 3 (Edit Items) and row c (Alternate Packages) – is in conflict with the requirement for an external Composition List.

The SMPTE 390M OP-ATOM (with internal essence) draft resolve these issues. The representation of alternates and synchronization of separate picture and sound reels, which are beyond the scope of OP-ATOM (with internal essence), are handled in the separate Composition Lists. Therefore, the present draft of this document defines the AS-DCP track files as a constraint on the standard OP-ATOM with internal essence.

Each AS-DCP track file will contain one File Package, representing the Track File in its entirety, and one Material Package. The latter shall not be used during playback and may be used to represent the original sync-locked reel. More precisely, this Material Package should reference, if available, all tracks created in the original version of the material. Such information can be used, a posteriori, to confirm the source of individual track files.

Track files shall carry the following Universal Label for the Operational Pattern:

Label	Meaning
`06.0e.2b.34.04.01.01.01.0d.01.02.01.10.00.00.00`	OP Atom: 1 Track, 1 SourceClip

Note: This label value is different than the registered version (registered version byte 8 = 02 instead of 01), but is being retained for the purpose of compatibility with existing practice.

4.3   Additional Constraints

4.3.1   Container

AS-DCP track files shall use the MXF Generic Container SMPTE 379M.

Number of Elements is described in the Picture, Sound, and Other sections below.

4.3.2   Interleaving

Essence in AS-DCP track files shall not be interleaved[1] between essence types.

Essence in each Generic Container (GC) in AS-DCP track files shall be frame-wrapped.

4.3.3   System Item

Essence in AS-DCP track files shall not include GC System Items.

4.3.4   Partitions

Essence in AS-DCP track files shall not be divided into Partitions. Each material file shall contain a single essence partition, per OP-ATOM (with internal essence).

The partition shall contain complete GOPs if MPEG II or other inter-frame coders are used. In other words, the partition shall start with an MPEG random access point I-frame, GOP header and Sequence Header.

When possible, Sound and other Essence Partitions should have the same extent as the Picture Essence Partitions to which they relate.

4.3.5   Index Tables

AS-DCP track files shall include standard MXF Index Tables per (proposed) SMPTE 377M. These Index Tables shall be divided into Index Table Segments.

Each Segment shall be carried in the Footer Partition of the file it indexes.

AS-DCP track files shall contain a Random Index Pack per (proposed) SMPTE 377M.

---

[1] From SMPTE 377M: Multiplexed means putting different partitions one after the other whereas "interleaved" means that the Essence Container itself has different components which are interleaved on a time division basis.

4.3.6   Encryption

Encryption of the essence contained in AS-DCP track files shall follow the AS-DCP Track File Essence Encryption specification.

Figure 2 illustrates the correspondence between a plaintext and an Encrypted Triplet[2]. The value *V* of a source plaintext *KLV* triplet is first encrypted to yield *E(V)*. The resulting *KLE(V)* triplet is wrapped in a encrypted *K'L'V'* triplet, identified by a unique and universal label *K'*. *V'* consists of the encrypted *KLE(V)* source triplet and the cryptographic information *I* necessary to decrypt *E(V)*. *L'* refers to the full length of *V'*. KLV Fill items may optionally be inserted in the source material prior to encryption to avoid regeneration of index tables.

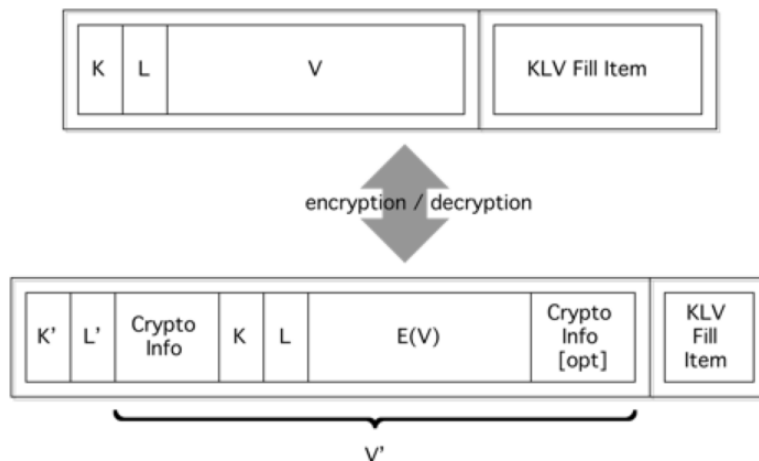Figure 2. Correspondence between Source and Encrypted KLV Triplets. Red hatching depicts the encrypted portion of the Encrypted Triplet; other items are left in the clear. Only the value of Source Triplet is encrypted, allowing the essence information to be encrypted prior to wrapping into Encrypted Triplets. The cryptographic information associated with the Encrypted Triplet is depicted by the two Crypto Info elements, with the trailing element being optional. The KLV Fill Items may be inserted to facilitate particular methods for KLV encryption, but are not otherwise required.

4.3.7   KAG Size

AS-DCP track files shall employ the default KLV Alignment Grid of 1. See (proposed) SMPTE 377M section 5.4.1.

4.3.8   Essence Types

The Essence types shall be selected from the repertoires listed in section 4 above.

Within an AS-DCP track file, all Essence Containers of the same kind (Picture, Sound, Subpicture, other) shall be of the same type.

4.4   Labeling

AS-DCP track files shall be labeled with a registered AS-DCP label, the GC and all other Essence Container Labels in every Partition Pack and every Preface Set.

---

[2] This specification does not require the essence to be wrapped in a KLV triplet to enable its encryption. In other words, essence may be encrypted prior to being wrapped in an Encrypted Triplet.

5   Essence Constraints

The MXF OP-ATOM files used by AS-DCP support a wide range of picture and sound essence schemes, and are, for instance, raster- and compression-agnostic. It is not yet determined whether AS-DCP will identify a single scheme, or a selection of sanctioned schemes, for each essence type. A number of constraints exist independently of this determination.

5.1   General

5.1.1   Native start and end points

The native start point, native end point and duration of a track contained in an AS-DCP track file are derived from the Start Position and Duration items found in the corresponding Source Clip (see SMPTE 377M at B.10). Unless otherwise specified, playback of an essence track shall start and stop at the native start and end points, respectively.

5.2   Picture

5.2.1   General

Each Picture Track File (Essence Container) in an AS-DCP track file shall contain one MXF GC Picture Item, each of which contains a single GC Picture Element.

5.2.2   Picture Packetization

The picture essence stream shall be encoded in KLV Packets using (proposed) SMPTE 381M, using frame by frame wrapping and indexed accordingly.

5.2.3   Picture Parameters

The picture shall be encoded according to the specifications in Annex A

5.3   Sound

5.3.1   General

5.3.2   Sound Sampling

All Sound Tracks in an AS-DCP track file shall have the same baseband sample rate and bit depth as specified in Annex A..

5.3.3   Sound Packetization

Sound essence shall be encoded in KLV packets and packetized in compliance with SMPTE 382M Audio in MXF GC, using Wave audio frame-by-frame wrapping.

5.4   Timecode

Timecode information is present in AS-DCP track files for compliance with the OP-ATOM specification and for informational purposes only. Timecode information shall not be required for the playback of AS-DCP track files. (note: This track is required by OP-ATOM)

5.4.1   General

In AS-DCP, MXF Timecode Tracks may be used to provide an indication of time relative to the start of a presentation.

AS-DCP track files shall contain synthetic time described by TimecodeSegments, i.e. continuously incrementing numbers from a specified starting point. AS-DCP track files shall not contain TimecodeStream data.

AS-DCP track files should have a starting time of 01:00:00:00 whenever possible.

5.4.2   Material Package Timecode

The Timecode of Material Packages in AS-DCP track files shall consists of a single continuous segment.

5.4.3   File Package Timecode

The Timecode of each File Package in AS-DCP track files shall consists of a single continuous segment, with a starting time that matches that of any historical Source Package if known, else 01:00:00:00.

5.4.4   Source Package Timecode

The Timecode of each historical Source Package in AS-DCP track files shall be consists of a single continuous segment, with a starting time that matches that of any preceding source package, or incoming master if known, else 01:00:00:00.

6   Header Metadata Constraints

6.1   General

Header Metadata of AS-DCP track files shall conform to (proposed) SMPTE 377M MXF Format and shall be labeled correctly in the Partition Packs.

Files shall be Closed and Complete.

Header Metadata shall be present in the Header Partition. No other Partitions shall contain copies of the Header Metadata.

6.2   Package IDs (UMIDs)

Every UMID in an AS-DCP track file shall be either a standard SMPTE 330M-2003 UMID or a DCP-mapped asset identifier.

UMIDs of AS-DCP track file Packages shall be generated using one of the regular SMPTE 330M methods.

UMIDs of AS-DCP historical Source Packages shall be obtained from the asset they identify.

7   Descriptive Metadata Constraints

7.1   General

AS-DCP track files shall contain (proposed) SMPTE 377M Header Metadata.

AS-DCP track files shall not make use of SMPTE 380M MXF DMS-1.

AS-DCP track files may contain additional metadata which conforms to the general rules for (proposed) SMPTE 377M DMS Frameworks. Additional metadata is currently defined in the KLV Encryption Spec.

7.2   DMS-DCP

DMS-DCP metadata shall be carried in AS-DCP track files in a DM Segment of a Static DM Track as defined by (proposed) SMPTE 377M.

DMS-DCP specifies the set of metadata that is to be packaged inside an MXF file for use within DCP.

7.3   Labeling

AS-DCP track files shall be labeled with a registered DMS-DCP label in every Partition Pack and every Preface Set.

7.4   File Names and Asset Identity

The names of AS-DCP track files may include the asset id or any other identifier meaningful to the originator.

However, the identity of the asset contained in the file shall be determined from the UMID of the File Package in the file, not from the filename.

8   MIME Types

8.1   Picture

The MIME type for a track file containing picture essence shall be `application/x-smpte-mxf;asdcpKind=Picture`.

8.2   Sound

The MIME type for a track file containing sound essence shall be `application/x-smpte-mxf;asdcpKind=Sound`.

9   Annex A – Additional Essence Constraints

9.1   Picture

The compressed picture shall be encoded using the MPEG II (ISO/IEC 13818-2:2000) picture codec. The 4:2:0 MP@HL profile shall be used unless otherwise specified in this section.

9.1.1   Picture Parameters

Raster resolution shall be 1920 by 1080.

9.1.2   Frame Rate

The nominal frame rate shall be 23.976 or 24.000 frames per second.

9.1.3   Compression

The playback of both constant bit rate (CBR) and variable bit rate (VBR) video stream shall be supported. For VBR, the maximum bit rates are considered as peak bit rates. The following constraints also apply to the playable portion of the elementary stream as defined by the associated composition playlist.

▪ The stream shall directly start with a sequence header;

▪ The values in the sequence header shall be correct and complete;

▪ If the stream is CBR, the bit_rate value of the sequence header shall be set to the bit rate of the stream;

▪ If the stream is VBR, the bit_rate value of the sequence header shall be set to the maximum bitrate of the stream;

▪ The time_code value of the GOP header shall be incremented all through the stream; meaning that the difference between the beginning and ending time codes will give the stream duration;

▪ At least the first GOP shall be a closed GOP meaning that its first P- or B-frame doesn't depend on the previous GOP, since there is none. This issue is important to avoid artifacts at the beginning of the stream. This problem could occur with some hardware encoders.

▪ The first picture of the stream shall be the beginning of a new GOP and include a GOP header.

▪ The last frame, in presentation order, shall be an I- or P-frame.

▪ The maximum bit rate shall be 80 Mbps for MPEG2 MP@HL.

▪ Only progressive sequences are allowed, Therefore, all progressive_sequence and progressive_frame flags shall be set to 1.

9.1.4   Packetization

▪ Compressed picture stream shall be packetized as an elementary stream in compliance with MPEG II ISO 13818-1:2000. MPEG-2 PES packetization shall not be used.

▪ Every GOP Header shall be directly preceded by a Sequence Header and optional sequence extension headers .[Must ensure that sequence header is wrapped in the same frame as its associated I-Frame]

9.1.5   Plaintext Offset[3]

Whenever the AS-DCP Essence Encryption specification is used to encrypt MPEG 2 picture essence as defined in this section, the Plaintext Offset parameter of each Encrypted Triplet shall be zero or, if non-zero, shall locate the first byte immediately following the first 4 byte slice start code..

9.2   Sound

9.2.1   Sound Sampling

Baseband sound shall consist of linear PCM samples at 48 kHz and 24 bit.

The sound sampling rate will result in the following relationship with the picture frame rate:

---

[3] The AS-DCP Track File Essence Encryption specification enables the encryption of individual essence frames within AS-DCP Track Files. For flexibility, a Plaintext Offset parameter is associated with each frame. This parameter is created at time of encryption and indicates the number of bytes, if any, that have been left unencrypted at the beginning of the frame. The AS-DCP Track File Essence Encryption specification does not however provide specific guidance for selecting a Plaintext Offset value, and compliant devices are expected to support any legal value the Plaintext Offset parameter may take. This section specifies the value of the Plaintext Offset parameter used when encrypting MPEG II essence. By leaving critical MPEG II header information unencrypted, the manipulation of the encrypted MPEG II elementary stream is facilitated. Standard MPEG II tools such as multiplexers may, for instance, be used with minimal effort.

For 24 fps picture, there shall be 2000 samples per frame

For 23.976 fps picture, there shall be 2002 samples per frame

9.2.2   Edit Rate

The edit rate of the sound track file shall be the same as the edit rate of the accompanying picture track.

9.2.3   Channel Mapping

5.1 channels is the base configuration. The input channels shall be mapped to presentation channels as shown in

Table **1**.

Table 1 - Six channel mapping

AES Pair#/Ch#	Channel #	Label / Name	Description
1/1	1	L/Left	Far left screen loudspeaker
1/2	2	R/Right	Far right screen loudspeaker
2/1	3	C/Center	Center screen loudspeaker
2/2	4	LFE/Screen	Screen Low Frequency Effects subwoofer loudspeakers
3/1	5	Ls/Left Surround	Left wall surround loudspeakers
3/2	6	Rs/Right Surround	Right wall surround loudspeakers

Samples shall be interleaved into the Generic Container in channel order as indicated in

Table **1**.

10 Bibliography

SMPTE, Proposed 378M, 2003. *For Television – MXF OP1a*

11 Change History

Ver	Date	By	Sect	Description
1	8 June 2004			
2.	15 June 2004			Changes from review
2.2	18 June 18, 2004			Changes from review. Plaintext interchange reference.
2,3	17 August 2004			Annex A: Entry point must reference the start of a GOP. Removed HP@HL
2.4	17 May 2005			Added ability to create encrypted picture track with zero plaintext offset. Documented Sound Essence UL. Clarified picture essence to be progressive.
2.5	18 May 2005			Removed specification for Sound Essence UL
2.6	26 May 2005			Minor editorial changes.
2.7	1 Nov 2005		4.2	Added note to specify OP-Atom UL (non std)