FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/IEC
FDIS
23001-7**

ISO/IEC JTC **1**/SC **29**

Secretariat: JISC

Voting begins
on: **2015-10-27**

Voting terminates
on: **2015-12-27**

Information technology — MPEG systems technologies —

Part 7:
Common encryption in ISO base media file format files

Technologies de l'information — Technologies des systèmes MPEG —

Partie 7: Cryptage commun des fichiers au format de fichier de médias de la base ISO

Reference number
ISO/IEC FDIS 23001-7:2015(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This third edition cancels and replaces the second edition (ISO/IEC 23001-7:2015), which has been technically revised.

ISO/IEC 23001 consists of the following parts, under the general title *Information technology — MPEG systems technologies*:

— *Part 1: Binary MPEG format for XML*

— *Part 2: Fragment request units*

— *Part 3: XML IPMP messages*

— *Part 4: Codec configuration representation*

— *Part 5: Bitstream Syntax Description Language (BSDL)*

— *Part 7: Common encryption in ISO base media file format files*

— *Part 8: Coding-independent code points*

— *Part 9: Common encryption of MPEG-2 transport streams*

— *Part 10: Carriage of timed metadata metrics of media in ISO base media file format*

— *Part 11: Energy-efficient media consumption (green metadata)*

— *Part 12: Sample variants in the ISO base media file format*

Introduction

Common Encryption specifies standard encryption and key mapping methods that can be utilized to enable decryption of the same file using different Digital Rights Management (DRM) and key management systems. It operates by defining encryption algorithms and encryption-related metadata necessary to decrypt the protected streams, yet it leaves the details of rights mappings, key acquisition and storage, DRM content protection compliance rules, etc., up to the DRM system or systems. For instance, DRM systems is intended to support identifying the decryption key via stored key identifiers (KIDs), but how each DRM system protects and locates the KID identified decryption key is left to a DRM-specific method.

DRM-specific information such as licenses, rights, and license acquisition information can be stored in an ISO Base Media file using a Protection System Specific Header box ('`pssh`'). Each instance of this box stored in the file corresponds to one applicable DRM system identified by a well-known `SystemID`. DRM licenses or license acquisition information need not be stored in the file in order to look up a separately delivered key using a `KID` stored in the file and decrypt media samples using the encryption parameters stored in each track.

The second edition of this part of ISO/IEC 23001 added XML representations of Common Encryption parameters for delivery in XML documents, such as an MPEG-DASH Media Presentation Description Documents (MPD). The second edition also defined the '`cbc1`' protection scheme using AES-CBC mode encryption.

The third edition added '`cbcs`' and '`cens`' protection schemes for pattern encryption, which encrypt only a fraction of the data Blocks within each video Subsample protected. Pattern encryption reduces the computational power required by devices to decrypt video tracks.

Information technology — MPEG systems technologies —

Part 7:
Common encryption in ISO base media file format files

1 Scope

This part of ISO/IEC 23001 specifies common encryption formats for use in any file format based on ISO/IEC 14496-12. File, track, and track fragment metadata is specified to enable multiple digital rights and key management systems (DRMs) to access the same common encrypted file or stream. This part of ISO/IEC 23001 does not define a DRM system.

The AES-128 symmetric block cipher is incorporated by reference to encrypt elementary stream data contained in media samples. Both AES counter mode (CTR) and Cipher Block Chaining (CBC) are specified in separate protection schemes. Partial encryption using a pattern of encrypted and clear blocks is also specified in separate protection schemes. The identification of encryption keys, Initialization Vector storage and processing is specified for each scheme.

Subsample encryption is specified for NAL structured video, such as AVC and HEVC, to enable normal processing and editing of video elementary streams prior to decryption.

An XML representation is specified for important common encryption information so that it can be included in XML files as standard elements and attributes to enable interoperable license and key management prior to media file download.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 14496–12, *Information technology — Coding of audio-visual objects — Part 12: ISO Base Media File Format*

ISO/IEC 14496–15, *Information technology — Coding of audio-visual objects — Part 15: Carriage of NAL unit structured video in the ISO Base Media File Format*

3 Terms, definitions, and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

NOTE Words used as defined terms and normative terms (SHALL, SHOULD and MAY) are written in upper case to distinguish them from the same word intending its dictionary definition.

3.1.1
constant IV
initialization vector (3.1.3) specified in a sample entry or sample group description that applies to all samples and *subsamples* (3.1.8) under that sample entry or mapped to that sample group

3.1.2
block
16-byte extent of sample data that may be encrypted or decrypted by the AES-128 block cipher, in which case, a cipher block

3.1.3
initialization vector
8-byte or 16-byte value used in combination with a key and a 16-byte *block* (3.1.2) of content to create the first cipher block in a chain and derive subsequent cipher blocks in a cipher block chain

3.1.4
ISO Base Media File
file conforming to the file format described in ISO/IEC 14496-12 in which the techniques in ISO/IEC 23001-7 may be used

3.1.5
NAL unit
syntax structure containing an indication of the type of data to follow and bytes containing that data in the form of an RBSP interspersed as necessary with emulation prevention bytes

3.1.6
NAL structured video
video streams composed of *NAL units* (3.1.5) of which the carriage is specified by ISO/IEC 14496-15

3.1.7
protection scheme
encryption algorithm and information defined in this part of ISO/IEC 23001 and identified by a four character code in an ISO Media track's Scheme Type Box ('schm')

3.1.8
subsample
byte range within a sample consisting of an unprotected byte range followed by a protected byte range

3.2 Abbreviated terms

AES	Advanced Encryption Standard as specified in Federal Information Processing Standards Publication 197, FIPS-197
AES-CTR	AES Counter Mode as specified in *Recommendation of Block Cipher Modes of Operation*, NIST, NIST Special Publication 800-38A
AES-CBC	AES Cipher-Block Chaining Mode as specified in *Recommendation of Block Cipher Modes of Operation*, NIST, NIST Special Publication 800-38A
AVC	Advanced Video Coding as specified in ISO/IEC 14496-10
HEVC	High Efficiency Video Coding as specified in ISO/IEC 23008-2
IV	Initialization Vector
NAL	Network Abstraction Layer, as specified in ISO/IEC 14496-10 and ISO/IEC 23008-2
URN	Unique Resource Name
UUID	Universally Unique Identifier

4 Protection schemes

4.1 Scheme type signaling

Scheme signaling SHALL conform to ISO/IEC 14496-12. As defined in ISO/IEC 14496-12, the sample entry is transformed and a Protection Scheme Information Box (‘sinf’) is added to the standard sample entry in the Sample Description Box to denote that a stream is protected. The Protection Scheme Information Box SHALL contain a Scheme Type Box (‘schm’) so that the scheme is identifiable. The Scheme Type Box SHALL have the following additional constraints:

— the `scheme_type` field SHALL be set to a value equal to a four character code defined in Clause 10;

— the `scheme_version` field SHALL be set to 0x00010000 (Major version 1, Minor version 0).

The Protection Scheme Information Box SHALL also contain a Scheme Information Box (‘schi’). The Scheme Information Box SHALL contain a Track Encryption Box (‘tenc’), describing the default encryption parameters for the track.

4.2 Common encryption scheme types

Four protection schemes are specified in this edition of Common Encryption. Each scheme uses syntax and algorithms specified in Clause 5 to Clause 9, as constrained in Clause 10. They are the following:

a) ‘cenc’ – AES-CTR mode full sample and video NAL Subsample encryption, see 10.1;

b) ‘cbc1’ – AES-CBC mode full sample and video NAL Subsample encryption, see 10.2;

c) ‘cens’ – AES-CTR mode partial video NAL pattern encryption, see 10.3;

d) ‘cbcs’ – AES-CBC mode partial video NAL pattern encryption, see 10.4.

5 Overview of encryption metadata

The encryption metadata defined by Common Encryption can be categorized as follows.

— Protection System Specific Data – this data is opaque to Common Encryption. This gives protection systems (i.e. key and digital rights management “DRM” systems) a place to store their own data using a common mechanism. This data is contained in the `ProtectionSystemSpecificHeaderBox` described in 8.1.

— Common encryption information for a media track – this includes default values for the key identifier (KID), Initialization Vector and vector size, protection pattern, and protection flag. This data is contained in the `TrackEncryptionBox` described in 8.2.

— Common encryption information for groups of media samples – this includes overrides to the track level defaults defined above. This allows groups of samples within the track to use different keys, a mix of clear and protected content, share a Constant Initialization Vector (for some schemes), etc. This data is contained in a `SampleGroupDescriptionBox` (‘sgpd’) that is referenced by a `SampleToGroupBox` (‘sbgp’). See Clause 6 for further details.

— Encryption information for individual media samples – this includes Initialization Vectors and Subsample encryption data. This data is sample auxiliary information, referenced by using a Sample Auxiliary Information Sizes Box (‘saiz’) and a Sample Auxiliary Information Offsets Box (‘saio’). See Clause 7 for further details.

6 Encryption parameters shared by groups of samples

Each sample in a protected track SHALL be associated with an `isProtected` flag, `Per_Sample_IV_Size`, `KID`, optional Block pattern information, and an optional `constant_IV`. This can be

accomplished by relying on the default values in the Track Encryption Box (‘tenc’) (see 8.2), and optionally specifying parameters by sample group. Encryption parameters specified in a sample group SHALL override the corresponding default parameter values for the samples in that group defined in the Track Encryption Box. Samples not mapped to any sample group SHALL use the defaults established in the Track Encryption Box.

When specifying the parameters by sample group, the Sample To Group Box (‘sbgp’) in the sample table or track fragment specifies which samples use which sample group description from the Sample Group Description Box (‘sgpd’). The format of the sample group description is uniform across all track types (as indicated by the handler type for the track). For fragmented files, it may be necessary to store both the Sample To Group Box and Sample Group Description Box in each track fragment to make them accessible for decryption of the samples they describe, e.g. when movie fragments are separately stored and delivered by streaming.

Tracks of all types SHALL use the CencSampleEncryptionInformationGroupEntry sample group description structure, which has the following syntax.

```
aligned(8) class CencSampleEncryptionInformationGroupEntry
   extends SampleGroupEntry( ‘seig’ )
{
   unsigned int(8)         reserved = 0;
   unsigned int(4)         crypt_byte_block = 0;
   unsigned int(4)         skip_byte_block = 0;
   unsigned int(8)         isProtected;
   unsigned int(8)         Per_Sample_IV_Size;
   unsigned int(8)[16]     KID;
   if (isProtected ==1 && Per_Sample_IV_Size == 0) {
      unsigned int(8)      constant_IV_size;
      unsigned int(8)[constant_IV_size] constant_IV;
   }
}
```


These structures use a common semantic for their fields as follows:

— isProtected is the flag which indicates the encryption state of the samples in the sample group. See the isProtected field in 9.1 for further details.

— Per_Sample_IV_Size is the Initialization Vector size in bytes for samples in the sample group. See the Per_Sample_IV_Size field in 9.1 for further details.

— KID is the key identifier used for samples in the sample group. See the KID field in 9.1 for further details.

— constant_IV_size is the size of a possible Initialization Vector used for all samples associated with this group (when per-sample Initialization Vectors are not used).

— constant_IV, if present, is the Initialization Vector used for all samples associated with this group. See the constant_IV field in 9.1 for further details.

— crypt_byte_block specifies the count of the encrypted Blocks in the protection pattern, where each Block is of size 16-bytes. See 9.1 for further details.

— skip_byte_block specifies the count of the unencrypted Blocks in the protection pattern. See 9.1 for further details.

In order to facilitate the addition of future optional fields, clients SHALL ignore additional bytes after the fields defined in the CencSampleEncryption group entry structures.

7 Common encryption sample auxiliary information

7.1 Definition

Each protected sample in a protected track SHALL have an Initialization Vector associated with it. Both Initialization Vectors and Subsample encryption information MAY be provided as Sample Auxiliary Information with `aux_info_type` equal to the scheme and `aux_info_type_parameter` equal to 0.

For example, for tracks protected using the 'cenc' scheme, the default value for `aux_info_type` is 'cenc' and the default value for the `aux_info_type_parameter` is 0, so content SHOULD be created omitting these optional fields. Storage of sample auxiliary information SHALL conform to ISO/IEC 14496-12.

The format of the sample auxiliary information for samples with this type SHALL be as follows:

```
aligned(8) class CencSampleAuxiliaryDataFormat
{
   unsigned int(Per_Sample_IV_Size*8) InitializationVector;
   if (sample_info_size > Per_Sample_IV_Size )
   {
      unsigned int(16) subsample_count;
      {
        unsigned int(16) BytesOfClearData;
        unsigned int(32) BytesOfProtectedData;
      } [subsample_count ]
   }
}
```

where

`sample_info_size` is the size of the sample auxiliary information for this sample from the Sample Auxiliary Information Size Box ('saiz');

`InitializationVector` is the Initialization Vector for the sample, unless a `constant_IV` is present in the Track Encryption Box ('tenc') (see the `InitializationVector` field in 9.1 for further details);

`subsample_count` is the count of Subsamples for this sample (see the `subsample_count` field in 9.1 for further details);

`BytesOfClearData` is the number of bytes of clear data in this Subsample (see the `BytesOfClearData` field in 9.1 for further details);

`BytesOfProtectedData` is the number of bytes of protected data in this Subsample (see the `BytesOfProtectedData` field in 9.1 for further details.

If Subsample encryption is not used (the size of the sample auxiliary information equals `Per_Sample_IV_Size`), then the entire sample is protected (see 9.4 for further details). In this case, all auxiliary information will have the same size and hence the `default_sample_info_size` of the Sample Auxiliary Information Sizes box ('saiz') will be equal to the `Per_Sample_IV_Size` of the Initialization Vectors. If `Per_Sample_IV_Size` is also zero (because constant IVs are in use) then the sample auxiliary information would then be empty and should be omitted.

NOTE Even if Subsample encryption is used, the size of the sample auxiliary information may be the same for all of the samples (if all of the samples have the same number of Subsamples) and the `default_sample_info_size` may be used.