

Recommendation

ITU-T J.1041 (03/2025)

SERIES J: Cable networks and transmission of television, sound programme and other multimedia signals

Conditional access and protection – Digital rights management for video and audio content distribution

Digital rights management for video and audio content distribution – System architecture

ITU-T J-SERIES RECOMMENDATIONS

Cable networks and transmission of television, sound programme and other multimedia signals

GENERAL RECOMMENDATIONS	J.1-J.9
GENERAL SPECIFICATIONS FOR ANALOGUE SOUND-PROGRAMME TRANSMISSION	J.10-J.19
PERFORMANCE CHARACTERISTICS OF ANALOGUE SOUND-PROGRAMME CIRCUITS	J.20-J.29
EQUIPMENT AND LINES USED FOR ANALOGUE SOUND-PROGRAMME CIRCUITS	J.30-J.39
DIGITAL ENCODERS FOR ANALOGUE SOUND-PROGRAMME SIGNALS – PART 1	J.40-J.49
DIGITAL TRANSMISSION OF SOUND-PROGRAMME SIGNALS	J.50-J.59
CIRCUITS FOR ANALOGUE TELEVISION TRANSMISSION	J.60-J.69
ANALOGUE TELEVISION TRANSMISSION OVER METALLIC LINES AND INTERCONNECTION WITH RADIO-RELAY LINKS	J.70-J.79
DIGITAL TRANSMISSION OF TELEVISION SIGNALS	J.80-J.89
ANCILLARY DIGITAL SERVICES FOR TELEVISION TRANSMISSION	J.90-J.99
OPERATIONAL REQUIREMENTS AND METHODS FOR TELEVISION TRANSMISSION	J.100-J.109
INTERACTIVE SYSTEMS FOR DIGITAL TELEVISION DISTRIBUTION (DOCSIS FIRST AND SECOND GENERATIONS)	J.110-J.129
TRANSPORT OF MPEG-2 SIGNALS ON PACKETIZED NETWORKS	J.130-J.139
MEASUREMENT OF THE QUALITY OF SERVICE – PART 1	J.140-J.149
DIGITAL TELEVISION DISTRIBUTION THROUGH LOCAL SUBSCRIBER NETWORKS	J.150-J.159
IPCABLECOM (MGCP-BASED) – PART 1	J.160-J.179
DIGITAL TRANSMISSION OF TELEVISION SIGNALS – PART 1	J.180-J.189
CABLE MODEMS AND HOME NETWORKING	J.190-J.199
APPLICATION FOR INTERACTIVE DIGITAL TELEVISION – PART 1	J.200-J.209
INTERACTIVE SYSTEMS FOR DIGITAL TELEVISION DISTRIBUTION (DOCSIS THIRD TO FIFTH GENERATIONS)	J.210-J.229
MULTI-DEVICE SYSTEMS FOR CABLE TELEVISION	J.230-J.239
MEASUREMENT OF THE QUALITY OF SERVICE – PART 2	J.240-J.249
DIGITAL TELEVISION DISTRIBUTION THROUGH LOCAL SUBSCRIBER NETWORKS	J.250-J.259
IPCABLECOM (MGCP-BASED) – PART 2	J.260-J.279
DIGITAL TRANSMISSION OF TELEVISION SIGNALS – PART 2	J.280-J.289
CABLE SET-TOP BOX	J.290-J.299
APPLICATION FOR INTERACTIVE DIGITAL TELEVISION – PART 2	J.300-J.309
MEASUREMENT OF THE QUALITY OF SERVICE – PART 3	J.340-J.349
IPCABLECOM2 (SIP-BASED) – PART 1	J.360-J.379
DIGITAL TRANSMISSION OF TELEVISION SIGNALS – PART 3	J.380-J.389
MEASUREMENT OF THE QUALITY OF SERVICE – PART 4	J.440-J.449
IPCABLECOM2 (SIP-BASED) – PART 2	J.460-J.479
DIGITAL TRANSMISSION OF TELEVISION SIGNALS – PART 4	J.480-J.489
TRANSPORT OF LARGE SCREEN DIGITAL IMAGERY	J.600-J.699
SECONDARY DISTRIBUTION OF IPTV SERVICES	J.700-J.799
MULTIMEDIA OVER IP IN CABLE	J.800-J.899
TRANSMISSION OF 3-D TV SERVICES	J.900-J.999
CONDITIONAL ACCESS AND PROTECTION	J.1000-J.1099
Renewable conditional access system	J.1000-J.1004
Digital rights management for cable television multiscreen service	J.1005-J.1009
Exchangeable embedded conditional access and digital rights management solutions	J.1010-J.1019
Downloadable system for multi-CA/DRM service of mobile broadcasting	J.1020-J.1025
Downloadable conditional access system for unidirectional networks	J.1026-J.1029
Downloadable conditional access system for bidirectional networks	J.1031-J.1035
Factual subscriber-base reporting and protected content delivery in conditional access systems	J.1036-J.1039
Digital rights management for video and audio content distribution	J.1040-J.1044
SWITCHED DIGITAL VIDEO OVER CABLE NETWORKS	J.1100-J.1119
SMART TV OPERATING SYSTEM	J.1200-J.1209
IP VIDEO BROADCAST	J.1210-J.1219
CABLE SET-TOP BOX – PART 2	J.1290-J.1299
CLOUD-BASED CONVERGED MEDIA SERVICES FOR IP AND BROADCAST CABLE TELEVISION	J.1300-J.1309
CLOUD-BASED SERVICES FOR IP DELIVERY OVER INTEGRATED BROADBAND CABLE NETWORK	J.1310-J.1319
TELEVISION TRANSPORT NETWORK AND SYSTEM DEPLOYMENT IN DEVELOPING COUNTRIES	J.1400-J.1409
ARTIFICIAL INTELLIGENCE (AI) ASSISTED CABLE NETWORKS	J.1600-J.1649

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T J.1041

Digital rights management for video and audio content distribution – System architecture

Summary

Recommendation ITU-T J.1041 is Part 2 of a multi-part deliverable composed of four parts. Part 2 describes the system architecture of digital rights management (DRM) for video and audio content distribution (DRMVACD), including system functional architecture, content encryption mechanism, license format, license acquisition protocol, and DRM server-side architecture, and certificate format.

The service requirements for the application of new generation video and audio technologies, along with the new video service models coupled with a diverse range of consumer devices used for the reception of content, bring new demands and opportunities to the video industry, especially broadcast and television video services, and raise higher requirements for DRM, described in Recommendation ITU-T J.1040. The DRM system not only needs to have well-defined security levels to meet the content protection requirements of different quality and different distribution window periods, but also needs to have sufficient platform compatibility to cover various terminal devices. In addition, higher encryption/decryption efficiency and encryption/decryption modes decoupled from content encapsulation are required to match the rapid development and transformation of video and audio technologies. In addition, an interoperability mechanism is required to achieve a more open DRM ecosystem.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T J.1041	2025-03-16	21	11.1002/1000/16262

Keywords

Digital rights management, DRM, system architecture, video and audio content distribution, video and audio content protection.

* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, and information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at <https://www.itu.int/ITU-T/ipr/>.

© ITU 2025

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

	Page
1 Scope.....	1
2 References.....	1
3 Definitions	2
3.1 Terms defined elsewhere	2
3.2 Terms defined in this Recommendation.....	2
4 Abbreviations and acronyms	2
5 Conventions	3
6 Architecture	4
6.1 Functional architecture	4
6.2 Content authorization	4
6.3 Key management	6
6.4 Security mechanisms	7
6.5 Trust model.....	7
7 Content encryption.....	8
7.1 Content encryption methods.....	8
7.2 Content encapsulation format.....	12
8 License format	16
8.1 License structure.....	16
8.2 License encoding	17
9 License acquisition protocol	24
9.1 License acquisition request.....	25
9.2 License acquisition response	26
9.3 Status information	29
9.4 Message signature mechanism	29
10 DRM server side	30
10.1 Key synchronization protocol.....	30
10.2 Key query protocol	33
Annex A – Formats of digital certificate, online certificate status protocol and certificate revocation list.....	37
A.1 Digital certificate format	37
A.2 Online certificate status protocol.....	47
A.3 Certificate revocation list (CRL)	49
Annex B – H.264/H.265 content encryption method.....	50
B.1 H.264	50
B.2 H.265	51
B.3 TS encapsulation method extension	51
B.4 Emulation prevention	51

	Page
Annex C – Cryptographic algorithm.....	52
Bibliography.....	54

Introduction

This Recommendation document is Part 2 of a multi-part deliverable composed of four parts. Part 2 describes the system architecture of digital rights management (DRM) for video and audio content distribution, including system functional architecture, content encryption mechanism, license format, license acquisition protocol, and DRM server-side architecture, and certificate format. The other parts are identified below:

Part 1: Requirements;

Part 2: System architecture;

Part 3: Client;

Part 4: Compliance and robustness rules.

Recommendation ITU-T J.1041

Digital rights management for video and audio content distribution – System architecture

1 Scope

This Recommendation is Part 2 of a multi-part deliverable composed of four parts. Part 2 describes the system architecture of digital rights management (DRM) for video and audio content distribution, including system functional architecture, content encryption mechanism, license format, license acquisition protocol, and DRM server-side architecture, and certificate format. The DRM requirements are described in [ITU-T J.1040].

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

- [[ITU-T J.1040](#)] Recommendation ITU-T J.1040 (2024), *Digital rights management for video and audio content distribution – Requirements*.
- [[ITU-T J.1042](#)] Recommendation ITU-T J.1042 (2025), *Digital rights management for video and audio content distribution – Client*.
- [ECMA 404] ECMA 404 (2017), *The JSON data interchange format*. <https://ecma-international.org/publications-and-standards/standards/ecma-404/>.
[Accessed 4 June 2025]
- [ISO/IEC 10118-3:2018] ISO/IEC 10118-3:2018, *IT Security techniques – Hash-functions – Part 3: Dedicated hash-functions*.
- [ISO/IEC 14496-12:2015] ISO/IEC 14496-12:2015, *Information technology – Coding of audio-visual objects – Part 12: ISO base media file format*.
- [ISO/IEC 14888-3:2018] ISO/IEC 14888-3:2018, *IT Security techniques – Digital signatures with appendix – Part 3: Discrete logarithm based mechanisms*.
- [ISO/IEC 18033-3/AMD1] ISO/IEC 18033-3:2010/AMD1:2021, *Information technology – Security techniques – Encryption algorithms – Part 3: Block ciphers – Amendment 1: SM4*.
- [ISO/IEC 23001-7:2016] ISO/IEC 23001-7:2016, *Information technology – MPEG systems technologies – Part 7: Common encryption in ISO base media file format files*.
- [ISO/IEC 23009-4:2013] ISO/IEC 23009-4:2013, *Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 4: Segment encryption and authentication*.
- [IEEE 1857-2023] IEEE 1857-2023, *IEEE Standard for Advanced Audio and Video Coding*.

[IEEE 1857.4-2018]	IEEE 1857.4-2018, <i>IEEE Standard for Second-Generation IEEE 1857 Video Coding</i> .
[IETF RFC 2045]	IETF RFC 2045, <i>Multipurpose internet mail extensions – Part 1: Format of internet message bodies</i> .
[IETF RFC 2104]	IETF RFC 2104, <i>HMAC:Keyed-Hashing for Message Authentication</i> .
[IETF RFC 5280]	IETF RFC 5280, <i>Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile</i> .
[IETF RFC 6960]	IETF RFC 6960, <i>X.509 Internet Public Key Infrastructure Online Certificate Status Protocol – OCSP</i> .

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 digital rights management (DRM) client [ITU-T J.1040]: A trusted entity in the digital media content playback device responsible for execution of DRM content-related permissions and restrictions.

3.1.2 digital rights management (DRM) content [ITU-T J.1040]: Digital media content managed by DRM technology.

3.1.3 digital rights management (DRM) server [ITU-T J.1040]: The entity that provides the license service to the DRM client.

3.1.4 encryption [b-ITU-T X.1367]: The cryptographic transformation of data to produce ciphertext.

3.1.5 decryption [b-ITU-T Y.101]: The decoding of encrypted information.

3.1.6 digital signature [b-ITU-T X.800]: Data appended to, or a cryptographic transformation (see cryptography) of a data unit that allows a recipient of the data unit to prove the source and integrity of the data unit and protect against forgery, e.g., by the recipient.

3.1.7 license [ITU-T J.1040]: A description of control information such as access permissions, usage rules, and keys for digital media content.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

CA	Certification Authority
CBC	Cipher Block Chain
CEI	Content Encryption Information
CEK	Content Encryption Key
CENC	Common Encryption
CRL	Certificate Revocation List
CTR	Counter Mode

DASH	Dynamic Adaptive Streaming over HTTP
DER	Distinguished Encoding Rules
DRM	Digital Rights Management
DRMVACD	DRM for Video and Audio Content Distribution
HLS	HTTP Live Streaming
HMAC	Hash-based Message Authentication Code
HTTP	Hyper Text Transport Protocol
IPTV	Internet Protocol Television
IV	Initialization Vector
JSON	JavaScript Object Notation
KID	Key Identifier
MPD	Media Presentation Description
NAL	Network Abstract Layer
OCSP	Online Certificate Status Protocol
PKI	Public Key Infrastructure
PMT	Program Map Table
SEI	Supplemental Enhancement Information
TS	Transport Stream
uimsbf	unsigned integer, most significant bit first
URI	Uniform Resource Identifier
UTC	Universal Time Coordinated
UUID	Universally Unique Identifier
VoD	Video on Demand

5 Conventions

In this Recommendation:

The keywords "**is required to**" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The keywords "**is recommended**" indicate a requirement which is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

The keywords "**is prohibited from**" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this Recommendation is to be claimed.

The keywords "**can optionally**" indicate an optional requirement which is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option, and that the feature can be optionally enabled by the network operator/service provider. Rather, it means that the vendor may optionally provide the feature and still claim conformance with the specification.

In this Recommendation, the words shall, shall not, should and may sometimes appear, in which case they are to be interpreted, respectively, as is required to, is prohibited from, is recommended, and can

optionally. The appearance of such phrases or keywords in an appendix or in material explicitly marked as informative are to be interpreted as having no normative intent.

6 Architecture

Based on cryptography, public key infrastructure (PKI) technology and authorization technology, this Recommendation defines the end-to-end functional architecture, content authorization, key management, security mechanisms, and trust model for digital rights management of video and audio content distribution. The functional architecture, technical mechanisms, basic formats, and protocols defined in this Recommendation can be used to construct an end-to-end digital rights management system for video and audio content distribution.

6.1 Functional architecture

The digital rights management system for video and audio content distribution is logically divided into two parts – digital rights management (DRM) server side and DRM client side, as shown in Figure 6-1. DRM server-side system includes core modules such as content encryption, key management, key gateway, and content authorization. The content encryption module uses the content encryption keys (CEKs) to encrypt the video and audio content. The key management module is responsible for synchronizing CEKs to the key gateway module after the generation of CEKs. The key gateway module stores the CEKs confidentially after receiving the synchronized keys and serves the key query of the content authorization module. The content authorization module generates license responses with a license containing the CEKs and key usage rules that are securely sent to the trusted DRM client defined in [ITU-T J.1042]. After the DRM client receives the license, it decrypts the CEKs according to the key usage rule, and decrypts the content with the CEKs to play.

DRM server-side modules and DRM client establish a trust relationship based on PKI technology and perform secure communication with each other based on this trust relationship.

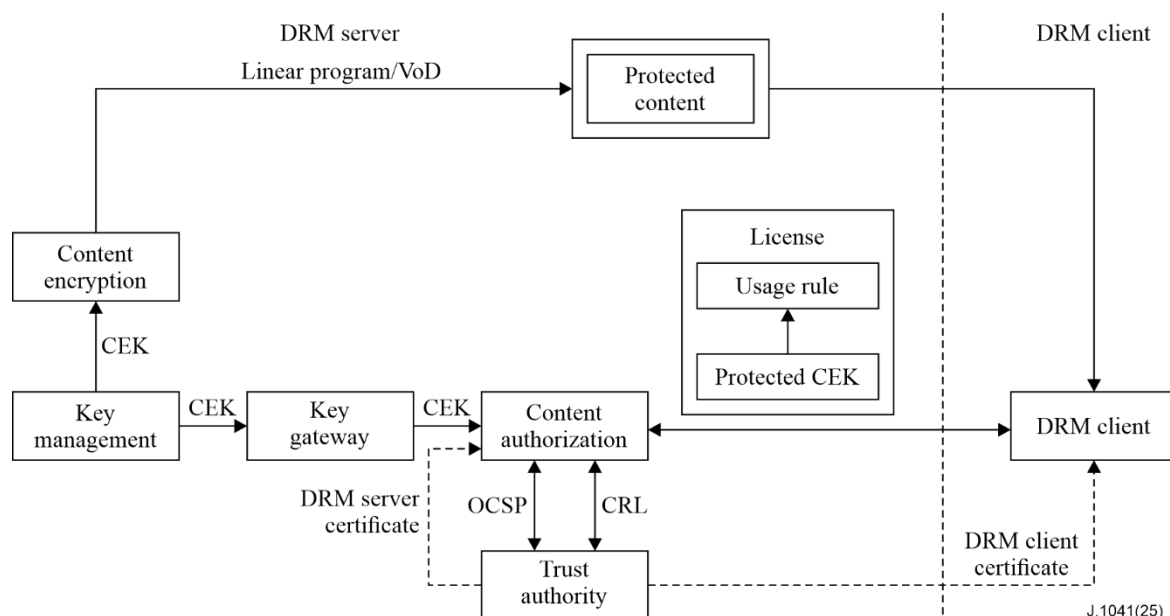


Figure 6-1 – Functional architecture of digital rights management for video and audio content distribution (DRMVACD)

6.2 Content authorization

The digital rights management system for video and audio content distribution (DRMVACD) performs content authorization based on the mechanism of associating hierarchical keys with key usage rules. Logically, the keys can be divided into multiple layers based on the sequence of

encryption, and the key that encrypts the current key is called the upper key. Each key has its own key usage rules. The current key can only be decrypted under the conditions specified in the upper key's usage rules. Each key may have multiple key usage rules, the key can only be used if it satisfies all of its usage rules. The mechanism of associating keys with key usage rules is shown in Figure 6-2.

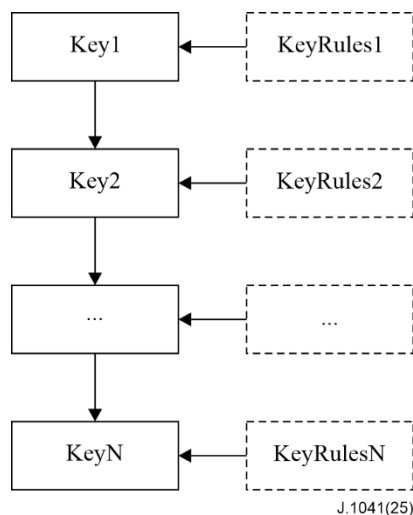


Figure 6-2 – The associating mechanism of keys and key usage rules

Key usage rules typically include start time, end time, time period, number of uses, cumulative time period, output rules, and client security level. Different types of keys may include special key usage rules.

The start time indicates that the key is only allowed to be used after the specified time.

The end time indicates that the key is not allowed to be used after the specified time.

A time period indicates that the key is only allowed to be used for a period of time after the first use.

"Number of uses" indicates the number of times the key can be used, counting 1 use when successfully using this key to complete a decryption.

The cumulative time period is the accumulation of the client player time from start to stop of a playback. If the accumulated time from multiple playback sessions is longer than the cumulative time period, the key usage is no longer allowed. The cumulative time period is typically used in a preview scenario, for example consider a video on demand (VoD) program with a total length of 90 minutes, then a cumulative time period of 600 seconds would only permit playback of up to 10 minutes from any scenes in the program.

The output rule specifies whether the decoded video and audio content is allowed to output to other devices, as well as the allowed output scope and mode. "No output rule" means no limitation on the output.

Client security level indicates that using CEK to decrypt content is only allowed on DRM clients at a specific security level. There are three defined client security levels: software security level, hardware security level, and enhanced hardware security level. Client security level information is included in DRM client certificate. A CEK can be used to decrypt content for playback or output only if the security level in the DRM client certificate is equal to or higher than the client security level specified in the key usage rules.

If there is no key usage rule, there is no restriction on the key usage.

The digital rights management system for video and audio content distribution may contain multiple types of keys, such as CEKs, session keys, hash-based message authentication code keys and device keys.

The content encryption key is the key that encrypts the digital media content. Each content item may have multiple content encryption keys. Key usage rules for CEK include: start time, end time, time period, number of uses, cumulative time period, output rules, and client security level.

A session key is a temporary key that is generated by the server when the client requests a license, which is used in the license to protect CEK.

A hash-based message authentication code key is a temporary key that is used to generate a hash-based message authentication code to protect the integrity of the license.

A device key refers to the key pair of the DRM client. The device key shall be an asymmetric key such that data encrypted with the device public key can only be decrypted by the DRM client.

The DRM server uses the CEK to encrypt the digital media content. It encrypts the CEK and other content-related keys (such as session key, hash-based message authentication code key, etc.) by using the hierarchical key encryption method, then packages these encrypted keys together with their key usage rules into a license and sends it to the DRM client. The session key and hash-based message authentication code key are encrypted using the device public key. The DRM client uses the keys according to the usage rules of the keys at different levels in the hierarchical key system to decrypt and play the digital media content. The content authorization mechanism of the digital rights management system for video and audio content distribution is shown in Figure 6-3.

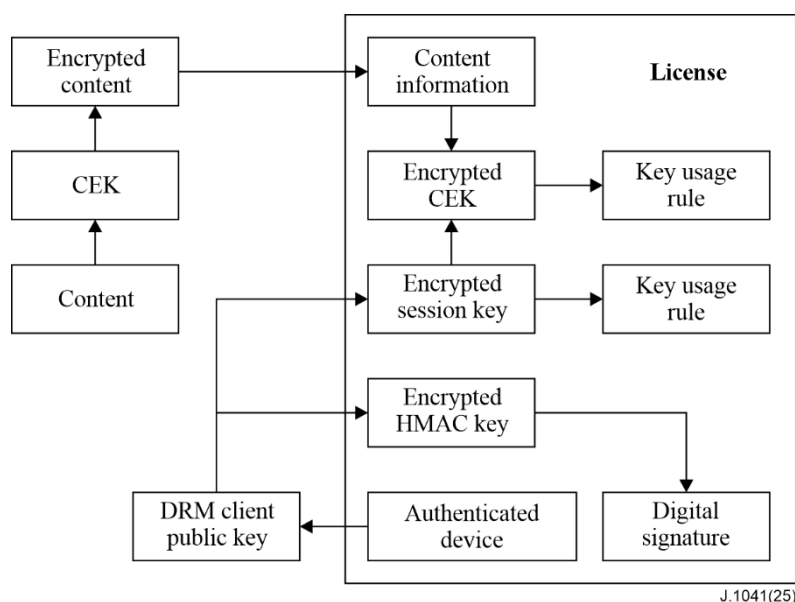


Figure 6-3 – Content authorization mechanism

6.3 Key management

DRM server and DRM client have their own asymmetric key pairs, based on an asymmetric cryptographic algorithm for license acquisition. The DRM server uses a symmetric cryptographic algorithm for content encryption, then encrypts the CEK and sends it to the DRM client in the license. Key management mechanism is shown in Figure 6-4.

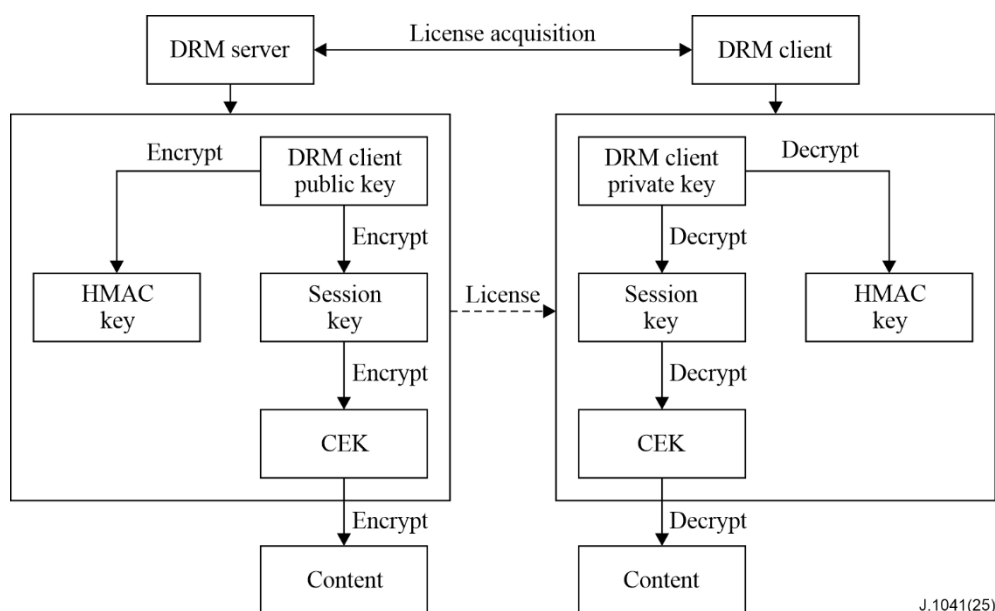


Figure 6-4 – Key Management mechanism

Certificates of the DRM server and the DRM client contain their own public keys. They are responsible for securing their own certificate and private keys separately.

The video and audio content is encrypted using the CEK. The DRM server generates a session key which is used to encrypt the CEK and then encrypts the session key with the DRM client public key. The DRM server encapsulates the encrypted session key and the encrypted content encryption key into a license and sends them to the DRM client. The license uses the hash-based message authentication code to guarantee its integrity, and the encrypted hash-based message authentication code key is encapsulated in the license.

After the DRM client receives the license, it decrypts the session key with the DRM client private key, then decrypts the CEK with this session key. Thus, the DRM client can decrypt the content with this CEK for decoding and playback.

6.4 Security mechanisms

The security mechanism of digital rights management for video and audio content distribution is specified as follows:

Data confidentiality

Confidentiality should protect sensitive data through encryption. Sensitive data should include at least the protected content and the CEKs;

Identity authentication

Identity authentication should be achieved by verifying the validity of each other's digital certificates between the DRM client and the DRM server;

Data integrity

Data integrity should be achieved by verifying the digital signature in the protocol message and the hash-based message authentication code in the license.

6.5 Trust model

The trust model of digital rights management for video and audio content distribution is based on the PKI system. In the DRM system, the core components of the DRM server side and the DRM client all apply a digital certificate from the trusted authority as the credential of their own identity. The

trust relationship between them is based on the validity of their digital certificates. If the DRM client's certificate is verified by the DRM server, the DRM server trusts the DRM client.

A digital certificate is the foundation of the DRM system to build a trust relationship. The DRM client should be associated with a digital certificate. Each DRM client should carry at least one digital certificate. Each DRM client must have a unique identifier which should be loaded from the digital certificate.

The trust chain of the digital rights management system for video and audio content distribution includes the root certification authority (CA) certificate, the DRM server sub-CA certificate, the DRM client sub-CA certificate, the DRM server certificate, the DRM client certificate, and the online certificate status protocol (OCSP) server certificate. The security trust mechanism is shown in Figure 6-5.

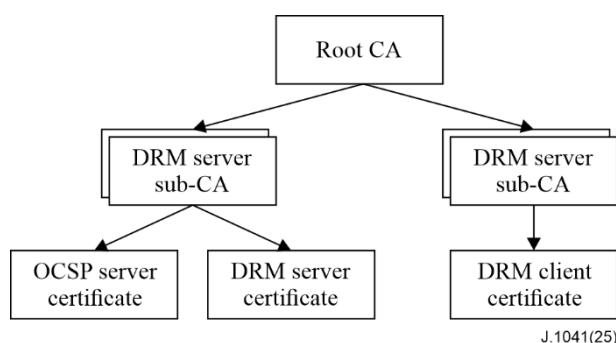


Figure 6-5 – Security trust mechanism

After the trust chain is established, the DRM server-side system securely stores the DRM server certificate and its private key, the OCSP server certificate, the DRM server sub-CA certificate, and the root CA certificate, and the DRM client system securely stores the DRM client certificate and its private key, the DRM client sub-CA certificate, and the root CA certificate.

The DRM server-side system determines the validity of the DRM client certificate based on the DRM client certificate revocation list (CRL) list, and the DRM client determines the validity of the DRM server-side certificate based on the OCSP response.

Annex A defines the digital certificates format, online certificate status protocol, and certificate revocation list format in digital rights management for video and audio content distribution.

7 Content encryption

This clause defines content encryption methods and encrypted content encapsulation formats for digital rights management used in the distribution of video and audio content. The encrypted content should contain the content identifier and the necessary information to obtain a license. Depending on the scenarios, different encrypted content encapsulation formats are defined.

7.1 Content encryption methods

7.1.1 Content encryption information

The DRM system shall encrypt the elementary stream of video content, and should add content encryption information (CEI) to the extended data of the elementary stream. The CEI is used to indicate how the subsequent video is encrypted. CEI includes the encryption flag, the current CEK identifier, the next CEK identifier, and the initialization vector. Before the next CEI appears, all video data is encrypted in the manner specified by clause 7.1.2, and the data that needs to be encrypted in each video frame uses the initialization vector in the current CEI. The CEI data syntax format is shown in Table 7-1.

Table 7-1 – CEI Data Syntax Format

Grammar	Number of bits
CEI_DATA () {	
encryption_flag	1
next_key_id_flag	1
Reserved	6
if (encryption_flag == 1) {	
current_key_id	128
}	
if (next_key_id_flag == 1) {	
next_key_id	128
}	
IV_length	8
IV	IV_length × 8
}	

encryption_flag: indicates whether the video data after this extended data is encrypted.

next_key_id_flag: indicates whether the CEI data contains the next key identifier.

current_key_id: indicates the key identifier used by subsequent video encryption.

next_key_id: indicates the next key identifier that video encryption will use. This allows the DRM client to request the key in advance from the server before the key is changed.

IV_length: The length of initialization vector of the encryption algorithm.

Initialization vector (IV): The initialization vector of the encryption algorithm.

7.1.2 Content encryption mechanism

For the video codec format specified by [IEEE 1857-2023] and [IEEE 1857.4-2018], CEI data is placed in the extension_data after sequence_header. The following uses the format of [IEEE 1857.4-2018] as an example, and the position of CEI is as follow:

```

extension_data (i) {
    while (next_bits (32) == extension_start_code) {
        extension_start_code
        if (i == 0) { /* after sequence header */
            if (next_bits (4) == ' 1101 ') /* CEI */
                CEI ()
        }
    }
}

```

The CEI syntax is shown in Table 7-2.

Table 7-2 – CEI syntax

Grammar	Number of bits	Note
CEI () {		
extension_id	4	' 1101 '
Reserved	4	
CEI_DATA ()		See Table 7-1
}		

The encryption for video content encoded in [IEEE 1857.4-2018] involves adding extended information that contains CEI data in extension_and_user_data (0), encrypting the slice in each frame, and retaining only the header information of the first slice after encryption. The encryption of each frame in the same video sequence uses the same initialization vector, that is, the initialization vector contained in the CEI extension information in the video sequence. [IEEE 1857.4-2018] specifies the following encryption syntax for video content:

```

video_sequence{
do{
sequence_header ()
extension_and_user_data (0)
do{
if (next_bits(32)) == intra_picture_start_code)
intra_picture_header ()
else
inter_picture_header ()
extension_and_user_data (1)
picture_data () //retains the first slice () header and the rest is encrypted.
}while (next_bits(32)) ==inter_picture_start_code| | next_bits 32() ==intra_picture_start_code)
}while (next_bits(32)) != video_sequence_end_code && next_bits (32) != video_edit_code)
if (next_bits(32)) == video_sequence_end_code)
video_sequence_end_code
if (next_bits(32)) == video_edit_code)
video_edit_code
}

```

The packaging of encrypted video content encoded in [IEEE 1857-2023] and [IEEE 1857.4-2018] is shown in Figure 7-1.

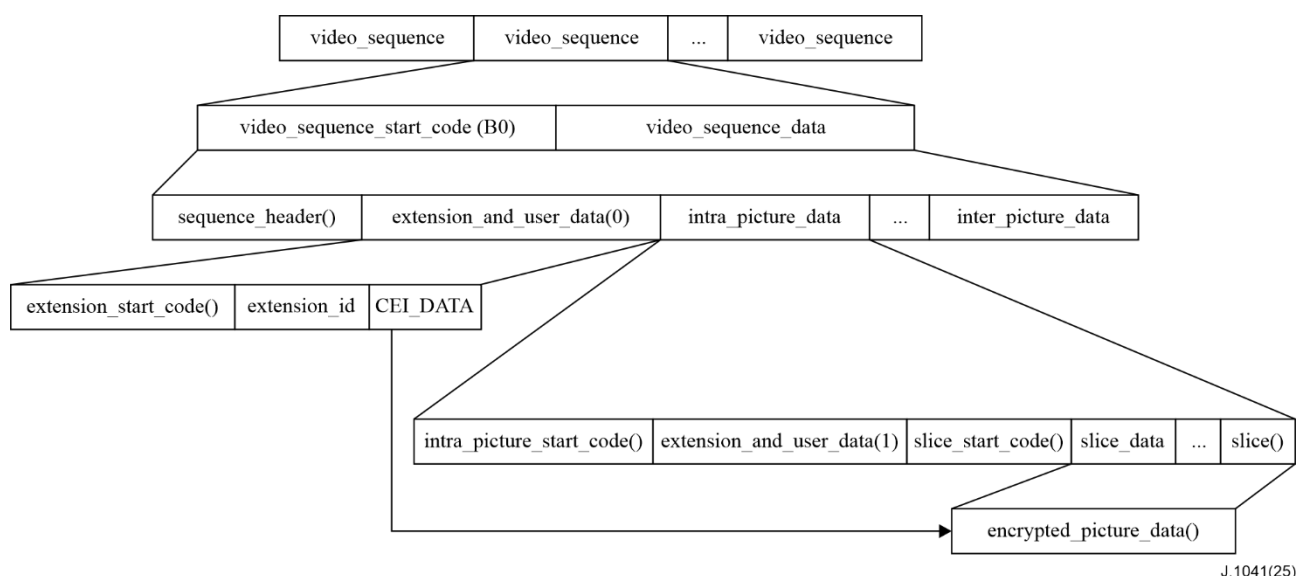


Figure 7-1 – Encrypted Content Packaging diagram

Content encryption is classified into full encryption mode and partial encryption mode.

In full encryption mode, the data part of the slice is encrypted, and last part that is less than or equal to 16 bytes is not encrypted. Encryption algorithm is SM4, as defined in [ISO/IEC 18033-3/AMD1]. Encryption mode is cipher block chain (CBC) or counter mode (CTR). The syntax is as follows:

```
encrypted_slice_data ()
{
  while (bytes_remaining () > 16)
  {
    protected_block // 16 bytes * n
  }
  unencrypted_trailer //1-16 bytes
}
```

Partial encryption mode encrypts 10% of the slice data, that is, after encrypting a 16-byte block, the following nine 16-byte blocks are not encrypted, and the last part that is less than or equal to 16 bytes is not encrypted. Encryption algorithm is SM4. Encryption mode is CBC or CTR. The syntax is as follows:

```
encrypted_slice_data ()
{
  while (Bytes_remaining () > 0)
  {
    if (bytes_remaining () > 16) {
      protected_block //16 bytes
    }
    unencrypted_trailer ///MIN (144, bytes_remaining ()) bytes
  }
}
```

7.1.3 Emulation prevention

The following conversion rules of the CEI_DATA and encrypted coded data prevent the emulation of start codes:

00 00 00 replace with 00 00 03 00

00 00 01 replace with 00 00 03 01

00 00 02 replace with 00 00 03 02

00 00 03 replace with 00 00 03 03

Before decryption, it is necessary to restore the CEI_DATA and encrypted coded data and then decrypt them accordingly. That is:

00 00 03 00 replace with 00 00 00

00 00 03 01 replace with 00 00 01

00 00 03 02 replace with 00 00 02

00 00 03 03 replace with 00 00 03

For the content encoded by H.264/H.265 video codec format, the content encryption method is specified in Annex B.

7.2 Content encapsulation format

7.2.1 TS encapsulation

Transport stream (TS) encapsulation refers to an encapsulation format of encrypted content when content is transmitted in a transport stream manner, such as digital television (TV) or Internet protocol television (IPTV). The distribution of TS-encapsulated content using HTTP live streaming (HLS) must comply with clause 7.2.4. To enable the client player to identify the encrypted stream and obtain DRM information in the encrypted stream, the DRM descriptor needs to be added to the program map table (PMT) table. The DRM descriptor in the PMT table includes encrypted content format, encryption method, license acquisition uniform resource identifier (URI) and so on. DRM descriptor syntax is shown in Table 7-3.

Table 7-3 – DRM descriptor syntax

Grammar	Number of bits
DRM_descriptor () {	
descriptor_tag	8
descriptor_length	8
video_format	4
video_encryption_method	4
audio_format	4
audio_encryption_method	4
DRM_data_bytes	(descriptor_length-2) × 8
}	

descriptor_tag: identifier for this descriptor containing the value 0xC0.

descriptor_length: length of the descriptor data.

video_format: indicates the video encoding format of this encrypted content specified in Table 7-4.