English Translation

IPTV STANDARD

VOD Specifications

IPTVFJ STD-0002 Version 1.1

Created on September 1, 2008	(Version 0.9)
Revised on November 28, 2008	(Version 1.0)
Revised on July 30, 2010	(Version 1.1)

IPTV Forum Japan

General Notes to the English Translation of IPTV Standards

3. The IPTV Standards are usually written in Japanese and approved by the IPTVFJ Technical Committee. This document is a translation into English of the approved document for the purpose of convenience of users. If there are any discrepancies in the content, expressions, etc., between the Japanese original and this translated document, the Japanese original shall prevail.

4. The establishment, revision and abolishment of IPTV Standards are approved at the IPTVFJ Technical Committee, which meets several times a year. Approved IPTV Standards, in their original language, are made publicly available through web posting, generally in about a few weeks after the date of approval. The original document of this translation may have been further revised and therefore users are encouraged to check the latest version at an appropriate page under the following URL:
   http://www.iptvforum.jp/en/

Table of Contents

Chapter 1 General Information

1.1 Introduction

This document is intended to define the specifications that are necessary for the supply of video-on-demand (VOD) services as part of the IPTV service—a video service utilizing an IP network.

Here, the term "VOD" refers to systems that deliver by streaming, on an on-demand basis, the video that the user wants to watch.

1.2 References

Table 1-1: References

Classification	Item	Standard
IPTV related		[1] IPTVFJ STD-0006 "IPTV Standards: CDN-scope Service Approach Specifications" [2] IPTVFJ STD-0007 "IPTV Standards: Internet-scope Service Approach Specifications" [3] "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group)
Digital broadcasting related		[4] ARIB STD-B1 "Digital Receiver For Digital Satellite Broadcasting Services Using Communication Satellites" [5] ARIB STD-B10 "Service Information for Digital Broadcasting System" [6] ARIB STD-B21 "Receiver for Digital Broadcasting" [7] ARIB STD-B24 "Data Coding and Transmission Specification for Digital Broadcasting" [8] ARIB STD-B32: "Video Coding, Audio Coding and Multiplexing Specifications for Digital Broadcasting" [9] ARIB TR-B14: "Operational Guidelines for Digital Terrestrial Television Broadcasting"
Monomedia	Video MPEG-2	[10] ISO/IEC 13818-2 ｜ ITU-T Rec. H.262
	Video H.264/MPEG-4 AVC	[11] ISO/IEC 14496-10 ｜ ITU-T Rec. H.264
	Audio MPEG2 AAC	[12] ISO/IEC 13818-7
	Audio MPEG1	[13] ISO/IEC 11172-3
	Caption	[14] ARIB STD-B24 "Data Coding and Transmission Specification for Digital Broadcasting"
Multiplex format	MPEG2 TS	[15] ISO/IEC 13818-1 ｜ ITU-T Rec. H.222.0
	TTS	[16] ARIB STD-B24 "Data Coding and Transmission Specification for Digital Broadcasting"
Encryption algorithm	AES	[17] FIPS publication 197 "Advanced Encryption Standard (AES)"
Streaming transmission/ control	RTP, RTCP	[18] RFC3550 "RTP: A Transport protocol for Real-Time Applications"
	RTSP	[19] RFC2326 "Real Time Streaming Protocol (RTSP)"
	SDP	[20] RFC4566 "SDP: Session Description Protocol"

Streaming transmission/file transmission	HTTP	[21] RFC2616 "Hypertext Transfer Protocol -- HTTP/1.1"
FEC	Pro-MPEG	[22] RFC2733 "An RTP Payload Format for Generic Forward Error Correction" [23] "Pro-MPEG Code of Practice #3 release 2" (Pro-MPEG Forum)
URI	URI	[24] RFC3986 "Uniform Resource Identifier (URI): Generic Syntax"
Playback control information transmission protocol	HTTP	[25] RFC2616 "Hypertext Transfer Protocol – HTTP/1.1"
Secure communication	SSL/TLS	[26] RFC2246 "The TLS Protocol, Version 1.0"
Communication protocol	UDP, TCP	[27] RFC768 "User Datagram Protocol" [28] RFC793 "Transmission Control Protocol"
	IP, ICMP	[29] RFC791 "Internet Protocol" [30] RFC792 "Internet Control Message Protocol"
	IPv6, ICMPv6	[31] RFC2460 "Internet Protocol Version 6 (IPv6) Specification" [32] RFC4443 "Internet Control Message Protocol (ICMPv6) for the Internet protocol Version 6 (IPv6) Specification" [33] RFC4291 "IP Version 6 Addressing Architecture"
DRM related	Marlin	[34] Marlin IPTV End-point Service Specification Version 1.0.2(or later version) [35] Marlin IPTV-ES Implementation Guidelines for VoD Version 1.2(or later version) [36] Marlin IPTV-ES/J Specific Compliance Rules for VoD Version 1.2(or later version) [37] Marlin Trust Management Document – for IPTV-ES Version 1.3(or later version)

1.3 Glossary

Table 1-2: Glossary

Terms	Description
3/1	This term refers to a multichannel stereo system that has 3 channels in the front and 1 channel in the rear. L (left), R (right) and C (center) channels are provided in the front and a monaural surround channel, MS, is provided in the rear.
3/2	This term refers to a multichannel stereo system that has 3 channels in the front and 2 channels in the rear. L, R and C channels are provided in the front and stereo surround channels, LS and RS, are provided in the rear.
5.1 channel	This term refers to a 3/2 multichannel stereo system with a low frequency enhancement (LFE) channel added. Sometimes described as 3/2 + LFE.
AAC	Advanced Audio Coding: AAC is an audio coding system that has been standardized by the International Standards Organization.
ADTS	Audio Data Transport Stream: ADTS is one of the audio stream data formats defined in MPEG-2 AAC.
AES	Advanced Encryption Standard: AES is the next-generation standard encryption system of the U.S. government that was approved in 2001 by the National Institute of Standards and Technology (NIST) of the U.S. Department of Commerce.
Aspect ratio	This term refers to the width-to-height ratio of the picture display area.
Audio mode	The format that is used to process audio signals including monaural, stereo, multi-channel stereo, dual audio and multi audio modes.
Basic registration	This term refers to the process of registering a user and the user's device with a service provider.
Best effort	The state or type of service provisioning in which Quality of Service (QoS) is not guaranteed when a communication network is used.
BML	Broadcast Markup Language: BML is an XML application language defined in ARIB STD-B24 Volume 2.
Caption	The service of superimposing video-related text on TV broadcast video.
Caption out-screen display function	This term refers to the function of displaying video at a reduced size to prevent the captions from overlapping the telop (television opaque projector) characters on the screen.
CAS/DRM client identifier	This term refers to an identifier that uniquely identifies a specific DRM client.
CBC	Cipher-Block Chaining: CBC is a block ciphering mode. In this mode, prior to the encryption of a specific block, an exclusive OR (XOR) operation is performed with the preceding ciphered block.
CDN	Content Delivery Network: CDN is a network designed for delivery of digital content.

CGMS	Copy Generation Management System: CGMS refers to the generation management information and system for copy control. Generation management is a method of content management using copy control information that divides content into three categories: "content that may be copied without limitation", "content that may be copied only for one generation" and "content that must not be copied".
CGMS-A	CGMS-A is a CGMS that is applicable to analog interfaces.
Communication network	A network that provides bidirectional transmission channels for receivers.
Content	This term refers to a collection of videos, audios, characters, and data, etc., which are intended to be played and viewed by users. In this document, the term refers to video content.
Content key (Kc)	Kc is the key for encrypting video content in VOD services. Each is unique to a specific video content.
Content playback control metafile	This term refers to the information file that the receiver uses when receiving/playing a stream. It consists of three XML documents—ERI, LLI and NCI—and contains information for reception/playback control and information relating to the DRM
Continuity index	This is a 4-bit area that is incremented for TS packets having the same PID to indicate the continuity of TS packets.
CRC	Cyclic Redundancy Check: A CRC is a type of error-checking code for detecting contiguous errors.
CRL	Certificate Revocation List: CRL is a list of the certificates of revoked nodes. In this document, the CRL held by the DRM client is a list of the certificates of revoked servers.
CRL server	This is the server that supplies the latest CRL.
Customer/contract management server	This term refers to the server used by a service provider to manage the users and their devices upon basic registration and the personal information of the users, including the services that the users have subscribed to.
Descriptor	This term refers to an area provided in a table for describing various types of information.
DHCP	Dynamic Host Configuration Protocol [RFC2131]: DHCP is the protocol that automatically allocates IP addresses to receivers. In addition to IP addresses, the parameters that are necessary for connection with an IP network, such as the subnet mask, gateway, and DNS server address, can be allocated to receivers at the same time.
DNS	Domain Name Service [RFC1034, RFC1035]: The protocol used for the services that map host names and IP addresses on a network
Downmix coefficient	Downmix coefficient is used when downmixing (converting) multichannel stereo to two-channel stereo, to calculate the individual components of the two-channel stereo from those of the multichannel stereo.
DRM	Digital Rights Management: DRM is a generic name for techniques to prevent unauthorized copying and distribution of digital content, such as audio and visual media, by means of encryption, etc. DRM includes means of keeping track of the distribution of content (e.g., digital watermarking).

DRM provider	This term refers to a subject that implements DRM. The CAS/DRM provider ID (drm_provider_id) for identification of a DRM provider occupies the upper two bytes of the license ID and represents the subject that manages the license ID.
DRM client	A function entity in a receiver to obtain and manage licenses and supply content keys when content are used.
DRM server	This term refers to a server that is in charge of the issuance and management, etc. of licenses.
DTS	Decoding Time Stamp: Time control information for stream decoding
Dual-mono	This is an audio mode for operation of two monaural audios within one ADTS.
ECG	Electronic Content Guide: ECG is the resident application for the navigation of VOD content.
ECG metadata	An XML document used for ECG, describing attribute information related to content, packages and licenses.
ECG metadata server	A server that provides ECG metadata.
Emphasis	This term refers to the process of placing emphasis on specific frequency components during signal modification in order to improve the SN ratio during demodulation.
Entity	A data set that is handled as one unit. A device or equipment that provides functions.
Entity (format)	This term refers to the information (format) that is transferred as a request or a response by the HTTP protocol, etc. It consists of an entity header field containing meta-information and an entity body holding the content of the information.
ERI	Entry Resource Information: ERI is the XML document that contains information about the attributes of object content. It includes the URL, etc. of the video content server.
ES	Elementary Stream: ES is a stream of encoded video, audio and caption data. One ES is transmitted in a TS packet having the same PID as the ES.
FEC	Forward Error Correction: FEC is an error correction system which, if an error or loss is detected in data sent out, restores the original data by using the redundant data that has been added to it, instead of retransmitting the original data.
GOP	Group of Pictures: GOP is a frame structure of MPEG video. It is a unit of encoding for a group of pictures consisting of one I frame and two or more P/B frames.
H.264/MPEG-4 AVC	MPEG-4 Part 10 AVC/H.264: This is the advanced encoding/decoding technique that has been co-developed by the Moving Pictures Expert Group (MPEG) of the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) and the Video Coding Expert Group (VCEG) of the International Telecommunication Union (ITU).
Home network	A network that connects devices in a household. An IP network is assumed.

HTML	HyperText Markup Language: HTML is a markup language for describing Web pages. It is possible to embed, in a document written in HTML, images, audio, video, and links to other documents.
HTTP	Hypertext Transfer Protocol [RFC2616]: The application layer protocol used for data transfer on the World Wide Web.
Identifier	This term refers to an assigned ID that is unique within a certain range. It is a value to identify a specific element in a table or descriptor.
Initialization vector (IV)	IV is a value that is applied when encrypting the leading block in CBC or OFB modes, etc.
IP	Internet Protocol [RFC791]: IP is a protocol which corresponds to the network layer of the OSI reference model. It allows for data transfer between nodes by using IP addresses allocated uniquely to the nodes on the network.
IPTV service	This is a generic name for content delivery services using an IP network.
IPv4	Internet Protocol Version 4: IPv4 is the international standard protocol that is used as the basis of the current LAN/Internet.
IPv6	Internet Protocol Version 6: IPv6 is the next-generation Internet protocol based on the current IP Protocol (IPv4). It features a number of improvements, such as the 128-bit IP address space, enhanced security functions and prioritized data transmission.
IP address	An ID number of a device that is connected to an IP network.
LFE	Low Frequency Enhancement: LFE is a channel which emphasizes low frequency sound in a multichannel stereo system.
Letter box	The system to place black areas to the top and bottom of the screen when video with an aspect ratio of 16:9 is displayed on a screen with an aspect ratio of 4:3.
License	License is the right to use a specific content or the substance of data that shows the right to use content and that permits using the content only when specified usage conditions are met. It includes a content key and information about the content usage conditions.
License ID	This is a unique identifier for specifying a particular license.
License reference information	In ECG metadata, this term refers to the data structure in which information relating to the license is stored.
LLI	License Link Information: LLI is an XML document for gaining access to the license for specific content.
MP@H14L	This is one of the MPEG-2 video coding methods, with the main profile at a high level (1440).
MP@HL	This is one of the MPEG-2 video coding methods, with the main profile at a high level. It is used to encode so-called HDTV of 1080i.
MP@LL	This is one of the MPEG-2 video coding methods, with the main profile at a low level. It is used to encode low-definition videos.
MP@ML	This is one of the MPEG-2 video coding methods, with the main profile at the main level. It is used to encode so-called SDTV of 480i.

MPEG-1	Moving Pictures Expert Group-1: The compressed encoding technology for data containing video and audio, which was standardized by International Organization for Standardization (ISO/IEC 11172).
MPEG-2	Moving Pictures Expert Group-2: The compressed encoding technology for data containing video and audio, which was standardized by International Organization for Standardization (ISO/IEC 13818).
MPEG-2 TS	Transport stream defined by the MPEG system standard (ISO/IEC 13818-1).
Multichannel stereo	This term refers to stereo sound systems consisting of three or more channels. The basic stereo channels (L and R) have center and surround channels, etc., added. Typical multichannel stereo systems are 3/1, 3/2 and 5.1 channels.
Multipart format	This term refers to an entity which is provided with one or more encapsulated entities in a single entity body.
Mute flag	This flag is to control the mute of the receiver from the transmitting side.
Mutual authentication	To mutually verify the validity of each component based on PKI.
NAT	Network Address Translation: The technical method that transparently interconverts between a single global IP address and shared multiple local IP addresses.
NCI	Network content Control Information: NCI is an XML document that contains information relating to the streaming reception of the object content.
Null packet	This is a TS packet which contains no useful information and which is used as padding, etc.
OFB	Output Feedback: OFB is one of the block encryption modes. In this mode, the result of an exclusive OR (XOR) operation on the encrypted initialization vector and the block to be encrypted is used as the encrypted block.
Package	Content billing unit. A package can indicate one or multiple content sets. Also, "package" is used to indicate the contract unit such as monthly. Packages are identified using purchase identifiers (PurchaseID).
PAT	Program Association Table: PAT specifies the packet ID of the PMT (program map table) that shows the content of programs.
Payload	Payload is an array of bytes that follows the header byte in a packet.
PCR	Program Clock Reference: PCR is stored in the header area of a TS packet.
PES	Packetized Elementary Stream: Packetized video, audio, independent data, etc. of variable length.
PID	Packet Identifier: PID is the identifier of a TS packet. It shows the attributes of the appropriate packet stream using 13-bit stream identification information.
PKI	Public Key Infrastructure: PKI is an infrastructure is designed to prevent the forgery, eavesdropping and tampering of communication data through use of public key encryption technology and digital signature.

PMT	Program Map Table: PMT is the table that specifies the packet identifier (PID) of a TS packet which transmits coded signals of video, audio, etc. comprising a program.
Portal	Entrance to a service
Portal server	This is a Web server that supplies the portal service. It supplies documents written in HTML or BML.
Portal service	A Web service that is operated by an IPTV service provider whose main objective is to enable content navigation for IPTV services.
Profile	Classification to limit the functions to be used in the MPEG2 coding method.
Protocol	Communication procedures and communication specifications.
Protocol stack	A stack of protocols that are required for processing on communication networks to describe the functions.
PSI	Program Specific Information: Information required to select a specific program, consisting of four tables: PAT, PMT, NIT and CAT.
PTS	Presentation Time Stamp: Information that manages the presentation output time.
Public key certificate	This term refers to data for certifying that a specific public key belongs to a specific entity. Every public key certificate has a signature affixed.
Receiver	This term refers to any receiver that is compatible with the IPTV service.
Renderer	This is the only functional block that is able to process encrypted content and content keys within the receiver. It consists of a decrypter and an AV decoder for the decryption of encrypted content and the decoding of content.
Resident application	This is an application that has been installed in the receiver.
Resource list	This is the entity that is put at the beginning of multipart data that contains index information about each individual entity (resource) of multipart format. It is defined in ARIB STD-B24.
RFC	Request For Comments: An RFC is a document concerning Internet engineering standards approved by the Internet Engineering Task Force (IETF)—a committee for standardization of Internet engineering.
RMPI	Rights Management and Protection Information: Conditions for use of content.
Rollup mode	This term refers to the caption service that sequentially adds and displays caption data transmitted as page data on a line-by-line basis in a preset area of three lines or so. The caption data is vertically rolled up during a line feed.
RTP/RTCP	Real-time Transport Protocol/RTP Control Protocol [RFC3550, RFC3551]: These are protocols for transmitting streamed data, such as audio and video. RTCP is used to notify the transmission line parameters, such as communication quality and delay time, from the receiver to the server.
RTSP	Real-Time Streaming Protocol: RTSP is the control protocol for delivering video and audio over an IP network on a real-time basis.

SAC	Secure Authenticated Channel: An encrypted channel based on mutual authentication.
Sampling rate	This term refers to the repetitive frequency at which sample values are taken from original analog signals of video, audio, etc. when those original signals are converted into digital signals.
Sequence	The processing procedures that are communicated between receivers and servers.
Service application	This term refers to a user's act of purchasing or contracting for a package offered by a service provider.
Service provider	This term refers to the body that provides a VOD service.
Session	A series of processes that are performed by a user.
SDP	Session Description Protocol: Session information that is notified when a multimedia session is started.
Side panel	The system to place black areas on both sides of the screen when video with an aspect ratio of 4:3 is displayed on a screen with an aspect ratio of 16:9.
SSL	Secure Socket Layer: SSL is a protocol for encrypting information and transmitting/receiving the encrypted information on the Internet. It is designed to prevent the forgery, eavesdropping and tampering of data.
Streaming	The viewing method in which data is reproduced in real time as a viewer receives the data.
Superimpose	The service to provide captions that are asynchronous to the main video, audio and data. It is used for up-to-the-minute news, changes in the program schedule, time signals, etc.
Table	This term refers to information transmitted in MPEG2-TS section format.
TCP	Transmission Control Protocol [RFC793]: TCP is a protocol that corresponds to the transport layer of the OSI reference model. Provided with a number of control functions (retransmission control, flow control, congestion control, etc.), TCP offers a reliable data communication link to higher-layer applications.
Time stamp	Time stamp indicates the time of a specific operation, such as the arrival of a data byte or the display of video or audio.
TLS	Transport Layer Security [RFC2246]: TLS is a protocol for encrypting information and transmitting/receiving the encrypted information on the Internet. This protocol, based on SSL which was developed by Netscape Communications, has been standardized by the IETF.
TS	Transport Stream: TS is a transport stream defined in ISO/IEC 13818-1.
TS packet	This is a fixed-length (188-byte) packet defined in ISO/IEC 13818-1.
TTS	The time-stamped TS defined in ARIB STD-B24 Volume 2. It is the transport stream packet defined in the MPEG system standard (ISO/IEC 13818-1) with a time stamp (32-bit) clocked at 27 MHz added.

UDP	User Datagram Protocol [RFC768]: This is a protocol that corresponds to the transport layer of the OSI reference model. Because of the absence of retransmission control and flow control functions, the reliability of communication data is not guaranteed. Even so, this protocol features a lighter processing load for data transmission/reception and is suitable for high-speed data transfer.
Unicast	This term refers to the mode of communication in which data is exchanged between nodes on a one-to-one basis.
URI	Uniform Resource Identifier: The description method to indicate where information is located. URI includes URL.
Video content server	This term refers to the server that supplies video content, either by the RTP/RTSP protocol or by the HTTP protocol.
VOD	Video On Demand: VOD is a system that delivers by streaming, on an on-demand basis, the video that the user wants to watch.
VOD license	This is a one-layer type license that is used in VOD services.
W3C	The World Wide Web Consortium: W3C is an international body that standardizes protocols for the supply and sharing of information on the Internet.
XML	Extensible Markup Language: XML is one of the markup languages. It is a meta-language.

Chapter 2 Overview

2.1 Service Requirements

The service requirements of this document are described below.

- The service provider shall be able to provide the VOD service using an IP network.
  Even though relatively stable connection with an IP network via a broadband is presupposed, the service provider shall be able to provide the means of coping with possible transmission jitters and packet losses in order to enable stable playback of content.
  The services shall be able to be provided over either the Internet or CDN (Content Delivery Network), or both.
- The receiver shall be able to be implemented relatively easily by extension from a digital broadcast receiver. As product types, TVs, STBs, DVRs (e.g., DVD recorders), etc. are presupposed.
- The service provider shall be able to provide video content service that contains following elements.
  - Video: SDTV and HDTV.
  - Audio: monaural, stereo, multi-audio (e.g., dual audio) and multi-channel (surround sound).
  - Caption.
- The following two video display modes shall be available.

  (a) Full-screen display

  The mode in which video is displayed on the entire screen.

  (b) Sub-window display (L-wrapped display)

  The mode in which video is displayed on a window opened in part of an HTML or BML document.
- For video content selection (content navigation), the following methods shall be available.

  (a) Content selection by a browser (HTML/BML)

  (b) Content selection by ECG (Electronic Content Guide) (option)
- The service shall allow for special modes of playback of video content.
  - In addition to the ordinary modes of playback, the service shall allow for pause, fast-forward playback, fast-rewind playback, and jump (playback from any chapter jump point).
  - In the event that the user stops viewing the video content midway, the service shall permit the user to resume viewing the video content from the point of interruption.
- The service provider shall be able to provide a means of protecting the copyright on video content.

2.2 System Model

A VOD service system model is shown in Figure 2-1. It should be noted that the functional components of servers in the diagram is a logical one: it does not specify a physical configuration. It should also be noted that the service provider need not necessarily provide all the server functions described here.

The individual servers in the system model are outlined below.

(1) Video content server

This server delivers video content.

As the streaming protocol, either RTSP/RTP or HTTP is used. The URI of the video content is notified to the receiver via a content playback control metafile that is supplied from the playback control information server.

(2) Playback control information server

This server supplies the playback control information that is necessary for playback of video content.

In VOD, only a content playback control metafile is used as the playback control information. The URI of the content playback control metafile is notified to the receiver via an HTML/BML document supplied from the portal server or by ECG metadata supplied from the ECG metadata server. The content playback control metafile describes the URIs of video content and the information that is necessary for the acquisition of licenses, including encryption keys for decrypting encrypted video content.

(3) ECG metadata server

This server delivers ECG metadata for video content navigation through an ECG application installed in the receiver. It should be noted that the supply of the ECG metadata server and the installation of the ECG application in the receiver are optional.

(4) DRM server

This server performs encrypted communication with the DRM client of the receiver and delivers to the receiver a license in which the key for decrypting encrypted content and the conditions for use of the content are described.

The receiver obtains the license from the DRM server using the URI described in the content playback control metafile and decrypts the encrypted video content.

(5) Portal server

This server provides HTML/BML documents to the receivers.

In addition to supplying the navigation screen, etc. for selection of video content, this server can, together with the customer/contract management server, perform various types of processing such as registration and authentication.

(6) Customer/contract management server and billing/settlement server

The server for customer management, including the management of contracts with customers, and the server for billing and settlement of accounts are outside the scope of this document.

These servers manage customer information, including the CAS/DRM client identifier (DRM_ID), which is involved in the basic registration, and, together with the portal server, perform the authentication processing when the receiver connects to the portal server and the billing and settlement processing for purchases made by customers. These servers do not directly communicate with the receiver. The receiver is provided with the functions of these servers via the portal server, etc.

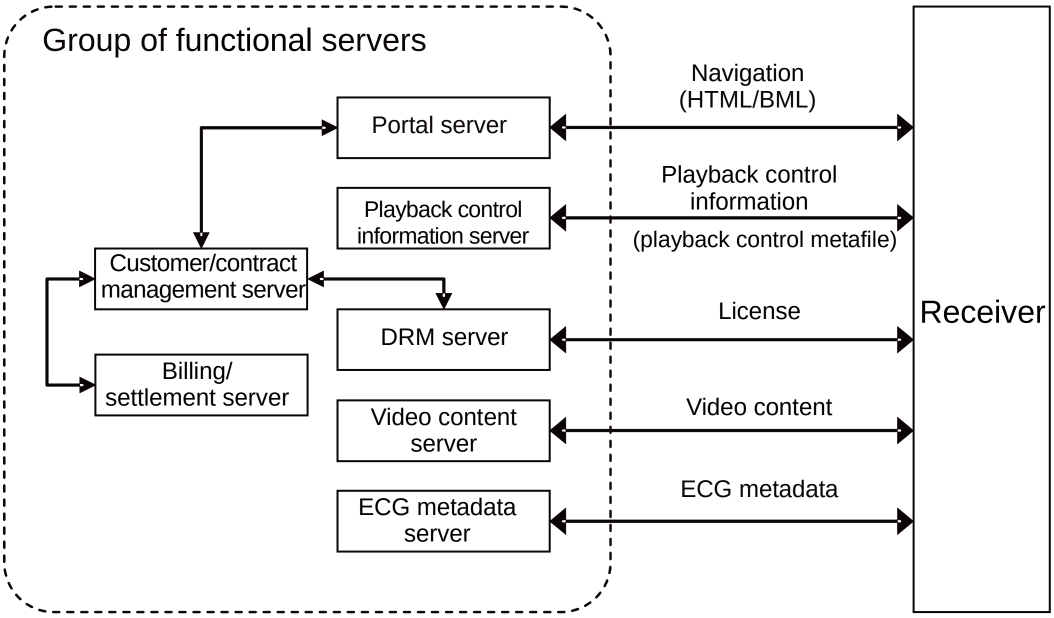

Figure 2-1 VOD System Model

2.3 System Flow

The basic sequence of the VOD service is explained below.

2.3.1 Selection from Browser

The sequence when video content is selected from the HTML browser or BML browser is shown in Figure 2-2. The outline of the processing performed is explained below.

(1) The receiver obtains from the portal server an HTML or BML document for selection of video content and presents it to the user via the HTML/BML browser.

(2) When the user initiates viewing of video content via the browser, the receiver uses the HTTP(S) protocol to request the content playback control metafile to the playback control information server via the URL specified in the HTML/BML document.

(3) The playback control information server transfers the specified content playback control metafile to the receiver.

(4) The receiver obtains the URI of the DRM server from the content playback control metafile and establishes encrypted communication channel with the DRM server to obtain a license that contains the decryption key for the video content.

(5) After obtaining the URI of the video content from the content playback control metafile, the receiver instructs the video content server to play the video content by RTSP or HTTP.

(6) The video content server transmits to the receiver the requested video content by RTP or HTTP. The receiver receives the video content, decrypts the video content using the decryption key that is contained in the license obtained from the DRM server, and plays the video and audio.

(7) Upon completion of playing the video content, the receiver asks the portal server for the HTML/BML document on the return screen. (The URI of the return screen is specified in the HTML/BML document in (2).)
When the video content playback is stopped mid-way, the receiver may send to the portal server the play-stop position information, together with the request for the HTML/BML document. The portal server can use that information to supply the resume function on the screen for video content selection.

For details on browser operation, see IPTVFJ STD-0006 "IPTV Standards: CDN-scope Service Approach Specifications" or IPTVFJ STD-0007 "IPTV Standards: Internet-scope Service Approach Specifications".

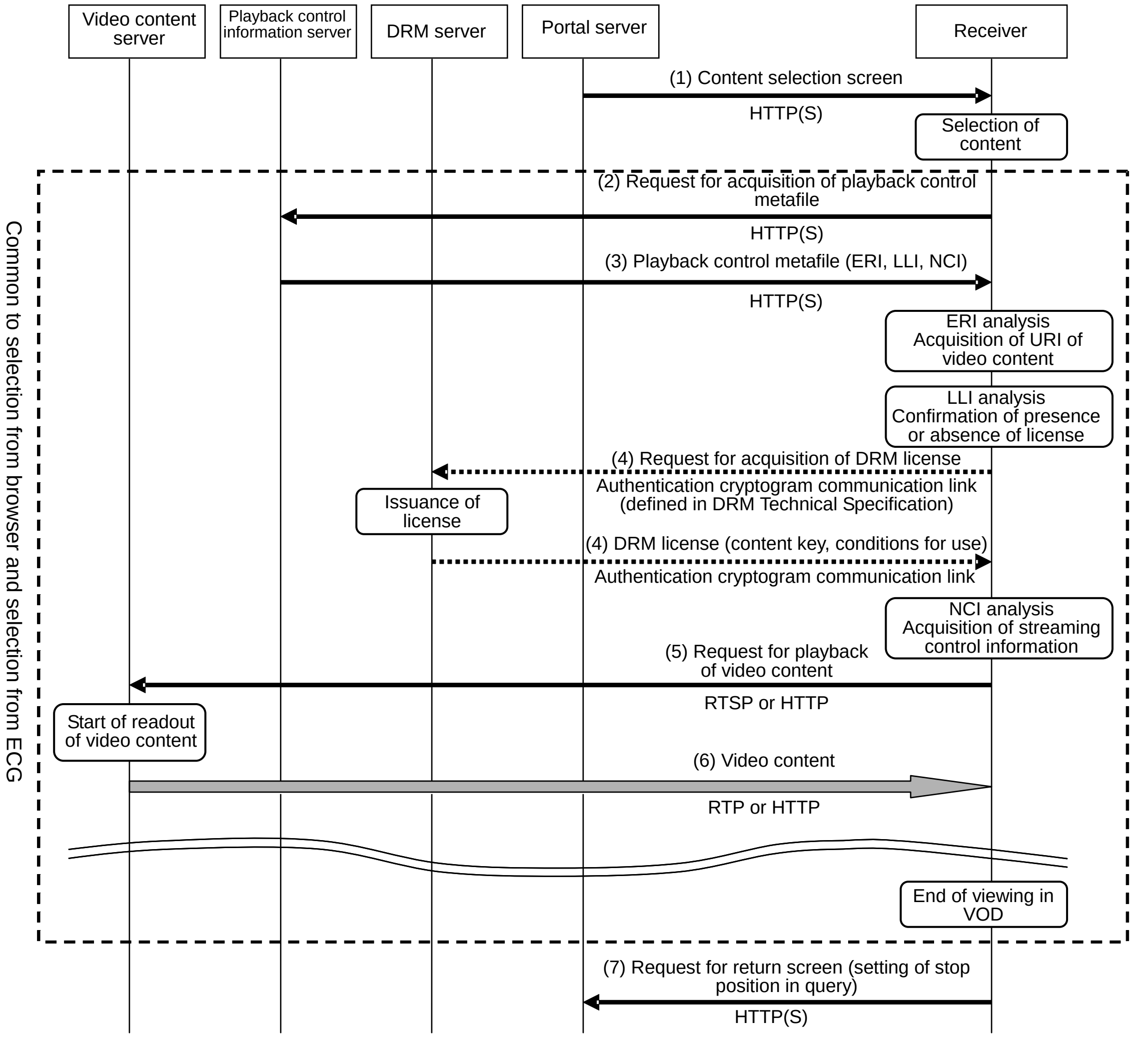

Figure 2-2 Basic Sequence of VOD (Selection from Browser)

2.3.2 Selection from ECG

The sequence when video content is selected from ECG is shown in Figure 2-3. The outline of the processing performed is explained below.

(1) The receiver obtains ECG metadata from the ECG metadata server and presents the video content selection screen to the user via the ECG application.

(2) When the user initiates viewing of the video content by the ECG application, the receiver uses the HTTP(S) protocol to requrest the content playback control metafile to the playback control information server via the URL specified by the ECG metadata.

(3) The playback control information server transfers the specified content playback control metafile to the receiver.

(4) The receiver obtains the URI of the DRM server from the content playback control metafile and establishes encrypted communication channel with the DRM server to obtain a license that contains the encryption key for the video content.

(5) The receiver obtains the URI of the video content from the content playback control metafile and asks the video content server for playback of the video content by RTSP or HTTP.

(6) The video content server transmits to the receiver the requested video content by RTP or HTTP. The receiver receives the video content and decrypts the video content using the decryption key that is contained in the license obtained from the DRM server to play the video and audio content.

(7) The action to be taken after completion of the video content playback depends on the receiver implementation.

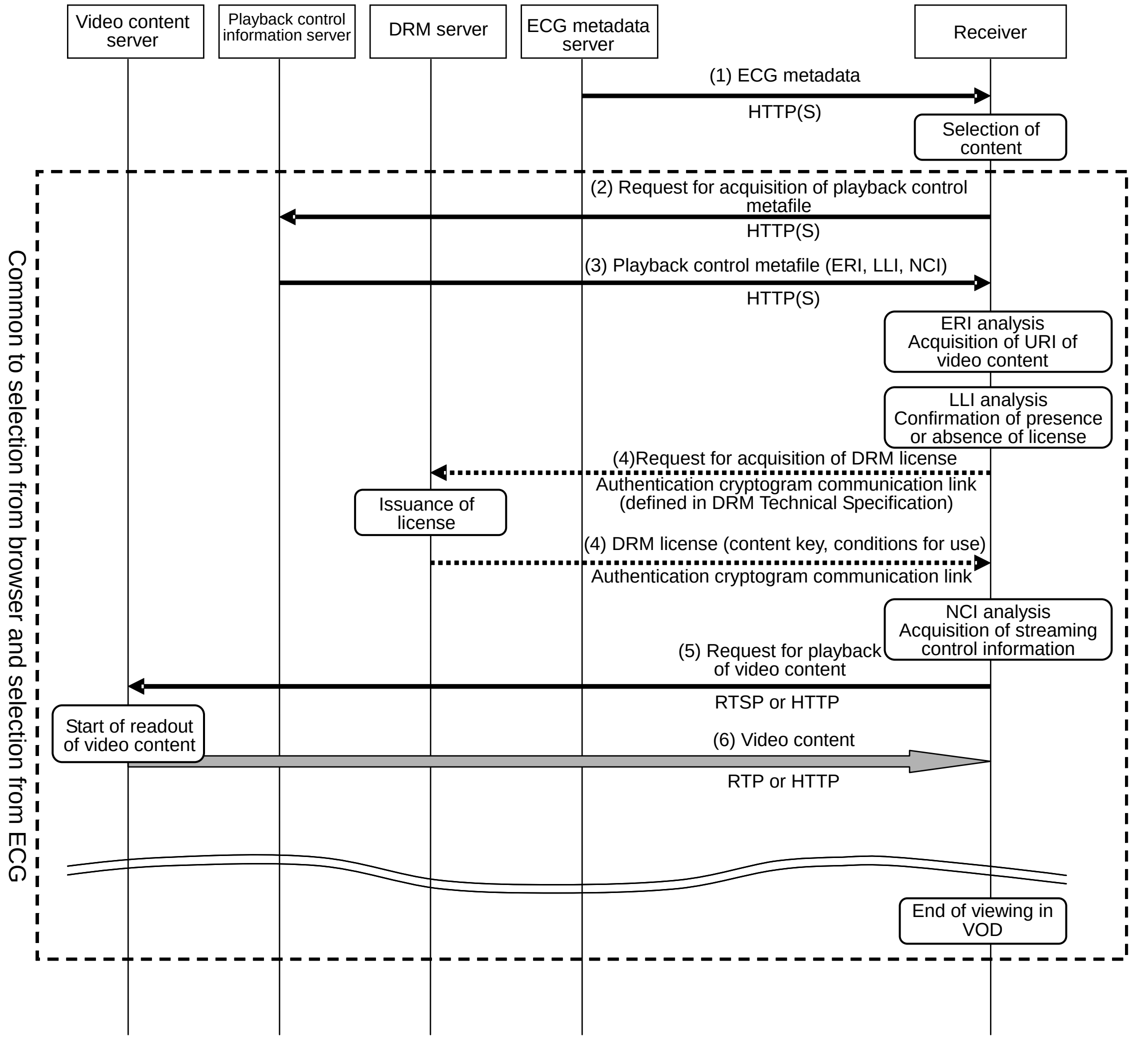

Figure 2-3 Basic Sequence of VOD (Selection from ECG)

For details about the operation of ECG and ECG metadata in the CDN scope, see IPTVFJ STD-0006 "IPTV Standards: CDN-scope Service Approach Specifications".

Chapter 3　Receiver Model

In this chapter, the concept of receivers is defined.

3.1　Receiver Model

3.1.1　Receiver Reference Model

A reference model of a receiver is shown in Figure 3-1. The individual components of the reference model are explained below.

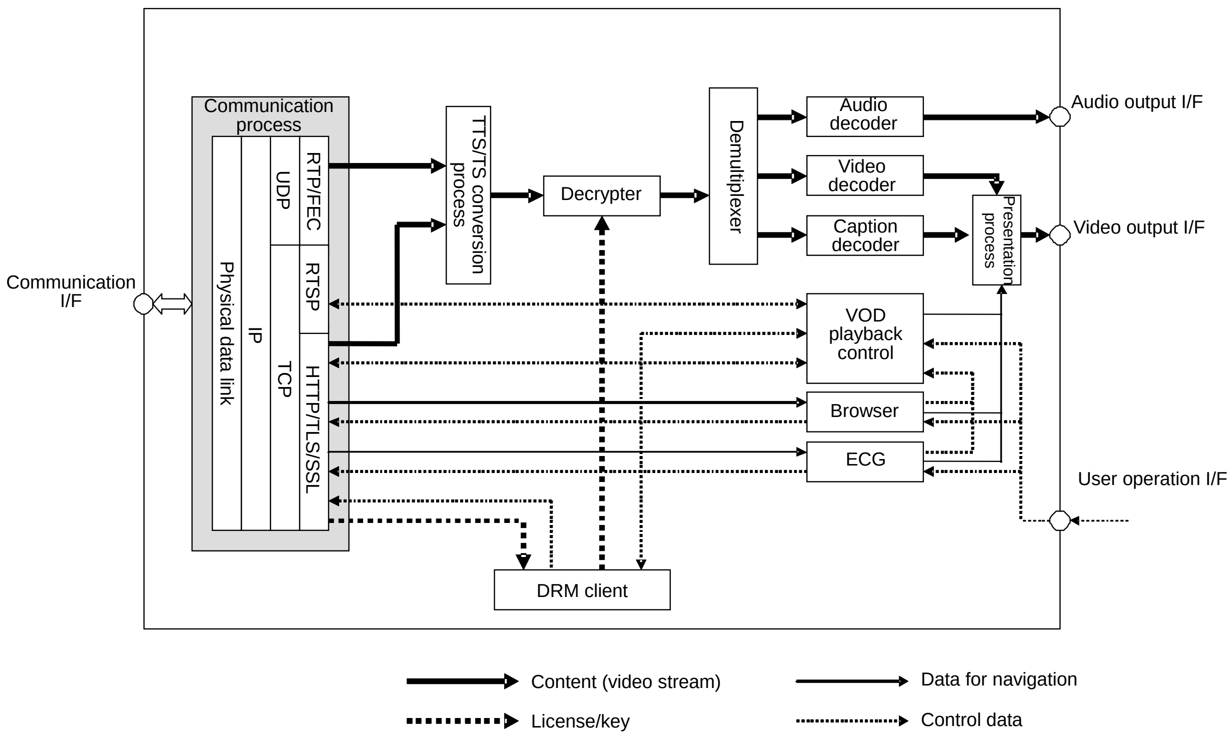

Figure 3-1　Receiver Reference Model

- Communication I/F
  This is the interface for the exchange of signals with the communication network.
- Communication process
  Various types of communication protocols are processed.
- TTS/TS conversion process
  TTS is converted into MPEG-2 TS.

  TTS/TS conversion process temporarily stores input TTS packets into a FIFO buffer, and extract them to make an output of MPEG-2 TS that is synchronized with the 27-MHz clock at the transmitter side by using the time stamp embedded in each TTS pakcet and the 27-MHz self-contained clock of the receiver.

- Decrypter
  The decrypter decrypts encrypted MPEG-2 TS by using the content key obtained from the DRM client.
- Demultiplexer
  The demultiplexer separates multiplexed MPEG-2 TS into packets, such as video, audio, caption and PSI/SI packets.
- Video decoder
  This decoder decodes video data.
- Audio decoder
  This decoder decodes audio data.
- Caption decoder
  This decoder decodes caption data.
- VOD playback control
  This control consists of the acquisition and analysis of content playback control metafile, the licensing processing via DRM client and the control of AV playback.
- Presentation process
  In this processing, visual screen data is created to present to the user by compositing stream data from the video decoder, caption data and, data displayed by the browser, ECG, etc.
- Video output I/F
  This is the interface for outputting video signals.
- Audio output I/F
  This is the interface for outputting audio signals.
- Browser
  This is an application that provides the function of executing and presenting an HTML or BML document obtained from a server.
  It is assumed that a list of contents offered by the VOD service, a list of purchased contents, detailed information about content, etc. are described in the HTML or BML document. The browser provides the function of navigating through the document.
- ECG
  ECG uses ECG metadata obtained from the ECG metadata server to permit navigating through a list of contents offered by the VOD service, a list of purchased contents, and detailed information about content, etc.
- User operation I/F
  This is the interface for receiving user-operated events. The events will occur by the operation of the user operation I/F such as buttons on the panel of the receiver, remote control, etc.

- DRM client
  This client obtains a license from the DRM server and supplies the content key extracted from the license to the decrypter.

3.1.2 Receiver Data Flow

The data flow when the VOD service is used is shown in Figure 3-2.

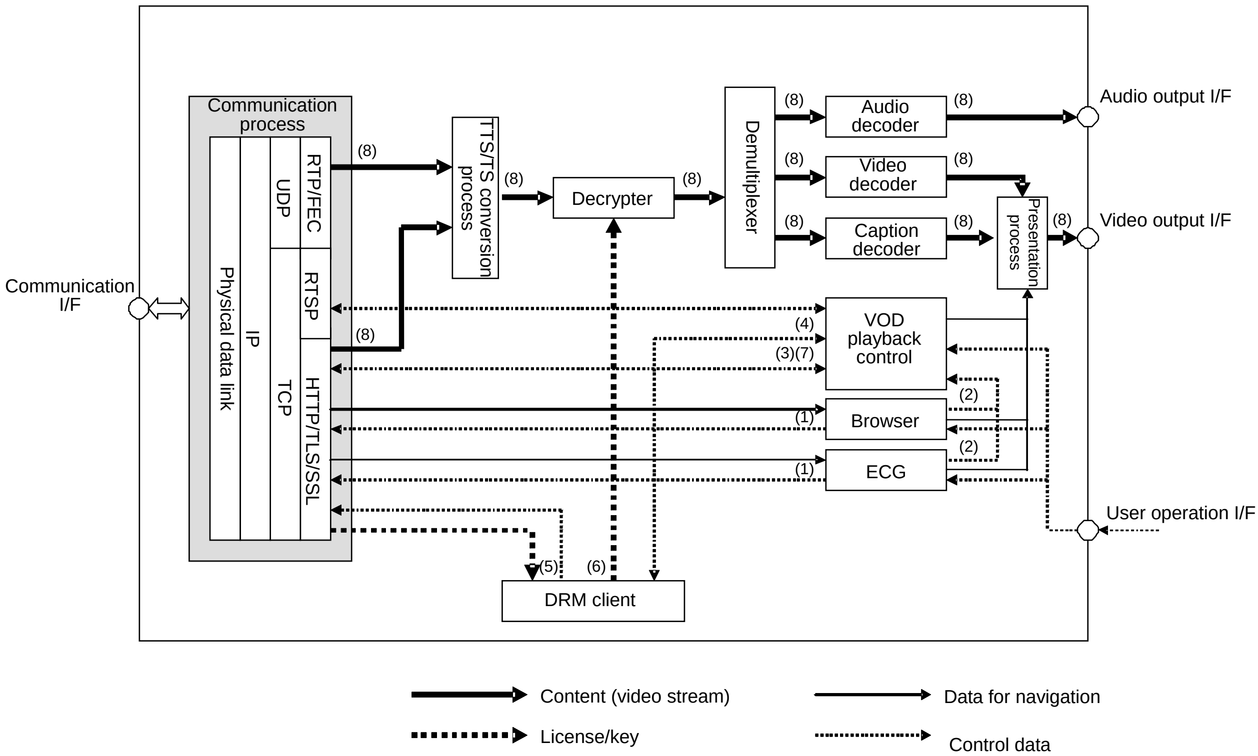

Figure 3-2 Data Flow in VOD Service

(1) Content is selected by the browser or ECG.

(2) The browser or ECG notifies the information related to the content playback control metafile of the selected content to the VOD playback control component.

(3) The VOD playback control component obtains and analyzes the content playback control metafile from the playback control information server.

(4) When the VOD playback control component detects that the selected content has been encrypted, it notifies the DRM client to obtain a VOD license.

(5) The DRM client obtains the VOD license from the DRM server.

(6) The DRM client sets the content key contained in the VOD license to the decrypter.

(7) The VOD playback control component uses the communications protocol (HTTP or RTSP) to issue a request for playback to the video content server.

(8) The video content that is received is subjected to communication processing, TTS/TS conversion processing and processing by the decrypter, demultiplexer, video decoder, audio decoder, and caption decoder so that the video and audio signals can be output.

The time from the instant at which the user selects content (1) till the instant at which the user starts viewing the selected content (8) shall be assumed to depend upon the receiver implementation. However, it may be presupposed that the time does not exceed 30 seconds.

In the sequence shown above, if the response from any of the servers results in an error or timeout, the receiver, after displaying an appropriate message, shall return to a prescribed state without attempting a retry.

3.2 Functional Requirements of Receivers

The presupposed requirements of the hardware and software of the receiver are described below.

3.2.1 Communication Processing

For specifications of the physical interface for communication processing, see ARIB STD-B21, Chapter 9, 9.2.1.1 "Physical interface specifications".

The receiver uses such protocols as IP, TCP, UDP, HTTP, RTP and RTSP and performs various types of communication processing. The protocol stack that is presupposed to be used in communication processing is shown in Figure 3-3, and the reference protocol specifications are shown in Table 3-1.

The VOD service to which the receiver connects decides the applicable streaming protocol (HTTP, RTP/RTSP). The receiver needs various types of protocol associated with VOD services. The HTTP protocol is used to obtain HTML/BML documents, content playback control metafile and ECG metadata that are necessary for selecting VOD content and displaying relevant information.

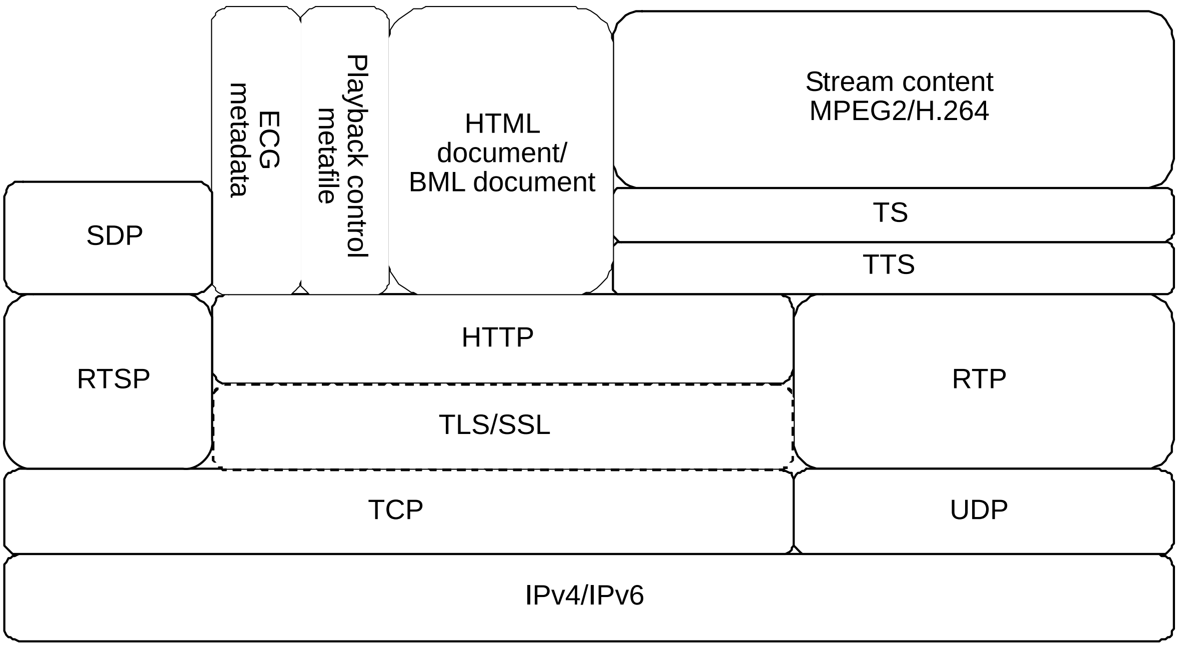

Figure 3-3 Protocol Stack Model

Table 3-1: Reference Protocol Specifications

Streaming transmission/control	
RTP, RTCP	RFC3550 "RTP: A Transport protocol for Real-Time Applications"
RTSP	RFC2326 "Real Time Streaming Protocol (RTSP)"
SDP	RFC4566 "SDP: Session Description Protocol"
HTTP	RFC2616 "Hypertext Transfer Protocol -- HTTP/1.1"
Communication protocols	
UDP	RFC768 "User Datagram Protocol"
TCP	RFC793 "Transmission Control Protocol"
IP	RFC791 "Internet Protocol"
ICMP	RFC792 "Internet Control Message Protocol"
IPv6	RFC2460 "Internet Protocol Version 6 (IPv6) Specification"
ICMPv6	RFC4443 "Internet Control Message Protocol (ICMPv6) for the Internet protocol Version 6 (IPv6) Specifications"
	RFC4291 "IP Version 6 Addressing Architecture"

3.2.2 Video/Audio Decoding Process and Output

3.2.2.1 Video Decoding Process and Output

See ARIB STD-B21, Chapter 6, 6.1.3 "Video-signal output". For the RGB analog terminal and digital video terminal, see 3.2.2.4.1 "RGB Analog Terminals" and 3.2.2.4.2 "Digital Video Terminals" of this document. For video output at a digital AV output terminal, see 3.2.2.5.5 "Digital AV Output".

When MPEG2 is used, see ARIB STD-B21, Chapter 6, 6.1.1 "Video decoding process" and ARIB STD-B21, Chapter 6, 6.1.2 "Video output signals". It should be noted, however, that 480p shall not be operated.

In the case of H.264/MPEG4 AVC, see 6.1.1.3 "Details of Operation of H264/MPEG-4 AVC Video" of this document.

3.2.2.2 Audio Decoding Process and Output

See ARIB STD-B21, Chapter 6, 6.2.3 "Audio output".

When MPEG2 AAC is used, see ARIB STD-B21, Chapter 6, 6.2.1 "Audio decoding process", ARIB STD-B21, Chapter 6, 6.2.2 "Audio mode discrimination and indication" and ARIB STD-B21 Appendix 4 "Down-mix processing in the AAC decoder". It should be noted that for a receiver which is provided with the optional downmix processing function for external pseudo surround sound processors and the optional downmix processing function for expansion of stereo sound fields (both functions are described in ARIB STD-B21, 6.2 "Audio decoding process and output"), arrangements shall be made so that the user can grasp the setting conditions (for example, displaying the downmix setting conditions).

In the case of MPEG1 Audio, see 6.1.2.1 "MPEG1 (Audio)" of this document.

When a digital audio output terminal is provided, it shall conform to the AAC extension defined in IEC 60958 and IEC 61937. When a Bluetooth interface is used for digital audio output, the Bluetooth logo authentication shall be obtained in order to guarantee its conformity to the Bluetooth Standard. For audio output at a digital AV output terminal, see 3.2.2.5.5 "Digital AV Output".

3.2.2.3 High-speed Digital Interface

As a high-speed digital interface, only the IP interface is defined. When an IP interface is provided as a high-speed digital interface, it shall conform to the following.

3.2.2.3.1 IP Interface Specifications

3.2.2.3.1.1 Protocol Stack Specifications

See ARIB STD-B21, Chapter 9, 9.2.1.2 "Protocol stack specifications" (T.B.D.).

3.2.2.3.1.2 Content Output Specifications

(1) Content transmission protocol

See ARIB STD-B21, Chapter 9, 9.2.2.1 "Content transmission protocol" (T.B.D.).

(2) Packet format

See ARIB TR-B14, Volume 2, 8.3.1 "Packet format".

(3) Stream format

The content transmission stream that is included in the packet payload defined in (2), above, shall be in time-stamped TS described in ARIB STD-B24, Volume 2, 8.1.4.

As a rule, with the following exceptions, video content obtained from the video content server shall be output without changing the TS packet configuration.

- Null packets may be deleted.
- When implementing copy control using DTCP, a DTCP_descriptor shall be inserted. For details, see the DTCP Specifications.

3.2.2.3.1.3 Content Description

The description of content shall meet the requirements defined in the DLNA Guidelines.

3.2.2.3.1.4 Content Selection Control

See ARIB TR-B14, Volume 2, 8.3.4 "Control of Content Selection" (T.B.D.).

3.2.2.4 Other Types of Output

3.2.2.4.1 RGB Analog Terminals

- Providing a VGA terminal is optional. When a VGA terminal is provided, it is indispensable to provide a connector which conforms to Chapter 4 "Physical Connections" of the Enhanced Display Data Channel Standard (Version 1) issued by VESA and to output signals in a format which conforms to Chapter 2 "VESA Video Signal Definition" of the "Video Signal Standard (Version 1, Rev. 1)" issued by VESA.
- Providing a DVI terminal for analog output is optional. When this type of terminal is provided, it is recommended that a connector which conforms to Chapter 5 "Physical Interconnect Specification" of the "Digital Visual Interface DVI (Revision 1.0)" issued by DDWG be provided as well. In this case, it is necessary to output signals in a format which conforms to Section 2.5 "Analog" of Chapter 2 "Architectural Requirements" of the "Digital Visual Interface DVI (Revision 1.0)".

3.2.2.4.2 Digital Video Terminals

- Providing a DVI terminal is optional. When a DVI terminal is provided, it is recommended that a connector which conforms to Chapter 5 "Physical Interconnect Specification" of the "Digital Visual Interface DVI (Revision 1.0)" issued by DDWG be provided. In this case, it is necessary to output signals in a format which conforms to Chapter 2 "Architectural Requirements" of the "Digital Visual Interface DVI (Revision 1.0)".
- For copyright protection technology, see 3.2.2.5 "Copy Control".

3.2.2.4.3 Digital AV Output Terminals

- Providing an HDMI terminal is optional.
  When an HDMI terminal is provided, it shall conform to the "High-Definition Multimedia Interface Specification" issued by HDMI Licensing, LLC.
- For copyright protection technology, see 3.2.2.5 "Copy Control".

3.2.2.5 Copy Control

When external output is implemented, copy control using the output control information contained in the license as defined in 7.1.2.3 "Constituent Elements of License" shall be implemented in accordance with Chapter 7 "DRM Specifications".

The above output control information contains control information corresponding to the various types of control information contained in the digital copy control descriptor and the content availability descriptor that are defined in ARIB STD-B10. In the paragraphs that follow, the control information that is used in various types of output control will be described as corresponding to the appropriate control information name in the above descriptors.

3.2.2.5.1 Analog Video Output

- Copy control for each analog video output format shall be in accordance with the specifications shown in Table 3-2.
- In terms of control information, the control information (copy control type information) corresponding to copy_control_type, the control information (digital copy control information) corresponding to digital_recoding_control_data, and the control information (analog output copy control information) corresponding to APS_control_data—all contained in the output control information in the license—shall be used.

Specifically, for Macrovision pseudo sync pulses and color stripes, the control information (analog output copy control information) corresponding to APS_control_data shall be used; for video ID signal CGMS-A, the control information (digital copy control information) corresponding to digital_recoding_control_data shall be used; and for video ID signal APS (analog output copy control information), the control information corresponding to APS_control_data shall be used.

Table 3-2: Copy Control for Each of Analog Video Output Formats

Analog video output**1		Macrovision **2	Video ID signal**3
	480i composite	Pseudo sync pulse / color stripe	CGMS-A APS
	480i component	Pseudo sync pulse	CGMS-A APS
	720p component	—	CGMS-A APS
	1080i component	—	CGMS-A APS
RGB analog output**4		—	—

**1) This includes the case in which received video signals are subjected to format conversion at the receiver side and output as analog video in a specific format.

**2) This requires the signing of an agreement between the broadcasting company and Macrovision. It is presupposed that parameters are not transmitted.

**3) Video ID signal is a signal transmitted using an identification signal waveform which is overlaid on VBI. It consists of CGMS-A and APS information, etc.

**4) For RGB analog output, see 3.2.2.4.1 "RGB Analog Terminals".

3.2.2.5.2 Digital Audio Output

Adopting digital audio output is optional. When it is adopted, copy control shall be implemented by using the control information (copy control type information) corresponding to copy_control_type, the control information (digital copy control information) corresponding to digital_recoding_control_data, and the control information (analog output copy control information) corresponding to APS_control_data—all included in the output control information in the license.

3.2.2.5.3 High-speed Digital Interface Output

The high-speed digital interface is optional. When a high-speed digital interface is provided, the following requirements shall be met.

- The control shall be implemented by using the control information corresponding to copy_control_type, which is included in the output control information in the license, and the control information corresponding to digital_recoding_control_data.
- The copyright protection system shall be DTCP. When DTCP is used, a DTCP_descriptor shall also be inserted. For details, see the DTCP Specifications.

3.2.2.5.4 Digital Video Output

- When DVI is provided to output video content, the protection of which is specified by the output control information in the license, protection technology shall be applied in accordance with the HDCP Specifications.

3.2.2.5.5 Digital AV Output

When HDMI is provided to output AV content, the protection of which is specified by the output control information in the license, suitable protection technology shall be applied in accordance with the HDCP Specifications.

3.2.3 DRM Processor/Decrypter

See 7.3 "DRM Related Functional Requirements for Receiver" of this document.

3.2.4 Browser

The browser interprets documents supplied by the portal server and offers a number of functions, such as the search and playback of VOD content and the simultaneous display of VOD content and browser screen.

The browser has the following functions.

- Displaying HTML or BML documents
- Serving as the interface for VOD playback control of VOD content
- Displaying VOD content and browser screen simultaneously

For details about the browser operation, see IPTVFJ STD-0006 "IPTV Standards: CDN-scope Service Approach Specifications" or IPTVFJ STD-0007 "IPTV Standards: Internet-scope Service Approach Specifications".

3.2.5 ECG

The ECG function is optional. Whether or not to provide it depends on the receiver implementatioin.

Using ECG metadata which is previously supplied by the content delivery service provider, ECG offers the functions of searchning content, displaying information, playing content, providing advice on the purchase of content and integrating with the portal.

It is presupposed that ECG has the following functions.

Funtion to obtain all ECG metadata or partial ECG metadall by specifying the conditions for search from individual service provider

- Function to obtain ECG metadata from more than one service provider
- Function to present to the user a list and details of ECG metadata obtained
- Function to go through the procedure for purchasing content and packages from the ECG metadata presented to the user
- Function to present the content and package purchased by the user
- Function to play the content that has been purchased.

For details about the operation of ECG and ECG metadata in the CDN scope, see IPTVFJ STD-0006 "IPTV Standards: CDN-scope Service Approach Specifications".

3.2.6 VOD Playback Control

VOD playback control has the following functions.

- Initiation processing

  Obtaining and analysis of content playback control metafile (Chapter 5 "Content Content Playback Control Metafile")

  Licensing processing via DRM client (Chapter 7 "DRM Specifications")
- AV playback control (Chapter 4 "Video Streaming Protocol")

  Starting and stopping of playback, trick playback (pause, fast forward, fast rewind, jump, etc.)
- Information display

  It is presupposed that the title, chapter, playback point (time code, progress), audio information, caption information, etc. are displayed. (An example is shown in Figure 3-4.)
- Switching between audios/captions

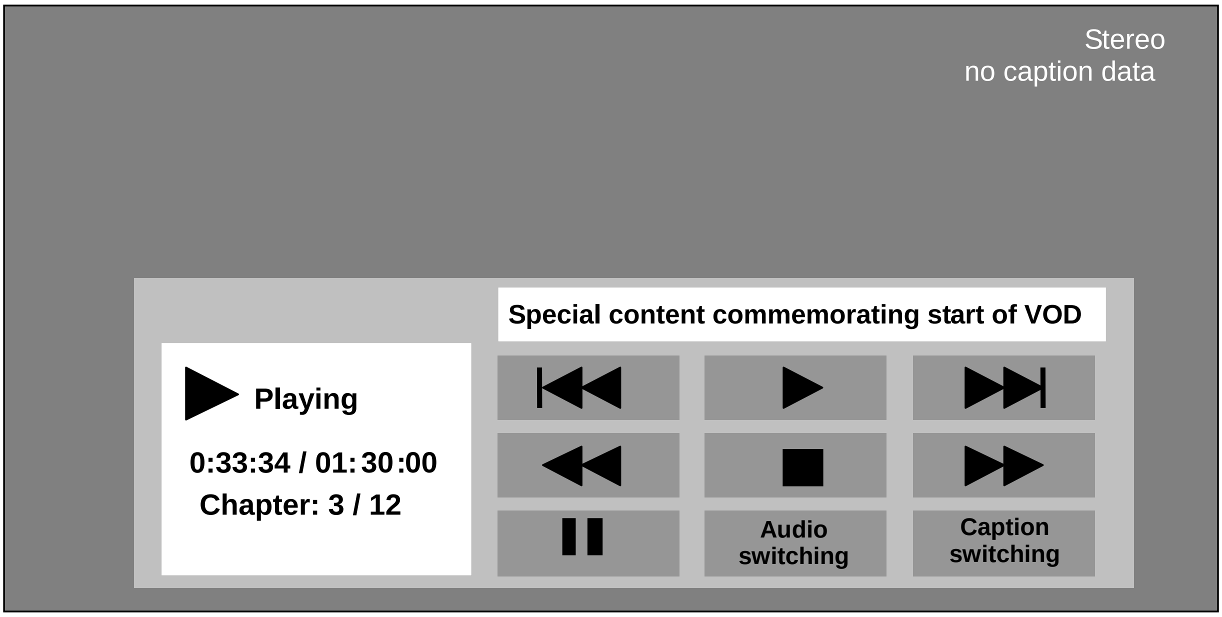

Figure 3-4　Example of Information Display

Chapter 4　Video Streaming Protocol

4.1　Video Transmission Protocol Based on RTP/RTSP

4.1.1　RTSP

4.1.1.1　Purpose

In the viewing of VOD content, RTSP is used as the protocol for controlling unicast streams. Here, the protocol is defined as the necessary stream control interface between the servers and receiver based on the specifications defined in RFC2326 "Real Time Streaming Protocol (RTSP)". Therefore, unless otherwise specified, the messages and headers that are defined as being indispensable by RFC2326 and the setting thereof shall conform to the RFC specifications.

As the lower-layer protocol, TCP shall be used.

4.1.1.2　State Transitions

The receiver as an RTSP client goes through the states and the transition of state that are defined in Appendix A of RFC2326. However, since these Specifications do not presuppose client-to-server transmission (RECORD), the transition of state is as shown below. It should be noted, however, that the receiver can send OPTION and DESCRIBE messages in any state and that those messages shall not cause the current state to change.

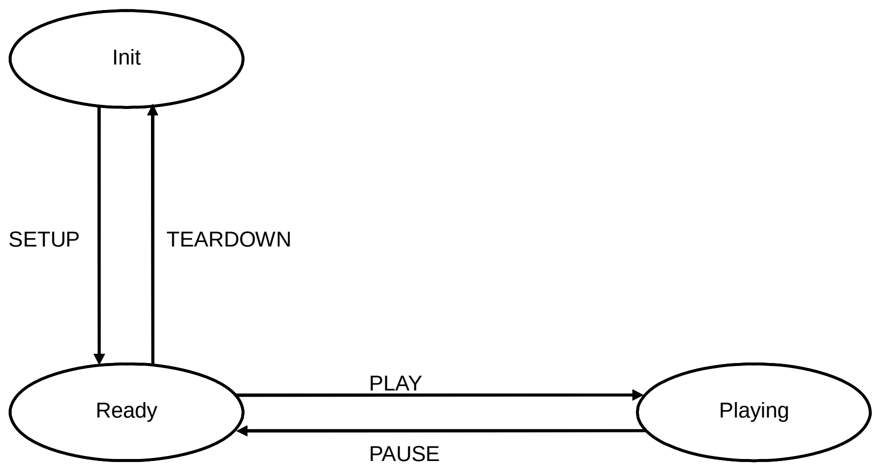

Figure 4-1　Transition of State of Receiver as RTSP Client

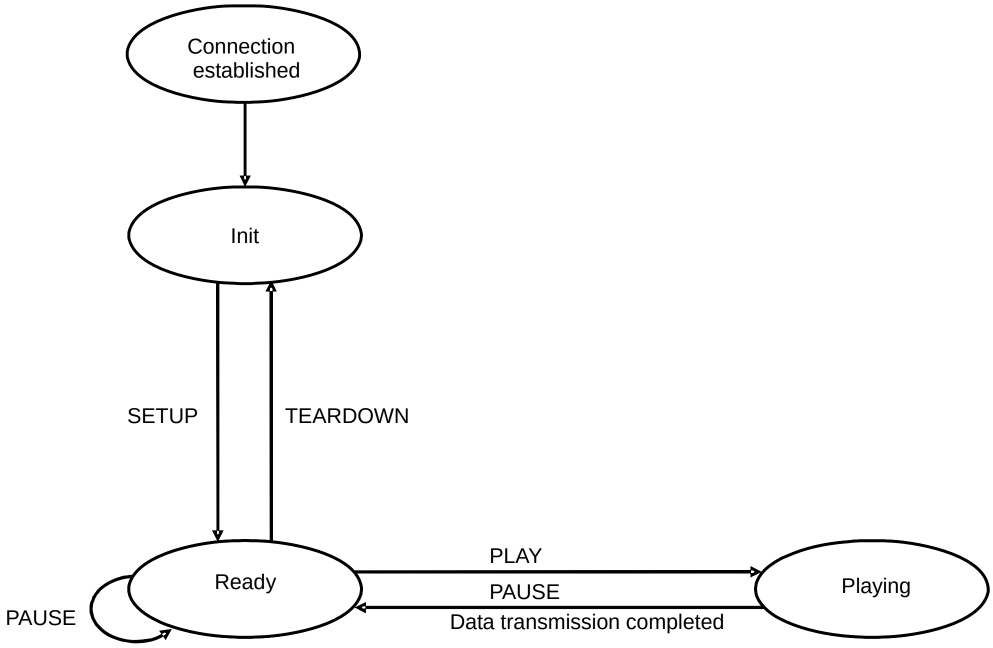

Figure 4-2 Transition of State of Server as RTSP Server

4.1.1.3 Indispensable Methods

Of the methods defined in RFC2326, those shown below shall be deemed either indispensable or optional for the video content server and the receiver. In the table, "transmission" in parentheses indicates transmission of the relevant method, and "reception" in parentheses indicates reception and interpretation of the relevant method.

Table 4-1: Indispensable Methods of RTSP

Method	Receiver	Server
DESCRIBE	Optional (transmission)	Indispensable (reception)
ANNOUNCE	Indispensable (reception)	Indispensable (transmission)
GET_PARAMETER		
OPTIONS	Optional (transmission)	Indispensable (reception)
PAUSE	Indispensable (transmission)	Indispensable (reception)
PLAY	Indispensable (transmission)	Indispensable (reception)
RECORD		
REDIRECT		
SETUP	Indispensable (transmission)	Indispensable (reception)
SET_PARAMETER		
TEARDOWN	Indispensable (transmission)	Indispensable (reception)

If the receiver receives a method which it does not support, it shall respond with Error Code 501 "Not Implemented" in accordance with the RFC (RFC 2326[10]) specifications.

4.1.1.4 Handling of Headers

In this document, the headers that are defined in RFC2326 ([12] Header Field Definition) shall be operated as shown in Table 4-2. The meanings of the terms "indispensable" and "optional" in the table and the handling of headers which are not defined in the table shall be as shown below.

- Transmitting side

  Indispensable: The header must be inserted.

  Optional: The header may be inserted. (The receiving side returns the appropriate response.)

  Not defined: The header need not be inserted. Even when it is inserted, there is no guarantee that the receiving side returns the appropriate response.

- Receiving side

  Indispensable: When the header has been inserted, the receiving side starts the appropriate action and returns a Response message. If the header has not been inserted, the receiving side returns the appropriate error code.

  Optional: When the header has been inserted, the receiving side starts the appropriate action and returns a Response message. When the header has not been inserted, the receiving side makes no action.

  Not defined: Even when the header has been inserted, the receiving side need not make any action. (It should be noted, however, that the receiving side is not prohibited from returning the corresponding Response.)

For the headers required for each RTSP messages and those handling, see 4.1.1.7 "Details of RTSP Messages". The FEC_Code header is one that is not defined in RFC. It is the optional header that the receiver uses to notify the supporting FEC to the server.

Concerning the Accept header, the receiver is not required to include it in DESCRIBE request. Thus, this header is optional.

The Location header shall not be used.

Table 4-2: Operation of RTSP Headers

Header	Type	Indispensable/optional	Method
Accept	R	Optional	DESCRIBE
Content-Length	r	Indispensable	DESCRIBE
Content-Type	r	Indispensable	DESCRIBE
CSeq	Rr	Indispensable	all
Notice	R	Indispensable	ANNOUNCE
Public	r	Indispensable	OPTIONS
Range	Rr	Indispensable	PLAY
Range	r	Indispensable	PAUSE
Scale	R	Optional*1	PLAY
Scale	r	Indispensable	PLAY
Session	R	Indispensable	PLAY, PAUSE, TEARDOWN, ANNOUNCE
Session	r	Indispensable	SETUP, PLAY, PAUSE, TEARDOWN, ANNOUNCE
Transport	Rr	Indispensable	SETUP
FEC_Code	R	Optional*2	DESCRIBE

Header: The name of a header.

Type: R: Header used in a Request message.

r: Header used in a Response message.

Rr: Header used in both a Request message and a Response message.

Method: The name of a method that uses the header. "All" indicates that the header is used in all methods.

*1: When this header is omitted, the server operates as "Scale: 1". "Scale: 1" must not be explicitly specified (see 4.1.1.7.4).

*2: This header is used when the receiver notifies FEC it supports (see 4.1.3.4).

4.1.1.5 Matters Common to Headers

4.1.1.5.1 Handling of Each Line

Each line of the format described below must terminate with a (CR+LF). Also, the boundary between message-headers and entity-body (e.g., SDP) is indicated by a (CR+LF). For example, when the server transmits a response with an SDP entity body, the response includes the last header + (CR+LF) + (CR+LF) + SDP body sequence. The end of entity-body (Message Length) depends on the entity-body and the Content-Length value.

When there are no entities, the message is terminated with a CR + LF. Namely, the message ends with the last header + (CR+LF) + (CR+LF). If the server receives an incomplete message, such as a final header + (CR+LF), it may start timeout processing after a certain time has elapsed. (When the server receives an additional CR+LF within that time, it processes the message as one complete message.)

4.1.1.6 Status Code of Response Message

The RFC status code (RFC2326 [7.1.1 Status Code and Reason Phrase]) shall apply. Unless otherwise specified, if the format of the description of a Request message is wrong, the appropriate error code (400 Bad Request) shall be returned.

4.1.1.7 Details of RTSP Messages

The formats of RTSP messages are defined below. The maximum size of an RTSP Response message shall be 4 KB.

4.1.1.7.1 OPTIONS

The receiver obtains a list of the methods that are supported by the server. The recevier is not required to inquire about methods supported by the server prior to the playback of content. The receiver can make a request under any condition, and the request does not affect the state of the server.

In the CSeq header, a sequence number defined in RFC (RFC2326 [12.17]) is set. In the Public header of a Response message, the list of the methods that are supported by the server is set in accordance with the RFC (RFC2326 [12.28]) specifications.

■ Request message

```
OPTIONS * RTSP/1.0
CSeq:　CSeq_Number
```

■ Response message

```
RTSP/1.0 Status_Code Reason_Phrase
CSeq: CSeq_Number
Public: (List of the set of methods supported by the server)
```

4.1.1.7.2 DESCRIBE

The receiver obtains from the server the information about the content specified by a Request message. The receiver can obtain said information in any state, and the server must interpret the request under any condition. The exchange of these messages do not affect the receiver state or server state.

The Accept header in a Request message shall follow the RFC (RFC2326 [12.1]) specifications. It is possible to explicitly request an SDP by including the Accept header in the message. When the Accept header is not included, both the receiver and the server must work as if "application/sdp" were specified by the Accept header. As the FEC_Code header, the FEC type/matrix parameter supported by the receiver and converted into a 4-digit hexadecimal value shall be set. For details, see 4.1.5.3.8 "Turning on/off FEC Function".

As the Content-Type header in a Response message, the media type of information about the content to be arranged in the entity body shall be set in accordance with the RFC (RFC2326 [12.16]) specifications. Since the information about the content is described by the SDP format, "application/sdp" shall be set. When the server receives a DESCRIBE request which does not contain the Accept header, it must work as if "application/sdp" were specified by the Accept header. In this case, "application/sdp" must be specified in the Content-Type header.

The Content-Length header shall be set in accordance with the RFC (RFC2326 [12.14]) specifications.

SDP is defined in 4.1.1.8 "SDP".

■ Request message

```
DESCRIBE   rtsp_URL    RTSP/1.0
CSeq:   CSeq_Number
[Accept:   application/sdp]
[FEC_Code: (Type of the FEC, FEC parameters)]
```

■ Response message

```
RTSP/1.0 Status_Code Reason_Phrase
CSeq: CSeq_Number
Content-Type: (Type of the message body)
Content-Length: (Length of message body)

SDP   body
```

4.1.1.7.3 SETUP

SETUP sets the transport parameter that is used during exchange of data between the receiver and the server.

In the Transport header, the format of the file to be delivered by the layer above the transport layer is set in accordance with the RFC (RFC2326 [12.39]) specifications. In the Session header in the Response message, the session ID that is managed at the server side is set in accordance with the RFC (RFC2326 [12.37]) specifications. In addition, in accordance with the RFC specifications, it is possible to set a timeout time using "timeout". The set timeout time shall be effective as long as the receiver remains in the Ready state or in the Playing state. Once it enters the Init state, the timeout time shall return to the prescribed 60 seconds. When "timeout" is not set, the receiver shall operate on the assumption that a timeout occurs in 60 seconds.

■ Request message

```
SETUP   rtsp_URL    RTSP/1.0
CSeq:   CSeq_Number
Transport: Transport_Parameter
```

rtsp_URL

The maximum length shall be 1,024 bytes.

Transport_Parameter in the Request message is defined as follows.

Transport_Parameter	=	"RTP/AVP" ["/UDP"] ";" parameter
parameter	=	unicast_parameter
unicast_parameter	=	"unicast" ";" "client_port" "=" port
port	=	1*5(DIGIT)

■ Response message

```
RTSP/1.0 Status_Code Reason_Phrase
CSeq: CSeq_Number
Session: Session_ID (;timeout=Timeout_time)
Transport: Transport_Parameter
```

Transport_Parameter in the Response message is defined as follows.

Transport_Parameter	=	"RTP/AVP" ["/UDP"] ";" parameter
parameter	=	unicast_parameter
unicast_parameter	=	"unicast" ";" uni_param
uni_param	=	["client_port" "=" port ";"] (Note 1)
		["source" "=" address ";"] (Note 2)
		"server_port" "=" port
address	=	host
port	=	1*5(DIGIT)

Note 1: The server must not change the port number that the receiver specifies as a request parameter.

Note 2: The "source" parameter shall be used when the address of the RTSP server is different from that of the media data (RTP) server.

4.1.1.7.4 PLAY

PLAY request gives an instruction to start playing the specified content or resume playing the content that has been stopped temporarily.

In the Range header, a value which consists of a maximum of five integral digits and one decimal digit shall be set in the npt-sec format of NPT (Normal Play Time) that is compatible with RFC2326 ([12.29]). The parameter "npt=" may be omitted.

In the Scale header, a viewing rate for trick playback which consists of a maximum of three integral digits and one decimal digit shall be set in accordance with RFC2326 [12.34]). For the value requested from the receiver, the server determines a value nearest to the viewing rate supported by the server for the content and returns the viewing rate at which the server can actually transmit the content. It should be noted, however, that a discrepancy between the value requested from the receiver and the value implemented by the content on the server does not normally occur since the receiver previously obtains the viewing rate supported by the server for the content from the Network content Control Information (NCI) defined in 5.3.3 "NCI" and sets that value in the Scale header. The receiver must not set 1 in the Scale header. In the case of normal playback, the Scale header shall not be included in the Request message. If the normal viewing rate (Scale value = 1) is requested, the server shall return the appropriate error code 406 "Not Acceptable", rather than respond with Scale: 1.0. If the Scale header is omitted from the Request message, the server operates with the normal viewing rate (Scale: 1).

■ Request message

```
PLAY   rtsp_URL    RTSP/1.0
CSeq:   CSeq_Number
Session: Session_ID
Range: Start_Time-[End_Time]
[Scale: Scale_Value]
```

■ Response message

```
RTSP/1.0 Status_Code Reason_Phrase
CSeq:   CSeq_Number
Session: Session_ID
Range: Start_Time-[End_Time]
Scale: Scale_Value
```

4.1.1.7.5 PAUSE

PAUSE request stops playing the specified content temporarily.

The stream position at which the server stops transmitting the stream shall be returned by the Range header in the Response message. If the receiver receives a header of the format "Range: x-y", the header shall be handled as "Range: x".

■ Request message

```
PAUSE   rtsp_URL    RTSP/1.0
CSeq:   CSeq_Number
Session: Session_ID
```

■ Response message

```
RTSP/1.0 Status_Code Reason_Phrase
CSeq:   CSeq_Number
Session: Session_ID
Range: Pause_Time
```

4.1.1.7.6 TEARDOWN

TEARDOWN request terminates the service of the specified content.

■ Request message

```
TEARDOWN    rtsp_URL     RTSP/1.0
CSeq:    CSeq_Number
Session: Session_ID
```

■ Response message

```
RTSP/1.0 Status_Code Reason_Phrase
CSeq:    CSeq_Number
Session: Session_ID
```

4.1.1.7.7 ANNOUNCE

ANNOUNCE request notifies the reason for stopping the stream from the server to the client receiver. This message is transmitted from the server asynchronously. It should be noted, therefore, that the value of CSeq in the Request message has nothing to do with the value of CSeq that has been transmitted from the receiver.

The Notice header in the Request message is an extended specification independent of the RFC2326 specifications. The Notice header is added to the ANNOUNCE request in order for the video content server to notify the reason for stopping the stream to the receiver. The format shall be as shown below.

```
Notice: Event_Code Event_Phrase

Event_Code     =   4DIGIT
Event_Phrase   =   *<TEXT, excluding CR, LF>
```

In this document, as Event_Code and Event_Phrase, any one of the combinations shown in the following table shall be described. If the receiver receives an ANNOUNCE request with Event_Code and Event_Phrase not shown in the table, it shall properly ignore the request.

Table 4-3: Event_Code and Event_Phrase in Notice Header

Event_ Code	Event_Phrase	Operation		Description
		Receiver	Server	
2101	End-of-Stream Reached	○	○	The content has ended.
2104	Start-of-Stream Reached	○	○	The content has returned to its beginning.
5401	Downstream Failure	○	○	The stream cannot be transmitted.
5404	Internal Server Error	○	○	A server error has occurred.

* In the "Operation" column, "○" indicates that Event_Code and Event_Phrase are operated.

■ Request message

```
ANNOUNCE    rtsp_URL     RTSP/1.0
CSeq:    CSeq_Number
Session: Session_ID
Notice: Event_Code Event_Phrase
```

■ Response message

```
RTSP/1.0 Status_Code Reason_Phrase
CSeq:    CSeq_Number
Session: Session_ID
```

4.1.1.8 SDP

The description of SDP for presenting the information that is necessary between the receiver and the video content server in the streaming of content shall conform to RFC4566 "SDP: Session Description Protocol". In this document, in particular, the following types of information are defined. When the parameter in the RTSP SETUP response contains information on the meaning equal with any of those types of information during streaming reception by the receiver, priority shall be given to that information.

(1) Information relating to presentation

i) Owner/creator and session identifier

o=username session_id version network_type address_type address

The video content server may set any value in this field

ii) Session name

s=session_name

The video content server may set any value in this field.

(2) Information relating to control

i) Time of availability (time the session is active)

t=start_time end_time

In this document, 0 shall always be specified for both start_time and end_time.

ii) Range of presentation (content length)

a=range:npt=start_time-end_time

The range over which the content can be supplied is referred to during jump playback, progress display, etc. Expressed in terms of time, the range is described as follows.

Each of start_time and end_time shall be specified by a value consisting of a maximum of five integral digits and one decimal digit of npt-sec format defined in RFC2326 (Normal Play Time). For start_time, 0 shall always be specified. For end_time, the duration (length) of the content shall be specified. The parameter "npt=" may be omitted.

(3) Information relating to decoding

i) Media description

m=media port transport_protocol fmt_list(payload_type)

When MPEG2 is used as the video coding system, the "m=" line shall be described as follows.

m=video 0 RTP/AVP[/UDP] 104

When H.264/AVC is used as the video coding system, the media shall be described as follows.

m=video 0 RTP/AVP[/UDP] 105

For fmt_list(payload_type) specified here, the a=rtpmap description shall be added. The description of a=rtpmap for MPEG2 and H.264/AVC, respectively, is as follows. In SDP, only one line of a=rtpmap relating to media description shall be described.

a=rtpmap:104 vnd.iptvforum.ttsmpeg2/27000000

a=rtpmap:105 vnd.iptvforum.ttsavc/27000000

ii) Transfer rate

a=bitrate:bitrate

As "bitrate", the MPEG2-TS rate shall be specified.

The "bitrate" shall be described as a positive integer in bps.

(4) Information relating to FEC

i) FEC type

When Pro-MPEG FEC is used, 96 shall be added to payload_type in the media description. Thus,

m=video 0 RTP/AVP[/UDP] 104 96

m=video 0 RTP/AVP[/UDP] 105 96

In addition, in the a=rtpmap description, a description relating to FEC type shall be added in the following format.

a=rtpmap:payloadtype fec_mode/clock frequency

For example, in the case of Pro-MPEG 1D (L*D = 10*10) in media description, the description relating to FEC type becomes as follows.

a=rtpmap:96 vnd.iptvforum.1dparityfec-1010/8000

For the FEC types that can be used and the method of describing them, see 4.1.3.

4.1.1.9 Detailed Specifications for Operation of RTSP Control

(1) Specification on position at which to start stream transmission at start of PLAY

The video content server shall start delivering the stream from the beginning of the GOP that includes Start_Time requested by the Range header in PLAY request from the receiver, even when resuming PLAY from PAUSE during normal playback.

(2) Specification of Range header (End_Time) in PLAY request/response

The receiver is not absolutely prohibited from specifying any value for End_Time of the Range header when making a PLAY request. It should be noted, however, that there are cases in which the video content server cannot properly respond to the specifications of any value for End_Time. In any case, the value specified for End_Time shall not exceed end_time of the range of presentation (content length) that is specified in "a=range" line of the SDP description.

When End_Time is specified in a PLAY request from the receiver, the video content server shall normally return a PLAY response to the receiver even if it fails to properly respond to the End_Time specification and simply delivers the content to the end. The value set for End_Time of the Range header in the PLAY response shall be such that even when the server transmits the content up to the specified point, the server can perform the same processing (i.e., sending out an ANNOUNCE request) as when it transmits the content to the end. The video content server that cannot properly respond to the specifications of any value for End_Time shall set for End_Time in the PLAY response either the same value as explicitly notified by end_time in the a=range line of SDP or a null and return it to the receiver in order to notify that the server is going to transmit the content to the end.

(3) Specifications on connection status during content transmission/reception

The TCP connection shall not be closed during the RTSP session. The video content server transmits VOD content only when the connection of the TCP connection in RTSP is maintained. When the TCP connection is disconnected, the server shall stop delivering the stream.

(4) Operation specifications on session close due to timeout

Once an RTSP session is established by the SETUP response, the receiver shall, in order to maintain the RTSP session, transmit a method or heartbeat to the server at intervals of time shorter than the timeout value specified by the Session header in the SETUP response. If the server receives neither a method nor a heartbeat from the receiver in a period of time longer than the timeout value, the server shall be allowed to close the session and stop the stream. As the heartbeat, the simplest text CR+LF (0x0d, 0x0a) shall be used. From a TCP Ack returned from the server, the receiver can confirm that the session is maintained.

(5) Guidelines on limitation to maintenance of PAUSE state

As long as the RTSP session is maintained, the server keeps reserving the resource for streaming even if the transmission has been stopped by PAUSE. In view of this, the time for which the PAUSE state can be maintained shall be limited to 60 minutes. The server must not force the connection to be closed within 60 minutes, except when it cannot receive proper heartbeat described in (4). It is to be desired that the receiver should release the PAUSE state within 60 minutes.

(6) Specification on limitation to PLAY from the Playing state

In this document, it is prohibited to issue a PLAY request directly in the Playing state. At the timing of a switch from normal playback to trick playback or vice versa, the receiver must shift from the Playing state to the Ready state via a PAUSE before it can issue a PLAY request.

(7) Specification on Restriction of TEARDOWN from the Playing state

In this document, it is prohibited to issue a TEARDOWN request directly in the Playing state. In the processing for stopping the viewing of content, the receiver must shift from the Playing state to the Ready state via a PAUSE and obtain the npt value at the time of stopping from the Range header in the PAUSE response before it can issue a TEARDOWN request.

(8) Specification on receiver operation during reception of ANNOUNCE method

When the receiver receives an ANNOUNCE request from the server, it shall send out a PAUSE request without delay after returning an OK as the response to the server so as to shift to the Ready state. After the PAUSE request is executed properly, the receiver shall send TEARDOWN request to shift to the init state and perform the termination processing properly. It should be noted, however, that when the receiver receives an ANNOUNCE request from the server at the same time as or right after the user's operation, the receiver may ignore the ANNOUNCE request (it need not respond to the request).

On the other hand, the video content server must operate properly even if it receives from the receiver PAUSE→TEARDOWN or PAUSE→PLAY at the same time as or right after the transmission of an ANNOUNCE that has not been responded to by the receiver.

(9) Specification on value that can be set to Scale header in PLAY request

As the value to be set for the Scale header of a PLAY request when making a request for variable-speed playback, the receiver shall select one from among the values of the elements of scale_value that the receiver has received as NCI defined in 5.3.3 NCI.

Regardless of whether or not the viewing rate specified by the Scale header of the PLAY request is one of the values of the scale_value elements notified as NCI to the receiver, the video content server shall select a value nearest to the specified value from among the values that actually allow for stream transmission, notify the selected value in the PLAY response to the receiver, and actually send out the stream at that value.

(10) Guidelines on limitation to frequent issuance of RTSP method

It is to be desired that the receiver implementation should be such that the video content server will not be overloaded by frequent issuance of RTSP requests from the receiver.

4.1.2 RTP

4.1.2.1 Purpose

The purpose of RTP is to define the media transport layer for transmitting audio, video and other media in the VOD service.

Here, the data format of the media transport layer is defined on the basis of the specifications of RFC3550 "RTP: A Transport protocol for Real-Time Applications". Unless otherwise noted, the data format shall conform to the RFC specifications.

As the RTP transport layer protocol, the UDP shall be used. As the network layer protocol, either IPv4 or IPv6 shall be used.

4.1.2.2 RTP Header

In general, RTP is presupposed to be used for transfer of audio, video and other media in the IP broadcast and VOD services. Therefore, RTP is intended to be used mainly for recovery of timing, detection of packet loss and correction of packets, and identification of payload and source.

As general matters defined in the RFC, the RTP header configuration and the header fields are explained below.

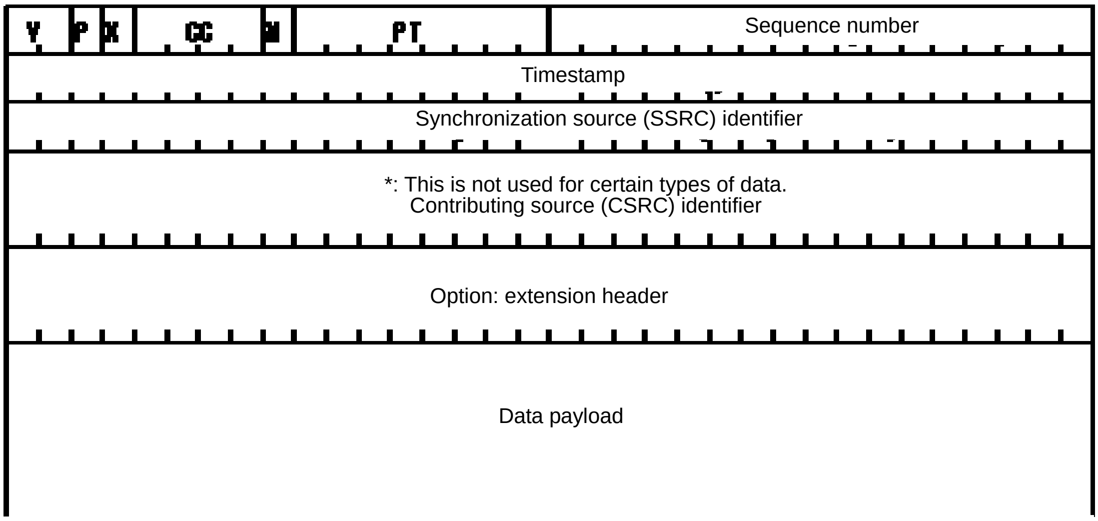

Figure 4-3 RTP Header Configuration

version (V): 2 bits

This field identifies the RTP version.

padding (P): 1 bit

When the padding bit is set, it indicates that the packet contains one or more padding bits at its end.

extension (X): 1 bit

When the extension bit is set, it indicates that the fixed header is followed by one extension header.

CSRC count (CC): 4 bits

The CSRC count indicates the number of CSRC identifiers which follow the fixed header.

marker (M): 1 bit

The marker indicates the boundary of frames shown in the packet stream.

payload type (PT): 7 bits

This field indicates the format type of the RTP payload.

sequence number: 16 bits

The initial value of the sequence number should be random. Incremented by 1 each time an RTP data packet is delivered, the sequence number indicates the sequence of packets. This field is used at the receiver side for packet sequence restoration, etc.

timestamp: 32 bits

The timestamp indicates the time at which the first octet sample of RTP data packet is obtained.

SSRC: 32 bits

SSRC is used as the identifier of each of RTP sessions which make up a single multimedia session (RFC3550 Section 3).

CSRC list: 0 to 15 items, 32 bits each

This list shows the identifiers of all data transmitters that are included in a single packet payload.

4.1.2.3 Specifications on RTP Header Transmission/Reception

The specifications on RTP header transmission/reception in the VOD service are shown in Table 4-4.

When an FEC is used, if the type of FEC used is subject to other restrictions on the RTP header operation in addition to those shown below, they shall also be observed.

Table 4-4: Specifications on RTP Header Transmission/Reception in VOD Service

Field	Field length	Transmitting specifications	Receiving specifications	
			Need for interpretation	Receiver processing specifications
version (V)	2 bits	Fixed to 0x2	Indispensable	0x2: The receiver judges that the transmitting operation is performed in accordance with these specifications. Other than 0x2: The receiver operatioins depend on the receiver implementation when it receives this version. The other receiver operations shall depend on the receiver implementation.
padding (P)	1 bit	Fixed to 0	Indispensable	0: The receiver judges that padding is not used. 1: The receiver operations depend on the receiver implementation when padding is used. The other receiver operations shall depend on the receiver implementation.
extension (X)	1 bit	Fixed to 0	Indispensable	0: The receiver judges that the extension header is not used. 1: The receiver operations depend on the receiver implementation when the extension header is used. The other receiver operations shall depend on the receiver implementation.

Field	Field length	Transmitting specifications	Receiving specifications	
			Need for interpretation	Receiver processing specifications
CSRC count (CC)	4 bits	Fixed to 0	Indispensable	0: The receiver judges that CSRC is not used. Other than 0: The receiver operations depend on the receiver implementation when CSRC is operated. The other receiver operations shall depend on the receiver implementation.
marker (M)	1 bit	Arbitrary	Optional	The receiver operation based on the marker bit shall depend on the receiver implementation. When the receiver doesn't support the bit, it shall properly ignore the bit.
payload type (PT)	7 bits	For the media packet (MPEG2-TTS), PT is 104 (when video coding system is MPEG2 video) or 105 (when video coding system is H.264/AVC). For PT of the FEC packet, see 4.1.3.	Optional	For a packet which is judged to be a media packet from the UDP port number, the receiver shall normally process the packet with PT 104 or 105 as an MPEG2-TTS. For a packet whose PT is neither 104 nor 105, the operation performed shall depend on the receiver implementation. For a packet which is judged to be an FEC packet from the UDP port number, see 4.1.3.
sequence number	16 bits	This field of the packet that is transmitted first after the PLAY method shall have any value. After that, the value shall be incremented by 1 each time a packet is delivered. When the value reaches FFFFh, it shall return to 0000h.	Indispensable	The receiver determines from the sequence number the presence or absence of missing RTP packets and the orderliness of RTP packets.
timestamp	32 bits	Basically, 90 kHz shall be used. However, this specification need not be strictly observed.	Optional	The operation to be performed based on timestamp shall depend on the receiver implementation.
SSRC	32 bits	A random value shall be set. The set value shall be reset each time the transmission of RTP is switched by the PLAY request.	Optional	The receiver shall detect from the change of SSRC the change of transmission by the PLAY method.

Field	Field length	Transmitting specifications	Receiving specifications	
			Need for interpretation	Receiver processing specifications
CSRC list	CC*32bit	CSRC shall not be operated.	Optional	When CSRC is operated, The receiver operations depend on the receiver implementation. The other receiver operations shall depend on the receiver implementation.

4.1.2.4 TTS Packet Carrying

The number of time-stamped TS packets carried over an RTP packet shall be variable, from 1 to 7. (It is to be desired that the number should be 7.)

4.1.2.5 RTCP

RTCP is a protocol that controls the session for data transmission and reception by RTP. By periodically transmitting an RTCP packet to the server, it is possible for the receiver to adjust the rate of data transfer, etc. from the server. In this document, the RTCP shall not be used.

4.1.3 FEC

If a packet loss occurs on the network, it can disturb the video and audio. It is, therefore, necessary to provide for the occurrence of packet losses. When a packet loss does occur, it needs to be restored using FEC.

Both the transmitting side and the receiver are able to implement FEC. As long as the same FEC is used in data transmission and reception, it can be used by both. Implementing FEC shall be optional.

When FEC is installed in the receiver, it shall as a rule conform to RFC2733 and its extension. In these specifications, "Pro-MPEG FEC Code of Practice #3 release 2" (hereinafter referred to as "Pro-MPEG FEC") shall be adopted. It should be noted, however, that it is allowable for both the server and the receiver to implement only FEC on column (hereinafter referred to as "Pro-MPEG 1D FEC") (see 4.1.3.5). When the server adopts an FEC system which is not currently defined in these specifications, as well as when the server operates Pro-MPEG FEC two-dimensionally (hereinafter referred to as "Pro-MPEG 2D FEC"), including FEC packets on row(see 4.1.3.5), the server shall employ an FEC system which is capable of data transmission and reception using different ports for media packets and FEC packets so that even a receiver which is not equipped with FEC can receive them. In this case, the receiver without FEC may ignore the FEC packets and process only the media packets. The FEC procedure is explained below.

4.1.3.1 Method of FEC Protection Processing

Figure 4-4 shows the structures of media packets and FEC packets in the IP layer.

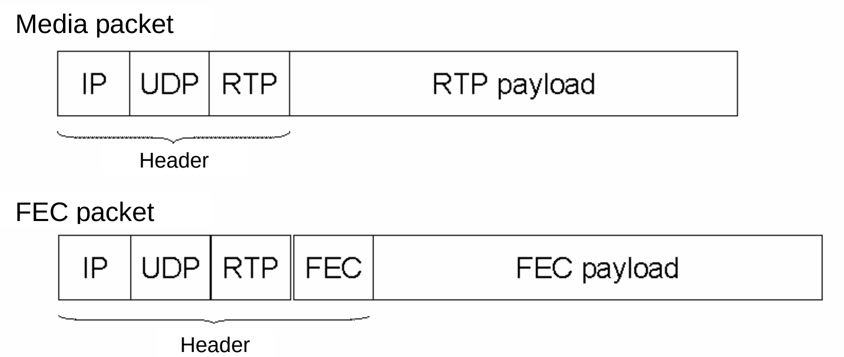

Figure 4-4 Structures of Media Packet and FEC Packet

For the media packet that is to be protected, the RTP header is referred to and the necessary calculations are performed to generate an FEC header and FEC payload.

The FEC header configuration defined in RFC2733 is shown in Figure 4-5. It should be noted that when Pro-MPEG FEC is used, this FEC header needs to be extended. For details about the extended FEC header when Pro-MPEG FEC is used, see 4.1.3.5(2).

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      SN base                  |         length recovery       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|E| PT recovery |                 mask                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          TS recovery                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4-5 FEC Header Configuration

The method of generating an FEC header is as follows. For details about the FEC header operation when Pro-MPEG FEC is used, see 4.1.3.5.

- SN base (16 bits): Minimum sequence number of the RTP packets associated with the FEC packet.
- Length recovery (16 bits): The XOR value for the total length of the payload, CSRC list, extension header and padding of the RTP packets associated with the FEC packet. This is used during packet recovery to determine the packet length of the packet to be recovered.
- E (1 bit): The bit indicates a header extension. Ordinarily, it is "0". It shall be set to "1" to indicate that the header is extended. When Pro-MPEG FEC is used, this bit shall be set to "1" because the FEC header is extended.
- PT recovery (7 bits): The result of XOR operation on the payload type value of the RTP header.

- Mask (24 bits): The flag that indicates the RTP packets associated with the FEC packet. When RTP packets with sequence number "N+i" is associated with FEC packet, the i bit is set to "1". When Pro-MPEG FEC is used, this flag is not used (all bits are "0".)
- TS recovery (32 bits): The result of XOR operation on the timestamp value of RTP header.

Generating an FEC packet involves concatenating a specific fields, payload, and padding with zeros(if any) from the RTP header of the media packets and then computing the XOR across the resulting bit string generated by the concatenation. This bit string, called an FEC bit string, is used to generate the FEC packet.

The fields that are concatenated to obtain a bit string are as follows.

- Padding Bit (1 bit)
- Extension Bit (1 bit)
- CC bits (4 bits)
- Marker bit (1 bit)
- Payload Type (7 bits)
- Timestamp (32 bits)
- Total length of CSRC list, header extension, payload and padding (16 bits)
- CSRC list (variable length) if CC is not 0
- Header extension (variable length) if X is 1
- Payload (variable length)
- Padding (variable length), if present

For each media packet, an FEC payload can be generated by concatenating the above values and then computing the XOR. The generation of an FEC packet is conceptually shown in Figure 4-6.

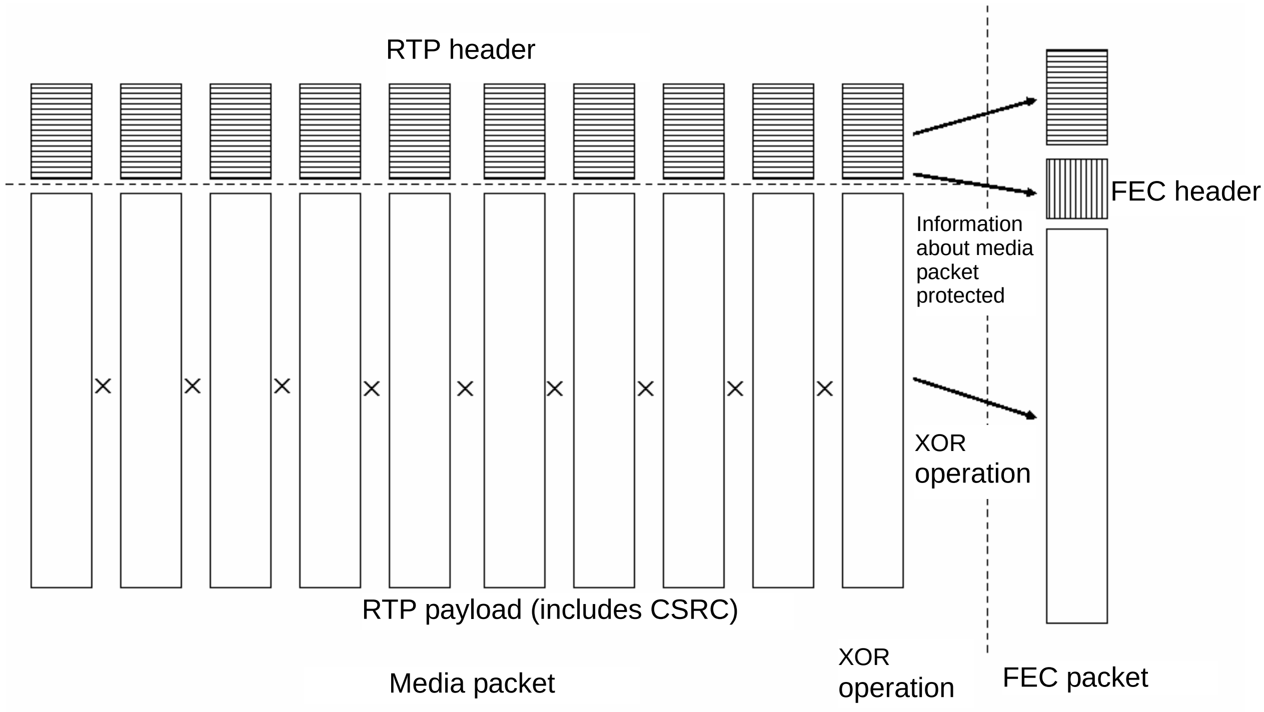

Figure 4-6 Concept of FEC Packet Generation

In order to permit even a receiver without FEC capability to receive an FEC packet stream, the stream shall be sent from a port number which is different from that for media packets. Any receiver without the FEC capability may ignore FEC packets and receive only media packets.

4.1.3.2 Method of FEC Recovery Processing

The procedure for recovering media packet xi from a total T of media packets and FEC packets is as follows.

(1) Generate the bit string for the media packets in T.

(2) Generate the bit string for the FEC packets in T. In this calculation, use PT Recovery, instead of Payload Type, and TS Recovery, instead of Timestamp, and set the CSRC list, extension and padding to null.

(3) If any bit strings generated from the media packets is shorter than that generated from the FEC packets, pad them to the same length as the bit strings generated from the FEC. (The padding MUST be added at the end of the bit string and may be any value.)

(4) Perform the exclusive or parity operation across the bit string to resulting in a recovery bit string.

(5) Generate a new packet with a standard 12-byte RTP header and no payload.

(6) Set the version of the new packet to 2.

(7) Set the Padding bit in the new packet to the first bit in the recovery bit string.

(8) Set the Extension bit in the new packet to the second bit in the recovery bit string.

(9) Set the CC field to the next four bits in the recovery bit string.

(10) Set the Marker bit in the new packet to the next bit in the recovery bit string.

(11) Set Payload Type in the new packet to the next seven bits in the recovery bit string.

(12) Set the SN field in the new packet to xi.

(13) Set the TS field in the new packet to the next 32 bits in the recovery bit string.

(14) Take the next 16 bits of the recovery bit string. No matter what unsigned integer this represent, take that many bytes from the recovery bit string as indicated by the integer (these correspond to the CSRC list, extension, payload and padding) and append them to the new packet.

(15) Set SSRC of the new packet to the SSRC of the media stream protected.

The concept of media packet recovery is shown in Figure 4-7

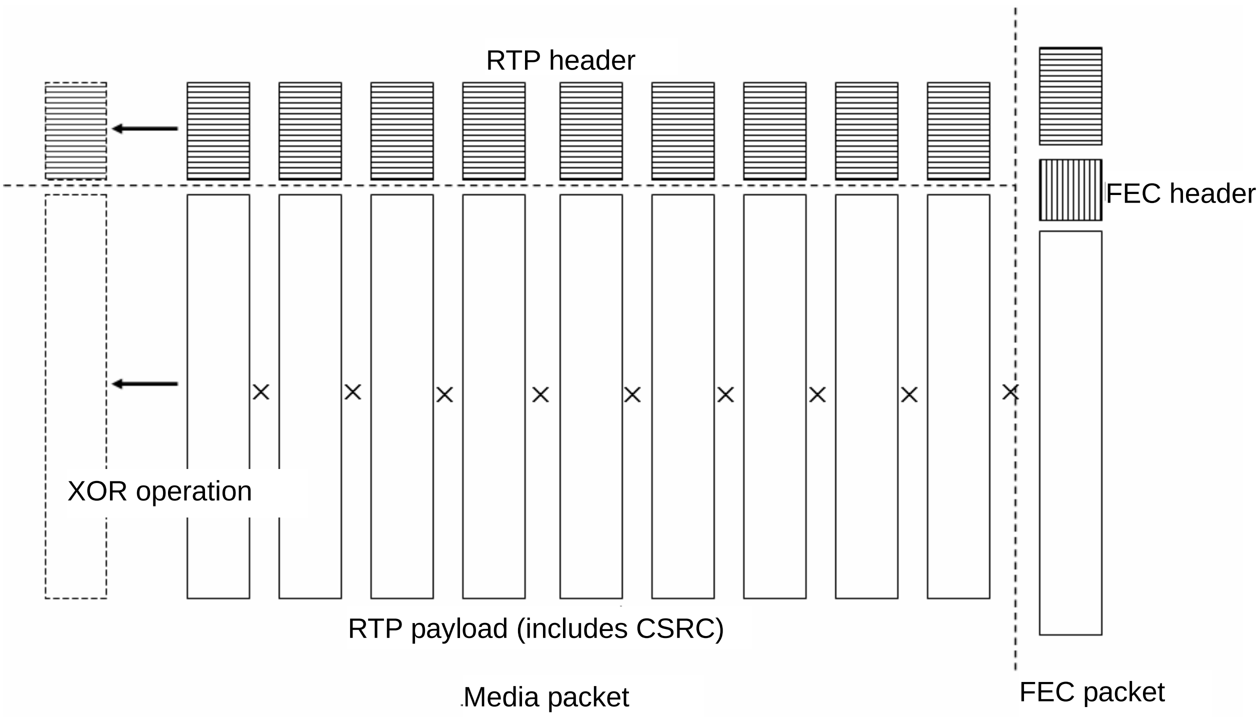

Figure 4-7 Concept of Media Packet Recovery

Although it is assumed that the method of determining whether or not a sufficient amount of data is available for the recovery of a lost packet depends on the available receiver implementation, examples of the algorithm and functions for packet recovery are shown in RFC2733 8.2 "Determination of When to Recover".

4.1.3.3 Method of Notifying FEC Information

The information that the receiver must know in order to receive an FEC packet and decode the FEC includes the following.

- FEC system

- FEC parameters
- FEC packet source IP address
- FEC packet destination port number

Concerning the FEC parameters, the number of parameters and the format differ from one FEC system to another. Therefore, the FEC type, including these, shall be defined and communicated to the receiver.

In the case of a unicast stream, a method of describing the FEC stream and the stream to be protected in SDP is given in RFC2733, Chapter 11 "Indicating FEC Usage in SDP". The FEC type shall be notified to the receiver using this method. The format is as follows.

a=rtpmap:[payload type] [FEC type]]/[clock frequency]

The designation of FEC types is shown in Table 4-5.

Table 4-5: Designation of FEC Types

Designation in SDP (FEC type)	FEC system	Parameters
vnd.iptvforum.1dparityfec-1010 (*)	Pro-MPEG 1D FEC	L*D=10*10 source IP address=same as media packet dest_port= Media packet + 2
vnd.iptvforum.1dparityfec-2005	Pro-MPEG 1D FEC	L*D=20*5 source IP address=same as media packet dest_port= Media packet + 2
vnd.iptvforum.2dparityfec-1010	Pro-MPEG 2D FEC	L*D=10*10 source IP address=same as media packet dest_port= Media packet +2, +4
vnd.iptvforum.2dparityfec-2005	Pro-MPEG 2D FEC	L*D=20*5 source IP address=same as media packet dest_port= Media packet +2, +4

* Pro-MPEG FEC CoP3 and CoP3 release 2 recommend that both 1D and 2D be designated as 2dparityfec. However, in order to discriminate 1D from the simple parity FEC and Pro-MPEG 2D FEC defined in RFC2733, it shall be defined as 1dparityfec.

An example of the description using the SDP is given below.

a=rtpmap:96 vnd.iptvforum.1dparityfec-1010/8000

The above description means that Pro-MPEG 1D (L*D=10*10) FEC is used with payload type 96.

In SDP, a=rtpmap relating to FEC shall be described in only one line. When a new FEC type is used, the designation of FEC shall be newly defined.

4.1.3.4 ON/OFF of FEC Function

In order to provide for a possible change in communication network quality and for an instruction for trick playback of VOD content, both the server side and the receiver side shall be able to turn on/off the FEC function.

When the FEC function is to be turned off at the server side, the FEC Type shall not be described in SDP. When the FEC function is to be turned off at the receiver side, the receiver shall ignore the FEC packet.

The receiver side shall be able to notify the server side which FEC system and matrix parameters it supports by using the FEC_Code header as the request header of DESCRIBE in RTSP. Table 4-6 shows the relationships between the FEC system, matrix size parameter and flag position.

Table 4-6: FEC System, Matrix Size Parameter and Flag Position

Flag position	FEC system	Matrix size parameter
1st bit	Pro-MPEG 1D FEC	10×10
2nd bit	Pro-MPEG 1D FEC	20×5
3rd bit	Pro-MPEG 2D FEC	10×10
4th bit	Pro-MPEG 2D FEC	20×5
5th to 16th bits	Reserved	Reserved

For each set of FEC system and matrix size parameter that the receiver supports, the receiver sets 1 as the flag in the appropriate bit, and for each set of FEC system and matrix size parameter that the receiver does not support, the receiver sets 0 in the appropriate bit, as well as in the reserved bits. This way the receiver generates a 16-digit binary number first. Then, the receiver converts the number into a 4-digit hexadecimal number and transmits it as the FEC_Code header to the server. Here are two examples. When the receiver supports Pro-MPEG 1D FEC 10×10 and 20×5, the 16-digit binary number becomes 1100000000000000. So, the receiver transmits FEC_Code: C000 by DESCRIBE request. When the receiver supports Pro-MPEG 1D, 2D FEC 10×10, the 16-digit binary number becomes 1010000000000000. In this case, the receiver transmits FEC_Code: A000. The FEC_Code header may be omitted. When it is omitted, however, the receiver is assumed to be incompatible with the FEC. This is also the case when the receiver omits DESCRIBE and starts with a SETUP in the RTSP sequence. Therefore, any receiver that supports FEC must always start the RTSP sequence with a DESCRIBE and transmit the FEC_Code header to the server.

Any service provider that wants to operate the FEC efficiently shall perform an AND operation on the FEC flag transmitted from the receiver and the flag of the FEC supported by the server and shall transmit the FEC system and matrix parameter for which the result of the AND operation gives flag 1. If there are no such flags, the provider shall turn off the FEC. If there is more than one such flag, the provider shall, on the basis of the order of priority determined previously, select the appropriate FEC system and matrix parameter, describe the relevant information in SDP as defined in 4.1.3.3, and notify it to the receiver. It should be noted that since some providers operate FEC in their own way without regard to the FEC information provided by the receiver, the receiver might find that the FEC information it supplies to the server is ignored. When the FEC system and matrix parameter that the receiver has notified that it supports by FEC_Code are returned by an SDP from the server, the receiver must receive the FEC packet.

4.1.3.5 FEC Specifications

(1) Additional limitations on RTP header when Pro-MPEG FEC is used

In accordance with the Pro-MPEG FEC specifications, the following limitations on the operation of the media packet RTP header shall be added to the specifications described in 4.1.2.3.

- padding (P): P shall always be 0.
- extension (X): X shall always be 0.
- CSRC count (CC): CC shall always be 0.
- marker (M): M shall always be 0.

These additional limitations, as well as the fields whose operation is fixed by these specifications, make it evident to the receiver that these fields are always 0 in the FEC packet too. They also make it possible to simplify the header recovery processing in the FEC recovery processing.

SSRC in the RTP header of the FEC packet shall be the same as SSRC of the media packet. As the payload type of the FEC packet, the value 96 as recommended in the Pro-MPEG FEC Specifications shall be used.

(2) FEC packet generation method/recovery method

First, the method of generating an FEC packet in Pro-MPEG FEC is explained below.

When Pro-MPEG 1D FEC or Pro-MPEG 2D FEC is used, an extended FEC header as defined in RFC2733 is used. The configuration of the extended FEC header is shown in Figure 4-8.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         SNBase low bits       |        Length Recovery        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|E| PT recovery |                     Mask                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          TS recovery                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|X|D|type |index|    Offset     |      NA       |SNBase ext bits|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4-8 Configuration of Extended FEC Header

The method of generating an extended FEC header is as follows.

- SNBase low bits (16 bits): The minimum sequence number of the RTP packets associated with the FEC packet.
- Length recovery (16 bits): The XOR value of the total length of the payload, CDRC list, extension header and padding of the RTP packets associated with FEC packet. This value is used during packet recovery to determine the length of the packet to be recovered.
- E (1 bit): The bit for header extension. To extend the FEC header, set this bit to 1.
- PT recovery (7 bits): The result of an XOR operation on the payload type of the RTP header.
- Mask (24 bits): Set all the bits to 0. (Alternatively, use the NA field.)
- TS recovery (32 bits): The result of an XOR operation on the Timestamp value of the RTP header.
- X (1 bit): Set this bit to 0. (This bit is reserved for header extension in the future.)
- D (1 bit): Set this bit to 0 for FEC on column (columns) and 1 for FEC on row (rows).
- Type (3 bits): To perform an XOR operation, set these bits to 0 (XOR = 0, hamming = 1, Reed-Solomon = 2).
- Index (3 bits): For the XOR method, set these bits to 0. (This field is used for more complex error protection code.)
- Offset (8 bits): Offset indicates the period of media packets associated with the FEC packet. Set Offset to L for FEC on columns and 1 for FEC on rows.
- NA (8 bits): NA indicates the number of media packets associated with this FEC. Set NA to D for FEC on columns and L for FEC on rows.
- SNBase ext bits (8 bits): Set all the bits to 0. (SNBase ext bits are used for protocols whose sequence number exceeds 16 bits.)

The operations that are performed for Pro-MPEG FEC ($L \times D = 10 \times 10$) are shown in Figure 4-9. In the case of Pro-MPEG 1D FEC, an operation is performed on the FEC header and FEC payload on column of the matrix to generate 10 column FEC packets. In the case of Pro-MPEG 2D FEC, an operation is also performed on row to generate 10 row FEC packets.

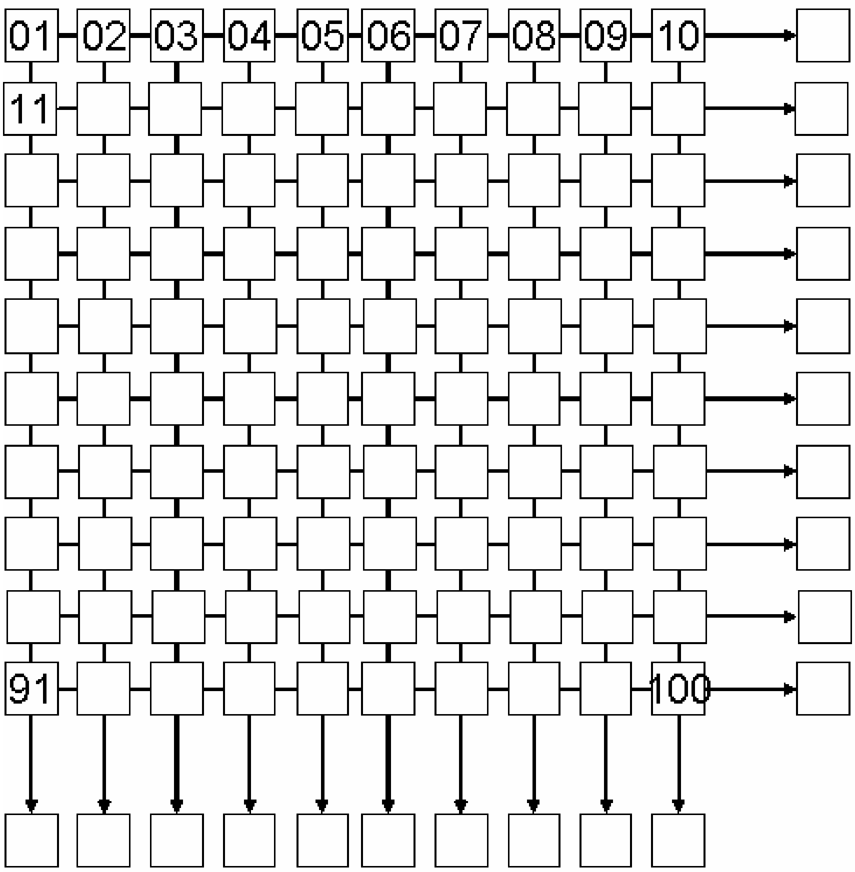

Figure 4-9 Operations Performed to Generate Pro-MPEG FEC Packets

For Pro-MPEG 1D FEC, the FEC packets on column are carried on the media packet port number + 2. For Pro-MPEG 2D FEC, the FEC packets on row, as well as those on column, are carried on the media packet port number + 4. Therefore, any receiver that has the FEC function needs to receive the FEC packets on row and column via the appropriate ports. Any receiver that does not support Pro-MPEG FEC may ignore FEC packets and needs to receive only media packets.

Next, the method of recovering a packet in Pro-MPEG FEC is explained below. In the case of Pro-MPEG 1D FEC, there are D media packets and one FEC packet on column. Even if one of the media packets is lost, it is possible to recover the lost packet by the method described earlier as long as the receiver receives a total of D packets. In the case of Pro-MPEG 2D FEC, there are D media packets and one FEC packet on column and L media packets and one FEC packet on row. Therefore, even if a packet is lost, it can be recovered as long as the receiver receives D packets on column or L packets on row. The recovery of a lost packet can be done by performing the calculation described in 4.1.3.2 (Method of FEC Recovery Processing) on column or row. Even if there are two or more lost packets in a row or in a column as shown in Figure 4-10, there are cases in which they can be recovered by first recovering the lost packets on column and then recovering the lost packets on row.

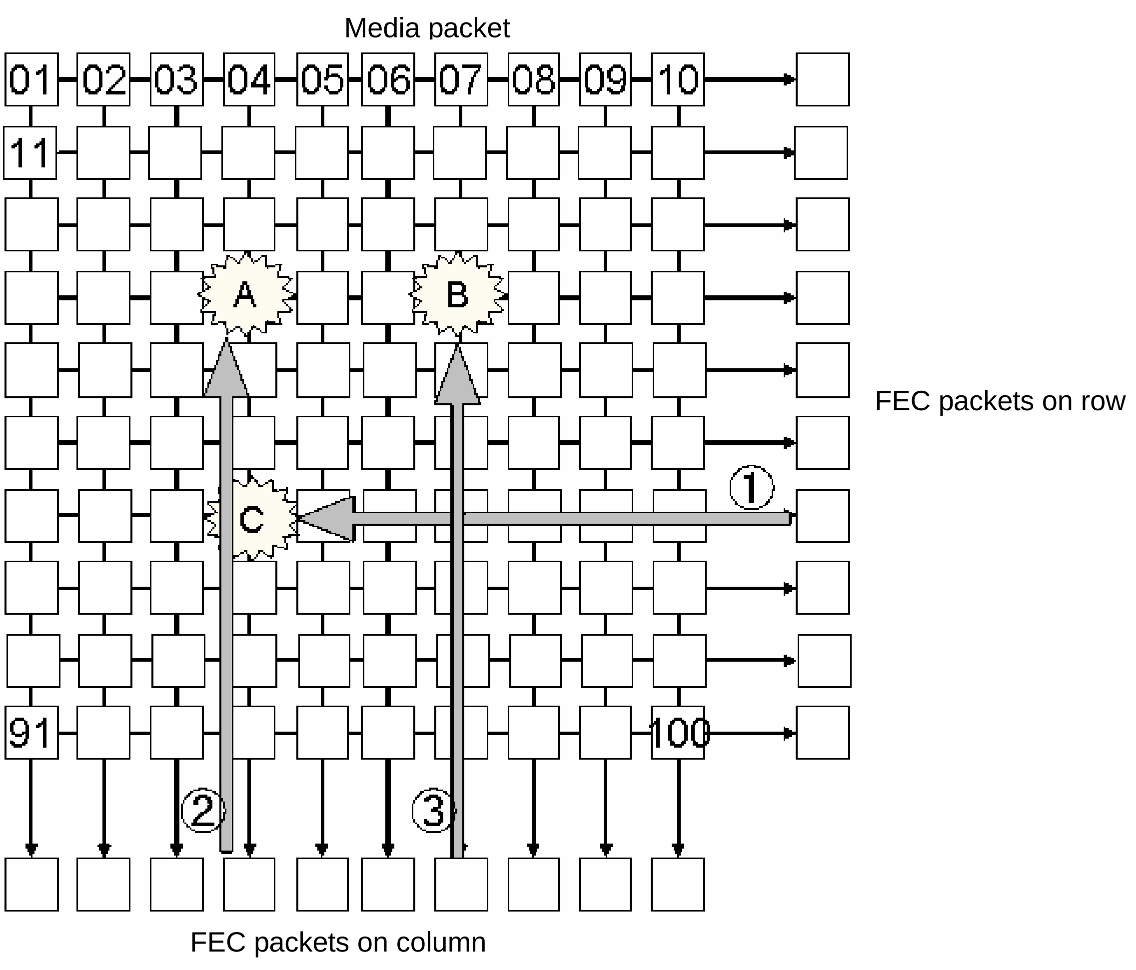

Figure 4-10　Recovery of Pro-MPEG 2D FEC Packets

In the above example, packet C is recovered first by operation (1) on row and then, packets A and B are recovered by operations (2) and (3) on column. This way, it might be possible to recovery two or more packets which are lost in the same row or column.

(3) Guidelines on FEC packet generation/transmission timings

As L × D matrices of media packets that are associated with FEC packet, only rectangular (square) ones shall be operated. Namely, FEC packets shall be generated and transmitted in such a manner that the SNBase numbers of L column FEC packets protecting the same matrix become consecutive.

In addition, all the FEC packets shall be operated in such a manner that they are completely transmitted before the media packet following the (L×D + L)th media packet is transmitted after completion of the transmission of the entire matrix associated with the FEC.

If a PLAY method is issued as a result of trick playback, etc. and the SSRC in the RTP header of the media packet is changed, media packets having a different SSRC shall not be mixed with the same matrix during the generation of FEC packets, and the generation of FEC packets using the former SSRC shall be limited to those FEC packets which can be generated at the time that the SSRC is changed. After that, a new matrix shall be generated.

4.1.4 Basic Communication Sequence (RTSP)

4.1.4.1 Basic Communication Sequence under Control of RTSP

As the basic communication sequence under the RTSP control of the video content server, the content play initiation sequence is shown in Figure 4-11. For details of the RTSP messages, see 4.1.1.

(1) When the receiver requires information about video content, it uses DESCRIBE method to obtain from the content server the information in SDP format. When the receiver needs to communicate to the content server the information about the FEC it supports, the receiver must use the FEC_Code header of DESCRIBE to do so.

(2) The receiver can issue a SETUP request to obtains transport parameters.

(3) The video content server generates session identifiers to the receiver's request and notifies the Session_ID via a Response message.

(4) The receiver makes a request to play the content via PLAY method.

(5) The video content server checks the Start_Time value of the Range header. When the video content server confirms that it can offer the content at the specified time, it starts delivering the video content data to the receiver in accordance with the specifications defined in 4.1.1.9(1).

When the server uses IPv4 to transmit the video content, consideration shall be given to preventing the receiver from failing to receive the video content due to the NAT function of a home router, etc.

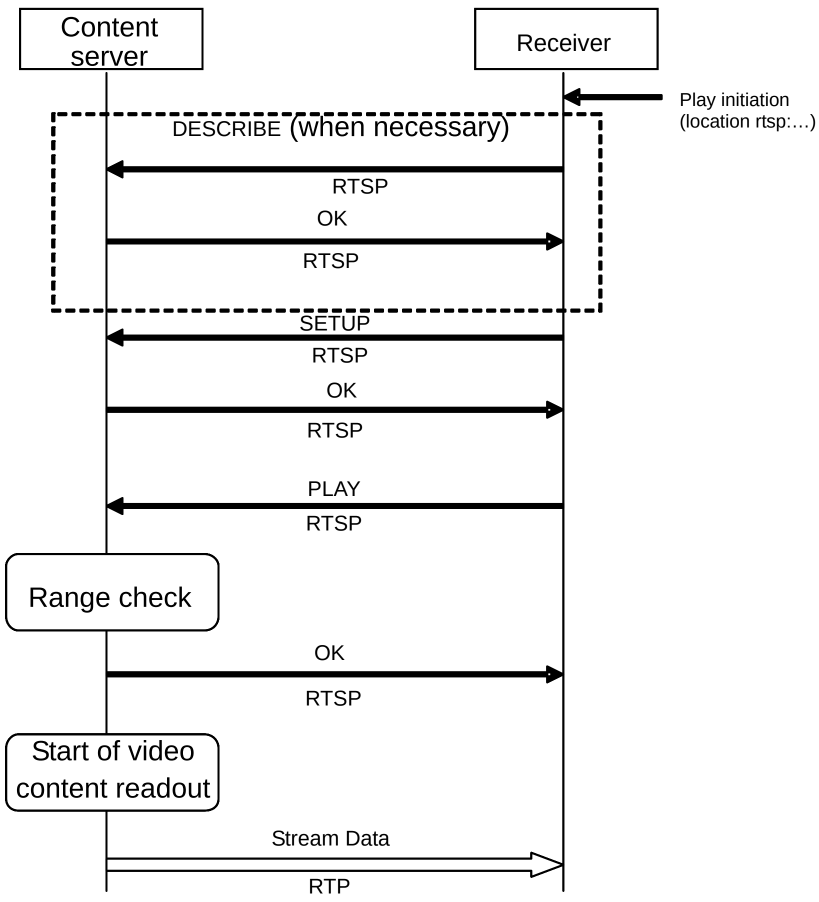

Figure 4-11 Content Play Initiation Sequence

For the operation to stop playing the video content, the applicable specifications in the following cases are described.

- Stop at the end of the content
- Termination of the content by the viewer
- Forced termination of the session by the server or termination of the playback due to a server error, etc.

(1) Stop at end of content

(i) The video content server notifies to the receiver the end of the content by transmitting, upon completion of the content readout, an ANNOUNCE request in which an ending code indicating the end of the content is set

(ii) The receiver that detects the end of the content via the ANNOUNCE request returns an OK to the video content server and then transmits a PAUSE request to the server.

(iii) The video content server returns to the receiver a PAUSE request response in which the content ending position is set in the Range header. The receiver that receives the response obtains the value set in the Range header as the ending position.

(iv) The receiver transmits a TEARDOWN request to the video content server and ends the RTSP session.

Concerning the notification of the ending position to the server, see IPTVFJ STD-0006 "IPTV Standards: CDN-scope Service Approach Specifications" or IPTVFJ STD-0007 "IPTV Standards: Internet-scope Service Approach Specifications".

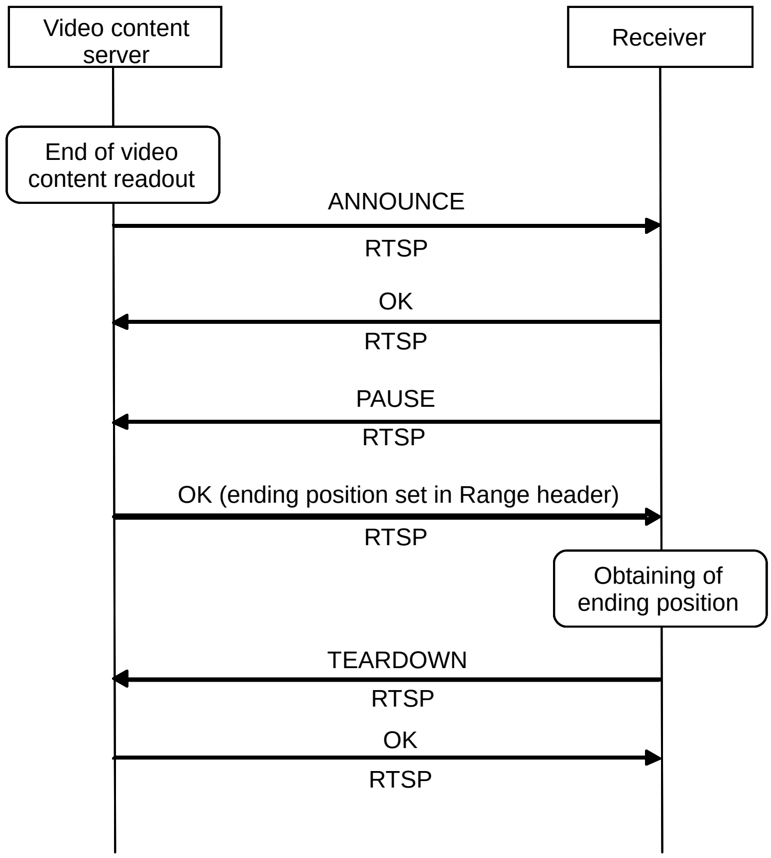

Figure 4-12 Sequence of Play Stop by Termination of Content by User

(2) Termination of content viewing by user

(i) When the receiver receives from the user an instruction to stop the content, it transmits a PAUSE request to the video content server.

(ii) The video content server returns to the receiver a PAUSE response in which the content ending position is set in the Range header. The receiver that receives the response obtains the value set in the Range header as the content ending position.

(iii) The receiver transmits a TEARDOWN request to the video content server and ends the RTSP session.

Concerning the notification of an ending position to the server, see IPTVFJ STD-0006 "IPTV Standards: CDN-scope Service Approach Specifications" or IPTVFJ STD-0007 "IPTV Standards: Internet-scope Service Approach Specifications".

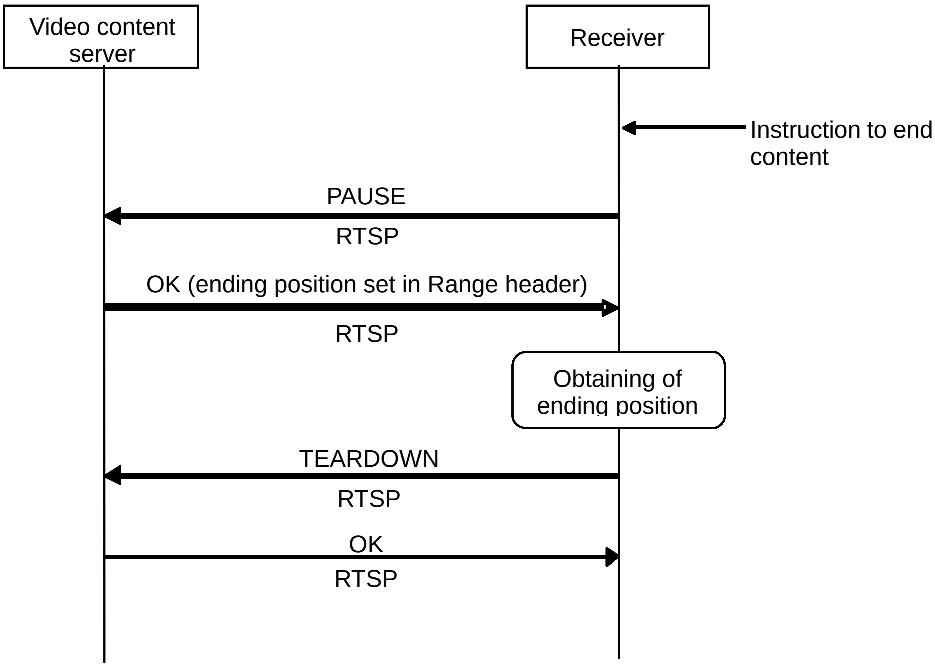

Figure 4-13 Sequence of Play Stop by User's Instruction to End Content

(3) Forced termination of session or termination of playback due to server error

(i) The video content server transmits an ANNOUNCE method to the receiver. The receiver confirms the Event_Code and Event_Phrase described in the Notice header of the ANNOUNCE method it has received and interprets the reason for stoppage of the stream according to Table 4-3.

(ii) If the receiver receives an ANNOUNCE method in which the Event_Code in the Notice header is 5***, it transmits to the server a PAUSE request followed by a TEARDOWN request, as in the case of (1), and terminates the session as a server error.

(4) Termination due to timeout of video stream reception

(i) If the receiver fails to receive the video stream within the preset time (timeout time) after reception of a PLAY response or if it fails to restart the reception of the video stream within the preset timeout time after interruption of the stream without notice thereof by an ANNOUNCE method, the receiver transmits to the server a PAUSE request followed by a TEARDOWN request and terminates the session as an error.

4.1.4.2 Trick Playback

By specifying suitable values in the Range header and Scale header in the Request header of a PLAY method, it is possible for the server to offer special modes of playback using a jump or variable-speed playback. The point of change in the RTP stream under the RTSP control can be identified by the change of SSRC in the RTP header (see 4.1.2.3). It should be noted that the SSRC is changed not only at the time of switching between normal playback and trick playback but also at the time normal playback restarts after a pause.

(1) Jump playback

By setting any value in Start_Time of the Range header during a Play request, it is possible for the receiver to request the video content server to transmit the video stream from a specified point of time. This permits the server to offer special modes of playback, such as viewing the video content again (resume) from the point at which the user stopped viewing the last time and jump playback to a chapter start point based on the chapter information described in ERI as defined in the content playback control metafile.

Jump playback is implemented using the following procedure.

(i) Upon receiving a jump playback instruction from the user, the receiver transmits a PAUSE to the video content server.

(ii) The video content server sets the current play position (NPT) in the Range header in response to the PAUSE request and at the same time, stops transmitting the RTP stream to the receiver.

(iii) The receiver specifies the jump playback position in the Range header and transmits a PLAY request to the server.

(iv) The video content server checks the value specified in the Range header of the PLAY request. As long as the specified value is within the length of the content, the server starts streaming from the new starting position.

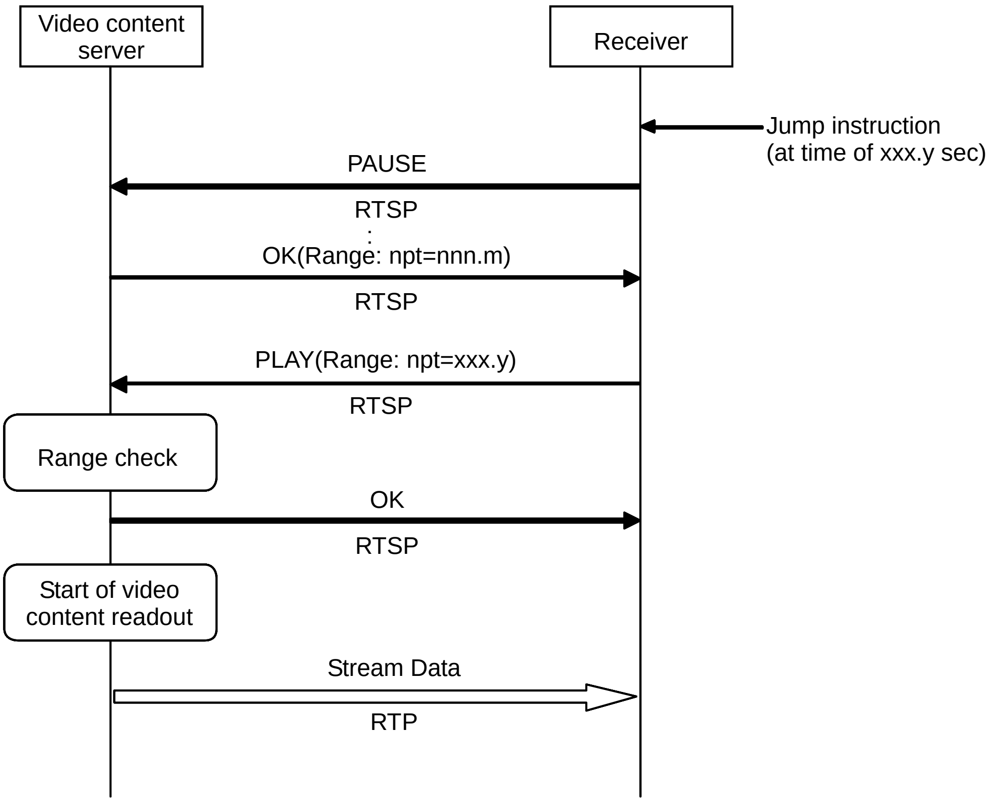

Figure 4-14 Sequence of Jump Playback

(2) Variable-speed playback

Variable-speed playback is performed when a viewing rate is specified by the Scale header during a PLAY request under the RTSP. To send out a video stream during variable-speed playback, two systems—the I-frame TS system and the ordinary TS system—are defined in Section 6.5. Concerning which of the transmission systems is used, it can be judged from the type attribute of the NCI that is described in the content playback control metafile. For the NCI, see 5.3.3 "NCI".

Variable-speed playback is implemented using the following procedure.

(i) Upon receiving a fast forward/rewind playback instruction from the user, the receiver first transmits a PAUSE to the video content server.

(ii) The video content server sets the current playback position (NPT) in the Range header to respond to the PAUSE request and at the same time, stops transmitting the RTP.

(iii) The receiver then specifies the desired viewing rate in the Scale header of a PLAY request and transmits it to the server. As the viewing rate specified in the Scale header, the receiver uses the viewing rate that it obtains from the NCI (Network content Control Information) defined in 5.3.3 "NCI". In the Range header, the receiver sets the NPT value obtained from the video content server in 0, above.

(iv) The video content server checks the Scale value in the PLAY request and decides a Scale value for streaming by comparing the above Scale value with the Scale values the server supports. The server starts the video streaming on the basis of the Scale value and Range value that have been decided.

(v) The receiver performs the playback processing according to the type of trick playback that is specified by the NCI (Network content Control Information) defined in 5.3.3 "NCI".

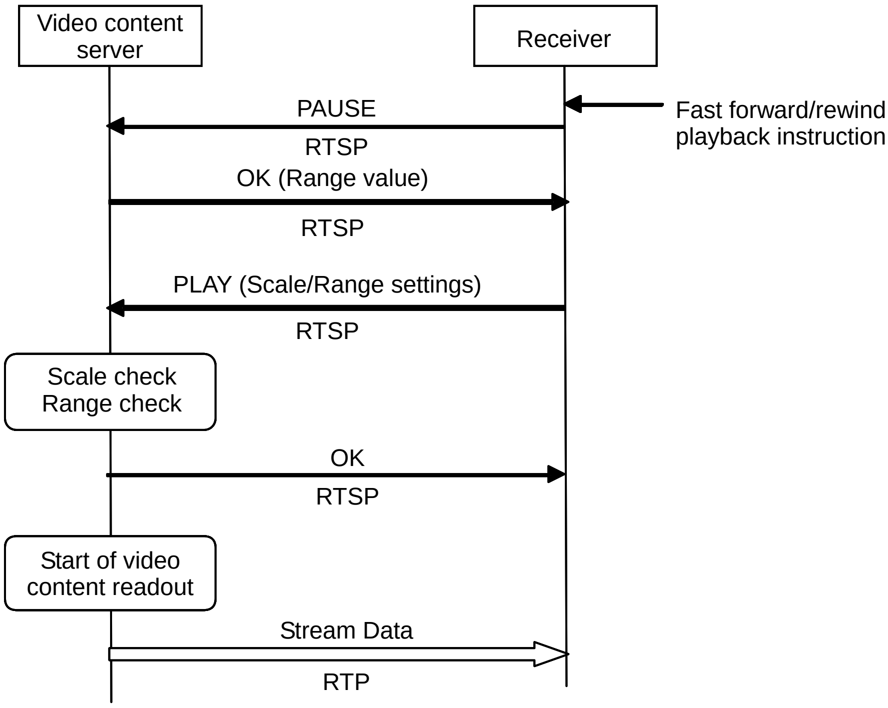

Figure 4-15 Sequence of Fast Forward/Rewind Playback

4.1.5 Compatibility with the Specifications of Digital TV Informatization Research Group

In the following paragraphs, the RTP/RTSP specifications defined in 4.1.1 through 4.1.4 are described in terms of references to and differences from the "Digital Television Network Functional Specifications of Digital TV Informatization Research Group." Thus, the content of the following paragraphs are essentially the same as those of 4.1.1 through 4.1.4.

4.1.5.1 RTSP

4.1.5.1.1 Reference RFC

See 8.3.1 "Reference RFC" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.2 Lower-layer Protocols

See 8.3.2 "Lower-layer protocols" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.3 Period of Existence of TCP Connection

See 8.3.3 "Period of existence of TCP connection" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.4 Timeout of RTSP Session

See 8.3.4 "Timeout of RTSP session" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.5 Maintenance of RTSP Session

See 8.3.5 "Maintenance of RTSP session" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.6 Transition of State

See 4.1.1.2 "State" in this document.

4.1.5.1.7 RTSP Method

See 8.3.6 "RTSP method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.8 RTSP Message Size

See 8.3.7 "RTSP message size" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.9 Handling of Headers

See 4.1.1.4."Handling of Headers" in this document.

4.1.5.1.10 Matters Common to Headers

See 4.1.1.5 "Matters Common to Headers" in this document.

4.1.5.1.11 Details of RTSP Messages

Details of the RTSP messages are described below.

It is desirable that the receiver should be so configured that it does not apply excessive load to the video content server through frequent transmission of an RTSP request, etc.

4.1.5.1.11.1 SETUP Method

See 8.3.10 "SETUP method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group). It should be noted, however, that there is no need to refer to 8.3.10.3 "Address parameters of Response message" in said specifications.

The server must not change the port number that the receiver specifies as a request parameter. The source parameter shall be used when the RTSP server and the media data (RTP) server at the server side have a different address.

It is possible to set a timeout time using "timeout" parameter in the Session header of a Response message. For details about the timeout of an RTSP session, see also 8.3.4 "Timeout of RTSP session" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group) that is referred to in 4.1.5.1.4 of this document.

4.1.5.1.11.2 PLAY Method

See 8.3.11 "PLAY method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group). It should be noted, however, that there is no need to refer to 8.3.11.4 "Server operation in response to request for trick playback" of said specifications.

Although it is not absolutely prohibited to specify End_Time in the Range header, reference shall be made to "(2) Specification of Range header (End_Time) in PLAY request/response" in 4.1.1.9 "Detailed Specifications for Operation of RTSP Control" of this document when operating the Range header.

4.1.5.1.11.3 PAUSE Method

See 8.3.12 "PAUSE method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.11.4 TEARDOWN Method

See 8.3.13 "TEARDOWN method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.1.11.5 ANNOUNCE Method

See 8.3.14 "ANNOUNCE method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

This message is sent from the server asynchronously. It should be noted, therefore, that the value of CSeq in a Request message has nothing to do with the value of CSeq that has been sent from the receiver.

4.1.5.1.11.6 OPTIONS Method

See 8.3.15 "OPTIONS method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

It is not indispensable to verify the server capacity by this method prior to playback of video content. The receiver is allowed to make a request in any state, and such request must not affect the state of the server.

4.1.5.1.11.7 DESCRIBE Method

See 8.3.16 "DESCRIBE method" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Note, however, that it is possible for the receiver to notify the FEC system and matrix parameters that the receiver supports to the server by describing the FEC_Code header in a Request message. For details, see 4.1.5.3.8 "Turning on/off FEC Function" in this document.

The receiver is allowed to transmit to the server a Request message of this method in any state, and the server must interpret the request in any state. The exchange of this message between the receiver and the server shall not cause their state to change.

If the server receives a Request message which does not contain the Accept header, it shall operate as though "application/sdp" were specified in the Accept header and must specify "application/sdp" in the Content_Type header of its Response message.

For details about the SDP sent by a Response message, see 4.1.1.8 "SDP" in this document. For the operation of the namespace in the description of information relating to FEC in 8.3.16.3 "Example of description of SDP body" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group), see 4.1.5.3.7 "Notification of FEC Information" in this document.

4.1.5.1.12 SDP

See 4.1.1.8 "SDP" in this document.

4.1.5.2 RTP

See 8.2 "RTP specifications" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.3 FEC

4.1.5.3.1 Positioning

See 10.1 "Positioning" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.3.2 Reference Standards

See 10.2 "Reference standards" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.3.3 FEC Data Transmission

See 10.3 "FEC data transmission" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

4.1.5.3.4 RTP Header Format

For operation of the RTP header of media packets when FEC is used, see 10.5.1 "RTP header format" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group). As the payload type for FEC packets, "96" as recommended in the Pro-MPEG FEC Specifications shall be used.

4.1.5.3.5 Method of FEC Protection Processing

See 10.5.2 "FEC header format" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group). For details about the operation of FEC, see 4.1.3.1 "Method of FEC Protection Processing" in this document, and (2) "FEC packet generation method/recovery method" and (3) "Guidelines on FEC packet generation/transmission timings" in 4.1.3.5 "FEC Specifications" in this document.

4.1.5.3.6 Method of FEC Recovery Processing

See 4.1.3.2 "Method of FEC Recovery Processing" and (2) "FEC packet generation method/recovery method" in 4.1.3.5 "FEC Specifications" in this document.

4.1.5.3.7 Notification of FEC Information

See 10.4 "Notification of FEC information" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group). Note, however, that "vnd.iptvforum" shall be used as the namespace of FEC type in SDP. For operation details, see 4.1.3.3 "Method of Notifying FEC Information" in this document.

4.1.5.3.8 Turning on/off FEC Function

See 4.1.3.4 "ON/OFF of FEC Function" in this document.

4.2 Video Transmission Protocol based on HTTP

For the video transmission protocol based on HTTP, see 9. "Video Transmission Protocol Based on HTTP" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Chapter 5 Content Playback Control Metafile

This section describes the communication protocol that the receiver uses to obtain the content playback control metafile from the playback control information server when the receiver starts playing the VOD content and the format of the content playback control metafile.

5.1 Outline

See 9.1 "Content playback control metafile" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

5.2 Transmission Protocol

5.2.1 Protocol Used and RFC Referred to

See 7.2.1 "Protocol used and RFC referred to" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

5.2.2 HTTP Specifications

See 7.2.2 "Server and client", 7.2.3 "Information transmitted", and 7.3 "HTTP specifications" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

5.2.3 Details of Communication Sequence

See 7.4 "Details of communication sequence" in the Protocol Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Although the timeout value for the receiver when obtaining the content playback control metafile shall be dependent on the receiver implementation, it may be presupposed that the content playback control metafile shall not be obtained in a time exceeding 15 seconds.

5.3 Configuration of Content Playback Control Metafile

For the transmission format, see 9.1 "Content playback control metafile" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

5.3.1 ERI

See 9.2 "ERI" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group)

5.3.2 LLI

See 9.3 "LLI" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

However, for the method of calculating the signature value in the signature elements and the method of verifying signatures at the receiver, see Chapter 7 "DRM Specifications" in this document.

5.3.3 NCI

See 9.4 "NCI" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Chapter 6 Video Content

6.1 Coding of Information Sources

6.1.1 Video

6.1.1.1 Specifications on Input Signals

See 7.1.1 "Specifications on input signals" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.1.1.2 Details of Operation of MPEG-2 (Video)

<Coding system>

See 7.1.2.1 "Coding system" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

<Restrictions on coding parameters>

See 7.1.2.2 "Restrictions on coding parameters" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

<Change of coding parameters>

In accordance with 7.1.2.3 "Change of coding parameters" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group), it is prohibited to change the coding parameters.

<Range of video coding rate>

See 7.1.2.4 "Range of video coding rate" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

<Other restrictions>

Concerning the GOP length, it is desirable to follow 7.1.2.4 "Other restrictions" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.1.1.3 Details of Operation of H264/MPEG-4 AVC Video

The "Application Application Specifications of H.264/MPEG-4 AVC" (［Appendix A］ of this document) shall be applied.

6.1.2 Audio

6.1.2.1 MPEG1 (Audio)

See 7.2.1 "MPEG1 Audio" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group)

6.1.2.2 MPEG2 (Audio)

See 7.2.2 "MPEG2 Audio" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Note, however, that the range of the audio coding rate shall be as follows:

Standard stereo: 96 kbps to 256 kbps
Multichannel stereo: Max. 384 kbps

6.1.3 Caption

See 7.2.3 "Operation of caption coding" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.2 Multiplexing

6.2.1 Multiplexing within Service

6.2.1.1 Restrictions on Change of Coding Parameters

- Concerning the video ES, see 7.1.4 "Other operation details" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).
- Concerning the audio ES, see 7.2.3 "Audio operation details" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group)

6.2.1.2 Identification of ES Type

See 5.2.1 "Definition of ES" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.2.1.3 Maximum number of ESes per service (1 content)

See 5.2.2 "Maximum number of ESes that can be processed simultaneously (per one transport stream)" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.2.1.4 Default ES

See 5.2.3 "Default ES" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.2.2 Details of Operation of MPEG-2 (system)

6.2.2.1 Synchronization of Video, Audio and Caption Presentation

See 5.3.1 "Synchronization of video, audio and caption" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

In the case of the TS for variable-speed playback, however, see 6.5 "Transmission of Stream for Variable-speed Playback" in this document.

6.2.2.2 Operation of PAT

See 5.3.2 "Operation of PAT" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Note, however, that the following specifications shall be applied.

(1) The progam_map_PID number that is described in the PAT shall always be 1.
(2) Whether or not to describe network_PID in the PAT is not defined. It should be noted, however, that even when it is described, the information provided by it shall be meaningless and shall not affect the receiver operation.
(3) The value of the program_number described in the PAT shall be used only to identify the PMT. It shall not serve as the service_id.

6.2.2.3 Handling of PMT and ES

See 5.3.3 "Handling of PMT and ES" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Note, however, that the following specifications shall be applied

(1) The value of the program_number described in the PMT shall be used to secure compatibility between PMT and PAT. It does not serve as the service_id.
(2) For the handling of the PMT and ES in the TS for variable-speed playback, see 6.5 "Transmission of Stream for Variable-speed Playback" in this document.

6.2.2.4 Maximum Rate of TS

See 5.3.4 "Maximum bit rate" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.2.2.5 Operation of PCR

See 5.3.5 "Operation of PCR" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

When variable-speed playback is implemented, however, see 6.5 "Transmission of Stream for Variable-speed Playback" in this document.

6.2.3 Time-stamped TS

In the VOD service, the TS with a timestamp shall be used to transmit PSI/SI, video, audio and caption signals.

6.2.3.1 Data Structure of Time-stamped TS

The data structure of the time-stamped TS is shown in Table 6-1.

Table 6-1: Data structure of time-stamped TS

Data structure	bit	Identifier
TimeStampedTS () {		
Do{		
timestamp	32	uimsbf

transport_packet() } }		

timestamp: The clock counter value for controlling the relative time of input of consecutive transport packets into the decoder. The timestamp is a 32-bit, 27-MHz, linearly advancing counter value, which shall repeat the range 0x00000000 to 0xFFFFFFFF.
transport_packet (): The transport packet as defined in ISO/IEC 13818-1.

6.2.3.2 Model for Generation of Time-stamped TS

The model for generating time-stamped TS by extending the ordinary TS generation model is explained below. The flow model is shown in Figure 6-1. A timestamp is added to the TS that is output from the multiplexer to generate a time-stamped TS.

(i) The 27-MHz clock that was used to generate the original TS is used to free-run the linear 32-bit counter. In an environment in which the 27-MHz clock cannot be used for that purpose, the PCR is extracted from the TS to configure an STC, from which a PLL is generated to reproduce a 27-MHz clock which operates the above counter.

(ii) When the part of the TS packet at which a timestamp is to be added arrives, the timestamp that is stamped with the above counter value is added to the TS to obtain a TTS packet.

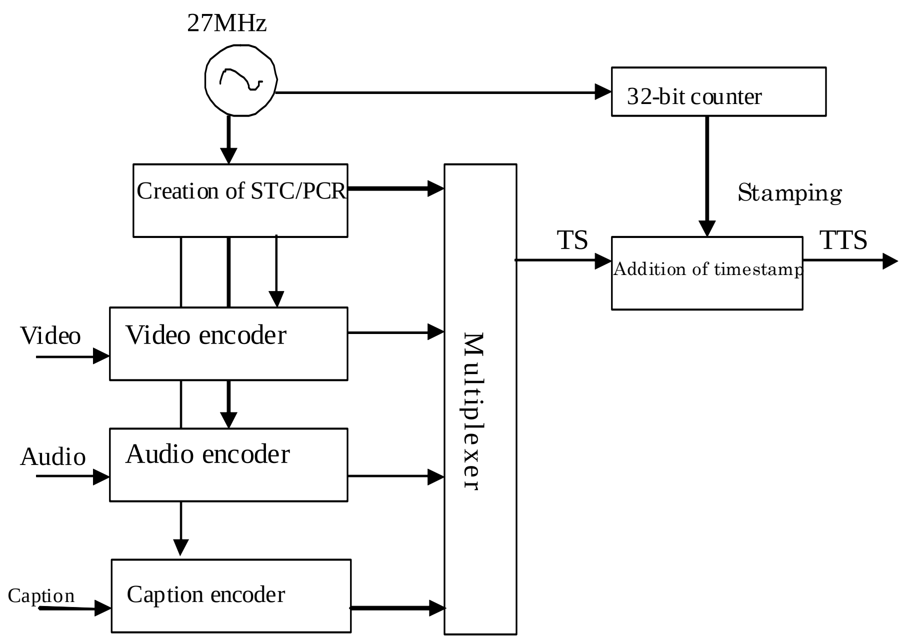

Figure 6-1 TTS Generation Model with Timestamp

6.3 Details of PSI

6.3.1 Tables and Descriptors Operated

See 6.1 "Tables and descriptors operated" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.3.2 TS Packetization and Transmission Specifications

See 6.2 "TS packetization and transmission specifications" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

6.3.3 Table Transmission Operation

See 6.3 "Table transmission operation" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Concerning the period of retransmission, it is not imperative to follow 6.3.3 "Period of PSI retransmission" in the above Codec Part. Even so, following the specifications defined in 6.3.3 is desirable.

6.3.4 PAT (Program Association Table)

See 6.4 "PAT (Program Association Table)" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Concerning the period of retransmission, it is not imperative to follow 6.3.3 "Period of PSI retransmission" in the above Codec Part. Even so, following the specifications defined in 6.3.3 is desirable.

6.3.5 PMT (Program Map Table)

See 6.5 "PMT (Program Map Table" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Concerning the period of retransmission, it is not imperative to follow 6.3.3 "Period of PSI retransmission" in the above Codec Part. Even so, following the specifications defined in 6.3.3 is desirable.

6.4 List for Allocation of Various Numerical Values

For allocation of the transport_stream_id, data_component_id and service_id, see 5.3.6 "Operation of various types of identifiers" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

Concerning the transport_stream_id, it is possible for the transmitting side to allocate any value to it. However, since the allocated value has no meaning, the receiving side shall safely ignore any value allocated by the transmitting side.

6.5 Transmission of Stream for Variable-speed Playback

6.5.1 Variable-speed Playback System

In the VOD service, when variable-speed playback (e.g., fast forward playback, fast rewind playback, etc. as trick playbacks) is implemented, the signals for each VOD content are transmitted in either I-frame TS mode or ordinary TS mode.

6.5.1.1 I-frame TS Mode

The I-frame TS mode is one in which, with the aim of decreasing the disc capacity required of the content at the transmitting side, the TS packets relating to the I-frame are extracted from the content of the ordinary MPEG-2 TS format (normal viewing rate content) for each request from the receiver and delivered to the receiver. The specifications on the transmission operation in I-frame TS mode and the standard processing by the receiver are described below.

[Specifications on transmission operation]

- In I-frame TS mode, P and B frames shall be discarded and only the I-frame information shall be transmitted. In the I-frame TS, the same PSI information as the original normal viewing rate content must be transmitted.
- In a single I-frame transmission section, the I frame shall be transmitted at the same bit rate as used in ordinary TS transmission.
- Different I frames shall not be transmitted by a single RTP.
- Every I frame shall be transmitted in such a way that the receiving side can reproduce an I frame of the duration that corresponds to the specified viewing rate.
- Every I frame must contain the PES header.
- When H264 is used, the I/P field pair shall be extracted as an I frame.

[Standard processing by receiver]

- Since the timestamp value and PCR value of a time-stamped TS are not always reliable, the receiver shall sequentially decode the data as it is received without regard to the timestamp value and PCR.
- Concerning the PID of video ES, the receiver shall use the same PID as used in playback of the normal viewing rate content in decoding the data it receives.
- The receiver shall not decode audio and caption packets (if any).

6.5.1.2 Ordinary TS Mode

The ordinary TS mode is one in which, with the aim of allowing for smooth decoding by the receiver, the transmitting side prepares content which are previously encoded for fast forward/rewind playback and delivers those content in ordinary MPEG2-TS format when requested from the receiver. The specifications on transmission operation and standard processing by the receiver in the ordinary TS mode are described below.

[Specifications on transmission operation]

- Content shall be transmitted using the same operation as for transmission of TS for normal playback.
- Content shall be transmitted at the same bit rate as used in the transmission of normal viewing rate content.
- The same PSI information as that of the original normal viewing rate content must be transmitted.
- The ES configuration shall be the same as that of the original normal viewing rate content.
- Audio and caption packets shall not be multiplexed for the purposes of playback and display of content at the receiver. Rather, they shall be multiplexed as packets for synchronization with consideration given to the decoding Process at the receiver side.
- Content shall be given the timestamp value, PCR, DTS and PTS of the prescribed time-stamped TS before it is transmitted.

[Standard processing by receiver]

- As in the case of normal viewing rate content, the receiver shall decode the content it receives according to the timestamp value, PCR, DTS and PTS of the time-stamped TS.
- Since audio and caption packets are not multiplexed for the purposes of playback and display of content at the receiver, they shall not be decoded at the receiving side.
- Audio and caption packets shall not be reproduced, and audio shall be muted.

6.5.2 Identification of Variable-speed Playback Mode

For each VOD content, either I-frame TS mode or ordinary TS mode is defined by the NCI. For the NCI, see 5.3.3 "NCI".

The receiver shall decide the appropriate method of decoding for fast forward playback and fast rewind playback based on the TS mode defined by the NCI.

Chapter 7 DRM Specifications

7.1 System Model and Functional Requirements of DRM

In this section, a system model of DRM in these VOD Specifications is defined and the functional requirements of the DRM system are stated clearly.

7.1.1 DRM System Reference Model

Here, the system reference model of the servers and the entities in the receiver relating to the DRM system are defined. The concept of the model is shown in Figure 7-1.

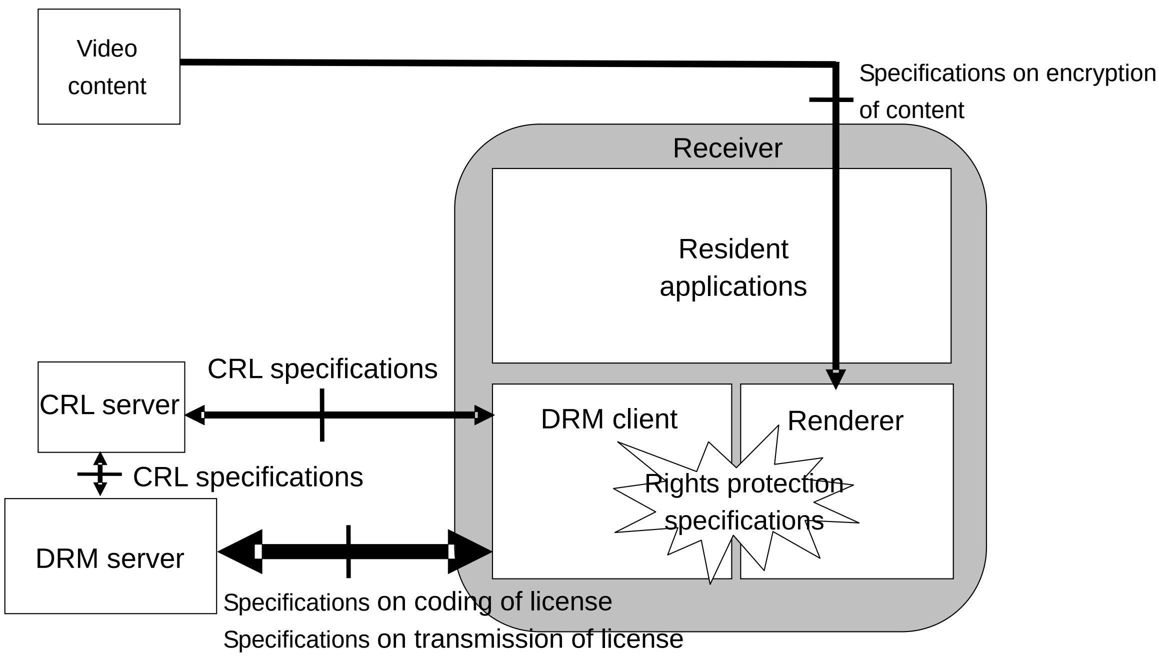

Figure 7-1 DRM System Reference Model

7.1.1.1 Definitions of Entities

- DRM server
  This is the functional entity that generates, manages and issues licenses. It has the following functions:
  - Generating and managing content keys and licenses
  - Establishing secure communication with the DRM client

- Judging whether or not to issue a license based on requests from the DRM client and transferring the license when issued
- Transferring trusted time information based on requests from the DRM client
- Updating and managing CRL

- CRL server
  This is the functional entity that generates, manages and issues CRL relating to the DRM server and DRM client. It has the following functions:
  - Generating, updating and managing CRL of the DRM server and DRM client
  - Transferring CRL based on requests from the DRM server and DRM client
- Content server
  This is the functional entity that sends encrypted streams in the VOD service. It has the following functions:
  - Generating and managing an encrypted stream of content by using the appropriate content key
  - Sending out an encrypted stream in response to a request from the receiver
- DRM client
  This is the functional entity inside the receiver that obtains and manages licenses and that supplies the appropriate content key when specific content is used. It has the following functions:
  - Establishing a secure communication link through mutual authentication with the DRM server
  - Obtaining licenses from the DRM server and managing them
  - Supplying content keys and information about the conditions for use of content to the renderer
  - Updating and managing CRL through communication with the CRL server
- Renderer
  This is the functional entity inside the receiver that receives and decodes encrypted streams in the VOD service and that reproduces content. It includes decrypter, demultiplexer, video decoder, audio decoder and caption decoder of the receiver model and presentation process defined in 3.1, Receiver Reference Mode, it has the following functions.
  - Decrypting encrypted streams by using content keys supplied from the DRM client
  - Decoding decrypted streams
  - Implementing output control during playback in accordance with the information about conditions for use of the license supplied from the DRM client

- Resident application
  This is the functional entity that controls the sequence of the entire process for VOD service at the receiver.

7.1.1.2 Items to be specified on DRM System

- License format
  These are the specifications on the encoding of a license as a substance. They consist mainly of content keys and information about conditions for use of licenses.
- Protocol for secure link and license delivery
  These are the specifications on the communication protocol for establishing a secure communication link between the DRM client and DRM server and for enabling the DRM server to safely deliver licenses. They are important technical specifications that underlie the DRM system and call for a high level of communication security through mutual authentication.
- Content encryption format
  These are the specifications on the encryption of content that are applied to content streams in the VOD service.
- CRL format and rules
  These specifications concern the encoding, transmission, updating and operation of the CRL that describes revoked servers and the encoding, transmission, updating and operation of the CRL that describes illegal receivers (DRM clients).
- Rights protection rules for the client
  These specifications concern the implementation rules of the DRM client and renderer and the playback output control and copy control based on the information about conditions for use of licenses.

7.1.2 License Model

7.1.2.1 Definition of License

The term "license" refers to the right to use specific content, such as playback of the content. Alternatively, it is defined as the substance of data that states the right to use specific content and that permits using the content only when specific conditions for use are met. The license includes a content key and information about the conditions for using the license.

7.1.2.2 License Delivery System

In VOD services, only a license that is issued for each specific content is delivered. A license for using the VOD service is defined as a VOD license. The concept of license delivery in the VOD service is shown in Figure 7-2.

The key for encrypting content (content key, Kc) is directly included into the VOD license to be delivered. The flow of delivery and use of a license is explained below.

(i) The DRM server generates a content key (Kc), by which the content that has been subjected to AV encoding is encrypted and set in the video content server.

(ii) Based on the prescribed conditions for use of the content, a VOD license, which includes the content key (Kc) and the information about the conditions for use of the content (RMPI: Rights Management and Protection Information), is generated and set in the DRM server.

(iii) When it comes to playing the VOD service content, the VOD license is first delivered from the DRM server to the DRM client in response to a request from the DRM client. Then, the DRM client sets the Kc in the decrypter of the renderer and transfers the RMPI to the renderer.

(iv) Next, the receiver accesses the video content server and receives the encrypted content by streaming. The encrypted content is decrypted using the Kc and subjected to AV decoding in the renderer. It is then output for playback in accordance with the RMPI (the output control information in particular).

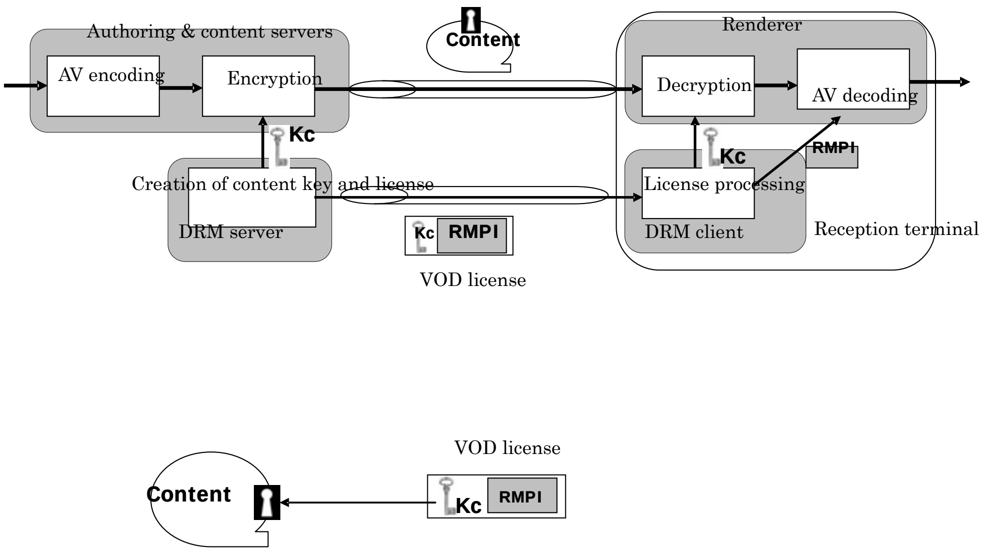

Figure 7-2　Concept of License Delivery System

7.1.2.3　Constituent Elements of License

The constituent elements of the VOD license used in the VOD service are shown in Table 7-1. Since a VOD license is issued each time a specific content is played, the RMPI (Rights Management and Protection Information) virtually consists of the output control information alone.

Table 7-1: Main constituent elements of VOD license

Constituent element	Explanation
Content key (Kc)	The key used to encrypt content.
Output control information	Information about restrictions on signal output/content copy (see Note)

Note: This information corresponds to the information provided by the digital copy control descriptor and the content availability descriptor that are defined in ARIB STD-B10.

7.1.3 Model of DRM Processing Operation

DRM processing in individual receivers and server entities based on these Specifications is outlined below. It should be noted that the DRM processing described here represents just model operations and that the DRM processing operations must not necessarily be the same as the model operations.

7.1.3.1 Basic DRM Processing Elements

First, for the DRM basic processing elements, the processing element operation model that forms the base of the DRM system is shown below.

(1) License acquisition processing element

A VOD license is delivered to the DRM client after the DRM server accepts the DRM client's request for it. In so doing, a high level of security is required of the communication link. In order to materialize the required security, it is necessary to establish a secure authenticated channel (SAC). The communication sequence between the DRM server and the DRM client is shown in Figure 7-3. Basically, the license acquisition processing is performed in this order: establishment of SAC → request for license → delivery of license → closing of SAC. Each of those processes involved in license acquisition processing element is explained below.

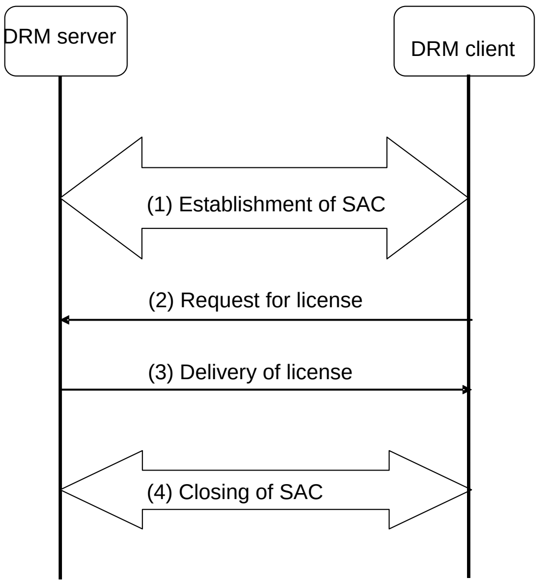

Figure 7-3 License Acquisition Sequence

(i) The means of establishing a secure authenticated channel (SAC) shall be based on mutual authentication using the PKI (Public Key Infrastructure). It is presupposed here that the DRM client holds the client certificate, root certificate and the DRM server's CRL and that the DRM server holds the server certificate, root certificate and the DRM client's CRL. As a result of this process, mutual authentication between the server and the client is completed and the keys for encrypting the messages in the processes of license request and delivery are shared by the server and the client.

(ii) Request for license

The DRM client requests a license by transferring to the DRM server a message including the license ID that identifies the license the DRM client wants to obtain.

(iii) Delivery of license

The DRM server judges whether the DRM client has the right to obtain said license. After confirming that the DRM client has said right, the DRM server delivers the license to the DRM client.

(iv) Closing of SAC

The DRM client and the DRM server cut off the SAC that has been established.

(2) Content encryption/decryption processing element

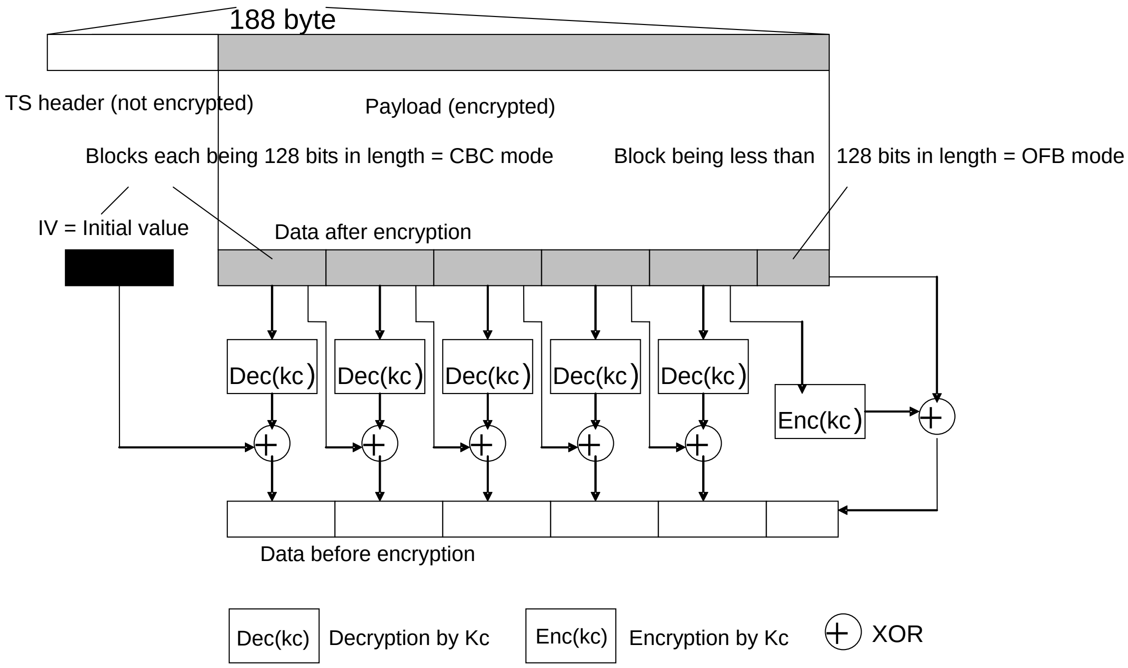

Figure 7-4 TS Packet Encryption/decryption Processing

The content is encrypted in the payload (excludes the adaptation field) of the TS packet by the 128-bit AES system in CBC mode and OFB mode (only for blocks less than 128 bits in length). The encryption and decryption processing is performed for each individual TS packet. The operation performed in the encryption/decryption processing of a single TS packet is shown in Figure 7-4.

(3) CRL update processing element

It is desirable that the DRM client should hold the latest CRL to implement proper server authentication in the license acquisition processing. Therefore, in accordance with the prescribed operational rules, the DRM client shall obtain the latest CRL from the CRL server and update it as required. For example, each time license acquisition processing is performed, the DRM client is supposed to check the date of the next update described in the CRL that has already been acquired at that time and to gain access to the CRL server on that date to update the CRL. It is desirable that the DRM server, too, should hold the latest CRL to implement proper client authentication in the license acquisition processing.

(4) Processing element for acquiring trusted time

Updating a CRL requires trusted time information. One means of obtaining trusted time information is that the DRM server and the DRM client establish a SAC, as in the case of license acquisition processing, so that the DRM client can obtain the time information from the DRM server.

(5) CAS/DRM client identifier (DRM_ID) registration processing element

The DRM server is required to judge whether or not to issue a license for each individual receiver. In order to implement this, it is necessary to previously register in the server a CAS/DRM client identifier (DRM_ID) associated with the information about a specific user. The method of registering DRM_ID shall be defined in separate operational specifications.

7.1.3.2 Entire sequence of DRM Processing Operations in VOD Service

The DRM-related processing operations in the individual phases in the entire sequence of VOD service are explained below. An example of a communication sequence between the server and receiver entities in the VOD service is shown in Figure 7-5. It should be noted that since general processing operations, including those which are outside the scope of the DRM specifications, are discussed here, the operation of peripheral functional entities which are not included in the DRM system model defined in 7.1.1 are also explained. Those functional entities are briefly defined below.

- Portal server
  This is a Web server operated by each service provider. It supplies HTML/BML documents to the receiver browser and offers the function of navigating the subscription to services and the playback of content.
- Customer management server
  This is a server by which the service provider manages the customer registration information and the information about subscriptions to services offered by the service provider. It is presupposed that this server functions as a database.
- Settlement server
  This is a server which has the function of billing and settling accounts involved in subscriptions to services offered by the service provider.
- Browser
  This is a module inside the receiver that implements the function of navigating the subscription to services, playback of content, etc. It refers to an HTML or BML browser.

(1) Basic registration

The term "basic registration" refers collectively to the user’s operation and the processing performed by the receiver/server to register the user’s personal information and information about the receiver with the service provider so as to permit the user to use any service whose copyright is protected.

As a rule, basic registration is implemented by communication between the portal server and the browser.

During basic registration, the user performs the prescribed operation on the basic registration page of the HTML or BML document that is supplied from the portal server. As a result, the CAS/DRM client identifier (DRM_ID) is registered with the customer management server as described in 7.1.3.1(5).

Although the method of transferring the DRM_ID to the server is not defined in this document, it is necessary to meet, at least, the following security requirements.

- The communication channel shall be protected against eavesdropping and tampering of content.
- Arrangements shall be made to prevent the DRM_ID from being transferred to some server other than that of the service provider that operates DRM.

(2) Subscription to service

The term "subscription to service" refers collectively to the user’s operation and the processing performed by the receiver/server to permit the user to use a specific content product (package). Subscription to service corresponds to ‘contracting’ when the package is "all-you-can-view" or "select" and to ‘purchasing’ when the package is a "single item" or "pack". (These package types are described later.) In the subscription to service, the user, following the user authentication, performs the operation for purchasing a specific VOD service content package on the ‘subscription to service’ page of the HTML or BML document that is supplied from the portal server. As a result, the appropriate package identification information is registered with the customer management server via the portal server. As long as the user’s personal information has been associated with the DRM_ID by the basic registration, the customer management server can communicate with the settlement server to collect the charge for the subscription to service. In addition, since the DRM_ID is associated with the corresponding license ID by the package identification information, the DRM_ID and license ID are automatically registered together with the customer management server.

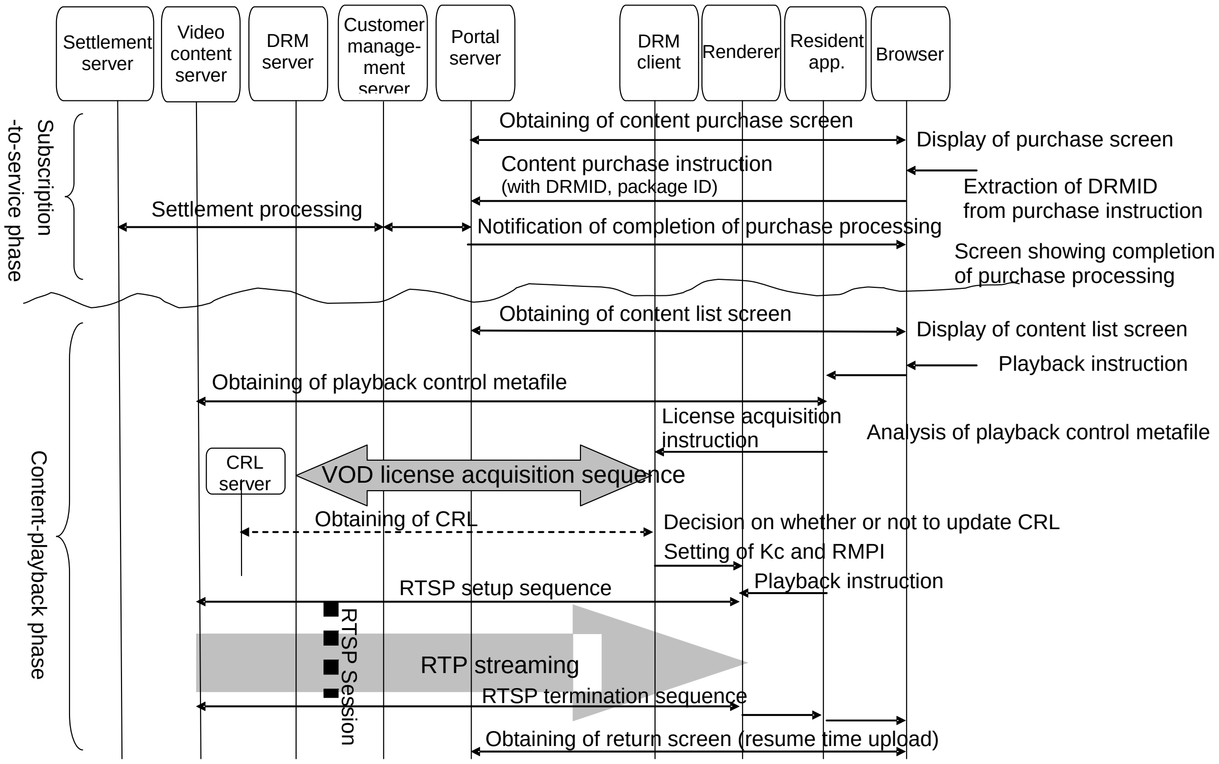

Figure 7-5 Sequence of Communication between Entities in VOD Service

(3) Playback of VOD content

The playback of VOD service content is initiated from either the HTML/BML document obtained from the portal server or the ECG. Figure 7-5 shows an example of the former. Namely, by calling the interface for initiating VOD playback in the HTML/BML document, the following types of processing, including DRM-related processing, are performed. The flow of content playback processing operations is shown in Figure 7-6.

(i) Processing of content playback control metafile

The resident application obtains the content playback control metafile from the video content server and analyzes the ERI, LLI and NCI. From the LLI, it extracts the license ID and DRM server URI for the content to be used and instructs the DRM client to obtain the license. At the same time, the resident application delivers to the DRM client the signature of the DRM server URI, the server certificate used for verification of the signature, and the certificate of the issuer thereof—all described in the LLI. For details about the content playback control metafile, see Chapter 5.

(ii) License acquisition processing

First, in order to prevent the VOD content from being directed to an illegal DRM server, the DRM client verifies the signature of the DRM server URI. This verification is implemented by using the server public key extracted from the server certificate for signature verification that has been delivered to the DRM client. In addition, the chain of server certificates is also verified by the root certificate held by the DRM client. Only after the signature is verified, the following types of processing are performed.

Next, the DRM client obtains the VOD license from the DRM server by performing the license acquisition processing described in 7.1.3.1. The content key (Kc) and output control information that are contained in the VOD license acquired are immediately delivered to the renderer. The content key is set in the decrypter of the renderer, and the output control information is used as required.

(iii) CRL update processing

The DRM client judges whether or not it is necessary to update the CRL. When necessary, the DRM client performs the CRL update processing described in 7.1.3.1.

(iv) Stream reception/playback processing

On the basis of the information contained in the ERI, the renderer gains access to the video content server that performs streaming in order to receive a content stream which is transmitted by RTP under the control of RTSP or a content stream which is controlled and transmitted by HTTP. The decrypter of the renderer decrypts the content stream using the content key (Kc) as described in 7.1.3.1. The decrypted content stream is then subjected to AV decoding. Where necessary, the content is output for playback in accordance with the output control information contained in the VOD license.

At the end of playback or when playback is interrupted due to error processing, etc., use of the content key (Kc) is prohibited.

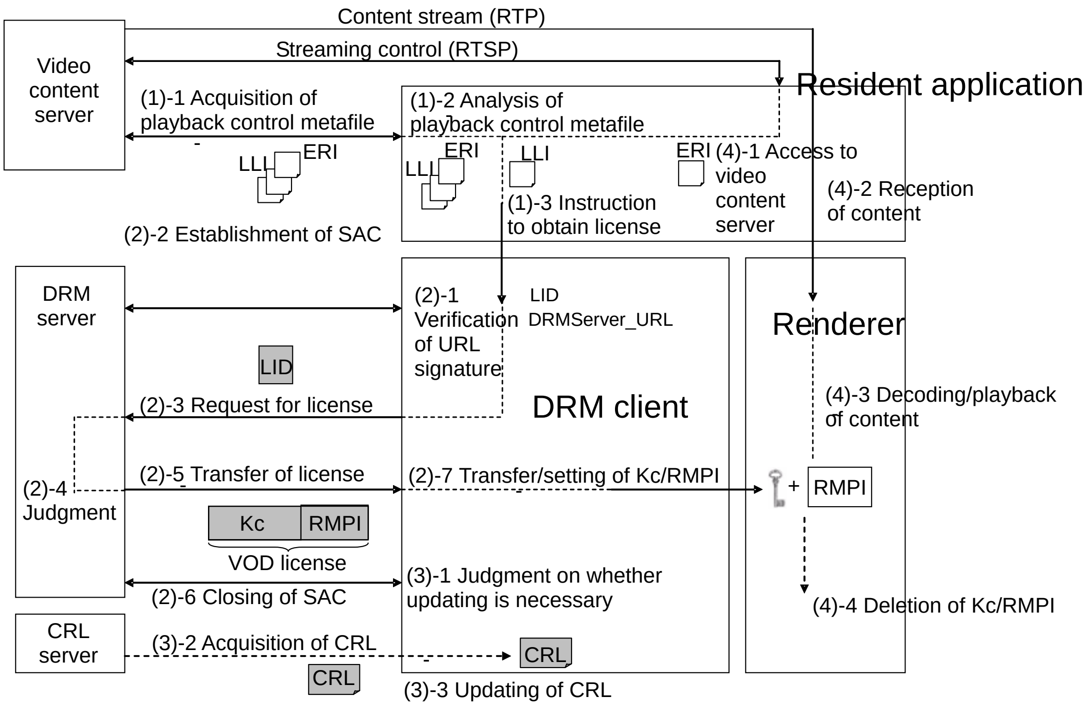

Figure 7-6 Flow of DRM Processing Operations in VOD Service

7.2 Detailed DRM Specifications

In this document, detailed specifications about the core of the DRM specifications are not defined. Instead, the method of adaptation of the specific DRM system shall be defined in Appendix B of this document. Here, detailed DRM specifications which are independent of the above specific DRM system are defined. It should be noted that the above specific DRM system must meet the DRM system model and functional requirements defined in 7.1.

7.2.1 Specifications on License ID

7.2.1.1 CAS/DRM Operator ID (drm_provider_id)

This is a 2-byte ID that identifies the specific operator of a DRM system. A DRM system operator is assumed to be associated with a service provider or a group of service providers. This ID is set in the upper two bytes of the license ID. (It identifies not only DRM system operators, but also CAS system operators in IP broadcasting.)

7.2.1.2 License ID

The license ID identifying a license obtained by the DRM client of the receiver contains information for identification of the type of service, such as VOD, download, or IP broadcasting, etc. An integrated numbering system is applied to all the licenses used by IPTV. The license ID of each VOD license must be unique in the scope of application of these Specifications. Each license ID consists of 8 bytes. The specifications on the numbering of license IDs are defined in Figure 7-7 License ID Structure

Table 7-2. The structure of the license ID for a VOD license is shown in Figure 7-7.

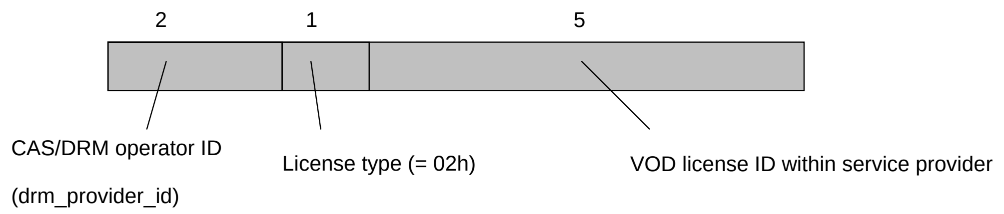

Figure 7-7 License ID Structure

Table 7-2: License ID of VOD license

Byte	Constituent element	Explanation
0-1	CAS/DRM operator ID (drm_provider_id)	This ID uniquely identifies the DRM operator.
2	License type	This indicates the type of license. For VOD licenses, the license type is always 0x02.
3-7	License ID within provider	This ID uniquely identifies the VOD license within the provider indicated by drm_provider_id.

7.2.2 Specifications on Encryption of Content

In encrypted content, TS packets containing a PAT or PMT shall not be encrypted.

7.3 DRM Related Functional Requirements for Receivers

In general, the receiver is supposed to perform the various types of processing that are performed by the operation model shown in 7.1.3. Specifically, the receiver is required to perform the types of processing that are defined in 7.2 and those which are defined by the specific DRM system.

Described below are the DRM process, the DRM-related specifications and matters worthy of special mention that are not shown above but must be performed by the receiver. The DRM functions shall be made available only in the services of a service provider with whom basic registration has been completed.

7.3.1 Response to Error in License Acquisition Processing

It is conceivable that the acquisition of a license fails due to the occurrence of any of the following events during the license acquisition process.

(i) Connection with the DRM server cannot be established. (The timeout value for the receiver during license acquisition shall be dependent on the receiver implementation. Even so, it may be presupposed that the acquisition of a license does not take more than 10 seconds.)

(ii) The SAC cannot be established due to an error in server or client authentication, etc.

(iii) The DRM client cannot obtain the required license as the DRM server rejects the client’s request for the license because the necessary subscription to service has not been made.

(iv) The session for license acquisition cannot be completed due to a network or server failure, etc.

If any of the above events occurs, the DRM client shall shift to the state that sets in when playback processing is completed, without actually performing the playback processing after the license acquisition described in 7.1.3.2.

7.3.2 Holding and Management of License

In order for the receiver to implement the DRM-related functions based on these Specifications, it is necessary to hold and manage the license that the receiver has obtained in the following manner.

Concerning the VOD license, the receiver obtains it at the start of content playback and holds it until completion of content playback. Therefore, the receiver need not consider holding more than one VOD license at the same time. On the other hand, the receiver is required to positively invalidate the VOD license at the end of content playback. Specifically, the receiver shall invalidate the VOD license at the time of reception of a TEARDOWN response when RTSP is used or at the time of completion of playback processing by VOD playback control when HTTP is used or when some system trouble occurs.

7.3.3 Encrypted and Non-encrypted Content in VOD Service

For encrypted content, the receiver performs the processing described in (a), below. For non-encrypted content, the receiver performs the processing described in (b), below. In VOD service content, whether a specific content has been encrypted or not is judged from the encryption element in the ERI of the content playback control metafile. Namely, when the value of the encryption element is 1, it can be judged that the content has been encrypted, whereas when the value is 0, it can be judged that the content has not been encrypted.

(a) Processing of encrypted content: The processing operation shall be as described in 7.1.3.2. During reception of encrypted content, the content shall be decrypted by referring to the scramble control flag of the TS packet header. The reason for this is that there are cases in which a non-encrypted component is included in the content even when the value of the encryption element in the ERI is 1. For example, an encrypted content may include a non-encrypted caption ES.

(b) Processing of non-encrypted content: It is possible to start streaming reception of non-encrypted content by RTSP/RTP or HTTP without performing the license acquisition processing based on the LLI. Whether or not to refer to the scramble control flag of the TS packet header during reception of non-encrypted content shall be left to the receiver’s judgment. (Processing whereby the receiver refers to the scramble flag and displays an error message if it receives a stream with scramble flag 1 shall depend on the receiver.)

7.3.4 Receiver Operation When Receiver is Revoked

If the receiver is revoked, the receiver can recognize the fact as it receives a notice from the DRM server that it has failed to verify the client certificate during the client’s attempt to establish a SAC for license acquisition.

In the VOD service, since a revoked receiver cannot acquire the license for playing content, the content itself cannot be played back. It is desirable, therefore, that a message stating that the receiver has been revoked should be displayed and the suitable action to take should be notified to the user. It is also desirable that if revocation of the receiver is notified during an attempt to acquire the license, the receiver should hold the relevant information and let it be known to the user where appropriate.

7.3.5 Copy Control and Output Control

For encrypted VOD content, copy control and output control shall be implemented as described in the VOD license. In the case of non-encrypted VOD content, the content can always be copied freely since there are no means of specifying copy control/output control.

7.3.6 Verification of Valid DRM System

If, at the start of playback of VOD content, the drm_system element does not exist in the LLI of the content playback control metafile or the DRM system is not one that is supported by the receiver, the receiver shall neither attempt to acquire the license nor perform the subsequent playback operation.

7.3.7 Signature Verification Processing for URI of DRM Server

If, at the start of playback of VOD content, the signature verification for the DRM server URI included in the LLI results in an error, the receiver shall neither attempt to acquire the license nor perform the subsequent playback operation.

7.3.8 Getting Reliable Clock Time

When passing judgment on the valid period in CRL update processing, the receiver must use trusted time information.

［Appendix A］　Application Specifications of H.264/MPEG-4 AVC

See Supplementary Specifications A "Operation of H.264/MPEG-4 AVC" in the Codec Part of the "Streaming Specification, Ver. 1.1: Digital Television Network Functional Specifications" (Digital TV Informatization Research Group).

［Appendix B］ Application Specifications of Marlin IPTV-ES System in DRM Specifications

The detailed DRM specifications when the Marlin IPTV-ES is applied as the DRM system described in this document are defined as follows.

B.1 Reference Specifications on Marlin IPTV-ES System

When the Marlin IPTV-ES system is applied in these VOD specifiations, reference shall be made to the following:

- Marlin IPTV End-point Service Specification Version 1.0.2(or later version)
- Marlin IPTV-ES Implementation Guidelines for VOD Version 1.2(or later version)
- Marlin IPTV-ES/J Specific Compliance Rules for VOD Version 1.2(or later version)
- Marlin Trust Management Document – for IPTV-ES Version 1.3(or later version)

B.2 License Encoding Specifications

B.2.1 License ID in request for license acquisition

In the Marlin IPTV-ES system, the method of specifying the license ID defined in 7.2.1 in the license acquisition request message shall be as follows.

The license ID shall be put in the upper eight bytes of UsageRuleReference in the Get Permission Request message as defined in 4.2.1 of the "Marlin IPTV End-point Service Specifications". The lower eight bytes of UsageRuleReference shall all be 0.

B.2.2 License

The body of the VOD license is positioned in the StatusExtension in the Get Permission Reply message as defined in 4.2.1.4 of the "Marlin IPTV End-point Service Specifications". The correspondence between the license constituent elements defined in 7.1.2.3 of this document and the parameters in the Marlin IPTV-ES Specifications are shown in Table B-1.

Table B-1: Correspondence between VOD License Constituent Elements and Parameters in Marlin IPTV-ES Specifications

License constituent element	Parameter name in Marlin IPTV-ES Specifications
Content key (Kc)	ContentKey
Output control information	RenderingObligation (Output Control Information)

The output control information contained in the VOD license is encoded as RenderingObligation in 4.2.1.4.1 of the "Marlin IPTV End-point Service Specifications". The parameters in the Marlin IPTV End-point Service Specifications correspond to the parameters in the digital copy control and content usage descriptors defined in ARIB STD-B10 as shown in Table B-2.

Table B-2: Correspondence of Output Control Information between ARIB STD-B10 Specifications and Marlin IPTV-ES Specifications

Descriptor parameter name	Parameter name in Marlin IPTV-ES Specifications
Digital copy control descriptor digital_recording_control_data	DigitalRecordingControlData
Digital copy control descriptor copy_control_type	CopyControlType
Digital copy control descriptor APS_control_data	APSControlData
Content usage descriptor image_constraint_token	ImageConstraintToken
Content usage descriptor retention_mode	RetentionMode
Content usage descriptor retention_state	RetentionState
Content usage descriptor encryption_mode	EncryptionMode

B.3 Specifications on License Transmission

As a rule, the acquisition of a license shall be implemented by using the service protocol defined in 4.2 of the Marlin IPTV End-point Service Specifications after a secure transmission link is established based on the SAC protocol as defined in 4.1 of the Marlin IPTV End-point Service Specifications. License acquisition by the DRM client is implemented as follows. First, the DRM client requests a license by specifying the appropriate license ID in UsageRuleReference in a Get Permission Request message. Then, the DRM server stores the license body in StatusExtension in a Get Permission Reply message in response to the above request.

For details about the communication protocol, including the SAC, between the DRM client and DRM server for the transmission of VOD licenses, see 4.1, 4.2.1 and 4.2.3 of the "Marlin IPTV End-point Service Specifications", Section 5 and 6.1 of the "Marlin IPTV-ES/J Specific Compliance Specifications: VOD Part" and Section 2, 3.1, 3.3 and Section 4 of the "Marlin IPTV-ES Operational Specifications: VOD Part".

B.4 Specifications on Content Encryption

For details about the specifications on content encryption in VOD, see 6.1.1 of the "Marlin IPTV End-point Service Specifications".

B.5 CRL Specifications

For detailed specifications on the CRL, reference shall be made to the following specifications and documents.

It should be noted that in the Marlin IPTV-ES system, the revocation list of DRM servers is defined as "CRL", whereas the revocation list of DRM clients is defined as "DRL".

- For the CRL/DRL data formats, see 5.2 and 5.3 of the "Marlin IPTV End-point Service Specifications" and 1.6 and 1.7 of the "Marlin Trust Management Document—for IPTV-ES".
- For DRL processing at the DRM server and CRL processing at the receiver, see 4.1.4.13 and 4.1.4.14 of the "Marlin IPTV End-point Service Specifications".
- For renewing operation of CRL/DRL, see Sections 6 and 7 of the "Marlin Trust Management Document—for IPTV-ES".
- To obtain a CRL, the DRM client shall access to the CRL server URI defined as CRL Distribution Points in 1.3 and 1.5 of the "Marlin Trust Management Document—for IPTV-ES".

B.6 Specifications on Signature Verification for DRM Server URI

For the method of signature verification for the DRM server URI, see 4.1.4.12 of the "Marlin IPTV End-point Service Specifications".

B.7 Specifications on Trusted Time

Concerning the trusted time information, reference shall be made to the following specifications.

For the protocol to obtain trusted time from the DRM server via SAC, see 4.2.2 of the "Marlin IPTV End-point Service Specifications" and 3.2 of the "Marlin IPTV-ES Operational Specifications: VOD Part".

For the operation of trusted time in DRM processing, see Section 3 of the "Marlin IPTV-ES/J Specific Compliance Specifications: VOD Part".

When the Get Trusted Time Protocol defined in 4.2.2 of the "Marlin IPTV End-point Service Specifications" is used to obtain trusted time, the relevant time information shall be obtained by a Packed Message with the license.

B.8 CAS/DRM Client Identifier (DRM_ID)

The CAS/DRM client identifier (DRM_ID) shall be the value of the Subject in the client certificate as defined in 5.1.1.4 of the "Marlin IPTV End-point Service Specifications" and 1.4 of the "Marlin Trust Management Document—for IPTV-ES "Copy Control and Output Control".

B.9 Copy Control and Output Control

The VOD reception/playback shall be controlled as described in the VOD license. In so doing, the copy control and output control shall be in accordance with Section 2 of the "Marlin IPTV-ES/J Specific Compliance Specifications: VOD Part".

B.10 DRM System Name

In the Marlin IPTV-ES system, ‘marlin_iptv_es’ shall be used as the value of the drm_system attribute of the LLI.

［Appendix C］ Explanation of Operation of I-Frame TS System

<Explanation 1>

In the operation of the I-frame TS system at the transmitting side, packets in the same I-frame section shall be continuously sent at the original rate so that the processing performed at the receiving side becomes the same as the processing performed during normal playback. In addition, those packets shall be transmitted in such a way that the receiver can naturally decode them along a time axis adjusted to the actual transmission rate, such as 2 times, 4 times or 10 times the original rate. Figure C-1 shows an example of TS during normal playback and an example of double-speed TS transmission in the I-frame TS system. It is considered possible to insert an I-frame in each section without transmission in order to implement trick playback smoothly. In this case too, the transmission rate shall be made the same as the original rate.

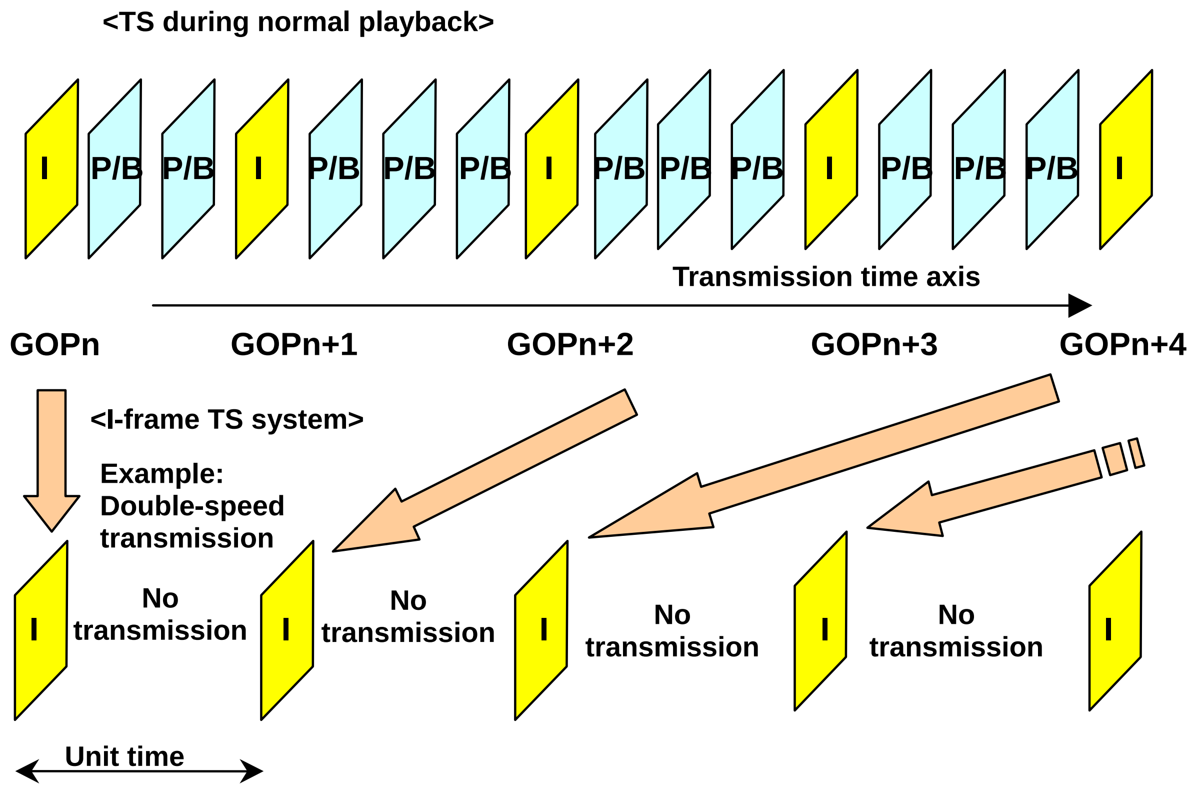

Figure C-1 Example of Double-speed Transmission in I-frame TS System

It should be noted, however, that in the MPEG2-TS layer, the PSI, PCR, audio, caption and NULL are multiplexed and hence, the transmission rate in unit time on the double-speed TS stream corresponds to original rate in the MPEG2-TS layer.

The I-frame TS system shall not be operated for HTTP streaming.

［Appendix D］ Explanation of DRM-related Operations

D.1 Concept of License in VOD Service

The VOD license used in VOD service is one that the DRM client obtains from the DRM server each time it plays a specific VOD content. It is valid only during the playback. On the other hand, the ECG metadata contains information of ValidityIntervalStart, ValidityIntervalEnd, Duration—usage conditions in real terms. The license ID is assumed to be associated with those usage conditions. In VOD, therefore, it is presupposed that the license ID is positioned and operated as an ID that is identified by the usage conditions, including the time-limit information, for specific VOD content. Thus, when the DRM client requests a license based on the license ID described in the LLI from the DRM server, the DRM server checks the time-limit information associated with the license ID and delivers the VOD license to the DRM client only when the client has been authenticated and the term of validity has not expired. Figure D-1 shows a case in which DRM clients obtain a license for the same content using license IDs based on different usage conditions. This case indicates that even when usage conditions differ, the body of the VOD license delivered to the DRM clients remains the same.

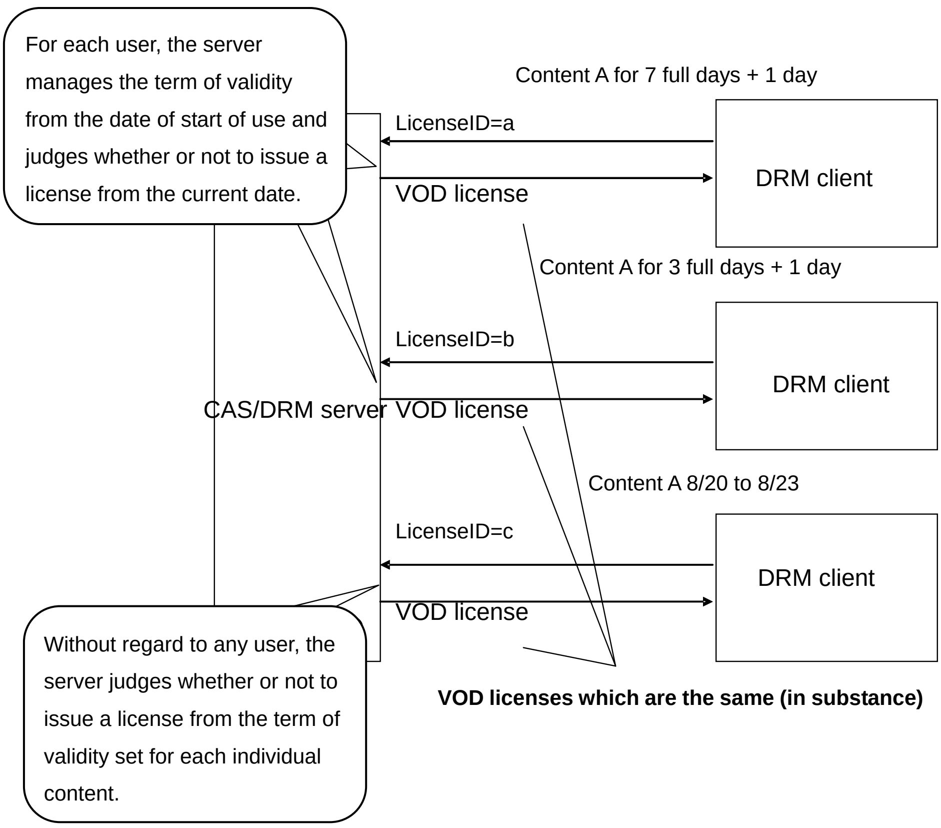

Figure D-1 Concept of License in VOD Service

D.2 Relationship between DRM-related IDs in VOD Service

The relationship between various DRM-related IDs in the VOD service, from the subscription to service to the playback of content, are explained below.

D.2.1 Definitions of related ID models

- User ID (UID)

  This is an ID that identifies a specific user. It is assumed to be set by the service provider not as an object of specifications but as an operation.
- CAS/DRM client identifier (DRM_ID)

  This is an ID that identifies a specific DRM client unique to a specific receiver as defined in this document.
- Content ID

  This is an ID that identifies a specific VOD content. It is assumed to be set by the service provider not as an object of specifications but as an operation.
- License ID (LID)

  This is an ID that identifies a license associated with a specific content defined in this document.
- Package ID (PID)

  This is an ID that identifies the unit of a specific content as a product. It is assumed to be set not as an object of specifications but as an operation.

D.2.2 Types of packages, and relationships between package, license and content

Each package of content product consists of one or more licenses that represent the right to use the content. Each license is always associated with a specific content and content key. On the other hand, packages can largely be divided into four types—"single item", "pack", "all-you-can-view" and "select". For each package type, the relationship between package ID and license ID/content ID is shown below.

(a) Single Item

A single item is a package consisting of only one license. An example of the relationship between package ID, license ID and content ID for single items is given in Figure D-2. Even single items of the same content become different packages consisting of different license when the usage conditions are different.

(b) Pack

A pack is a package consisting of a finite number of fixed licenses. The relationship between package ID, license ID and content ID for packs is shown in Figure D-3.

(c) "All-you-can-view"

This is a package whereby the user can view, without limitation, multiple unspecified content included in a specified scope by signing a monthly contract, etc. with the service provider. In this case, the package consists of multiple unspecified licenses. Thus, licenses as the constituent elements of the package vary every month. The relationship between the package ID, license ID and content ID for "all-you-can-view" is shown in Figure D-4.

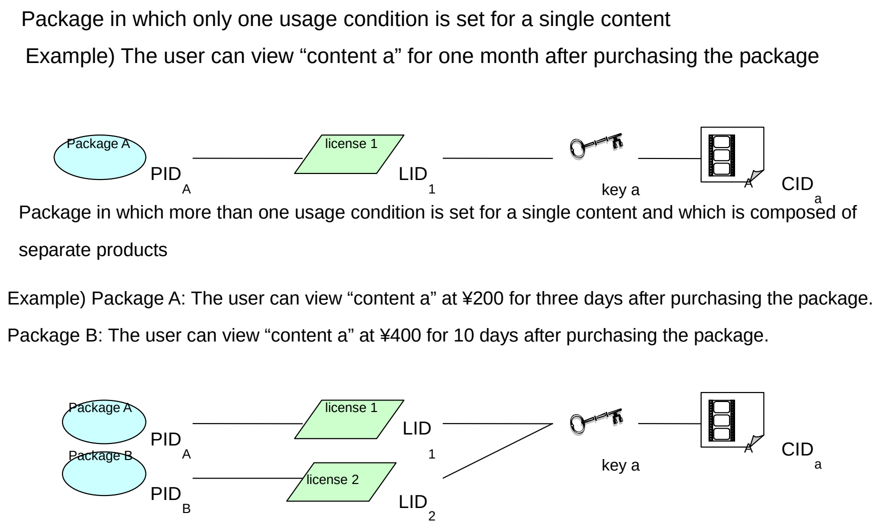

Figure D-2　Relationship between Package ID, License ID and Content ID for Single Items

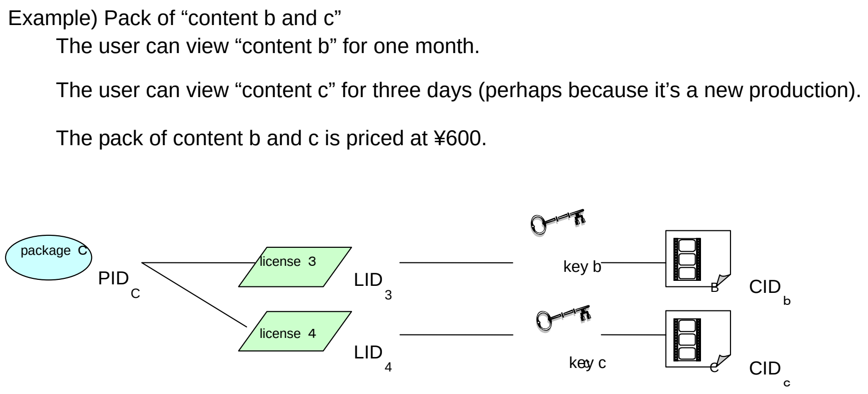

Figure D-3　Relationship between Package ID, License ID and Content ID for Pack

Example) “All-you-can-view movies” for ¥1,500 per month

- The user can view 20 movies in one month from April 1, 2007.
- From May 1, 2007, the user can view 20 movies in one month (10 of those movies are replaced with other ones).
- For example, the user can view the “content *d*” for one month from April 1, 2007.
- By renewing the contract, the user can view “content *d*” for another one month from May 1, 2007.

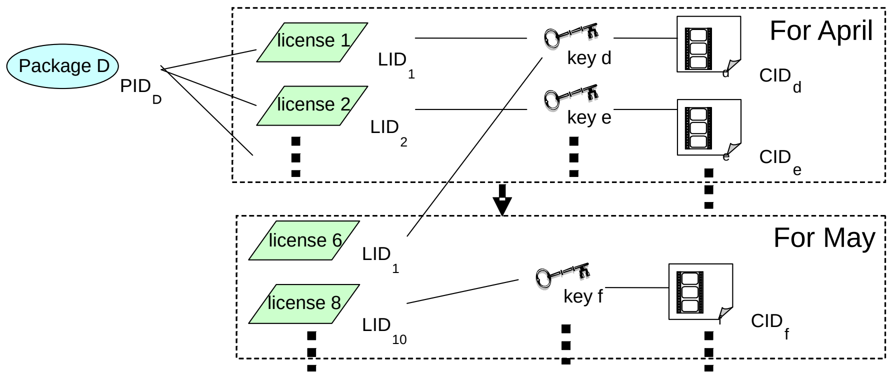

Figure D-4 Relationship between Package ID, License ID and Content ID for "All-you-can-view"

Example) “Selected movies” for ¥500 per month

- In one month from April 1, 2007, the user can view two out of 20 movies.
- In one month from May 1, 2007, the user can view two out of 20 movies (10 of those movies are replaced with other ones).

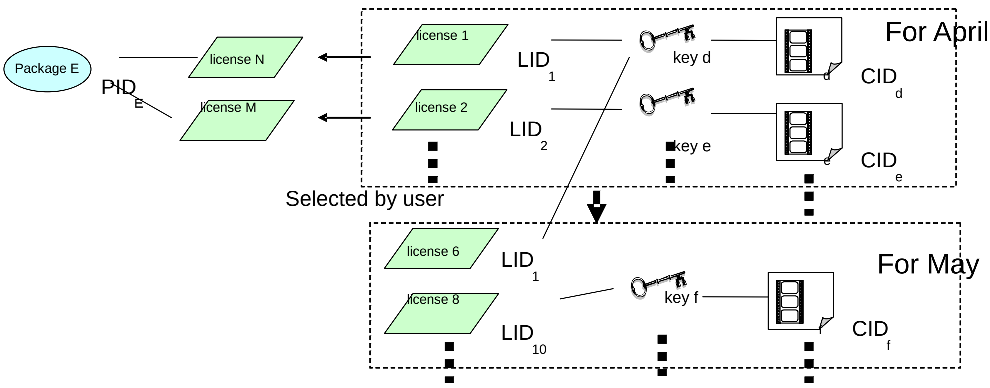

Figure D-5 Relationship between Package ID, License ID and Content ID for Select

(d) Select

"Select" is a package whereby the user can view a finite number of content which the user selects from among multiple unspecified content included in a specified scope. In this case, the multiple licenses which the user can select from are assumed as the constituent elements of the package. As in the case of "all-you-can-view", it is presupposed that the user will renew the contract on a monthly basis, etc. Therefore, the licenses that form the constituent elements of the package change as the content that make up the population change and the user selects specific content. The relationship between package ID, license ID and content ID is shown in Figure D-5.

D.2.3 Registration of user and device in basic registration

The basic registration is the process whereby the user makes necessary registrations with the service provider before using the service. By this, the user identifier (UID) that identifies the user and the CAS/DRM client identifier (DRM_ID) that identifies the user's device are registered in the customer management server. Here, it is conceivable that there are a number of variations in operation according to how UID and DRM_ID are bound to each other. Typical variations in operation are shown in Figure D-6

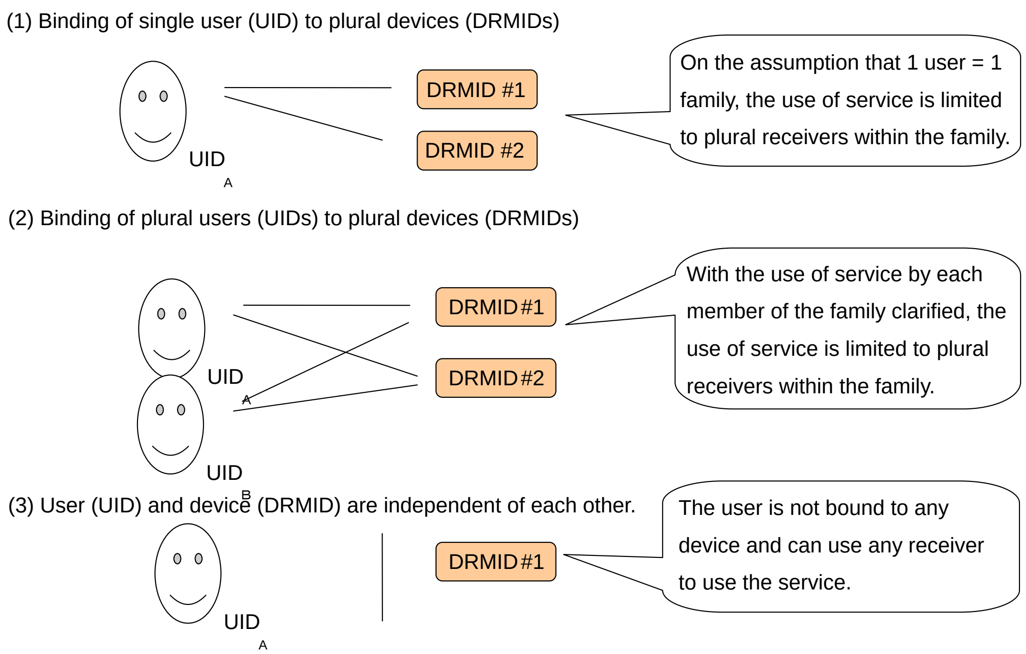

Figure D-6 Concept of Binding of User to User's Device

In the case of (1) in Figure D-6, there are two variations in operation, which is giving the license to the user or to the device in the subscription to the service. Operation (3) is excluded from the scope if these Specifications.

D.2.4 Connection between IDs by subscription to service

When the registration of a package as a product and the basic registration of the user are completed, the connection between the IDs shown in D.2.2 and D.2.3 is built by the servers of the individual service providers. On this assumption, when the user subscribes to services, the IDs of the two systems are connected together. At this time, the operation differs according to whether the license for use of the content is given to the user or the user's device as a result of the subscription to service. If the license for use of content is given to the user when UID and DRM_ID during the basic registration are bound together as shown in (1) in Figure D-6, the operation of a so-called domain is implemented, making it possible for plural devices within the family to use the service via a single subscription to service. If the license is given to a specific device, the user can use the service only on the device with which a subscription to service is made, even when the user owns more than one device. Looking at the connection between IDs during subscription to service, in the former case the user ID (UID) and package ID are bound together, whereas in the latter case the CAS/DRM client identifier (DRM_ID) and package ID are bound together. Figure D-7 shows the concept of connection between IDs during subscription to service.

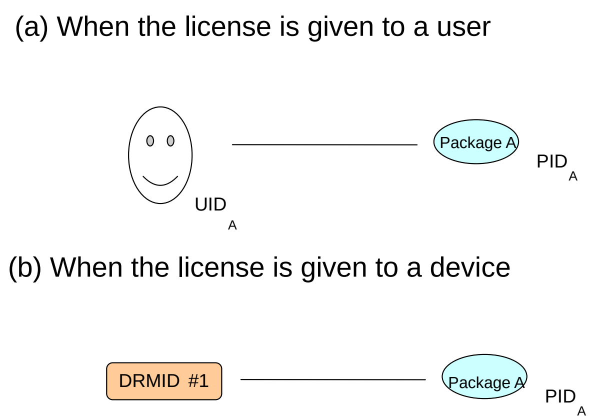

Figure D-7 Connection between IDs during Subscription to Service

D.2.5 Use of relationship between IDs during acquisition of license

When the subscription to service is completed, the connection between IDs shown in Figure D-8 is built on the server of the service provider. On the basis of this assumption, the DRM client obtains a license during playback of the content. As shown in Figure D-9, from DRM_ID transferred from the DRM client to the DRM server during establishment of a SAC and the license ID contained in the license request message, the server of the service provider confirms the bind between the DRM_ID and LID with the aid of the DRM server, customer management server, etc. and transfers the license only when they have been properly bound together and the usage conditions, including the term of validity, are met.

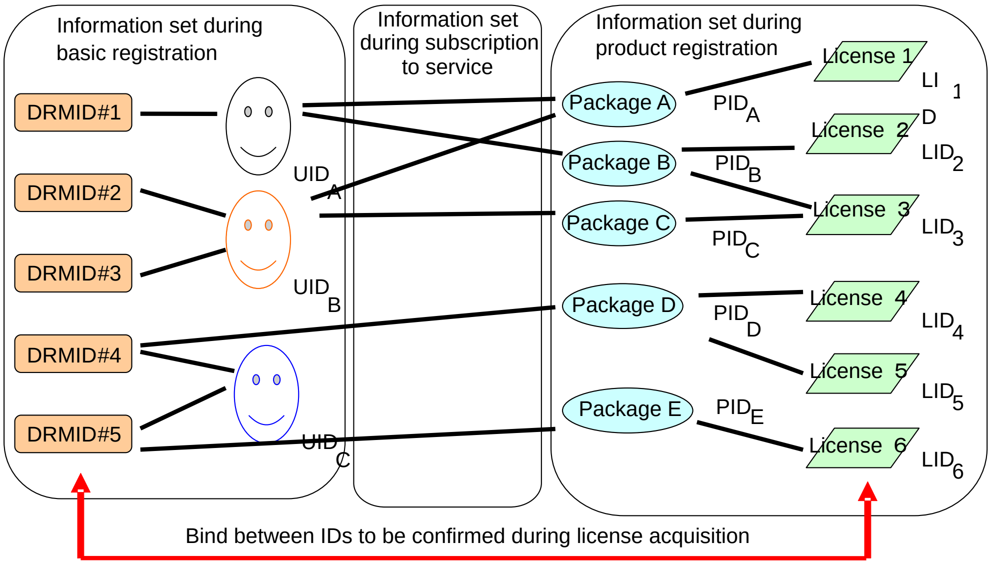

Figure D-8　Connection between IDs after Subscription to Service

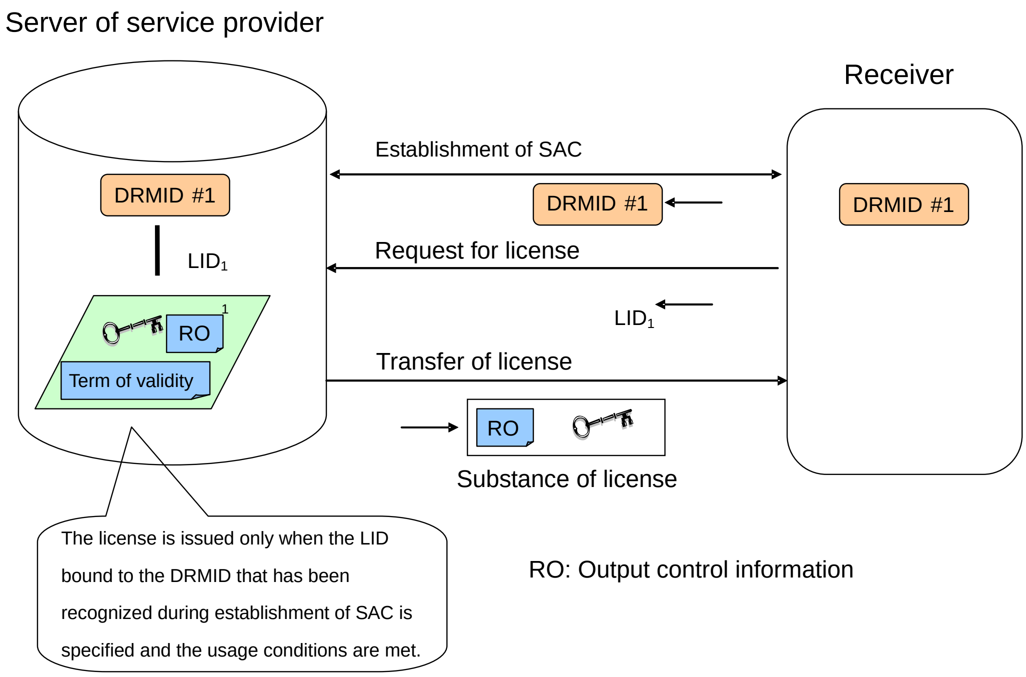

Figure D-9　Coordination of IDs for License Acquisition

IPTV Standard

VOD Specifications
IPTVFJ STD-0002 Version 1.1

September 11, 2008 Version 0.9 issued
December 8, 2008 Version 1.0 issued
August 6, 2010 Version 1.1 issued

IPTV Forum Japan

3F, Futaba Akasaka Bldg.,
8-5-43, Akasaka, Minato-ku, Tokyo 107-0052, Japan
Phone: 03-5858-6685
Fax: 03-5858-6675
e-mail: sec@iptvforum.jp