Technical Report on
HIKARI Service Network Architecture

Ver. 1.2

March 26, 2002

HIKARI Service Architecture Consortium
Technical Committee
### Authors (1/2)

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aoyagi Yasuhiro</td>
<td>Inktomi Japan K.K.</td>
</tr>
<tr>
<td>Ariizumi Takeshi</td>
<td>KDDI Corporation</td>
</tr>
<tr>
<td>Demura Tatsuhiko</td>
<td>Intel Corporation</td>
</tr>
<tr>
<td>Hasegawa Satoshi</td>
<td>NEC Corporation</td>
</tr>
<tr>
<td>Hashimoto Junichi</td>
<td>Nippon Telegraph and Telephone East Corporation</td>
</tr>
<tr>
<td>Hayashi Kazuhiro</td>
<td>Pioneer Corporation</td>
</tr>
<tr>
<td>Hayatomo Satoshi</td>
<td>NTT Communications Corporation</td>
</tr>
<tr>
<td>Igarashi Tatsusi</td>
<td>Fujitsu Limited</td>
</tr>
<tr>
<td>Inami Makoto</td>
<td>Fujitsu Limited</td>
</tr>
<tr>
<td>Jinzenji Hiroshi</td>
<td>Nippon Telegraph and Telephone West Corporation</td>
</tr>
<tr>
<td>Kobayashi Akira</td>
<td>NEC Corporation</td>
</tr>
<tr>
<td>Kuhwara Hiroshi</td>
<td>Fujitsu Limited</td>
</tr>
<tr>
<td>Mizuhara Noboru</td>
<td>Hitachi, Ltd.</td>
</tr>
<tr>
<td>Morita Hideyo</td>
<td>Nippon Telegraph and Telephone Corporation</td>
</tr>
<tr>
<td>Morita Shinya</td>
<td>NTT Communications Corporation</td>
</tr>
<tr>
<td>Nagamatsu Hideko</td>
<td>Fujitsu Limited</td>
</tr>
<tr>
<td>Nakano Go</td>
<td>Matsushita Electric Industrial Co., Ltd.</td>
</tr>
<tr>
<td>Nakata Toshiyuki</td>
<td>NEC Corporation</td>
</tr>
<tr>
<td>Niki Teruki</td>
<td>Matsushita Electric Industrial Co., Ltd.</td>
</tr>
<tr>
<td>Nishi Hiroyuki</td>
<td>Nippon Telegraph and Telephone Corporation</td>
</tr>
<tr>
<td>Nomura Tetsuya</td>
<td>Nippon Telegraph and Telephone East Corporation</td>
</tr>
<tr>
<td>Ohta Naohisa</td>
<td>Sony Corporation</td>
</tr>
</tbody>
</table>

(in alphabetical order)
### Authors (2/2)

<table>
<thead>
<tr>
<th>Name</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>OKEYA Katsuyuki</td>
<td>Hitachi, Ltd.</td>
</tr>
<tr>
<td>SOUI Hiroyuki</td>
<td>Matsushita Electric Industrial Co., Ltd.</td>
</tr>
<tr>
<td>SUGIHARA Motooki</td>
<td>Pioneer Corporation</td>
</tr>
<tr>
<td>TAKAHASHI Eichiroy</td>
<td>Fujitsu Limited</td>
</tr>
<tr>
<td>TAKAYAMA Katsunori</td>
<td>Fujitsu Limited</td>
</tr>
<tr>
<td>TAKAYAMA Kazuho</td>
<td>Hitachi INS Software, Ltd.</td>
</tr>
<tr>
<td>TANABE Yoshinori</td>
<td>Nippon Telecommunications Consulting Co., Ltd.</td>
</tr>
<tr>
<td>TANI Mikiya</td>
<td>NEC Corporation</td>
</tr>
<tr>
<td>TANNO Hajime</td>
<td>Intel Corporation</td>
</tr>
<tr>
<td>URA Yasuo</td>
<td>Sharp Corporation</td>
</tr>
<tr>
<td>YOSHIURA Hiroshi</td>
<td>Hitachi, Ltd.</td>
</tr>
</tbody>
</table>

*(in alphabetical order)*
1. Introduction
2. Basic Model
3. Functional Model and System Architecture
4. Case Studies : System Configuration Examples
5. Technical Issues on HIKARI Service Platform
6. Technical Issues on Security
7. Technical Issues on Copyright Protection
8. Technical Issues on HIKARI Service Terminals
9. Technical Issues on Protocols and Content Formats
1. Introduction

Contents of Chapter 1

1-1. Background
1-2. Objectives of HSAC
1-3. Business Development
1-4. Operational Organization of the Consortium
1-5. How the Study will Proceed(in Consortium)
1-6. How the Study will Proceed(in Technical Committee)
HIKARI Service Architecture Consortium (HSAC) is in the process of implementing a high degree of secure, new IT business platform for the 21st Century called the HIKARI Service Platform, which can be used by both content providers and content users for business and personal applications. For example, if there is an environment that allows easy access to TV level quality bi-directional video communication service, this high quality service will be easily available to the local community.

However, there are uncertainties about the emerging copyright protection and information security issues that continue to become more serious in the broadband age. Technologies such as content ID, electronic watermark, and data encryption have been studied as safety measures for such issues. Applying these technologies effectively to the HIKARI Service provided to the general user would require consistent interface conditions from the sending side to receiving side. To enable video distribution service providers to provide video content that satisfy the required high standards, the service quality for stream data also must be assured.
HSAC is contemplating an optical fiber network, which is expected to be the mainstay of broadband access. It is analyzing service/player models, functional models, and system models from a standpoint of providing high-security services that can be safely used by content providers and users. The analysis results are also disclosed widely to set down the interface conditions.
HSAC studies and defines the interface conditions required in order to provide the HIKARI Service. This Consortium intends to collaborate as required with other related governing bodies for standards, and will conduct its studies based on a policy of applying all existing interface conditions that can be utilized for the HIKARI Service. It will carry out activities aimed at using the interface conditions between the systems defined in these tasks as a basic model for each industry.
HSAC has set up two committees: the Service Committee and Technical Committee under the President, Directors Meeting, and Administrative Office, and is holding discussions by the working groups (WG) of each committee. The Technical Committee is conducting ad hoc studies focused on services and technologies respectively, and is carrying out a short-term intensive analysis before further studies by the three WG system shown in the above chart.
1-5. How the Study will Proceed (in Consortium)

The investigative process by the Consortium is shown in the above figure. The Service Committee and Technical Committee are carrying out joint studies at HSAC. With support from the Service Committee, the Technical Committee selects a specific service from trends in the outside world that has the potential of becoming a popular feature of HIKARI Service in the future, and then defines the player models related to this service. Also, the Technical Committee will reveal system models and interface conditions for implementing a specific service.

The main results of these studies are outlined below.

1. **Summary of HIKARI Service Concepts**

   Service Specification defines the services that can be incorporated in HIKARI Service. It describes the features, attributes, and requirements of these services.


   Technical Report defines the service network architecture for implementing a service defined in Item 1 (Service Specification). It describes the interface requirements that must be defined by Item 3 (Interface Requirements for HIKARI Services).

3. **Interface Requirements for HIKARI Services**

   System Interface Requirements defines the interface conditions for each interface requirement extracted by Item 2 (Technical Report on HIKARI Service Network Architecture).

The process of studies by the Technical Committee is shown in the above figure.

The Technical Committee proceeds in four steps shown in the above figure to establish the details of the HIKARI Service Platform. This “Technical Report on HIKARI Service Network Architecture” summarizes the results of the first three study steps (from “Study on HIKARI service mode/architecture” to “Study on system model/architecture”).
This chapter describes the services, players and functions related to basic services that are the objects of studies by HSAC for the HIKARI Service.

In this chapter, the services are placed in two categories based on the relationship between service providers and users so that the contents of the studies can be used as general-purpose models. More specifically, we have classified the services into two basic models: one for B2B2C services and the other for C2C services. A company provides B2B2C services to a general user via another company player. C2C services are used for service or information exchange between general users.

However, because B2B2C and C2C services are too wide-ranging, we have created specific models as objects of services that seem to be most representative among these services.
2-1. Given Conditions on Network Infrastructure (1/2)

The HSAC Technical Committee does not study details of the network infrastructure. Instead, it studies higher technological issues such as the technology for video content distribution/delivery service on the network infrastructure, technology for creating models, and various management control technologies related to the content itself and content distribution/delivery. Therefore, the network infrastructure is presented as a prerequisite condition for the studies conducted by the Technical Committee. The network infrastructure conditions are defined in the above figure before the models are created. Each network infrastructure condition is described below.

(1) IP based network
We have selected the Internet Protocol (IP) as the communication protocol for the network layer to be used for these services.

(2) Always-on NW
Basically, it is assumed that a user is always connected to the network. Note, however, that “Always-on NW” here means “a state where connection is always enabled (fixed charge)” or “a state where the volume of use is close to the always connected status”. Always-on NW is similar to the flat rate service which is currently and widely used in a public network.

(3) IP reachability through the Internet
To implement the HIKARI Service studied by HSAC, a high-speed communication network (such as CDN) is required as an alternative to the current best-effort Internet. However, IP reachability all over the world is also essential. Therefore, connection to content providers via the Internet must be enabled. The data flow from “Content providers” and back via “The Internet” in the above figure indicates how the Internet provides reachability to IP content all over the world. If it is combined with the data flow from “Content providers” to “Users” in the above figure, video content located in Paris can be delivered to Japan via the Internet and temporarily stored in a server, and then delivered in a form that is suitable for a user over the high-speed optical environment provided by HIKARI Service Platform.
The following issues regarding the Internet, which is the current service platform, must be addressed to implement the HIKARI Service studied by HSAC. These issues are handled as prerequisite conditions and approached as technical subjects in the research for the implementation method. Chapters 5 and 6 provide the details.

(4) **Bi-directional broadband NW**
Access to a bi-directional broadband service is a prerequisite for HIKARI Service. Broadband here mainly indicates the 10Mbps to 100Mbps range. Other ranges can also be included as required (between 1.5Mbps and several Gbps).

(5) **Security, quality**
Security and quality assurance is a prerequisite for the HIKARI Service Platform.

(6) **Guaranteed transmission speed (End-to-end)**
Guaranteed end-to-end transmission speed is a prerequisite for the HIKARI Service platform. Also, it must be possible to assure different end-to-end speeds according to the rates charged.

(7) **1:N multicast**
It is assumed that the HIKARI Service Platform will provide the 1:N multicast function as a system to provide live content to a large number of viewers. Chapter 4 describes the above live content provider service as a net-live service. The implementation technology of multicast is discussed in Chapter 5.
2-2. Basic Model for B2B2C Services

Contents of Section 2-2

2-2-1. Focusing Services
2-2-2. Player Model
2-2-3. Definition of Players
2-2-4. Functions on each Player
2-2-1. Focusing Services

HIKARI Service Platform

Examples
- Online Shopping
- Content Delivery (Software, music, video, etc.)

Focusing on “Video Delivery Service”
- It requires broadband HIKARI Service Platform.
- Large demand can be expected.

The service provided by a company (provider) to a general consumer (user) is defined as a B2C service. This section discusses the B2B2C service, which is provided by one company to a general user via another company so that the service can be even more versatile. The B2C or B2B2C service includes the following facilities.

- Online shops (electronic shops) that sell products
- Content distribution/delivery such as software, video, and music

This section studies the video distribution/delivery service from two viewpoints: one is that this service requires a broadband HIKARI Service Platform and the other is that a wide demand by general users is expected. “Video distribution” here is a service that is used to send content to local content providers (i.e. network service company such as ISP) over a wide area from an original owner company. “Video delivery” is defined as a service that is used to send content to a general user from a local content provider. Therefore, video distribution is a B2B service and video delivery is B2C service. The combination of these services is defined as the B2B2C service.
The above figure is an example of players in the video delivery service.

[From creation of content to distribution]
(1) “Video Material Holder” creates video materials and provides the created material to a video content creator.
(2) “Video Content Creator” uses (composition and edit) the material to create a single integrated content.
(3) “Copyright Management Provider” certifies the copyright and assigns a content ID for the video content created by a video content creator.
(4) “Video Content Provider” receives the content provided by a video content creator and provides this original content to a distribution service provider. It also requests the distribution service provider to manage and distribute the content.
(5) “Distribution Service Provider” manages and distributes the original content provided by a video content provider (including distribution of the content to cache/mirror servers).

[From user request to delivery]
(6) “Authentication Service Provider” performs user authentication when a user (terminal) logs on to a network.
(7) After user authentication, the service menu of the video service is provided from the Service Portal.
(8) When the user selects a video content from the service menu, a request is sent to a distribution service provider that manages the selected content. The distribution service provider locates an access point that is optimal for the selected content and returns this access point information to the terminal.
(9) The delivery service provider will deliver the video content according to the user’s request. At this time, if the selected content is not stored in cache/mirror servers of the delivery service provider, the delivery service provider will request a distribution service provider to distribute the video content.
(10) After the service is provided, the user is billed for the content service charge by each provider. “Billing Proxy Service Provider” charges and collects payments on behalf of the providers. “Customer and Terminal Service Provider” provides application codes download, and manages a terminal for a user.
2-2-3. Definition of Players

<table>
<thead>
<tr>
<th>Classification</th>
<th>Players</th>
<th>Definition of Players</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIKARI Service Platform</td>
<td>Distribution Service Player</td>
<td>The provider who distributes original content to the caches/mirrors. This provider decides the optimal allocation of caches/mirrors and has the request-routing function.</td>
</tr>
<tr>
<td></td>
<td>Delivery Service Player</td>
<td>The provider who delivers content to the end-users.</td>
</tr>
<tr>
<td></td>
<td>Service GW</td>
<td>The gateway that selects an appropriate provider based on the information above layer 3.</td>
</tr>
<tr>
<td>Content related Players</td>
<td>Video Material Holder</td>
<td>The provider who makes and sells video materials. (This provider is out of scope in this document)</td>
</tr>
<tr>
<td></td>
<td>Video Content Creator</td>
<td>The provider who makes and sells video content. (This provider is out of scope in this document)</td>
</tr>
<tr>
<td></td>
<td>Video Content Provider</td>
<td>The provider who sells video content to end-users. For the distribution and delivery of content, this provider uses Distribution Service Player and Delivery Service Player.</td>
</tr>
<tr>
<td>User</td>
<td>Terminal</td>
<td>Terminal at the end-user.</td>
</tr>
<tr>
<td>Value Added Service Player</td>
<td>Network Management Service Provider</td>
<td>The player who provides network management services for content distribution and delivery.</td>
</tr>
<tr>
<td></td>
<td>Billing Proxy Service Provider</td>
<td>The provider who charges for the end-user and collects charge from the end-user instead of other providers.</td>
</tr>
<tr>
<td></td>
<td>Service Portal</td>
<td>The portal site that provides service menu to end-users.</td>
</tr>
<tr>
<td></td>
<td>Authentication Service Provider</td>
<td>The provider who provides authentication services.</td>
</tr>
<tr>
<td></td>
<td>Copyright Management Provider</td>
<td>The provider who certifies content copyright, and provides/manage content IDs.</td>
</tr>
<tr>
<td></td>
<td>Customer and Terminal Service Provider</td>
<td>The provider who maintains terminal and provides application codes.</td>
</tr>
</tbody>
</table>

The above table indicates possible players for video distribution services and their definitions.

“Main Players” in the former part of the table provide the basic functions that are indispensable for implementation of services.

On the other hand, “Value Added Service Players” in the latter part of the table perform a part of the functions of other players by proxy. Therefore, if an original player can provide the services including the corresponding functions, the value added service players are not required. Also, if a corresponding function is outsourced to the value added service players, the original player is not required to have this corresponding function.
This section describes the functions that each provider should be able to provide. Among the providers shown in the above figure, “Video Material Holder” and “Video Content Creator” are not within the scope of the studies of HSAC. Therefore, their functions are not described below.

**Functions provided by “Content Provider”**

**Video content coding:** Function for coding the video content into the required format.

**Copyright protection:** Prevents replication and tampering of content by coding and inserting an electronic watermark in the video content.

**Advertisement insertion:** Advertisement insertion function. A content provider often inserts advertisements as a part of the video content.

**Security:** Firewall function for encryption and prevention of unauthorized access.

**Functions provided by “Distribution Service Provider”**

**Video content distribution:** Function for sending original video content to cache/mirror servers.

**Billing:** Function for billing a provider. “Billing” of a distribution service provider means charging a content provider under contract for content distribution. The billing may be a fixed-rate or pay-per-use system, which is a charge proportional to the amount of use according to frequency of access.

**Path connection for content distribution:** Communication path setting function for distributing the video content to cache/mirror servers.

**Cache/mirror allocation management:** Function for managing the data on allocation of video content to the cache/mirror servers.

**Receipt management:** Function for managing payments from a contract provider.

**Multicast:** Function for providing the same content simultaneously to multiple specific cache/mirror servers on a network. The distribution service provider uses multicast to distribute the video content to cache/mirror servers.

**Request-routing:** Function for providing a user with a route to an optimal site from the cache/mirror server allocation data and congestion data of each cache/mirror server.

**Security:** Firewall function for encryption and unauthorized access prevention.

**Advertisement insertion:** Advertisement insertion function. A distribution service provider often inserts an advertisement as a part of the video content.

**Customer management:** Customer information management function related to a provider.
Functions provided by “Delivery Service Provider”

- **Video content delivery**: Function for sending video content to a user.
- **Path connection for content delivery**: Communication path setting function for delivering the video content to users.
- **Cache/mirror**: Function as a cache/mirror of original video content.
- **Multicast**: Function for providing the same content simultaneously to multiple specific nodes on a network. The delivery service provider uses multicast when the video content is delivered to a user.
- **Access log management**: Function for managing access information to the video content allocated in a cache/mirror.
- **Video content control**: Functions for controlling the video content such as start, stop, slow, and fast-forward during video playback.
- **Video content ticket management**: Function for detecting copyright infringement of a video content and for controlling viewing restrictions within the scope of the license.
- **Billing**: Function for billing a user.
- **Customer management**: Function for customer information management related to a provider. Customer management of a delivery provider means the management of contract user information. The information to be managed is a user’s address, name, age, sex, user ID, and password, and details of the user’s service contract.
- **Receipt management**: Function for managing payments from a contract user.
- **Transcoding**: Function for encoding and converting the video content in accordance with the terminal conditions of a user.
- **Security**: Firewall function for encryption and unauthorized access prevention.
- **Advertisement insertion**: Function for inserting advertisements. Advertisement insertion by a delivery service provider may be implemented in the form of switching and displaying between video content and video advertisement, as a banner or text that appears on a part of the menu screen or viewer screen, or in other forms.

Functions provided by “Service GW”

- **Connection and disconnection with terminal**: Function for connecting and disconnecting to/from a network in accordance with the request from a terminal.
- **Data flow control**: Function for assigning providers (specify the connection destination) based on the information of layer 3 or higher.

Functions provided by “Terminal” (Note that a part of these functions can be handled by the network side. See Chapters 5 and 8.)

- **Video control**: Function for decoding received encoded video content in accordance with the coding rules being used and then displaying it on a screen.
- **Storage of video content**: Function for saving received video content on a storage device.
Connection and disconnection to the network: Function for connecting and disconnecting to/from network.

Security: Function that provides a firewall for encryption to prevent eavesdropping and unauthorized access.

Terminal security: Function that provides a firewall to prevent unauthorized access to home equipment.

**Functions provided by “Copyright Management Provider”**

Certification and management of copyright: Function for certifying the content copyright and for issuing/managing a content ID.

**Functions provided by “Network Management Service Provider”**

Network service management: Function for performing management tasks, such as network operations to provide the proper services to a provider, and to correct network troubles.

Bandwidth control: Function for securing the network band required for transferring video content and performing Quality of Service (QoS) control.

**Functions provided by “Service Portal”**

Management and provision of service menu: Function for providing a service menu to a user.

Information search: Video content search function for information requested by a user.

**Functions provided by “Authentication Service Provider”**

Authentication: Function for authenticating a user who is attempting to access a network or provider.

**Functions provided by “Billing Proxy Service Provider”**

Billing delegation: Function for billing a user in place of a provider.

**Functions provided by “Customer and Terminal Service Provider”**

Terminal maintenance: Function for managing a terminal and providing application software to a terminal.
2-3. C2C Service Basic Model

Contents of Section 2-3

2-3-1. Focusing Service
2-3-2. Player Model
2-3-3. Definition of Players
2-3-4. Functions on each Player
2-3-1. Focusing Service

Focusing on “Bi-directional video communication services”
- It requires broadband HIKARI Service Platform.

C2C service is defined as a service that is used for exchanging information between users. The various C2C services may be as follows. (Sometimes a company (provider) provides a site for C2C services. Such C2B2C services are also referred to as the C2C service.)

- Auction over the Internet (A company provides a system and “site” for conducting an auction over the Internet.)
- Chat
- Bi-directional video communication
- Network game
- Personal telecasting

This document focuses on the bi-directional video communication service that requires the broadband HIKARI Service Platform.
2-3-2. Player Model

The above figure shows examples of players that are assumed in the bi-directional video communication service and flow of services.

(1) “Authentication Service Provider” authenticates a user (terminal) who connects to the network.

(2) After user authentication, “Service Portal” provides a communication service menu.

(3) When a user selects the bi-directional video communication service from the menu and specifies a communication destination, a connection request is sent to “Session Management Provider”.

(4) When a response is received from the destination terminal, a bi-directional video communication session is established.

(5) “Delivery Service Provider” starts bi-directional video communication service to the destination terminal requested by a user.

(6) After providing the service, the Session Management Provider bills the user with the bi-directional video communication service charge. The billing and collection of the charges are taken care of by “Billing Proxy Service Provider”.

“Customer and Terminal Service Provider” provides application codes download for a user, and manages the terminal.
### 2-3-3. Definition of Players

<table>
<thead>
<tr>
<th>Classification</th>
<th>Players</th>
<th>Definition of Players</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIKARI Service Platform</td>
<td>Session Management Provider</td>
<td>The provider who provides bi-directional video communication service to the end-users.</td>
</tr>
<tr>
<td></td>
<td>Delivery Provider</td>
<td>The provider who delivers video stream information bi-directional.</td>
</tr>
<tr>
<td></td>
<td>Service GW</td>
<td>The gateway that selects an appropriate provider based on the information above layer 3.</td>
</tr>
<tr>
<td>User</td>
<td>Terminal</td>
<td>Terminal at the end-user.</td>
</tr>
<tr>
<td>HIKARI Service Platform</td>
<td>Network Management Service Provider</td>
<td>The player who provides network management services for content distribution and delivery.</td>
</tr>
<tr>
<td></td>
<td>Billing Proxy Service Provider</td>
<td>The provider who asks charge for the end-user and collects charge from the end-user instead of other providers.</td>
</tr>
<tr>
<td></td>
<td>Service Portal</td>
<td>The portal site that provides service menu to end-users.</td>
</tr>
<tr>
<td></td>
<td>Authentication Service Provider</td>
<td>The provider who provides authentication services.</td>
</tr>
<tr>
<td></td>
<td>Copyright Management Provider</td>
<td>The provider who certifies content copyright, and provides/manages content IDs.</td>
</tr>
<tr>
<td></td>
<td>Customer and Terminal Service Provider</td>
<td>The provider who maintains terminal and provides application codes.</td>
</tr>
</tbody>
</table>

The above table indicates and defines possible players for bi-directional video communication services.

“Main Players” in the former part of the table provide the basic functions for implementation of services.

On the other hand, “Value Added Service Players” in the latter part of the table performs a part of the functions of other players by proxy. Therefore, if an original player provides services including corresponding functions, these value added service players will not be required. Also, if a corresponding function is outsourced to the value added service players, the original player is not required to have the corresponding function.
This section describes the functions that each provider should be able to provide.

**Functions provided by “Session Management Provider”**

Customer management: Function for managing a series of connection processes between users based on the customer management information (including terminal attribute information). This is the most important function of the Session Management Provider. This function includes management of a conference session as well as 1:1 user management.

Access log management: Function for managing the access information of a user.

Billing: Function for billing a provider. “Billing” of a session management provider is the billing for a contract user. The target of billing may be (1) a charge for using the service and (2) a charge for using the network line. In some cases, (2) is applied for billing an individual provider.

Receipt management: Manages the payment from a contract user.

Security: Firewall function for encryption and unauthorized access prevention.

Advertisement insertion: Advertisement insertion function. Advertisement insertion by a session management provider may be in the form of a banner or text appearing on a part of the menu screen, or in some other form.

**Functions provided by “Delivery Service Provider”**

Video content delivery: Function for sending the one’s own video shot to others.

Path connection for bi-directional video communication: Communication path setting function for establishing bi-directional video communication.

Multicast: Function for providing the same content simultaneously to multiple specific users in a network. This function is required for bi-directional video communication between three or more users.

Video content control: Control function required for video content playback.

Transcoding: Function for encoding and converting video content in accordance with the terminal settings of a user.

Security: Firewall function for encryption and unauthorized access prevention.

Advertisement insertion: Advertisement insertion function. Advertisement insertion by a delivery service provider may be in the form of a banner or text appearing on a part of the communication screen, or in some other form.
Functions provided by “Service GW”

Connection and disconnection with the user terminal: Function for connecting and disconnecting to/from network in accordance with a request from a terminal.

Provider selection: Function for assigning providers (specify the connection destination) based on the information of layer 3 or higher.

Functions provided by “Terminal”  (Note that a part of functions can be handled by the network side.  See Chapters 5 and 8.)

Video content control: Function for decoding a received encoded video content in accordance with coding rules being used, and then displaying it on a screen.

Video content storage: Function for saving a received video content on a storage device.

Video content coding: Function for coding video in accordance with the coding rules being used.

Video content transmission: Function for sending video.

Connection and disconnection to the network: Function for connecting and disconnecting to/from a network.

Security: Firewall function for encryption and unauthorized access prevention.

Terminal security: Function such as a firewall for preventing unauthorized access to home equipment.

Home terminal accommodation: Function for accommodating home equipment and providing a connection interface to a network.
2-3-4. Functions on each Player(3/3)

Functions provided by “Service Portal”
Management and provision of service menu: Function for providing a service menu to a user.

Functions provided by “Authentication Service Provider”
Authentication: Function for providing user authentication for access to a site.

Functions provided by “Network Management Service Provider”
Network service management: Function for performing management tasks, such as network operations to provide the proper services to a provider, and to correct network troubles.
Bandwidth control: Function for securing the network band required for transferring video content and performing QoS control.

Functions provided by “Customer and Terminal Service Provider”
Terminal maintenance: Function for managing a terminal and providing application software to this terminal.

Functions provided by “Billing Proxy Service Provider”
Billing delegation: Function for billing a user in place of a provider.
### Contents of Chapter 3

3-1. Functional Architecture Model  
3-2. System Architecture Model

This chapter defines an architecture model that is used as the basic system configuration derived from the studies described in Chapter 4.
In this section, a functional architecture model is created based on the functions provided by the players described in Chapter 2. The players are first separated into three categories: user-related players, HIKARI Service Platform-related players, and content-related players. The functions provided by each player are re-sorted.

The functions provided by “Terminal” and “Customer and Terminal Service Provider” are sorted into the frame of User-related Functions. The functions provided by “Video Content Provider” and “Copyright Management Provider” are sorted into the frame of Content-related Functions. The functions provided by other players are sorted into the frame of HIKARI Service Platform-related functions. Furthermore, HIKARI Service Platform-related functions are sorted into two categories: “Information Sharing Management & Control Functions” for controlling services and applications and “Information Sharing Networking Functions” for content distribution and delivery.
3-1. Functional Architecture Model (2/2)

The above figure shows the functions (see Chapter 2) that are provided by these players based on the classification of players described in previous section. In the above figure, the functions are further classified into functional groups within the main framework.
3-2. System Architecture Model

The above figure shows the system architecture model from Content Provider to End User based on the functional architecture and functional groups described in section 3-1. The target range of studies by the HSAC Technical Committee is indicated with a thick dotted line in the figure. As mentioned in Chapter 2, the studies of this Technical Committee do not cover the physical infrastructure network. An interface for functions such as QoS control and multicast communication control is necessary between a content server and the underlying network in order to actually perform content distribution/delivery. Therefore the infrastructure network is studied up to a certain degree. The application service itself is basically studied by the Service Committee and not by the Technical Committee. However, the Technical Committee does examine this service to preserve the consistency of the interface.

The End User is placed on the left and Content Provider on the right in the model shown in the above figure. In this model, the configuration elements shown in the upper portion have a higher logical level. Basically, this figure shows a general model that consists of the system configuration elements (servers, GW, etc.) used to implement each functional group. In this figure, there are two types of service gateways: one that connects the End User and HIKARI Service Platform and another that connects the Content Provider and HIKARI Service Platform. These service gateways can be individually defined as more specific system elements when the details of a service are established. In the modeling phase, however, they are simply defined as Service Gateway. In the case studies in Chapter 4, more detailed configuration elements are studied based on this model.

The points where interface specification seems to be necessary between system configuration elements are also shown in the above figure.
4. Case Studies: System Configuration Examples

Contents of Chapter 4

4-1. Broadband Net-Live Services
4-2. Bi-directional Video Communication Services

This chapter defines the system configurations and functional models for a specific service. The service sequence between functional models is explained to clarify the interface specification.

As a specific service example, we have selected “Broadband Net-Live Services” from the B2B2C service category and “Bi-directional Video Communication Services” from the C2C service category.
4-1. Broadband Net-Live Services

Contents of Section 4-1

4-1-1. What is Broadband Net-Live Services
4-1-2. Example of Network Configuration
4-1-3. Example of Players
4-1-4. Example of Service Sequence
4-1-5. Interface Reference Points
4-1-1. What is Broadband Net-Live Services(1/3)

Service Requirements

**Essential Requirements**

- Video/audio quality higher than that of broadcasting media
- Content selection by user demand

**Optional Requirements**

- Providers can deliver content to the specified or unspecified users by using QoS guaranteed CDN
- Users can select their favorite camera angles

The following requirements are assumed for broadband net-live services.

The following two requirements are crucial for the service.

- **Video/audio quality higher than that of broadcasting media**
  
  It is essential to offer high-quality video/audio by fully utilizing the features of the HIKARI Service Platform in order to differentiate the broadband net-live services from current broadcasting media.

- **Content selection by user demand**
  
  A variety of content is supplied from various Content Providers for the HIKARI Service Platform. The users can search and view the preferred content.

Although the following two items are not indispensable, they are assumed as optional items.

- **Providing services for a closed group of users, such as members-only services or local community services**
  
  Unlike ordinary broadcasting, there are merits such as a secure QoS and precise advertisement targets via a distribution service that focuses on specific users.

- **Special operation by interactive control**
  
  A user is allowed to control and view a video as desired. For example, the user can select the camera angle.
The advantages of broadband net-live services are listed above from the viewpoint of users and content providers.

<table>
<thead>
<tr>
<th>For users</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Users can watch the live program from the beginning even after the start time of the live program</td>
<td>My favorite time (using time-shift function) ..........</td>
</tr>
<tr>
<td>Users can select their favorite camera angle</td>
<td>My favorite scene ......</td>
</tr>
<tr>
<td>Users can select their favorite live programs among the large number of programs</td>
<td>My favorite program ......</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For content providers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Providers can specify the users, which would bring the AP variation</td>
<td>Specific content distribution depended on the aim and purpose</td>
</tr>
<tr>
<td>Providers can charge users on a content by content basis because the high quality of content</td>
<td>Increase of content distribution channels</td>
</tr>
</tbody>
</table>
What is the time-shift function?

The time-shift function is provided as a feature of broadband net-live services.

The time-shift function is defined as “a function to store or cache live content at content distribution sites and delivery sites, and deliver this content for a user who requests the viewing after the live feed has already started.”

The issue regarding this function is the implementation of a scheme for delivering the requested content at a different timing while storing this live content at both distribution sites and delivery sites.
Example of system configurations of B2B2C broadband net-live broadcasting type services (Example)

The following six segments are assumed:

1) **Live stage**

   Encodes live material onsite, and feeds it to “Distribution server” on the Distribution Platform via the network equipment, such as a router. Or, feeds the live material to “Content Provider” and/or “Distribution Platform” via an optical video transmission service without encoding the live material.

2) **Content provider**

   Encodes live material and advertisements, performs copyright protection and encryption, and feeds this data to “Distribution server” on the Distribution Platform.

3) **Distribution Platform**

   Encodes video/audio sent from a live stage, performs copyright protection and encryption, and feeds this data to “Distribution server”. The Distribution Platform consists of “Content storage server” for handling the viewing time-shift of users, “Service portal server” for receiving a request from a user and providing menus, “Authentication server” for user authentication, “Distribution management server” for managing the video distribution schedule, distribution log, request routing, distribution band, and others.

4) **CDN Backbone Network**

   Provides a high speed and broadband connection between Distribution Platform and Access Edges.

5) **Delivery Platform**

   Accommodates and collects the optical access lines of HIKARI Service subscribers (=users), and delivers content to a user from a “Cache server”.

   The delivery platform uses the live split function to deliver live content to multiple users.

6) **Home**

   The home LAN mainly consists of “ONU”, used for terminating the optical access line, “HGW”, used for performing security functions and IP address management of a terminal, and “STB”, used for providing video/audio playback and browsing by connecting to a TV set.
Example of players for B2B2C broadband net-live services (Example)

We can expect multiple players to provide the system, especially for the HIKARI Service Platform segment. Based on the reasons below, however, we assume the following four main players.

Reasons:

- It seems unlikely that a charge-based, MPEG-2 6Mbps class, high-speed stream QoS will be provided over a system of mutual connection of multiple providers on an open network. It is more natural for a single distribution service provider to utilize a closed network between the distribution platform and access edge that accommodates subscribers.

- Although there could be individual players in the value added portion such as authentication, customer management, and billing proxy service providers, a single provider is expected to provide the system because of the number of combinations that exist.

Main four players:

- Content provider
- Distribution/delivery provider
- Service GW: For both B-B portion and B-C portion
- User (viewer)
Service Sequence (Example)

An example of the service sequence assuming the following five phases is given below.

1. Purchase of ticket
2. Distribution of content
3. Viewing
4. Completion of viewing
5. Payment

This example also assumes that a content provider will ask a distribution/delivery provider to handle user authentication and proxy billing of information fees.

In the example, this service sequence starts from a user purchasing a live ticket to view the “Japan vs. Italy soccer game”.

1. Purchasing the live ticket

Before the user buys a live ticket, the Content Provider registers a content ID and information charges (content price) in the Content Distribution Function.

The user follows instructions on the screen to search for a specific content and purchases a ticket. The screen is processed by the menu display function of a Service Portal. The source information (content list, information charge, ID, password), however, is provided through the Content Distribution Function.

The information (log) of the ticket purchase is sent to the Content Provider who will provide the content.
### 2. Distribution of content

The live video from the Content Provider is distributed to a cache/mirror of the Content Delivery Function.

First, the Content Provider performs the required tasks such as advertisement insertion, encoding, and encryption for the content.

Meantime, the Content Distribution Function registers attributes, such as content ID and decode key, based on the information from the Content Provider. Also, the distribution schedule and content URL are registered and a remote edge server, which will distribute the content, is specified.

When the content video is uploaded from the Content Provider, the Content Distribution Function sends it to the specified remote edge server. When the content is being distributed in real time, the uploaded video is sent directly. For time-shift viewing, however, the content is stored by the Content Distribution Function and Content Delivery Function at the distribution destination.

---

**Real time live Distribution Process**

- **Live distribution request**
  - Ad. insertion
  - Encoding
  - Copyright protection

- **Live streaming (Distribution)**
  - (IP multicast)

- **Live streaming (feed)**
  - OK

- **Cache**
  - Upload completion notification

- **Distribution completion notification**

**Time-shifted Distribution Process**

- **Live content distribution**
  - (w/ IP multicast)
  - Store

- **Distribution completion notification**

---

**Technical Report on HIKARI Service Network Architecture**

Copyright © 2001-2002 HIKARI Service Architecture Consortium. All rights reserved.
3. Viewing

The user follows screen instructions to select the content offered for the ticket that was purchased in advance.
3. Viewing (continued from previous page)

When the user enters an ID and password in response to an inquiry from the Content Distribution Function, the authentication process is performed. After the user is authenticated, the decode key is sent to the user terminal. At the same time, the Content Distribution Function uses the request routing function to determine the IP address of the optimal delivery server. This IP address is also sent to the user terminal.

The user terminal requests delivery of the content through this address, and the Content Delivery Function starts delivering the content. When the live content is distributed in real time, the video is directly distributed and delivered to a user from the Content Provider. For time-shift viewing, however, the video content is delivered from the video stored in cache/mirror of the Content Delivery Function.
4. Completion of viewing

When the live distribution finishes, the Content Distribution Function notifies both Content Provider and user about the completion of distribution. If a user stops viewing the live distribution before the ending time, a distribution end request is sent before the notification of completion from the Content Distribution Function.

After confirming the end of distribution, the Content Distribution Function releases the sessions with both Content Provider and user.

After the session is released, the live end screen is displayed and the main menu reappears. A user must then end the service or select another service.

Most likely, it will be necessary to have a function of monitoring whether the video received by a user is of the demanded quality level. This is not covered in this service sequence, however, the function is explained in Section 5-7.
5. Payment (Monthly-based billing proxy)

When a Content Provider asks a Content Distribution/Delivery Provider to handle the user authentication and billing of information charge by proxy, billing will be performed as follows. Note that the billing pattern varies widely depending on a business model. The example below is just one of many billing patterns.

- **User → Content Distribution/Delivery Provider**
  Monthly communication charge and information charge are paid.

- **Content Distribution/Delivery Provider → Content Provider**
  Information charge collected from a user is paid with the billing log.

- **Content Provider → Content Distribution/Delivery provider**
  Distribution service charge and billing proxy charge are paid.
The above figure shows an outline of information exchanged between the functions based on the service sequence mentioned earlier.

1. Purchase of ticket
   Basically, information is exchanged between three entities: Terminal, Service Portal, and Content Distribution Function. The Content Provider is also involved in the content registration sequence.

2. Distribution of content
   Basically, information is exchanged between two entities: Content Distribution Function and Content Provider.

3. Viewing
   The information required for viewing is mainly exchanged between three entities: Terminal, Service Portal, and Content Distribution Function.
   
   Actual content distribution/delivery involves four entities: Terminal, Content Delivery and Content Distribution Functions, and Content Provider.

4. Completion of viewing
   Basically, information is exchanged between three entities: Terminal, Service Portal, and Content Distribution Function. The Content Provider is also partly involved because it must be notified of the completion of viewing.

5. Payment
   Four entities are involved: Terminal, Service Portal, Content Distribution Function, and Content Provider. Note, however, that the flow of payment depends on a business model.
We have identified the interface reference points between functions based on the service sequence. Note that the above figure indicates only the main functions to be implemented by an interface at each corresponding interface point. It does not list all information.

Also note that the content provider’s in-house interface is not within the scope of HSAC studies.

1) Interface between Content Provider and Content Distribution Function
- Information required to request a Content Distribution Provider for a distribution.
- Data to be used for feeding the content video to a Content Distribution Provider.
- Information required for processing the live starting and ending.
- Payment-related information

2) Interface of the Content Distribution Function
There are a number of servers in the Content Distribution Function. Therefore, the interface between these servers must be defined eventually. It is not studied in detail in this report.

3) Interface between Content Distribution Function and Service Portal
- User authentication
- Content information, ID, password
- Information required for processing the start and end of viewing.
- Payment-related information

4) Interface between Content Distribution Function and Content Delivery Function
The Content Delivery Function uses this interface to receive the video data distributed from the Content Distribution Function. Note that information such as the delivery server operating statuses and user viewing log are also exchanged in addition to the video data.

5) Interface between Service Portal and Terminal
- Various information such as menu screens, user authentication result, content information, ID, and password.
- Information required for processing the start and end of viewing.
- Payment-related information

6) Interface between Content Delivery Server and Terminal
A terminal uses this interface to receive the video data delivered from the Content Delivery Function. Note that video playback control and other information are also exchanged besides the video data.
4-2. Bi-directional Video Communication Services

Contents of Section 4-2

4-2-1. What is Bi-directional Video Communication Services
4-2-2. Example of Network Configuration
4-2-3. Example of Players
4-2-4. Example of Service Sequence
4-2-5. Interface Reference Points
4-2-1. What is Bi-directional Video Communication Services

Service Requirements

<table>
<thead>
<tr>
<th>Essential Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Peer-to-Peer Video Conference Connection (Services)</td>
</tr>
<tr>
<td>- Various providers that offer video conferencing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Optional Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Independent from the User Terminal Types</td>
</tr>
<tr>
<td>- Video conference server will set up the conference</td>
</tr>
<tr>
<td>(SIP protocol is assumed to be used)</td>
</tr>
</tbody>
</table>

The following requirements are assumed for bi-directional video communication services.

The following two requirements are crucial for the service.

- **Implementation of video conference connection**
  The basic service specification is same as that of a standard video conference.

- **Various providers that offer video conferencing**
  A system that implements a bi-directional video communication service between terminals only, may be one of the possible forms of videoconferencing. In this section, however, we assume that the Video Conference Service Provider (the main entity that will provide the service), CDN provider (enables sending and receiving of broadband video), and other providers are involved in the video conference service.

Although the following two items are not indispensable, they are assumed as required items.

- **Implementation of communications between different types of terminals**
  Enables communications between different types of terminals by confirming the terminal conditions and performing the required transcoding, even when the user terminals are dissimilar.

- **Invitation from Video Conference Server**
  We assume two methods for joining in a video conference. One is that the participant accesses the server to join in a conference, and the other is that the server invites a participant’s terminal. In this section, we have assumed the latter (server invites participant).
4-2-2. Example of Network Configuration

The above figure shows an example of the system configuration for implementing a peer-to-peer video conference connection.

The main configuration elements are given below.

(1) **Home**

We have considered three types of terminals: dedicated terminal, STB + TV set, and PC. In addition to a terminal, the home system configuration includes the “ONU” for terminating the optical access line and the “HGW” for performing the security function and IP address management of a terminal.

(2) **Delivery Platform**

Accommodates and collects the optical access lines of HIKARI Service subscribers (= users), and mutually delivers the video of users via the “Delivery Server”. In some cases, the transcoder function that adjusts the speed between terminals may be required.

(3) **CDN Backbone Network**

Used for transferring information including video. The “Network Management Server” is used for setting and management of the band and QoS.

(4) **Session Setup Platform**

This platform consists of the “Directory Server” and “Session Management Server”.

We assume the use of a SIP server (SIP proxy server) as session management server. If data such as “name” is used to specify the video conference destination, the directory server is used to convert the name into an address (such as URL) to be used by SIP (given in a sequence later in this document). The DNS and ENUM for converting a URL to the IP address are also included in the directory server.

(5) **Video Conference System**

Consists of the “Video Conference Server” and “MCU”.
4-2-3. Example of Players

- Delivery Platform
- Access Edge #1
- Access Edge #2
- Home #1
- PC
- ONU
- STB
- Home #2
- Network Management Server
- Delivery Server
- Terminal Management Server
- Session Management Provider
- Service Portal
- Billing Server
- Service GW
- CDN Backbone Network
- Video Conference System

If the video rate is MPEG-4 2Mbps maximum, the CDN requirements are not so strict. Therefore, different players may be able to provide the Access Edge and Backbone Network.

We assume that each player can provide authentication, payment, and customer management as integrated functions, instead of these functions being provided by a specific player.

We have assumed that the following players are required to implement the bi-directional video communication service.

1) User
2) Service GW
3) Delivery Provider
4) CDN Provider
5) Service Portal
6) Video Conference Service Provider
7) Session Management Provider
4-2-4. Example of Service Sequence(1/7)

The above figure shows an example of the video conference service sequence between three parties using P2P.

The service sequence consists of the following five phases.
1) Reservation of video conference
2) Starting a video conference
3) Connecting the participants
4) Disconnecting the participants
5) Ending the video conference

1) **Reservation of video conference**

Before starting a video conference, a user makes a reservation for the conference.

A user registers the number of conference participants, time, and image quality level in accordance with the screen information. The display on a screen may be provided either by the menu display function of the service portal or by the Video Conference Function, which directly sends the display information to a terminal. (In the above service sequence example, the service portal displays the “Main Menu Screen” and “ID and Password” confirmation screen, and the Video Conference Function displays the “Video Conference” conditions setting screen.)

We have considered using SIP as the session protocol.
2) Starting a video conference

When a video conference starts, the person who reserved the conference (Member A) specifies the participants to be connected to the video conference (Members B and C), enters the ID and password obtained when the reservation was made, and requests connection. When the connection is requested, the directory server of the Session Management Function converts the IP addresses of the connection destinations. The conference participants are specified by ALSO command of the INVITE signal of SIP protocol.
2) Starting a video conference (continued from previous page)

The Session Management Function processes the connection request. To request a connection, the user authentication (confirmation of ID and password) of the Video Conference Function is necessary. Next, the connection request and user authentication (Member A) are also performed in the same manner for the Delivery Function.

After the above two requests are approved, the connection conditions are confirmed between a terminal and Video Conference Function and video communication between the video conference system and user (Member A) is set up.
3) Connecting the participants (Connecting to Member B)

The other participants (Member B) are connected when a request is sent to the Session Management Function from the Video Conference Function, which holds the information of the participants. The Session Management Function calls up the terminal of a participant and confirms the response. After the response is confirmed, the Session Management Function requests connection to the Delivery Function and user authentication (Member B) is performed.

After the above process, the connection conditions are confirmed between the terminal of Member B and the Video Conference Function. When the video communication between the video conference system and user (Member B) is set, communication between two participants (Member A and Member B) will start.

The connection to the next participant (Member C) is performed in the same manner as the connection to Member B (the connection to Member C is omitted in this service sequence).
4) Disconnecting the participants

The above service sequence shows how a participant leaves the conference while communication is in progress. Basically, the disconnection procedure will be the same as the connection procedure, but in reverse. The disconnection information of Member B’s terminal is sent to the Video Conference Function via the Session Management Function. When the Video Conference Function receives the disconnection information, it will notify the end of user communication (Member B) to the Delivery Function, and requests a release of the connection.

After the connection is released, the charge information of the user (Member B) is displayed on the terminal. This model assumes that the video conference service charge is paid by the initiator (Member A) of the conference and that each participant (Member A, B, and C) pays for their own communication charge. Therefore, only the communication charge is displayed on the terminal of Member B.
5) Ending the video conference when the reservation time passes

The above service sequence shows when the video conference ends after its reservation time expires. Since the Video Conference Function manages the reservation time, the disconnection information of a participant’s terminal does not have to be sent to the Video Conference Function via the Session Management Function, unlike the case mentioned earlier, where a participant leaves the conference earlier than the predetermined time.

After the Video Conference Function sends the communication end request to each terminal, it will request the Delivery Function to release the connection and the connection is released. After the connection is released, the video conference service charge and communication charge are displayed on the terminal of the conference initiator (Member A). Only the communication charge is displayed on the terminal of the other participant (Member C).
4-2-4. Example of Service Sequence (7/7)

6. Settlement (Monthly-based billing proxy)

Billing will be performed as indicated below when a Delivery Provider asks a Video Conference Service Provider to handle the billing proxy of communication charge. Note that the billing pattern varies widely depending on a business model. The following example is just one of the many billing patterns.

- **User → Video Conference Service Provider**
  Monthly communication charge and service charge are paid.

- **Video Conference Service Provider → Delivery Provider**
  Communication charge collected from a user is paid with the billing log.

- **Delivery Provider → Video Conference Service Provider**
  Billing proxy charge is paid.
4-2-5. Interface Reference Points

We have identified the interface reference points between functions based on the service sequence. Note that the above figure indicates only the main functions to be implemented by an interface at each corresponding interface point, and does not list all information.

(1) Interface of Session Management Function

There are a number of servers in the Session Management Function. Although the interface between these servers must be defined eventually, it is not studied in detail in this report.

(2) Interface between Session Management Function and Video Conference Function

This interface is used for the main functions of a video conference and the tasks that involve connection/disconnection of conference members. Also, these functions are often linked to perform a single task (for example, approval and connection of conference members). Therefore, all the interfaces cannot be listed.

(3) Interface of Video Conference Function

Details are not studied in this report as mentioned in (1) above.

(4) Interface between Video Conference Function and Service Portal

• User authentication result, ID, password, and other various information
• Information required for processing the start and end of a video conference.

(5) Interface between Video Conference Function and Delivery Function

This interface is used for sending and receiving of video between conference participants with the Video Conference Function as the core.

• Information required for processing the start and end of a video conference.
• Video information
• Payment-related information

(6) Interface between Service Portal and Terminal

• Menu screen, user authentication result, content information, ID, password, and other various information.
• Information required for processing the start and end of a video conference.

(7) Interface between Delivery Function and Terminal

This interface is used for sending and receiving of video data between the Delivery Function and terminals.

(8) Interface between Video Conference Function and Terminal

This interface is used to establish the connection conditions by confirming the attributes of a terminal.

(9) Interface between Session Management Function and Terminal

This interface establishes and releases the sessions between participants and Video Conference Function at the start and end of the conference.
5. Technical Issues on HIKARI Service Platform

Contents of Chapter 5

5-1. Transcoding
5-2. Cache/Mirror/Request Routing
5-3. Multicasting
5-4. Network Quality and QoS Assurance
5-5. Time-shift Function
5-6. Metadata/EPG
5-7. Viewing Quality Management
5-1. Transcoding

Contents of Section 5-1

5-1-1. Why Transcoding Is Necessary?
5-1-2. Classification of Transcoding
5-1-3. Issues for Transcoding at Terminal (User)
5-1-4. Issues for Content Delivery Provider Transcoding
5-1-5. Issues for Content Provider Transcoding
5-1-6. Issues for Transcoding Related to Bi-directional Video Communications
5-1-7. Conclusion Regarding Transcoding
5-1-1. Why Transcoding Is Necessary?

**Reason:**
Required when the video format created by the content sender and video format requested by a content recipient are different.

**Typical cases:**
- When a terminal can decode a specific format only.
- When the allocated network bandwidth is restricted.

**Reason:**
The main reasons of transcoding are:
- Diversity of terminal capabilities
- Diversity of available bandwidth (including fluctuations)

At present, we cannot focus on a single video format because there are too many types. A user cannot be satisfied if the resolution drops below the terminal capacity.
(The user may accept SD resolution when HD is displayed on an SD TV set.)
5-1-2. Classification of Transcoding

Classification according to transcoding type

1) Transcoding of compression format
   This transcoding type changes the compression system itself.
   Example: Transcoding between MPEG-4 and MPEG-2
           Transcoding between DV and MPEG-2/MPEG-4

2) Transcoding of compression parameters
   This transcoding type uses the same compression format and changes only the parameters.
   Example: Transcoding of frame rate and image size.
           Transcoding between MPEG-2 TS and MPEG-2 PS.

Classification according to players that transcode data

1) Terminal (User)
2) Content delivery provider
3) Content provider

Transcoding performs the following two tasks:
   Changing the compression format
   Changing the compression parameters

Changing the compression parameters is comparatively easy. Changing the compression format, however, is normally a highly complicated task.

Note that this document does not cover possible cases of transcoding between non-compression formats (a comparatively easy task).

Technically, transcoding can be performed at any location from a content provider through the terminal (user).

Since transcoding at a content distribution provider is not much different than transcoding at a content delivery provider, we have described the latter only. (In terms of services, it is more likely that a content delivery provider performs the transcoding.)
5-1-3. Issues for Transcoding at Terminal (User) (1/2)

Reasons:
1) Exceeding the receiving capacity/processing capacity (throughput) of a terminal.
2) Exceeding the display capacity (resolution, etc.) of a terminal.
3) When a target compression format cannot be decoded by a terminal.

Measures and issues (1):
1) Regarding “Exceeding the receiving capacity/processing capacity (throughput)”
   A) The case is not of transcoding but of packet loss processing.
   The packet loss recovery is a differentiating function to be executed in a decoder within each terminal.
   B) There is a system that processes only important data set up by hierarchical coding.
   → Although item B is a fundamental solution, hierarchical coding is not widely used yet.

Technical issues:
The recovery algorithm poses a problem with respect to item A. However, it is not a target of the standardization being carried out by HSAC. Technical matters related to item B have been resolved already.

Transcoding at a terminal (user) basically takes place because there are gaps between the received stream and the stream that can be processed by the terminal.

The gaps include:
1) Throughput of reception/processing
2) Display capacity
3) Decoding capacity

Items 1) and 3) above fall under the same problem category when decoding is performed by a CPU.

The measures for each problem comply with the above, but there are technical issues.

Note, however, that this method is an element that does not affect compatibility even when the coding methods of terminal manufacturers are different. This element can also be instrumental in differentiating terminal manufacturers. We believe that it is not appropriate for standardization by HSAC.
5-1-3. Issues for Transcoding at Terminal (User) (2/2)

Measures and issues (2):

2) Regarding “Exceeding the display capacity of a terminal (resolution, etc.)”
   A) Display by dithering, frame omission, and interpolation are performed
      normally even under the current circumstances.
      → The algorithm itself is a yardstick that distinguishes a manufacturer and is not
      an issue of study by HSAC.

3) Regarding “Compression format that cannot be decoded by a terminal”
   A) Generally, processing and implementation will be easier if a terminal is
      equipped with a decoding function of a specific compression format, to avoid
      the need for transcoding data into a compression format that can be decoded
      by the terminal. There could be an implementation that is valid only when a
      similar compression algorithm is used (for example MPEG-2 to MPEG-4).
      → The solution is to provide a high-speed algorithm that can transcod data
      without reverting it to its source image.
      The expansion of quantization error must be suppressed.
      This factor distinguishes a manufacturer and is not an issue of study by
      HSAC.

Transcoding between compression formats can result in expansion of the quantization error and
picture quality shall unquestionably deteriorate.
The symptoms that a user can detect, such as color differences and out-of-focus, may vary.
A transcoding method that does not affect picture quality is a yardstick that distinguishes a
manufacturer. However, there is no transcoding method that is optimal for all types of video data.
(The optimal transcoding method varies depending on the content.)
5-1-4. Issues for Content Delivery Provider Transcoding

Reason:

1) Capacity exceeding.
   - Exceeding the receiving/processing capacity (throughput) of a terminal.
   - Exceeding the network band capacity

Measures and issues:

Regarding Item 1)

A) Sending data by reducing the frame rate of received data and resolution.
   → There is no Technical problem. It is possible also to implement a dynamic band change.

B) Resending only the important hierarchically coded data.
   → This method is an established system and does not pose any Technical problem.

In either case, the main concern is facility costs and not the tasks. **When the data is encrypted by a content provider, implementation of the above methods is difficult.**

The handling of encrypted data is a problem in transcoding by a content delivery provider. If a content provider authorizes decoding, the only problem will be the facility costs. When the compression format is changed dynamically in accordance with the bandwidth, the load will be higher.
5-1-5. Issues for Content Provider Transcoding

Reason:
1) There are various terminals with a variety of bands/compression formats.

Measures and issues:
1) It is assumed that a content provider cannot manage bands/compression formats of each terminal when content is distributed using 1:N multicast.

→ Sends streams with multiple bit rates/compression formats for each to a content provider. Then terminal selects a suitable stream among them. This method is more economical under the current circumstances than preparing hierarchically coded streams. (It is the dominant method for the stream service of the current Internet conditions.)

At present, this method is practical for handling the diversity of compression formats.

Under the current circumstances, it is most common for a content provider to deliver data in all possible compression formats. (Normally, the content providers provide RealAudio/RealVideo, Windows Media and narrowband/broadband data.)

The current technology can show the menu of the compression formats that the terminal can process. However, automatic selection is difficult because the available bandwidth fluctuate in a practical situation.
For bi-directional video communication:
Transcoding is necessary when multiple compression formats are used for encoding and decoding functions of all participating terminals.

The procedure would be as follows:
1) MCU tries to find a universal compression format that can be handled by all terminals.
2) If no common compression format is found, MCU will transcode the data.

Video format that can be used by bi-directional video communications
MPEG-2 or MPEG-4 is used in accordance with bit rate that can be applied and the terminal capacity. The use of MPEG-4 is common due to band restrictions. MPEG-2 is used when a high bit rate is guaranteed. The terminals to be used are bi-directional video terminals (MPEG-4), STB (MPEG-2), and PC (MPEG-4 and MPEG-2).

It is assumed that the MCU will provide support for different types of terminals and a transcoding service for some narrow-band users. These MCU services, for example, would include transcoding between the MPEG family and H.263, and transcoding between MPEG-2 and MPEG-4.
5-1-7. Conclusion Regarding Transcoding

**Live/VoD**

- It is practical to have a content provider provide multiple compression formats.
- Hierarchical coding is a better method if it becomes widely accepted.
- The portion depending on terminal capacity is a yardstick that distinguishes a terminal manufacturer, and thus has low priority in HSAC’s standardization process.

**Bi-directional video communication**

- Encoding and decoding formats must both be considered.
- It is practical to have the MCU determine and transcode a compression format.
5-2. Cache/Mirror/Request Routing

Contents of Section 5-2

- 5-2-1. Cache/Mirror/Request Routing and CDN
- 5-2-2. Role of CDN in the Hikari-Service Platform
- 5-2-3. Function of CDN
- 5-2-4. Interface among Players of CDN
- 5-2-5. Issues for CDN (Organization of the Network)
- 5-2-6. Issues for CDN (Request Routing)
- 5-2-7. Issues for CDN (Controlling Mirror/Cache)
- 5-2-8. Issues for CDN (Live Casting and Mirror/Cache)
5-2-1. Cache/Mirror/Request Routing and CDN

Definition of CDN:

- Cache/mirror (Surrogate) which duplicates and delivers content
  + An efficient content distribution system
  + A request routing system which guides the request to the surrogate (incl. A mechanism to avoid bottlenecks)
  + An accounting system (Logs and reports access logs)

Cache/Mirror/Request Routing are important components which comprise a CDN (Content Distribution Network)

That is to say, a CDN is composed of the four components shown in the figure.

1) The cache/mirror (generally known as ‘Surrogate’) is installed near a terminal to store content frequently accessed on that terminal.
2) A distribution system efficiently delivers video content from a Video Content Distribution Service Provider to a surrogate.
3) A request routing system finds a surrogate that satisfies the optimal conditions for a user (surrogate closest to user, and satisfies the required quality).
4) An accounting system compiles access logs and collects billing information.
5-2-2. Role of CDN in the Hikari-Service Platform

This diagram shows the role of CDN within the Hikari-Service Platform. CDN corresponds to the part surrounded by dashed line and consists of the following players.

1) Surrogates (Cache/Mirrors) who retain copy of the original content. (So called Video Content Delivery Service Provider)

2) Request Routing System which leads the user to the best suited (conditions such as nearest/matches the quality etc.) surrogate.

3) Video Content Distribution Service Provider who distributes the content to the surrogates.

4) A Bypass network to let the Video Content Distribution Service Provider efficiently distribute the content.

The network function might use the Hikari-service platform or it could also use other networks such as satellite networks.
5-2-3. Function of CDN

- Hiding Network Delay due to the following factors:
  - Unmanaged best-effort network/physical distance/complex network topology etc.
  - Architecture based on economical/historical factors rather than performance
  - Network bottlenecks such as IXPs (Internet eXchange Point)
  - Packet Loss

- Easing Server Bottleneck
  - Becomes more important role in the Hikari-Service Platform
  - Serve for Live Feeds

There are two main problems solved by the CDN in the Hikari-Service Platform namely;

1) Hiding Network Delay (bottleneck)
   This is due to the fact that THE internet is a best effort network in which QoS cannot be guaranteed. The network bottleneck will certainly become less of a problem when a world-wide optical core networks becomes available. However, guaranteeing QoS in a managed network is certainly very important.

2) Easing Server Bottleneck
   When the network bottleneck becomes less due thanks to Hikari-service Platform, the server bottleneck of the original content holder will become much more prominent. Especially, for live broadcasts of very popular events.
We summarize the Player Interface among the CDN players. This is just a rough summary and a more detailed study is necessary.

However, as can be seen a great number of API’s have to be specified among the players.

Terminal ↔ Distribution Provider (Request Routing System)
Delivery Provider (Surrogate) ↔ Distribution Provider (Request Routing System)
Distribution Provider (Distribution Server) ↔ Distribution Network
Distribution Provider (Distribution Server) ↔ Delivery Provider (Surrogate)
Content Provider ↔ Distribution Provider (Distribution Server)
5-2-5. Issues for CDN (Organization of the Network)

Issues No.1: Network Organization

- How to build the distribution network:
  - Create Overlay network on top of conventional Network
  - Use the optical backbone provided by the Hikari-service platform
  - Satellite networks

- Points to solve in building the distribution network
  - How to support live broadcast.
    - Define Surrogate to include more functions suited for streaming broadcast. (Role of CDN becomes more important.)
  - How to resolve policies at distribution time.
    - Guaranteeing QoS=Decrease packet loss:
      » Trade off between necessary and available bandwidth.
      » How to provide for high access bandwidth (e.g. 10Mbps)

- Supporting C2C.
  - There is a need.
  - A business model to provide for C2C has to be made: e.g. An ASP which provides Internet Video Telephone uses CDN.

The issues for CDN involve the entire network, request routing, control of cache/mirror, and cache during live broadcasting.

The first issue for CDN concerns the entire network, essentially the construction of an efficient CDN.

1) Choices among several methods for constructing a distribution system network

The following methods may be used to construct a backbone for a further improved video rich environment on the HIKARI Service Platform.

- Application level multicasting by overlay network.
- Broadband network (high-speed network) support.
- Satellite network distribution to stations.

2) Issues regarding construction of distribution system network

The following points must be considered for constructing a distribution system network regardless of the above choices.

a) Live broadcast system support: Unlike an on-demand broadcasting system, a live broadcasting system involves a number of issues related to QoS. One way to solve this problem is by installing a surrogate as a relay server that performs a live feed efficiently, instead of a simple cache/mirror/splitter. When a surrogate is ranked as relay server, the role of CDN becomes more important.

b) The policy of band allocation during video content distribution is important in order to guarantee QoS (i.e. reduction of packet loss). Especially, when one considers video rich content of 10Mbps to 100Mbps, it is necessary to resolve the tradeoffs with the bandwidth that is available.

3) C2C support

So far, we have been discussing CDN in terms of B2B2C. It is also necessary to consider whether CDN should be capable of supporting C2C and whether there is a demand for such support. In reality, it is desirable to use CDN in some form when video rich content is delivered by C2C. A new business model is needed for this purpose. For example, the Internet Video Telephone Service provider ASP uses CDN for video distribution.
5-2-6. Issues for CDN (Request Routing)

Issue No.2: Request Routing

How to determine the Best Surrogate

• Criteria for the Optimal Surrogate
  – Closest to User = Based on Minimum RTT or Minimum Hop Count
  – Minimum Packet Loss
  – Predetermined mapping between surrogate and User.
  – Choose based on Content Type
  – Surrogates’ Load Status

The second issue for CDN is the issue for the Request Routing System that is “How to determine the optimal Surrogate?”. Ways to realize this issue is based upon each vendor’s know-how. So, we would not go into the details. Rather, we discuss the criteria condition for the optimal surrogate. The conditions sited below can be used alone, or can be combined.

1) The surrogate nearest to the user: Here, ‘nearest’ can be defined as having the minimum RTT (Round Trip Time), the minimum hop count, geographical conditions etc. It should be noted that the surrogate nearest to the Request Routing System is not necessarily the surrogate nearest to the user. So, in order to find the nearest surrogate, various information such as the surrogate last selected by the user is necessary. On the other hand, if the surrogate is distributed or localized, then choosing the surrogate nearest to the Request Routing System would also suffice.

2) Something similar to 1) but which also reflects the condition of the network is to choose the surrogate with the minimum packet loss. As in 1) statistics to gather the packet loss between the surrogate and the user is necessary.

3) One other way to choose is to statically define the surrogate based on user’s IP address and given information about the user.

4) Another way is to determine the surrogate (or the server) based on the requested content type. This will reflect the Delivery/Distribution Provider’s policy.

5) It is also important to consider the load status of the surrogate itself. Statistics to determine the surrogate’s load are CPU Load, number of connections etc.
5-2-7. Issues for CDN (Controlling Mirror/Cache) (1/4)

Issue No.3: Controlling Mirror/Cache

Controlling Mirror/Cache to reflect users’ preferences in order to attain higher hit ratio

Characteristics of Cache and Mirror

- **Cache:**
  - Basically Pull based (Store what the user needs)
  - Difficult to reflect CDN Provider’s Policy
- **Mirror:**
  - Basically Push based. (Store what the original provider wants to distribute)
  - Difficult to reflect user’s preferences

The third issue for CDN is how to control cache and mirror in order to attain higher hit rate. (This is because high hit ratio is crucial for user’s satisfaction and is a very important factor in the Hikari-Service Platform.)

The main technology for CDN is a request routing that provides efficient content distribution by connecting an access request from a user to the nearest mirror/cache.

The content is retained in a cache/mirror normally by pull based storing for cache and push based storing for mirror. We believe that video data can be distributed more efficiently if the user’s viewing frequency is taken into consideration.

In this section, we discuss the considerations required for the user’s viewing frequency.

The advantages and disadvantages of a cache and mirror depend on the method used to achieve this goal.

1) Basically, the pull based data as required by a user is retained in a cache. Therefore, it is comparatively easy to reflect the users’ viewing rate. However, it is difficult to determine the policies of content providers and therefore controlling the content distribution is not an easy task. (For example, when new content is delivered based on a projection of users’ needs).

2) The content stored in a mirror, on the other hand, are push based. Therefore, it is easy to determine the intentions of content providers regarding content distribution. The disadvantage, however, is the difficulty of gauging the users’ viewing rate using the original mechanism alone.
Issues Related to both Mirror and Cache: where to provision them:

- **Close to the user**
  - Quick Response Time (Advantage)
  - Cache Capacity has to be relatively small (Disadvantage)
  - Difficult to reflect many users’ preferences (Disadvantage)
  - Possible to meet each user’s preferences (Advantage)

- **A little distant from the user (and supporting many users)**
  - Response time gets longer (Disadvantage)
  - Cache capacity can be larger (Advantage)
  - Possible to reflect the preference statistics of many users (Advantage)

Another Issue that is important for Caches and Mirrors is to where to provision them, either close to the user, or a little further from the user and supporting many users. Merits and demerits of the two strategies are summarized on the slide.

1) The advantages and disadvantages of a cache installed near a user are summarized below.
   - Fast response due to the close proximity to a user (advantage).
   - Relatively small cache capacity (disadvantage).
   - Difficult to reflect many users’ viewing rate (disadvantage).
   - Potential of satisfying individual user preferences (advantage).

2) The advantages and disadvantages of a cache installed at a little distance from a user (located to support many users) are summarized below.
   - Although the response speed will be slower than case 1) above, it is possible to reflect the viewing rate of a greater number of users to a certain extent.
5-2-7. Issues for CDN (Controlling Mirror/Cache) (3/4)

An approach taken in the Caches

- Combination of Large scaled cache and small caches
  - Provider’s policy is reflected in how the data is managed in the large scaled cache (how much space to allocate for a specific content etc)

Considering the characteristic described in the previous slide, one realistic way for caches is to combine large scaled cache and small caches as shown in the figure. In fact there is an IETF draft which describes how to manage caches using the configuration. In this configuration, it also becomes possible to reflect the providers’ policy via the data management strategy for the large scaled cache.
5-2-7. Issues for CDN (Controlling Mirror/Cache) (4/4)

Methods to reflect the users’ preferences in mirror/cache:

- **Selective Pull: (Mirror)**
  - Make it possible to fetch the content from the origin server when the data is not present as in a cache

- **Gather statistics about the miss hits. (Mirror)**
  - Gather statistics and periodically send them to the origin.
  - Content Provider decides which content to mirror based on the statistics

- **Personalization (Cache)**
  - Prefetch based on user’s preferences
  - Easy to realize if the cache is near to the user
  - Also possible to reflect a region’s demographic preferences (Baseball match of a local team etc)

To overcome the disadvantages of mirrors some methods such as selective pull and making use of miss hit statistics exist.

(We have surmised that the CDN policy for cache-based content distribution methods that have disadvantages can be handled by data management with a large-scale cache (for example, the amount of space allocated in a cache for a specific content) as shown in the previous slide.)

Also, Personalization in caches can be realized using the methods shown on the slide.
5-2-8. Issues for CDN (Live Casting and Mirror/Cache) (1/2)

Issue No.4: Live casting and Mirror/Cache
   How to make the most use of cache/mirror for live casting

Current Status: (Supplements IP Multicasting)
   • Request routing:
     – Necessary to choose the optimal splitter
   • Caches
     – Some caches support splitting/multicasting via so called multi-media
       extensions.
   • Services
     – CDN providers are beginning to announce support for live casting

Usages
   • Caching of pre-recorded content
     – Material to be used in educational programs
     – Commercial Messages
   • Supporting Trick Play
     – Use Cache’s recording mechanism

The 4th issue for CDN is how to efficiently support live casts using cache/mirror/request routing.

A surrogate is useless to a distribution network that is configured for IP multicast only. When both
multicast and unicast broadcasting coexist (or live distribution is performed by unicast only), the role
of CDN becomes important in the following senses.
1) When the definition of a surrogate is extended to a splitter, request routing will be necessary to
select an optimal splitter.
2) There are cache server products that provide splitting function and multicasting support by a
media extension name. Therefore, these caches may be considered for the same framework.
3) A leading CDN company has already announced that it will provide a live broadcasting service.

Besides the basic splitting function, practical methods for combining cache/mirror functions and
splitter functions are listed below.
1) Caches and uses content such as education/training materials and commercials that can be
distributed in advance.
2) Provides trick play support such as rewinding by the cache server recording function.
### 5-2-8. Issues for CDN (Live Casting and Mirror/Cache) (2/2)

Issues related to realizing Trick Play

- Issues related to recording the live on the cache while giving a live feed:
  - Copyright issue has to be settled
  - Various Information has to be sent from the content provider to the surrogate such as
    - "Till what time can the program be sent from the surrogate"
    - "How many minutes can the past image be retained and cast"
  - Method to provide accounting log for the late comers to be sent to the provider

- Issues related to the fact that a new stream is created for each user who requests a trick play

Supporting Trick play mentioned in the last slide however is not easy to realize. There are many issues which have to be solved as shown in this slide.

1) The issues to be considered for recording on a cache during a live video content feed are as follows.

   1.a) The copyright issue must be resolved since recording is permitted.
   1.b) There must be a way to provide instructions on metadata, such as on-demand enable/disable information and conditions, from a source content distributor to a surrogate.
   1.c) There must be a function that sends a log of the users viewing the content after a live feed is started to a source content distributor for billing. (This function is provided in a standard cache.)

2) A further complication for a cache server is the necessity of creating a new stream for each user who requests a trick play (special playback).
   A practical solution must be found, for example, by restricting the types of trick play.
5-3. Multicasting

Contents of Section 5-3

5-3-1. Necessity of Multicasting
5-3-2. Applicable Field for Multicasting Delivery/Distribution
5-3-3. Interface Between Content Provider and Delivery/Distribution Provider
5-3-4. Interface Within Delivery/Distribution Provider
5-3-5. Interface Between User and Delivery/Distribution Provider
5-3-6. Conclusion for Multicasting Interface
5-3-7. Multicasting and IP Multicasting Technology
5-3-8. Protocols List Related to IP Multicasting
5-3-9. Overview of Protocols Related to IP Multicasting
5-3-10. Approaching for Selection of Protocol Related to IP Multicasting
5-3-11. Whole Map Applied IP Multicasting Protocol to HIKARI Service Platform
5-3-12. Discussion of Multicasting Security
5-3-13. Conclusion of Multicasting
5-3-1. Necessity of Multicasting

- Technology to remove network bottlenecks in distributing same data to a lot of users
- Multicasting is the best technology for rich content such as live casting

Necessity of multicasting

Normally, unicasting on the Internet is conducted as one-to-one communication. The figure shows an example of MPEG-2 (6 Mbps) live unicast broadcasting that distributes streaming data to 1000 recipients (terminals). In other words, 1000 packet data containing the same video data is distributed. In this example, a 6 Gbps band (1000 users multiplied by 6 Mbps) will be occupied at a server side.

In the case of multicast broadcasting, all recipients (terminals) can view a data packet sent from a single live video source. If a node determines that there are recipients in its downstream and then transfers a packet data to these users, a 6 Mbps band will be required at the server side even when all 1000 recipients (terminals) are sent the 6 Mbps stream data.

Therefore, the multicast broadcasting is the optimal video-rich live broadcasting service.
5-3-2. Applicable Field for Multicasting Delivery/Distribution

Applicable field for multicasting

[Position to provide multicast distribution function]
1) Multicast transmission is automatically executed when a multicasting system is supported by the network of delivery/distribution providers (platform) and a content provider sends content data to distribution/delivery providers.
2) Under the conditions indicated above, the content provider controls the multicast “indirectly” or the multicast information is delivered and received between distribution/delivery providers.

[Multicast distribution applicable content]
- (Live video content) : Application of multicast distribution is effective.
- (VoD archives) : Unicast distribution/delivery is compulsory.
- (Broadcasting content) : Application of multicast delivery/distribution is effective.

[Precondition for division of functions]
- (Content provider) : Provider of original content
  - Edits materials sent from a content provider
  - MPEG and Real encoding may or may not be provided.
- (Delivery/distribution providers) : Provider of streaming (distribution, relay) servers.
5-3-3. Interface Between Content Provider and Delivery/Distribution Provider (1/5)

The following four interfaces are extracted, as a result of discussing the necessary interface between content provider and delivery/distribution provider for realizing multicasting.

<table>
<thead>
<tr>
<th>#</th>
<th>Function name/Interface name</th>
<th>Study for issues of interface between “Content Provider” and “Delivery/Distribution Provider”</th>
</tr>
</thead>
</table>
| (1) | Suppress live casting content delivery/distribution | • In case that a content provider has a lot of channels for live casting, there are any possibility that not forwarding all streaming of channels but a part of streaming to delivery/distribution provider.  
• For example, we may suppress delivery/distribution streaming to channel that there is no active user.  
• It needs making interface regulation to inform user's join/leave from delivery/distribution provider to content provider. |
| (2) | Access log (viewing/listening information) collection of broadcasting channel | • It is necessary to forward join/leave information (that viewing start/end to live casting of the multicasting group) for multicasting group.  
• It needs making interface regulation to inform the above-mentioned information from delivery/distribution provider to content provider. |
| (3) | Illegal receiving guard for no contracted content | • Content delivery/distribution permission is given by authentication of the user login. The phase is realized with a protocol between the user and the content provider.  
• It is needed for delivery/distribution provider to suppress illegal join because user which not given permission by the above-mentioned authentication phase can also receives live casting.  
• It needs making interface regulation to inform joinable user information from delivery/distribution provider to content provider. |
| (4) | Additional operation of a broadcasting channel | • When a content provider makes the live casting channel increase, it is necessary to notify delivery/distribution provider of the addition of a multicasting group.  
• It needs making interface regulation for not enough of office notification (off-line work). |

Interface between content provider and delivery/distribution providers

We have extracted the following four interfaces from our study of the interface required between a content provider and delivery/distribution providers for multicasting.

1. Suppressing live broadcasting content delivery/distribution.
2. Collecting broadcasting channel access logs (viewing information).
3. Protection against illegal viewing of content by a user who does not have a contract.
4. Operation when a new broadcasting channel is added.
5-3-3. Interface Between Content Provider and Delivery/Distribution Provider (2/5)

Suppress live casting content delivery/distribution

1) When a content provider offers several hundreds of live distribution channels and must have a basic control for the continuous flow of all channel streams to delivery/distribution providers, the load is spread over the entire CDN.

2) The effective method is to suppress the delivery/distribution load streams of channels that do not have a viewer.

3) The interface specification is required for sending notifications from the delivery/distribution provider to a content provider when a user (viewer of live broadcasting channel associated with a multicast group) logs in or out to/from a multicast group.
5-3-3. Interface Between Content Provider and Delivery/Distribution Provider (3/5)

Access log (viewing/listening information) collection of broadcasting channel

Collecting broadcasting channel access logs (viewing information)

1) The information on a join or leave to/from a multicast group (start and end of viewing a live broadcasting channel associated with this multicast group) must be handed over between providers.

2) An interface specification is required for notification of the above information from the delivery/distribution provider to a content provider.
Protection against illegal viewing of content by a user who does not have a contract

1) The content delivery/distribution is permitted upon authentication when a user performs the portal access. This process is implemented using the interface between user and content provider.

2) A guard system is required to block a user who was not authorized in the above authentication process from performing an illegal join to a multicast group associated with a live broadcasting channel.

3) An interface specification is required for notification of information from a content provider to the delivery/distribution provider regarding users who can log in to a multicast group.
5-3-3. Interface Between Content Provider and Delivery/Distribution Provider (5/5)

Additional operation of a broadcasting channel

1) An interface specification is required for notification of information from a content provider to the delivery/distribution provider regarding the addition of a new multicast group that supports a new live broadcasting channel.

2) This issue can be handled by offline contract.
5-3-4. Interface Within Delivery/Distribution Provider (1/6)

- The following three interfaces are extracted, as a result of discussing the necessary interface within delivery/distribution provider for realizing multicasting.

<table>
<thead>
<tr>
<th>Function name/Interface name</th>
<th>Study for issues of interface within content delivery/distribution provider (network provider)</th>
</tr>
</thead>
</table>
| (1) How to prevent the performance degradation resulting from scale expanding of content delivery/distribution | • Concerning of the performance degradation resulting from increasing users and spreading service area.  
  • Discussing of disposing where multicasting function and making it cooperate.  
  • Necessity for a mechanism to avoid getting over the capability of content delivery/distribution provider. |
| (2) Multicasting address management | • Necessity for address management mechanism for distinguishing content in content delivery/distribution provider.  
  • Address management method has 2 ways. One is a static allocation(content provider allocates in advance), the other one is a dynamic allocation(on registering content).  
  • Opening the address to a content provider or not open (In case of not open, content delivery/distribution provider must have a table between content and address). |
| (3) Content delivery/distribution model | As a result of discussing multicasting processing model within content delivery/distribution platform(delivery/distribution provider + network provider), there are the following three suggestion.  
  • Depending on processing capability of center server  
  • Depending on processing capability of edge/relay server  
  • Depending on processing capability of network |

Interface within delivery/distribution provider organization

The following three interfaces have been extracted from our studies on interfaces required for multicasting within the delivery/distribution provider organization.

1. How to prevent the performance degradation resulting from scale expanding of content delivery/distribution ⇒ Optimal multicasting routing
2. Multicasting address management
3. Content delivery/distribution model
5-3-4. Interface Within Delivery/Distribution Provider (2/6)

Optimal multicasting routing

- The performance issue becomes more critical when the number of users or service area is increased. It is necessary to study the location for setting up the multicast function and provide a suitable interface.

- A system that does not exceed the capacity of a delivery/distribution provider is required. A study on the measures to be taken when the capacity is exceeded is also necessary.
5-3-4. Interface Within Delivery/Distribution Provider (3/6)

• An address management system is required for a delivery/distribution provider to identify the content of each content provider.

• Address management includes static allocation (allocation of address by a content provider in advance) and dynamic allocation (allocation of address when a content is registered).

• Two methods of address management are one that discloses the addresses to a content provider and another that does not. (When addresses are not disclosed to a content provider, the association between content and address must be maintained within a delivery/distribution provider organization.)
5-3-4. Interface Within Delivery/Distribution Provider (4/6)

Content delivery/distribution model
(Multicasting control depending on processing capability of center server)

Multicasting by center server

High-function and value-added feature service oriented

- The multicast function is provided by a distribution provider.
- Advantageous model for providing high-function services such as editing of a content being delivered and distributed.
- Model for detailed subscriber management support.
- For example, this model has the upper hand in implementing a service, such as inserting an advertisement according to the preference of each user, for a multicast group that is distributing the same live video program.
- There are no restrictions on each provider of delivery, distribution, and network.
5-3-4. Interface Within Delivery/Distribution Provider (5/6)

Content delivery/distribution model
(Multicasting control depending on processing capability of edge/relay server)

Multicasting according to edge/relay server capacity

- Broadband and large-capacity service oriented

- The multicast function is provided by both delivery provider and distribution provider.

- Advantageous model for expanding the scale of CDN.
  This model allows the content delivery and distribution area to be expanded by linking up providers when a large-scale CDN for a country or area is constructed.

- There is an advantage that each local area provider can control the content delivery/distribution independently when the delivery conditions are different for a country or area (for example, the timing for inserting an advertisement is different, depending on area).
5-3-4. Interface Within Delivery/Distribution Provider (6/6)

Content delivery/distribution model
(Multicasting control depending on processing capability of network)

Multicasting according to network capacity

High-speed, high-precision service oriented

- The multicast function is provided by a network provider.
- This model puts importance on the content delivery and distribution capacity. This model is advantageous for high-precision live video distribution that requires broadband/large-capacity data transfer.
- The IP multicast system of the network layer must be interfaced for switching live video content (channel switching). The prospect of implementing this model is high in an integrated management CDN where a content provider also serves as the network provider.
5-3-5. Interface Between User and Delivery/Distribution Provider (1/3)

- The following two interfaces are extracted, as a result of discussing the necessary interface between user and delivery/distribution provider for realizing multicasting.

<table>
<thead>
<tr>
<th>Case</th>
<th>Function name/Interface name</th>
<th>Study for issues of interface between “user” and “delivery/distribution provider”</th>
</tr>
</thead>
</table>
| (1)  | View start (Join)/End (Leave) | • In case that content provider has a lot of channels, it is desirable that not forwarding streaming data of all channel to the user but suppressing to forward streaming data except for content to the user.  
• It needs making interface regulation to inform user’s join/leave to delivery/distribution provider. |
| (2)  | Illegal receiving guard for no contracted content | • Content delivery/distribution permission is given by authentication of the user login. The phase is realized with a protocol between the user and the content provider.  
• It is needed interface to suppress illegal join because user which not given permission by the above-mentioned authentication phase also can receives live casting. |

---

Interface between user and delivery/distribution provider

The following two interfaces have been extracted from our study of the interface required between user and delivery/distribution provider for multicasting.

(1) Viewing start (join) and end (leave)

(2) Protection against illegal viewing of content by a user who does not have a contract.
Viewing start (join) and end (leave)

- When a content provider offers several hundreds of live broadcasting channels, it is desirable to suppress the delivery/distribution of content other than those being viewed by users instead of sending stream data of all channels to users at all times.

- The interface is required for notifying a delivery/distribution provider when a user (viewer of live broadcasting channel associated with this multicast group) who has joined a multicast group logs in or out.
5-3-5. Interface Between User and Delivery/Distribution Provider (3/3)

Illegal receiving guard for no contracted content

Protection against illegal viewing of content by users who do not have a contract

1) Content delivery/distribution is permitted by authentication when a user performs the portal access. This process is implemented using the interface between user and content provider.

2) A guard method is required to prevent a user who was not authorized in the above authentication process from performing an illegal join to a multicast group associated with the live broadcasting channel.

3) An interface specification is required for notification of information from a content provider to the delivery/distribution provider regarding users who can log in to a multicast group.
Conclusion on multicasting interface

Interface between content provider and delivery/distribution provider
(1) Starting or suppressing stream distribution to broadcasting channel.
(2) Report on corresponding statuses of broadcasting channel login (viewing) to collect broadcasting channel access logs (viewing information).
(3) Notification of information of users who are denied access to a multicasting group because they do not have contract, for protection against illegal viewing of content.
(4) Exchange of information on a new multicasting group associated with a newly added broadcasting channel for efficient operation.

Interface within delivery/distribution provider organization
(1) Generation of optimal information on multicasting routing
(2) Multicasting address management
(3) Content delivery/distribution model

Interface between user and delivery/distribution provider
(1) Viewing start (join) and end (leave)
(2) Protection against illegal viewing of content by a user who does not have a contract.
5-3-7. Multicasting and IP Multicasting Technology

• Technologies to actualize multicasting
  Multicasting is a mechanism to simultaneously delivery a content to a lot of users (terminals), there are the following means.
  – Use the satellite communications
  – Use the IP multicasting (IETF RFC 1112, 2236, 2710, 2117 etc.)
  – Use the splitter
  – Use the Xcast (Explicit Multicast)
    (IETF Draft xdraft-ooms-xcast-basic-spec-02.txt)
  – Others

Below, we had selectively discussed IP multicasting technology that is expectable to expand and to be scalable within the above mentioned technologies in the future.

Study on multicasting technology issues

The multicasting is a system for simultaneous distribution of a single content source to multiple users (terminals). The following distribution methods may be used.
  - Multicasting distribution by communication satellite
  - Distribution by IP multicasting
  - Multicasting distribution using splitter
  - Multicasting distribution by Xcast (Explicit Multicast)

In our study here of the above methods, we have focused on IP multicasting because it is scalable and has scope for future development.
5-3-8. Protocols List Related to IP Multicasting

The protocols of IP multicasting have been studied and proposed mainly by IETF. Group management protocol, address allocation, routing protocol, data transfer protocol, Ethernet SW solution, and high-reliability file transfer are being studied and proposed (RFC/Draft). In this section, we mainly discuss group management protocol, address allocation, routing protocol, and data transfer protocol.
### 5-3-9. Overview of Protocols Related to IP Multicasting (1/5)

- **Group management protocol**
  Protocols to dynamically manage join/leave to multicasting group between host and router.
  - IGMP v1 (RFC1112)
  - IGMP v2 (RFC2236) Selection of Querier. Existence confirmation of the specified group member.
  - IGMP v3 (Draft) Packet filtering from the specified host.
  - MLD (RFC2710) IGMPv2 to IPv6
  - MLD v2 (Draft) IGMPv3 to IPv6

<table>
<thead>
<tr>
<th>Protocol stack</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP</td>
</tr>
</tbody>
</table>

**Overview of protocols related to IP multicasting (1)**

Group management protocol

A group management protocol is used between a user terminal and router for dynamic control of join and leave to/from a multicasting.

The main group management protocols are IGMP (ipv4) and MLD (ipv6).
5-3-9. Overview of Protocols Related to IP Multicasting (2/5)

- Data transfer protocol
  Protocols to transfer movie/audio/image made by content provider
  - RTSP/RTP (RFC2326/RFC1889, 1890) Standard(*)
  - PNA/PNA (Original/Original) adopted by RealNetworks (Old)
  - RTSP/RDT (RFC2326/Draft) adopted by RealNetworks (New)
  - MMS/MMS (Original/Original) adopted by Microsoft

* There is a case to capsule RTSP/RDT with HTTP to forward through the FireWall

<table>
<thead>
<tr>
<th>Protocol stack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
</tr>
<tr>
<td>RTSP</td>
</tr>
<tr>
<td>TCP</td>
</tr>
<tr>
<td>IP</td>
</tr>
</tbody>
</table>

Overview of protocols related to IP multicasting (2)

Data transfer protocol
A data transfer protocol is used for transferring video, audio, and image data created by a content provider. Several protocols have been proposed.
The standardized protocol is RTSP/RTP.
The original protocols are PNA/PNA, RTSP/RDT, and MMS/MMS.
### 5-3-9. Overview of Protocols Related to IP Multicasting (3/5)

#### Routing protocols

<table>
<thead>
<tr>
<th>Type of routing protocol</th>
<th>Distribution tree structure</th>
<th>Path select algorithm</th>
<th>Member join procedure</th>
<th>Suited number of member</th>
<th>Other feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>DVMRP (RFC1075)</td>
<td>Distance vector</td>
<td>Source Routed SPT</td>
<td>RPM + RPF</td>
<td>Many members</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Original unicast routing protocol</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Recognition of suitable interface by Poison Reverse.</td>
</tr>
<tr>
<td>MOSPF (RFC1584)</td>
<td>Link state</td>
<td>Source Routed SPT</td>
<td>RPM + RPF</td>
<td>Seldom influences</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not dependent on a unicast routing protocol</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Remove path with Prune message after Flooding</td>
</tr>
<tr>
<td>PIM-DM (I-D)</td>
<td>(Any type)</td>
<td>Source Routed SPT</td>
<td>RPM + RPF</td>
<td>High density of members</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not dependent on a unicast routing protocol</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Remove path with Prune message after Flooding</td>
</tr>
<tr>
<td>PIM-SM (RFC2117)</td>
<td>(Any type)</td>
<td>Source Routed SPT</td>
<td>RPM + RPF</td>
<td>Low density of members</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not dependent on a unicast routing protocol</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Shift to SPT by communication rate surveillance.</td>
</tr>
<tr>
<td>PIM-SSM (I-D)</td>
<td>(Any type)</td>
<td>Source Routed SPT</td>
<td>RPM + RPF</td>
<td>Low density of members</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not dependent on a unicast routing protocol</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Suitable for 1: N broadcasting</td>
</tr>
<tr>
<td>CBT (RFC2189)</td>
<td>(Any type)</td>
<td>Source Routed SPT</td>
<td>RPM + RPF</td>
<td>Low density of members</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Not dependent on a unicast routing protocol</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Allocating tree core</td>
</tr>
</tbody>
</table>

---

### Overview of protocols related to IP multicasting (3)

**Routing protocol**

A routing protocol is used for control and optimal delivery of a multicasting address packet to viewers who log in.

Several methods are available.
5-3-9. Overview of Protocols Related to IP Multicasting (4/5)

Overview of protocols related to IP multicasting (4)

<table>
<thead>
<tr>
<th>Protocol Stack</th>
<th>Each protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DVMRP</td>
<td>Distance Vector Multicast Routing Protocol</td>
</tr>
<tr>
<td>IDMR</td>
<td>Inter-Domain Multicast Routing</td>
</tr>
<tr>
<td>MOSPF</td>
<td>Multicast Open Shortest Path First</td>
</tr>
<tr>
<td>PIM-DM</td>
<td>Protocol Independent Multicast Dense Mode</td>
</tr>
<tr>
<td>PIM-SM</td>
<td>Protocol Independent Multicast Sparse Mode</td>
</tr>
<tr>
<td>CBT</td>
<td>Core Based Tree</td>
</tr>
<tr>
<td>PIM-SSM</td>
<td>Protocol Independent Multicast Source Specific Mode</td>
</tr>
<tr>
<td>SPT</td>
<td>Shortest Path Tree</td>
</tr>
<tr>
<td>RPM</td>
<td>Reverse Patch Multicasting</td>
</tr>
<tr>
<td>RPF</td>
<td>Reverse Path Forwarding</td>
</tr>
</tbody>
</table>

Copyright © 2001-2002 HIKARI Service Architecture Consortium. All rights reserved.
Address allocation protocols

Multicasting address static allocation
- SAP (Session Announcement Protocol) RFC2974
  - It has been determined address to announce a group address used as a multicasting address every scope to stand for arrival area of packet. In case of taking the address, use of a unused one is announced after grasping whether which one is used by receiving one for a while. Announce information is notified in the SDP (Session Description Protocol)(RFC2327).
- GLOP addressing RFC2770
  - Each AS(Autonomous System) is allocated a fixed address that generally called Class A from IANA. This is temporary counter.

Multicasting address dynamic allocation (World Wide)
[Under draft by IETF MALLOC WG. Dynamically allocating a global multicasting address by using the following three protocols corresponded to the three classes.]
- MASC (Multicast Address-Set Claim) RFC2909
  - Protocol to class of among domain (AS)
- AAP (Address Allocation Protocol) IETF draft
  - Protocol to class of within domain (AS)
- MADCAP (Multicast Address Dynamic Client Allocation Protocol) RFC2730
  - Protocol to allocate multicasting address to clients.

Overview of protocols related to IP multicasting (5)

Address allocation protocols
Multicasting address allocation is of two types: static and dynamic.
Addressing protocols such as SAP and GLOP use static allocation. Addressing protocols such as MASC, AAP, and MADCAP use dynamic allocation.
5-3-10. Approach for Selection of Protocol Related to IP Multicasting

- **Group management protocol**
  - IGMP (IPv4), MLD (IPv6)
  - IGMPv2 (RFC2236) is currently major. IGMPv3 is better to efficiently distribute broadcasting multicasting. (set with PIM-SSM)

- **Data transfer protocol**
  - Determination is difficult, as there are reasons why there are various mounting.
  - Recommended protocol is RTSP/RTP.

- **Routing protocol**
  - Need to determine an optimal protocol by the type of multicasting.
  - PIM-SSM for 1:N (video, live, TV program etc.)
  - PIM-DM for 1:N (most of users receive, such as urgent distribution)
  - PIM-SM for N:N (meeting, games needed bidirectional communication)

- **Address allocation**
  - Static address allocation is realistic. Because there is no shortage of the multicasting address in the present though dynamic address allocation is better for effective use.

---

**Approach to selection of IP multicasting protocol**

**Group management protocol**

IGMP (IPv4) and MLD (IPv6) are mainstream protocols.

Although v2 (RFC2236) is currently mainstream for IGMP, v3 is recommended for efficient distribution of broadcast type multicasting. (Implementation with PIM-SSM as a set is recommended.)

**Data transfer protocol**

There are numerous data transfer protocols, which makes it difficult to select a specific one. We recommend RTSP/RTP, however.

**Routing protocol**

There are several protocols, and therefore the optimal protocol must be selected according to the type of multicasting.

PIM-SSM is recommended for 1:N type (video distribution, live broadcasting, TV, etc.).

PIM-DM is recommended for 1:N type (if this content is distributed to most of the users, such as an emergency multicast).

PIM-SM is recommended for N:N type (conferences, bi-directional video communication games).

**Address allocation**

The dynamic allocation provides a more effective usage of addresses. However, there is no shortage of multicasting addresses at the present time. We recommend static allocation.
Entire map of IP multicasting protocol applied to HIKARI Service Platform

This slide shows the entire map of locations where the IP multicasting protocol is applied.

Interface between content provider and delivery/distribution provider

1. Delivery of live broadcasting content and suppressing of distribution
   • We believe that the delivery/distribution provider can abandon content data if there is no viewer. Therefore, no interface is required.
2. Collecting broadcasting channel access logs (viewing information)
   • Optimal use of RTSP protocol/IGMP is recommended for transferring viewing information of users.
3. Protection against illegal viewing of content by a user who does not have a contract
   • Many methods can be used to identify a contracted user. For example, providing the delivery/distribution provider with a contracted user ID list, or sending an encrypted key of a content to contracted users. The detailed interface is a matter that must be studied in the future.
4. Operation when a new broadcasting channel is added
   • Application of SAP/GLOP addressing is recommended, since this operation is equivalent to setting a multicasting address.

Interface within delivery/distribution provider organization

1. Method for preventing degradation of performance due to expansion of content delivery/distribution scale
   [Optimal multicasting routing information]
   • Since an increase in broadcasting type multicasting, such as live broadcasting, is expected for HIKARI Service Platform, PIM-SSM multicast routing is recommended because it is optimal for broadcasting.
2. Multicasting address management
   • Although dynamic management will be required in the future, static management is good enough at present.
   • SAP/GLOP addressing is recommended.
3. Content delivery/distribution model
   • IP multicasting is a model where a network provider delivers and distributes a content.

Interface between user and delivery/distribution provider

1. Viewing start (join) and end (leave)
   • IGMPv3/MLDv2 is recommended (to be implemented with the PIM-SSM routing protocol as a set)
2. Protection against illegal viewing of content by a user who does not have a contract
   • The authentication function must be integrated in the IGMP protocol when the user join to a multicasting group is processed.
   • A proposal (draft) on an expanded IGMPv2 has been submitted to IETF.
   • Studies by the IETF Multicast Security Working Group (MSEC-WG) on multicasting security systems are in progress.
   ⇒ The detailed interface is a matter to be studied in the future.
5-3-12. Discusson of Multicasting Security (1/3)

Issues of IP multicasting security and necessary functions

- **User access control**
  - Need a user authentication function (ex. take it into IGMP)

- **Streaming data concealment**
  - Need ciphering IP multicasting data (Ciphered by common key reason why a lot of active users simultaneously receive same data)
  - Need suppressing alternation of IP multicasting data (Communicate through IPsec)

- **Peculiar issue to IP multicasting**
  - Cipher key management (There are the following issues due to sharing common key among users.)
    - How to delivery common key
    - How to change common key

To clear the above issues, under being discussed by IETF MSEC (Multicast Security Working Group), SMuG (The Secure Multicast Research Group)

The issues for assuring IP multicasting security are summarized below.

- **User access control**
  A user authentication function must be provided so that only contracted users can log in to a multicasting group. (Example: Integration of user authentication function in IGMP)

- **Streaming data concealment**
  IP multicasting data must be encrypted. Since a number of people will look up the same multicasting data, a common encryption key is generally used.
  Furthermore, protection against tampering of IP multicasting data by unauthorized persons must be provided. The IPsec communication method is available for this purpose.

- **IP multicasting-specific issues**
  Encryption key management (The following issues must be addressed when a common key is shared among a large number of receiving users.)
  - Distributing a common key.
    A common key must be distributed by a secure method to contracted users.
  - The common key must be changed when a new user signs up or a user leaves the multicasting group.
    This measure is required so that a) a user who leaves a multicasting group will be no longer able to view data, b) a new user will not be able to view data distributed before this user signed up with the multicasting group, or c) the user authentication is performed in the time zone within a single content.
    Also, since a number of users share the common key, there is a high risk of disclosure.
    Therefore, a key must be changed periodically.

The IETF Multicast Security Working Group (MSEC) and The Secure Multicast Research Group (SMuG) have carried out studies to resolve these issues.
Group key management architecture by IETF (MSEC)

The IETF (MSEC) group key management architecture <draft-ietf-msec-gkmarch-01.txt> has been proposed.

A group key management server is located between sender and receiver.

User authentication and exchange of a data encryption/decryption key are performed between sender (receiver) and group key management server by REGISTRATION PROTOCOL.

RE-KEY PROTOCOL reissues a data encryption key/decryption key as an option.

DATA SECURITY PROTOCOL encrypts the streaming data between sender and receiver.
Overview of the IP multicasting security system for HIKARI Service Platform (Draft)

This slide shows an overview of a system using the group key management architecture of IETF (MSEC) on the HIKARI Service Platform.

A group key management server is located at the delivery/distribution provider end.

Definition of interface between content provider and delivery/distribution provider
- Generating the group key and obtaining the encryption key
- Redistirution of encryption key

Definition of interface between user and delivery/distribution provider
- User authentication and obtaining the decryption key
- Redistirution of encryption key

We believe that the IP multicasting security can be safely implemented by the above methods.
5-3-13. Conclusion of Multicasting

1. Multicasting is necessary to provide rich live casting services.

2. The following functions and names of I/F are required to realize multicasting.
   - Between content provider and delivery/distribution provider
     - Start/Suppress streaming distribution to the broadcasting channel
     - Refusal notification to MCG joining by no contracted user
     - Exchange the generating information of MCG to new broadcasting channel
   - Between user and delivery/distribution provider
     - View start(Join)/end(Leave)
     - Illegal receiving guard for no contracted content
     - Notify status(join/viewing) of the MC(broadcasting channel) every change
   - Within content delivery/distribution provider
     - Generating optimal routing information
     - Multicasting address management

3. IP multicasting is technology that is expectable to expand and to be scalable in the future, and it is mostly realizable with the protocol set considered by IETF etc.

4. As a future issue, there are user authentication, data encryption (key management) etc. to realize the charged content broadcasting.
5-4. Network Quality and QoS Assurance

Contents of Section 5-4

5-4-1. Entire Flow of This Section
5-4-2. Issues for QoS (B2B2C Service)
5-4-3. Issues for QoS (C2C Service)
5-4-4. Technologies for Guaranteeing QoS
5-4-5. Issues for Broadband Net-live Services
5-4-6. Current QoS Assurance Technology
5-4-7. QoS Assurance Protocol
5-4-8. Other Measures
5-4-9. Conclusion
The above diagram shows the relationship between information given in Section 5-4.

The first half of Section 5-4 provides an overview of requirements, issues, and technical measures for the entire HIKARI Service.

The latter half provides a detailed study of the broadband net-live services that was selected between two services as a specific HIKARI Service in Chapter 4.
5-4-2. Issues for QoS (B2B2C Service)

The above table indicates QoS issues for each service that is classified as a B2B2C service. HIKARI Service can be placed in four categories from the following three viewpoints.

1. Characteristic of content (Live broadcasting or storage content)
2. Method of receiving a service (Streaming or downloading)
3. Bi-directional/unidirectional broadcasting

Moreover, the issues are presented with consideration for the following four parameters.

(1) Packet loss
(2) content encoding rate
(3) Maximum allowable delay (standby time + transmission delay time)
(4) Maximum allowable fluctuation
5-4-3. Issues for QoS (C2C Service)

<table>
<thead>
<tr>
<th>Name of service</th>
<th>Live broadcasting</th>
<th>Storage content</th>
</tr>
</thead>
<tbody>
<tr>
<td>VHS level</td>
<td>1.5 Mbps</td>
<td></td>
</tr>
<tr>
<td>NTSC level</td>
<td>3 Mbps</td>
<td></td>
</tr>
<tr>
<td>High quality</td>
<td>6 Mbps</td>
<td></td>
</tr>
</tbody>
</table>

Packet loss
No resend protocol is available because content is viewed in real time. There is no recovery method for packet loss besides the resend protocol. A packet loss will appear as noise in the picture quality. Therefore, packet loss is the most important issue for real-time services.

Content encoding rate
- VHS level (1.5 Mbps) good enough?
- NTSC level (3 Mbps) good enough?
- High quality (6 Mbps) good enough?

Allowable delay
Approximate time required for web access
Relays widely on the access operation performed by a user.
Approximate time required for web access

Transmission delay time
The delay will be no more than several seconds because of the importance put on live broadcasting.
User-friendly bi-directional communications must be enabled.
The delay is especially critical for interactive games.

Fluctuation
Determined by the network specification (delay, fluctuation, etc.) and terminal specification (buffer, etc.) based on a service request. The allowable buffer capacity will be different for unidirectional or bi-directional broadcasting. Also, it must be noted that measures for fluctuation will become more critical for bi-directional competition games.

The above table indicates QoS issues for each service that is classified as a C2C service.
5-4-4. Technologies for Guaranteeing QoS

<table>
<thead>
<tr>
<th>Characteristic of content</th>
<th>Live broadcasting</th>
<th>Storage content</th>
<th>Download</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of receiving a service</td>
<td>Streaming</td>
<td>Streaming</td>
<td>Streaming</td>
</tr>
<tr>
<td>Bi-directional/Unidirectional broadcasting</td>
<td>Unidirectional</td>
<td>Bi-directional</td>
<td>Unidirectional</td>
</tr>
</tbody>
</table>

### Example of specific services

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TV relay broadcast</td>
<td>Personal net broadcasting station</td>
<td>Network type - JuckBox</td>
<td>Music download</td>
</tr>
<tr>
<td>Concert live relay broadcast</td>
<td>Video monitoring</td>
<td>VoD</td>
<td>Video download</td>
</tr>
<tr>
<td>Remote monitoring</td>
<td>Monitoring/At-home care</td>
<td>Video clip, CM</td>
<td>File download</td>
</tr>
</tbody>
</table>

### (1) Packet loss
- The allowable packet loss depends on the extent of real-time broadcasting requirement. In case of some latency is allowed, retransmission can be applied.
- Since services that demand critical real-time broadcasting such as video phones and network games fall under this category, no measures are available at the present time.
- A time latitude for retransmission can be ensured by storing data in a cache at the terminal end.
- Not required.

### (2) Content encoding rate
- It is impossible to guarantee the encoding rate on open network in practical terms. It is possible to guarantee it to a certain degree using the rate shaping function (upper/lower limit) of L2/3-SW.

### (3) Allowable delay
- Main Technical measures:
  - Storing content at the local site
  - Content caching/request routing
  - Bypasses
  - Rate shaping function (upper/lower limit) by L2/3-SW

### (4) Fluctuation
- Fluctuation must be corrected as much as possible (suppressing the symptoms of fluctuation at user end) by buffering at the terminal end or using the IP shaper function of server and network equipment.

The above table indicates the techniques that assure QoS for each service type.
1. Prerequisites for studies
Among the “specific” services covered in Chapter 4 (“broadband net-live services” and “bi-directional video communication service”), the following sections describe the “broadband net-live services”, the nature of which demands very high importance on assuring QoS over the network.

Reasons why broadband net-live services is covered here

<table>
<thead>
<tr>
<th>Bandwidth of video</th>
<th>Broadband net-live</th>
<th>Bi-directional video communications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of requested picture quality</td>
<td>This service enables users to enjoy videos. The level of the requested picture quality is high.</td>
<td>Communications is the main purpose of this service and therefore the level of the requested picture quality is low. (The level of demand for delay is high.)</td>
</tr>
<tr>
<td>Number of simultaneous accesses</td>
<td>This service distributes live content. Therefore, a large volume of simultaneous access is expected.</td>
<td>The number of users who login to this service at one time (distance learning students) is several to several tens.</td>
</tr>
</tbody>
</table>

2. Network section that may be a problem
The above figure shows an example of the broadband net-live services. In this figure, we have analyzed the causes for problems in network quality and bandwidth assurance by dividing a network into the following three sections. Among these three, “distribution/delivery section” and “access section” need to be resolved. The “feed section” is exempted from our studies, as shown in the figure.

3. QoS parameters within the scope of our study
We have mentioned earlier that the QoS parameters include (1) packet loss, (2) encoding rate of content, (3) maximum allowable delay (standby time + transmission delay), and (4) maximum allowable fluctuation. Minimizing packet loss is important for the network quality of broadband net-live services. Therefore, we studied bandwidth assurance and priority control to find a way to minimize packet loss.
5-4-6. Current QoS Assurance Technology

<table>
<thead>
<tr>
<th>Current state of bandwidth assurance</th>
<th>Solution at present</th>
</tr>
</thead>
<tbody>
<tr>
<td>At present, it is impossible to guarantee the bandwidth for a network that extends over multiple vendor equipment and multiple ISP.</td>
<td>① Provide service in a manageable range by limiting the number of users and areas.</td>
</tr>
<tr>
<td>Several protocols for bandwidth assurance and priority control are currently being reviewed.</td>
<td>② Construct a system using equipment of a single vendor.</td>
</tr>
<tr>
<td>The function is integrated in actual equipment.</td>
<td>③ Install mirror/cache near a user. (Effective when the mirror/cache is used for the time-shift service.)</td>
</tr>
<tr>
<td>The detailed matters comply with each vendor’s original specification and there is no mutual connectivity.</td>
<td></td>
</tr>
<tr>
<td>A way to maintain service quality by methods different from bandwidth assurance is also being studied.</td>
<td></td>
</tr>
</tbody>
</table>

The current state of bandwidth assurance is indicated on the left side of the above figure.

Under the current circumstances, it is difficult to assure the bandwidth for a network that extends over multiple vendor equipment and multiple ISP. This is due to the following problems because the details of bandwidth assurance protocol, which is on the way toward standardization, have not been defined.

- No mutual connectivity between different vendor equipment
- Packet transfer control under the same conditions is not possible because operation rules are not standardized.
- Management of entire network is not enabled.

Therefore, to provide services by assuring the bandwidth of a network, it is necessary to implement measures ① to ③ indicated under “Solution at present” in the above figure.

In the following pages, we will describe on the related bandwidth assurance/priority control protocols. We will also overview the techniques used to maintain service quality by methods different from bandwidth assurance.
An overview of Integrated Services (Intserv) and Resource Reservation Protocol (RSVP) is provided as an example of the bandwidth assurance protocol under review at the present time.

(1) Overview of Intserv and RSVP
In this system, an individual application requests a network for the required bandwidth, which is reserved and assured in accordance with this request. Intserv specifies the QoS parameters and defines a parameter exchange format, and RSVP performs the bandwidth reservation accordingly.

(2) Characteristics of RSVP
- The receiver side initiates the reservation of resources.
- Application to multicasting communications is also enabled in addition to unicasting.
- Since RSVP applies soft-state control (control for periodically updating resource reservation messages), it can dynamically handle the changing of a route and multicasting group members during communications.

(3) Standardization trends
The IETF standardization work related to Itserv and RSVP is nearly complete.
http://www.ietf.org/html.charters/intserv-charter.html
http://www.ietf.org/html.charters/rsvp-charter.html

(4) Issues
- These protocols do not provide the function that can transfer data using the requested communication quality using an actually allocated resource. Such architecture must be implemented by network equipment such as a terminal and router using standalone technology.
- Since bandwidth control is performed per flow, the control is complicated and scalability is a problem.
The outline of Differentiated Services (Diffserv) is described as an example of the bandwidth assurance protocol being studied at the present time.

(1) Outline

This method divides the IP packets that are fed into a network into categories at the gateway and transfers a packet in accordance with the buffer and bandwidth allocated for each class.

(2) Characteristics

- The result of assigned classes is imbedded in the header of an IP packet as Diffserv Code Point (DSCP).
- Complicated processing such as DSCP setting is performed at the gateway. Therefore, the load on the network can be less and this method is suitable for a large-scale network.
- Unlike Intserv, Diffserv does not require the status management for each flow. Therefore, it can be implemented by a simple control operation.

(3) Standardization trends

- The IP packet portion to be used for DSCP setting and the transfer method at each node based on the set DSCP are defined.
  
  [Link: http://www.ietf.org/html.charters/diffserv-charter.html]

(4) Issues

 Expedited Forwarding (EF: bandwidth assurance type) is also available for Diffserv in addition to Assured Forwarding (AF: relative priority control type). When the traffic with high priority becomes congested, the quality cannot be guaranteed when AF is used. Whether the requested bandwidth can be assured is confirmed at the time of acceptance when EF is used. Therefore, the complication in control, such as resource management, is the issue for EF.
5-4-8. Other Measures

Optical video distribution system (SCM-PON): CATV technology is used on optical fiber.

• System for transmitting the video signals used for coaxial CATV by converting them as is into optical signals.
• Converts analog video or digital QAM video into optical signals using FM conversion.
• Different from the standpoint of service, which assumes an IP network.
• Effective for a simple multicast-type service (The time-shift service and special playback are difficult.)
• This method has been already provided as optical CATV service in some areas.

Content switch with cache function: Technology to store live stream data in cache

• This technology has a conventional high-order layer switching function and cache server function.
• The video packet delay is compensated by applying cache technology for the distribution of a live stream.
• Note, however, that a delay of several seconds is unavoidable.

Besides the concept for assuring the bandwidth on an IP network and also assuring the video service quality, there are ways to ensure quality using different technologies. Technologies currently available for this purpose are discussed below.
5-4-9. Conclusion

**Implementation of RSVP is a difficult matter at the present time.**
- Since bandwidth control is performed per flow for RSVP, control is complicated and scalability is a problem.
- Detailed studies on the practical use of RSVP are not being carried out at present.

**Conclusion**

It is practical to implement priority control using Diffserv at present. Note, however, that there are cases where priority control may not be effective if there is an increase in the ratio of traffic (such as video data) that should be taking high precedence over other information among the entire volume of data. Therefore, it is also necessary to construct a distribution and delivery network that can assure sufficient bandwidth.

RSVP and Diffserv are bandwidth assurance protocols. Bandwidth control is performed per flow for RSVP, and therefore control is complicated and scalability is a problem. Also, detailed studies on the practical use of RSVP are not being carried out at present.

Due to the above, it is practical to implement priority control using Diffserv at the present time. Note, however, that there may be an increase in the ratio of traffic (such as video data) that should take higher precedence over other information among the entire volume of data due to the increase of the access network bandwidth and wider expansion of video services. In such cases, priority control may not be effective. Therefore, it is also necessary to construct a distribution and delivery network that can assure sufficient bandwidth.
5-5. Time-shift Function

Contents of Section 5-5

5-5-1. How to Implement the Time-shift Function
5-5-2. Implementing the Time-shift Function by Storing Data in Server
5-5-3. Issues for Time-shift Function by Storing Data in Server
5-5-1. How to Implement the Time-shift Function

1) Implementation by means of a storage function in a cache server in CDN

2) Implementation by means of a storage function in the HDD of a terminal

Two methods for implementing the time-shift function are:

1) Storing data in CDN
2) Storing data in a terminal

To store data in a terminal, a hard drive or memory may be used.

See Chapter 8 for the time-shift function implemented by storing data in a terminal.
5-5-2. Implementing the Time-shift Function by Storing Data in Server

Issues for studies as server architecture

The above figure shows an example of a server configuration that has a storage function for implementing the time-shift function.

As shown in the figure, there are three stream flows for implementing the time-shift function.
1) Receiving a live stream from a content provider and storing data in a hard drive.
2) Live playback → Transferring of live stream
3) Time-shift playback → Content playback from hard drive

A delivery server switches live playback and time-shift playback in accordance with a request from a terminal.

Time-shift playback can perform a special playback operation of the content stored using the VoD method.
(Special playback operations such as jump backward, rewind, and fast forward playback within the hard drive space where time-shift playback is enabled.)

A method for reading the live playback from the last ending location on a hard drive and a method for transferring received data on memory are enabled for time-shift playback.
5-5-3. Issues for Time-shift Function by Storing Data in Server

Issues for studies as server architecture

(Issue 1) Increase of load on server and network caused by time-shift playback

The server (VoD) processing and network bandwidth are required for the number of terminals that are simultaneously used for viewing time-shift playback.

(Issue 2) Increase of delay due to storage operation to a hard drive

To realize time-shift function, VoD processing by unicasting for each terminal is necessary.

Therefore, when time-shift playback is performed, a server that has a VoD server processing (independent stream playback) capacity for a number of terminals is required. Also, a unicasting network bandwidth must be allocated between terminals and server.

When the live playback is performed from storage hard drive in order to implement seamless live playback, the delay will increase due to the hard drive storage operation. (The increase of delay amount is less when the live playback is performed from the memory.)

Note, however, that a delay of several seconds is unavoidable due to the encoding operation for the played back live video. Therefore, this delay is not a critical problem.
5-6. Metadata/EPG

Contents of Section 5-6

5-6-1. What Is Metadata?
5-6-2. Standardization of Metadata
5-6-3. Flow of Metadata in HIKARI Service Platform
5-6-4. Type of System for Providing Metadata to Users
5-6-5. Conclusion on Technical Issues Related to Metadata and EPG
5-6-1. What Is Metadata? (1/3)

Metadata: This data defines various attributes of a content. It is provided to a user separately from the content in order to improve the facility of viewing.

Metadata defines the characteristics of a content in order to facilitate and improve the viewing (for example, making it easier to find and view a desired content from multimedia such as audio and video) and to restrict the viewing for specific viewers only. Normally, content includes a wide variety of material ranging from a text-based to video and audio. To find specific content from the information distributed all over the HIKARI Service Network, a user needs to perform a search. It is comparatively easy to search for a text-based content. Extracting the characteristics of audio or video content from this data, however, is harder. Metadata, which can define the characteristics of specific information, can be very useful for this purpose. In broadcast services, an Electronic Program Guide (EPG) can be compiled on the user terminal using metadata attached to a program. It also enables a user to obtain detailed information of a program in the form of text and image data. Metadata includes: (a) information directly related to a content, such as the program name, category, outline, and remarks, (b) information associated with a content, such as location, broadcasting time, access restriction, and distribution parameters, and (c) information associated with a user, such as access history.

Broadband bi-directional communications is enabled for HIKARI Service Network. Therefore, each user can request and obtain metadata. However, the EPG data must be updated due to a change in the content distribution/delivery time by broadcast communication, and the load on a network will increase. Therefore, some type of system is required to reduce the data volume, like taking advantage of multicast communication or updating only changed portion of data. Although unidirectional broadcasting is used, the availability of broadband multi-address transmission allows a large volume of metadata to be distributed to all users at once. The EPG data can also be updated in real time. The number of channels and therefore the entire volume of data are also limited.

- Displaying Electronic Program Guide (EPG)
- Searching a content
- Creating a content library
- Viewing the program digest

- Content identification information
  - Content ID
- Information related to the details of a content
  - Program name, category, outline, and review
- Information associated with a content
  - Location, broadcasting time, viewing restrictions
- Information related to ownership rights
  - Author, producer
5-6-1. What Is Metadata? (2/3)

TV-Anytime Forum and ARIB (see 5-6-2) have carried out mainly the following standardizations of metadata.

✓ Metadata on program information
   Information provided as an aid to search for content to be viewed and reservation of storing.
   • Program Information
     Attributes related to a program such as program title, outline, keyword, and category.
   • Group Information
     Information related to groups configured based on associated content.
   • Program Location
     Information used to show the scheduled content distribution time.
   • Service Information
     Information related to a service provided by such as Content Distributor.

✓ Segmentation metadata
   Information attached to indicate a category and keyword for each segment, such as a program scene, and used to search and playback a specific scene in a program.
### 5-6-1. What Is Metadata? (3/3)

Example of metadata description (Program Information)

Source of data: Proposals of General Broadcasting System based on Home Servers (ARIB)

This is an example of a description of Program Information in a program information metadata. XML is used as description language.

```xml
<ProgramInformationTable>
<ProgramInformation ProgramId="crid://hbc.com/foxes/episode11"/>
<BasicDescription>
  <Title type="main">The one where Fox jumps in the Potomac</Title>
  <Synopsis>Fox goes to Washington and jumps in the Potomac</Synopsis>
  <Keyword>Fox</Keyword><Keyword>Washington</Keyword><Keyword>Potomac</Keyword>
  <Genre type="main">Comedy</Genre>
</BasicDescription>
<OtherIdentifier>102330002211</OtherIdentifier>
<MemberOf>crid://hbc.com/foxes/all</MemberOf>
</ProgramInformation>
```

The above example of Program Information describes the following information of a content:

- **Title**: The one where Fox jumps in the Potomac
- **Outline**: The one where Fox jumps in the Potomac
- **Search keyword**: Fox, Washington, Potomac
- **Genre**: Comedy
5-6-2. Standardization of Metadata

<table>
<thead>
<tr>
<th>Standardization body</th>
<th>ISO/IEC JTC1 SC29/WG11 (MPEG)</th>
<th>TV-Anytime Forum</th>
<th>ARIB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Related publications</td>
<td>MPEG-7</td>
<td>SP003v1.1, SP004v1.1</td>
<td>Proposals of General Broadcasting System based on Home Servers</td>
</tr>
</tbody>
</table>

- **Definition of metadata**
  - Information associated with program
  - Details of program
  - Access conditions
  - Location
  - Access record of a user
  - Segment information, etc.
  - Method for looking up content
  - Resolution of location
  - CRID

- **Metadata**
  - Metadata description language
  - Metadata description format
  - Metadata transmission coding system
  - Metadata transmission with an electronic watermark
  - Content identification data
  - Storage control method for stream type content
  - Storage control method for metadata
  - Mutual interfacing method for content

The ARIB Specification, which provides consideration for compatibility with other standards and migration capability, is applied by HSAC as a reference standard.

The above table indicates representative standardization bodies and publications related to metadata. MPEG is in an ongoing process of standardizing the compression of audio/video data. MPEG-7 standardizes the definition of attributes to be searched for in multimedia data, instead of standardizing the multimedia data itself. The MPEG-7 standardized data includes: Descriptor (D), Description Scheme (DS), and Description Definition Language (DDL). A descriptor is a basic tool for describing a characteristic of multimedia content, and MPEG-7 defines the depiction method and meaning of a descriptor. A description scheme is a framework that defines a structure between multiple description tools and relationship in terms of meaning. In the same way as the descriptor, MPEG-7 defines the depiction method and meaning of a description scheme. The Description Definition Language defines the descriptors and depiction method of a description scheme.

TV-Anytime Forum is a non-profit organization established in September 1999 to carry out standardization. The Forum is in the process of standardizing metadata to broadcast programs to the terminals with storage and providing audio/visual services without depending on transmission systems, such as the Internet. The scope of the technologies to be standardized includes the specifications of metadata, content request, and ownership rights management. The metadata specification is defined in “Specification Series: S-3 Specification (SP003v1.1)”.

Association of Radio Industries and Businesses (ARIB) is a corporate body that performs studies, researches, and developments related to the effective usage of radio waves in the communication/broadcasting sector. ARIB also performs joint technological studies with various standardization bodies.

A proposal of a system for using metadata is included in a part of the document “Proposals of General Broadcasting System based on Home Servers” submitted by ARIB based on Draft No. 2003 of the Committee for Broadcasting System based on Home Servers in Information Communication Council. This document defines the metadata description language, metadata description format, and metadata transmission coding system, including considerations for compatibility issues with MPEG-7 and TV-Anytime Forum. The proposals of the stream type content storage control system, metadata storage control system, and link between contents also have considerations for migration with existing broadcasting systems. Due to the above, HSAC has decided to use the ARIB Specification as reference standard.

Besides the above, Content ID Forum (cIDf) is studying the metadata called content ID as a part of standardizing the attribution set aimed at content management.
The above figure shows the expected flow of metadata on HIKARI Service Platform. For the sake of convenience, the flow of metadata is explained below by separating the content into two types: non-storage and storage.

(1) Non-storage type content

A user can select and view a desired program in Electronic Program Guide (EPG) for programs that are distributed live according to a time schedule. In existing broadcasting (example: BS digital broadcasting) systems, the service information (SI) created by program production companies (broadcast station) is collected at the SI Collection and Delivery Center. SI of all stations are subsequently returned to each program production company and distributed as EPG data from each program production company. The net-live distribution services of HIKARI Service Platform also require a function that can compile SI of each content provider and distribute them to a user (terminal) as EPG data. In this report, this function is defined as “SI Collection and Delivery Provider”. The content provider-specific SI uploaded to the SI Collection and Delivery Provider do not have the physical constraints of service area, unlike existing broadcasting systems. Therefore, the maximum viewer target can be extended to the entire world. Targeting the entire world would require vast volumes of SI data if SI of all stations are prepared, and it is certainly not practical to distribute EPG data covering all programs to all users. Therefore, we consider it necessary to provide a system to screening the user profile data (such as language and area), user preference data (such as content genres and personalities), and terminal attribute information (such as playback file format and frame rate), and create an EPG in accordance with distribution area and individual user preferences. The update method of a user-specific or area-specific EPG also must be studied. Such a configuration will enable a user to view widely diverse content that were never provided by conventional broadcasting services. There is no need to return the final form of created EPG data to a content provider. Instead, the EPG data can be directly distributed to a user from an SI Collection and Delivery Provider taking full advantage of the IP network. In this case, an EPG data should not be referred to as metadata any longer.

(2) Storage type content

On the other hand, server storage type content on a network are distributed according to the viewing request of a user. On-demand type services like this one requires a content search method that is different from the time-schedule-base EPG search method. TV-Anytime Forum has proposed a content locating method in the SP004 Specification. We conclude that the specification of content based on this proposal is possible also for HIKARI Service Platform.
5-6-3. Flow of Metadata in HIKARI Service Platform (Cont’d)

The flow of metadata is explained below with the corresponding serial numbers shown in the figure.

A metadata creator adds metadata that indicates attributes and locations of a content owned by a content provider. The database function for collecting metadata groups that extend over multiple content providers is provided by “Content Search Service Provider”. A Content Search Service Provider accepts the content attributes and registration of content location from the Content Provider. Sometimes, a mechanical search of content on the network may be extensively performed. The information collected thus is stored in a database and part of this database is provided to a user as content information. (①)

A user uses the Content Reference Identifier (CRID) to specify a desired program from the content information acquired from a Content Search Service Provider or other media. (②)

A Content Search Service Provider will return a physical location (location data) of the content matching up with the specified CRID to the user. (③)

When a CRID that specifies multiple contents such as a serial program is specified, the CRID group may be returned to the user. Even in such cases, the location data is eventually provided to a user after several communication exchanges. The user will be able to view the content by specifying a content provider based on the location data and requesting content distribution. (④ to ⑥)

As mentioned above, we have defined SI Collection and Delivery Provider for the time schedule-based content search and Content Search Service Provider for an on-demand content search. A seamless content access at the user terminal from a service provider who integrates both functions is a more natural choice. Construct such an integrated user interface environment is an issue for HIKARI Service.
5-6-4. Type of System for Providing Metadata to Users

Basically, a broadcasting service is a unidirectional broadcast type service without any intermediate storage in the path of distribution. At present, studies are being carried out for the application of the PDR system, which involves installing a storage device at a user’s home and stores content and metadata in this device. As the digitalization progresses in the future, bi-directional broadcasting services will become commercially available. In anticipation, studies are being directed toward the NDR system, which allows metadata to be stored in a network for access and use from a terminal side.

On the other hand, bi-directional video communication is a basic requirement for HIKARI Service. There are also the following issues:

1. Ensuring synchronization between content and metadata
   Since the content and its associated metadata are managed separately, a means must be provided to synchronize them securely. For this purpose, a link must be provided between the metadata management system and dynamic management of the content location and pairing must be performed.

2. Assuring latest information
   Since there will be a large volume of NDR access, we deduce that NDR must be installed in a decentralized manner like the mirror and cache for storing content. The metadata stored in a NDR system must be updated instantaneously when a new data is added or updated so that the latest information is always available.

3. Assuring easy access to content
   It is desirable that a user can quickly search and find the required information from the vast volume of content provided. For this purpose, the details of a content (semantics) also must be defined based on some unified standard in addition to the description format of metadata, and the content must be organized into a systematic form accordingly.

We expect that application of the NDR system for HIKARI Service will improve the service portal (easier content search for user), improve ownership rights protection, make it easier to access content located worldwide, and promote greater content distribution.

The following issues need to be considered to achieve the above goal.

1. Ensuring synchronization between content and metadata
   Since the content and its associated metadata are managed separately, a means must be provided to synchronize them securely. For this purpose, a link must be provided between the metadata management system and dynamic management of the content location and pairing must be performed.

2. Assuring latest information
   Since there will be a large volume of NDR access, we deduce that NDR must be installed in a decentralized manner like the mirror and cache for storing content. The metadata stored in a NDR system must be updated instantaneously when a new data is added or updated so that the latest information is always available.

3. Assuring easy access to content
   It is desirable that a user can quickly search and find the required information from the vast volume of content provided. For this purpose, the details of a content (semantics) also must be defined based on some unified standard in addition to the description format of metadata, and the content must be organized into a systematic form accordingly.
5-6-5. Conclusion on Technical Issues Related to Metadata and EPG

- Detailed construction method of Content Search Service Provider
- Measures for marked increase of programs and content associated with the expansion of service area
- Detailed construction method of Service Information (SI) Collection and Delivery Provider
- How the EPG should be for an on-demand service

ARIB and TV Anytime Forum are also reviewing metadata on an IP network. We believe that we can make the most of their study results. On the other hand, since a high-speed IP network infrastructure is used for HIKARI Service Platform, a service such as streaming distribution of rich content using CDN must be considered in addition to storage type services. We have summarized the issues below for a study on the detailed construction method of a Content Search Service Provider in HIKARI Service Platform using metadata and EPG (which uses the metadata).

(1) Unlike broadcast wave including today’s satellite broadcasting, IP network broadcasting does not have restrictions regarding distribution areas and therefore the maximum target of content and viewers can be extended worldwide. In that case, the amount of metadata related to program information alone will be vast. It is not practical to equally distribute EPG data covering all programs to all users. Therefore, we conclude that a system is required to screen the user profile information (such as language and area), user preference information (such as content genre and personalities), and terminal attribute information (such as playback file format and frame rate), and an EPG should be created in accordance with distribution area and individual user preferences. Also, the update method of a user-specific or area-specific EPG must be considered.

(2) In general, Service Information (SI), which is a component of an EPG, is created by each program production company (broadcast station). It is then collected at the SI Collection and Delivery Center and distributed to users as EPG data. It is necessary to study the kind of SI Collection and Delivery Center to be constructed on an IP network for HIKARI Service Platform, including its service operation methods.

(3) In the case of on-demand content distribution services, the concept of a channel-specific timetable for an EPG has no use. Instead, a content guide personalized for each user preferences is required. Therefore, we must study the user interface of a standard content guide. It is desirable that these user interfaces should be able to access time-schedule-based content (program) guide in a seamless manner.
5-7. Viewing Quality Management

Contents of Section 5-7

5-7-1. Why is Viewing Quality Management Necessary?
5-7-2. Targets of Quality Surveillance and Detection of Degradation
5-7-3. Detection of Degradation at Server
5-7-4. Detection of Degradation at Network
5-7-5. Detection of Degradation at Terminal
5-7-6. Viewing Quality Management Method
5-7-7. Conclusion and Future Issues
5-7-1. Why is Viewing Quality Management Necessary?

A paid viewing service should be able to provide the user with a video stream without noticeable degradation in picture quality. Viewing quality can be managed in three steps as shown in the above figure, namely “detecting quality degradation”, “identifying degraded item”, and “control for eliminating degradation”.

When a user makes a claim for degradation of video quality, if the quality surveillance data is retained at all times, it will be easy to determine whether the cause of degradation was due to a problem on the network or user side. If there is a service level agreement (SLA) between a user and provider, such surveillance data will be required information.
The three probable targets of quality surveillance in detecting degradation are summarized below.

1. Detecting degradation at the server
   - Detection at the server may be performed using the two methods outlined below.
     1) Surveillance of the in-use and performance statuses of server resources
        This system monitors the in-use and performance status of server resources such as CPU, memory, HD, and NIC and the server operating status such as the number of logged in clients. When the rate of use has exceeded a certain value or there are many packets and transactions waiting for processing in the queue, the distributed video might start to degrade, or a new request for video distribution may be denied.
     2) Surveillance by means of a server log
        If the server or user terminal software (for video reception) detects poor quality in each stream sent from a server or an abnormal disconnection, this event is recorded in a server log.

2. Detecting degradation in the network
   - There are various parameters that identify network quality. Among these, the following parameters are immediately related to degradation of the video quality and must be monitored.
     1) Packet loss rate
        Packet loss affects the video quality significantly and is the most important surveillance parameter.
     2) Delay (or delay fluctuation)
        Delay (or delay fluctuation) is the next most important surveillance parameter. Fluctuation is absorbed to some degree by the terminal buffer. If the delay is longer than the designated period or if there is a delay fluctuation, the following phenomena may occur:
        (a) the video packet required for decoding and playback at the terminal is not delivered and (b) the packet resent after detection of packet loss is not received on time for playback.

3. Detecting degradation at the terminal
   - The quality of the video received or status of a packet delivered to a terminal must be confirmed to verify whether the video ultimately received by a viewer has degraded compared with the original.
### 5-7-3. Detection of Degradation at Server (1/2)

**Surveillance of access status/performance of server resource**

#### Example: Surveillance items at server

<table>
<thead>
<tr>
<th>Category</th>
<th>Item</th>
<th>Main reasons why a relevant item is under surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>Amount of available memory</td>
<td>This item is related to the number of clients that can be connected and bit rate of transmission files.</td>
</tr>
<tr>
<td>NIC</td>
<td>Total throughput</td>
<td>This item is related to the volume of transmission to a connected client. Especially in the case of a TCP stream, there may be quality degradation in the flow control due to an insufficient bandwidth.</td>
</tr>
<tr>
<td>CPU</td>
<td>CPU utilization (%)</td>
<td>This item is related to a connection request and distribution processes.</td>
</tr>
<tr>
<td></td>
<td>Number of threads in a processor queue</td>
<td></td>
</tr>
<tr>
<td>HDD</td>
<td>Number of requests in a disk queue</td>
<td>These items are related to an access delay due to processing of the remaining connection requests.</td>
</tr>
<tr>
<td></td>
<td>Number of delayed retrievals completed per second</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Speed at which a server reads data from HDD</td>
<td></td>
</tr>
<tr>
<td>Streams</td>
<td>Number of clients in the server connection standby status</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bandwidth calculated based on the number of connected clients</td>
<td>These items are related to an access delay due to processing of the remaining connection requests.</td>
</tr>
<tr>
<td></td>
<td>Total number of error occurrences per second</td>
<td>This item means that an internal packet cannot be generated in a server and indicates degradation of the quality.</td>
</tr>
<tr>
<td></td>
<td>Number of connected clients</td>
<td>If the threshold is set with a sufficient safety margin, the surveillance of this item may be good enough in some cases.</td>
</tr>
<tr>
<td></td>
<td>Number of clients to whom a content is being distributed</td>
<td></td>
</tr>
</tbody>
</table>

We have provided an example of the surveillance items in server resources as the primary detection method at the server. The items in the above table must be monitored using (a) a performance monitor supplied with commercially available servers or (b) a server management tool used for a wider range of server surveillance.

The performance monitor or the server management tool may not always be able to monitor the same item. Moreover, there are cases where the definition, monitoring cycle, and accuracy are different. Therefore, there must be a benchmark that extracts beforehand the items relative to video quality degradation.

Also, it is necessary to set threshold values for each surveillance item where server congestion or degradation of the distributed video is detected. The preset benchmark must take into account that a threshold depends on the server type, OS, and hardware configurations to a certain degree.
5-7-3. Detection of Degradation at Server (2/2)

Surveillance by server log

We have suggested using surveillance server logs as the second method for detecting degradation at a server. The table lists the logs likely to be related to quality management and problem detection among logs that can be retrieved by commercially available servers. (The logs that can be retrieved depend on the server specifications.) Some of the logs in the table are not directly related to quality management and problem detection. However, all logs that are even marginally involved are listed up because the data content and accuracy vary according to the server equipment. Note that the volume of data will increase significantly and therefore the intervals between data retrieval must be suppressed to a minimum within the range where degradation detection is not affected.

<table>
<thead>
<tr>
<th>General item</th>
<th>Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server information</td>
<td>Name of file being played back</td>
<td>Server’s DNS</td>
</tr>
<tr>
<td></td>
<td>Number of bytes sent to a client from a server</td>
<td>Number of clients connected to a server</td>
</tr>
<tr>
<td></td>
<td>Number of packets sent from a server</td>
<td>Average load (%) of server processor</td>
</tr>
<tr>
<td></td>
<td>IP address</td>
<td></td>
</tr>
<tr>
<td>Client information</td>
<td>IP address</td>
<td>Type of CPU used by client</td>
</tr>
<tr>
<td></td>
<td>DNS name</td>
<td>Number of bytes sent from a server and received by a client</td>
</tr>
<tr>
<td></td>
<td>Time stamp of a data stream when an entry is generated in a log file</td>
<td>Number of packets sent from a server and properly received by a client during the first attempt.</td>
</tr>
<tr>
<td></td>
<td>Client status code</td>
<td>Number of packets lost during the transfer from server to client, and could not be recovered by the error correction of a client layer or UDP re-transmission of network layer.</td>
</tr>
<tr>
<td></td>
<td>Player’s global identifier</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Player’s version</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Client language</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Browser type</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Host application</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Host application version</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Client OS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Client OS version</td>
<td></td>
</tr>
<tr>
<td>Stream information</td>
<td>Length of data stream that has been played back until client event</td>
<td>Average bandwidth used when a client connects to a server</td>
</tr>
<tr>
<td></td>
<td>Made effective when the last command event was sent (Stop, rewind, fast-forward, etc.)</td>
<td>Protocol used for accessing stream</td>
</tr>
<tr>
<td></td>
<td>File length (sec.)</td>
<td>Transfer protocol used for a stream</td>
</tr>
<tr>
<td></td>
<td>File size</td>
<td>Audio codec used for a stream</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Video codec used for a stream</td>
</tr>
</tbody>
</table>
5-7-4. Detection of Degradation at Network

The evaluation graph above is an example of the relationship between packet loss (discarding rate) and video quality. The vertical axis, “Overall quality of video”, indicates the DMOS evaluation value. DMOS evaluation value is the level of dissatisfaction that a viewer feels upon comparing the video (evaluated video) degraded due to packet loss with the coded video (original video) that has no packet loss.

The findings from this evaluation example indicate that: (a) the video quality will not be significantly affected by a small amount of packet loss and (b) if the amount of packet loss exceeds a certain threshold, the quality will degrade considerably. This graph is obtained when retransmission control is provided. If there is no retransmission, the video quality will degrade even at a packet discarding rate lower than that in this example. Since the numeric values depend on the software being used, the receiving buffer size, and the coding method, they cannot be determined uniformly.

Comparing the graph of a 384 kbps video transmission with that of a 1Mbps transmission, it is clear that the video quality of the 1 Mbps graph degrades even when the discarding rate of the 1 Mbps graph is lower than that of 384 kbps. By this comparison, we can conclude that the surveillance of packet loss is more important for high-quality video distribution.
5-7-5. Detection of Degradation at Terminal

Method used to detect degradation at terminal

To determine whether the video ultimately received by a user has been degraded compared with the original, degradation must be checked at the terminal. Ideally, degradation should be checked using a received video. To evaluate the quality of a received video, dedicated surveillance equipment for detecting freeze-repeat symptoms of a video and block noise occurrence can be installed at the terminal (as shown in the figure). (Besides this, a system (or equipment) that compares a received video with its original can be also applied. However, the latter method is generally implemented when a highly accurate quality evaluation is required, such as evaluation of a server or network performance. The method indicated in the above figure is more popularly used to monitor video quality.)

Note, however, there is the issue of the cost involved in providing all viewers with a function for detecting degradation. Therefore, in practical terms, the actual solution seems to be the method of detecting quality degradation based on response time, throughput, or packet loss at the terminal, instead of checking the quality of the received video itself.
5-7-6. Viewing Quality Management Method (1/2)

Example: Procedure for handling viewing quality management

1. Detection of quality degradation
   - Detects degradation of quality based on response time, throughput, or packet loss.

2. Route search
   - Determines a target of management during URL-to IP-address conversion, route search, and detection route.

3. Isolating a degraded segment
   - Isolates a segment where the performance has degraded by analyzing the alarm correlation and checking each network section.

4. Network bottleneck analysis
   - Measures the MIB information, CPU in-use rate, number of packets, packet loss, and delay to identify a problem.

5. Server bottleneck analysis
   - Measures the system log, CPU in-use rate, memory in-use rate, and process status to identify a problem.

6. Analysis result report
   - The content of an analysis result report would be the quality degradation amount, cause of degradation, affected range, measures for improvement, and the know-how available.
5-7-6. Viewing Quality Management Method (2/2)

Example of viewing quality management system

The above figure indicates the concept of a viewing quality management system.

1. Function for detecting quality degradation
   Detects packet loss at a terminal and sends terminal/server/degradation parameters to the quality management function.

2. Function for analyzing degraded section
   Performs a route search at the sender and receiver sides using ‘Traceroute’. The degraded section can be isolated by monitoring throughput or packet loss.

3. Network/Network Element’s control function
   Since priority control is performed at the router in accordance with predetermined rules, the quality management function sends the Network/Network Element’s control function command.

4. Quality management function
   Performs quality management using network configurations, IP address, user information, and routing information.
5-7-7. Conclusion and Future Issues

Conclusion

[Detection of quality degradation]
- Detection of degradation at server: Monitors the CPU and memory use rate
- Detection of degradation at network: Monitors packet loss and line use rate
- Detection at terminal: Response time, throughput, packet loss, etc.

[Viewing quality management method and system]
- Required functions:
  - Function for detecting quality degradation
  - Function for analyzing degraded section
  - Network/network equipment control functions
  - Quality management function

Future issues

- Method for detecting the degradation of network quality

[Conclusion]

We believe that the three methods in this slide can be applied for detecting viewing quality degradation. We have summarized the items that should be monitored by each facility and equipment. We have also introduced an example of the procedure for handling viewing quality management and a concept for the viewing quality management system.

[Future issues]

Most of the building blocks of the viewing quality management system (presented this time) can be constructed by combining commercially available technologies. There are, however, the following future issues.

- Method for detecting degradation of network quality

Finding the packet loss rate and delay by capturing a packet of each individual distribution stream is not a practical implementation method. Furthermore, the quality value of an individual stream packet and that of a network may not match exactly because the quality changes depending on the user terminal and operation.

Therefore, the degradation of network quality should be detected by setting up a test terminal, performing distributing/receiving stream test, and sending/receiving a test packet. In this case, there must be a way to determine: (a) specification of a test packet that simulates the characteristics of a distribution stream, (b) the amount of test required in terms of statistics, and (c) a method for determining the degradation criterion.
6. Technical Issues for Security

Contents of Chapter 6

6-1. Protection Against Server Tampering
6-2. Protection Against Terminal Attack
6-1. Protection Against Server Tampering

Contents of Section 6-1

6-1-1. Possibility of Server Attack
6-1-3. Server Attacks on Current Networks
6-1-4. DoS (Denial of Services)
6-1-5. DDoS (Distributed DoS)
6-1-1. Possibility of Server Attack

Issues on preventing intrusions and illegal operations

<table>
<thead>
<tr>
<th>#</th>
<th>item</th>
<th>detail</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intrusion</td>
<td>➢ Tampering content or access rights</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>➢ Leakage or falsification of personal information</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Leakage</td>
<td>➢ Falsification of important data such as account information</td>
<td>Should be cared upon authentication</td>
</tr>
<tr>
<td>3</td>
<td>Falsification</td>
<td>➢ Making the system crash</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Spoofing</td>
<td>On accessing content or accounting</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Eavesdropping</td>
<td>Listening or recording content being distributed on the network</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>DoS</td>
<td>Making the system unusable or overloaded by sending illegal or a high volume of data</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Offensive material</td>
<td>Distributing harmful content such as adult material that a child should not view</td>
<td></td>
</tr>
</tbody>
</table>

As the services such as B2B, B2C, and C2C become more diverse, security issues such as authentication and protection against unauthorized access and malicious attacks will become increasingly critical for HIKARI Network. Also, with the rising use of bi-directional connections and broadband communications, dangers such as unauthorized access and malicious attacks will increase much more compared with the Internet now. (See the above table.)

Although the servers used in a conventional IP network are exposed to various types of malicious attacks, we basically discuss the most significant types and their measures for HIKARI Network in this section.

The measures for each type of attack as indicated in the table must be provided for server security. Also, there must be a clearly defined security policy for the entire system, and this policy must be implemented with utmost seriousness.

[Probable issues]

With the use of services such as B2B, B2C, and C2C growing as a result of an increasing variety of players and server applications, the security issues of the Internet today will be much bigger for HIKARI Network. Probable issues such as authentication and protection against unauthorized access and malicious attacks are indicated in the table. (See the above table.)

The security for the above items must be taken into consideration at a server end, as is the case with a general user. In particular, the security policy must be very clear on the protective measures for server security.
The ISO definition of evaluation criteria for the security of an information network system is indicated below. An administrator must take the criteria into account for system construction and operation.

**[Evaluation Criteria]**

- Information Security Management System (ISMS): ISO17799
  A standard on the security policy, management objectives, management methods and system operation of management systems; and management of system documents such as information security policy.

- IT Security Evaluation Criteria: ISO15408 (JIS5070)
  A standard on the security features of a product and their warranty

- Guidelines for the Management of IT Security (GMITS): ISO TR 13335
  A series of technical reports which provides guidance of security management of the whole system. It consists of concepts and models, managing and planning, management techniques, selection of safeguards, and safeguards for external connections.

For deployment of the HIKARI service, we should consider applying standards as above at the server (distributor) side. Actions will be needed at the client (ordinary user) side since there is no such standard.

Technology such as user authentication, encryption, and firewall are available as protection from unauthorized access. Providing high security will be an issue of ever-increasing concern for the implementation of HIKARI Service because the quality and volume of its information database will continue to expand. Therefore, the guidelines for security products and operation systems and the management at the server will become extremely critical. Application of industrial standard rules must be considered for these guidelines. Agencies such as Ministry of Economy, Trade and Industry are taking the initiative now to carry out the studies on information security authentication and evaluation systems.

To implement HIKARI Service, the server end (distributor side) must study the application of standards such as above. At the present time, there are no standards on security at the terminal end (general user side) that might form a pair with a server.
Malicious attacks on a server, such as unauthorized rewriting of a web page, has been a relentless threat in the past. The following table summarizes the main attack methods and relevance to the bi-directional connection and broadband communications features of HIKARI Service.

<table>
<thead>
<tr>
<th>#</th>
<th>Attack methods</th>
<th>Bi-Directional Connection</th>
<th>Broadband</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Port Scanning</td>
<td>B</td>
<td>B</td>
<td>Illegal access</td>
</tr>
<tr>
<td>2</td>
<td>Remote Stack Overflow</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Spoofing</td>
<td>B</td>
<td>B</td>
<td>Spoofing</td>
</tr>
<tr>
<td>4</td>
<td>Bucket Brigade Attack</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Social Engineering</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Sniffing</td>
<td>B</td>
<td>B</td>
<td>Eavesdropping</td>
</tr>
<tr>
<td>7</td>
<td>Denial of Service (DoS)</td>
<td>B</td>
<td>A</td>
<td>DoS</td>
</tr>
<tr>
<td>8</td>
<td>Distributed Denial of Service (DDoS)</td>
<td>A</td>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>

Relevancy: A - Large, B - Medium

As indicated in the table, Denial of Service (DoS) and Distributed Denial of Service (DDoS) especially may cause great harm to HIKARI Service. The measures for DoS and DDoS are given next.

Malicious attacks on conventional network servers including broadband network

[Classification of attacks according to malicious intent]

- Directed at a server account. Most of these attacks can be stopped with adequate measures.
- Directed at harming the operation of a server. Measures against this type of attacks for DDoS are under study even now, and there is no absolute protection at the present time. These attacks may lead to major degradation in performance even when they do not cause a service shutdown.

[Classification of attacks according to the attacking method]

See the table above.

[Issues for HIKARI Network]

DoS and DDoS attacks are critical and will harm the HIKARI Service Network because it uses bi-directional connection and broadband communications. (See the table above.)
6-1-4. DoS (Denial of Service)

DoS (Denial of Service)

[Actions]
- Upgrading a firewall and correcting the security hall in an OS/application
- User authentication, watching the number of access per user
- Separate content server and user request processing server

[What is DoS?]
Sends unauthorized data to a server for the following purposes:
- Stops server functions.
- Increases the traffic load to exceed the limit and disables service.
Examples of DoS attack methods are Smurf, SYN flood, and mail bomb.

[DoS characteristics for HIKARI Service 1: Malicious attack on broadband]
An attacker does not necessarily have to connect to a broadband network to direct conventional DoS attacks. DoS attacks can be quite potent, even when the attacker uses a network with a narrow band. When DoS is attempted on a broadband network, however, the number of possible attacks per unit time will increase, and the damage will be far more critical. There is also a type of attack that sends a vast volume of data using a wide bandwidth.

[DoS characteristics for HIKARI Service 2: Heavy-load/wide bandwidth service]
A typical application of video distribution services under review by HSAC carries a heavy load and uses a wide bandwidth. Due to this reason, DoS can be executed using a simple method such as performing multiple accesses. Moreover, since real-time transmission is very important for a live-net video distribution service (a major service of HIKARI Network Service), the damage caused by a service shutdown would be very high.

[Measures]
As long as a network depends on IP, it is crucial to adequately implement traditionally used methods as measures against server attacks, including DoS.
- Blocking unnecessary access: Block unnecessary access to a service by a tool such as a firewall.
- Avoiding a security hall: Correct the security halls of OS and applications quickly. For this task, always exercise caution regarding the information about organizations that perform security alerts. Handle the problem together with the OS manufacturer if any kind of hall is detected.
In addition to the above, the following measures that focus on the characteristics of HIKARI Service also seem appropriate.
- Monitoring the number of accesses: Provides restrictions on the number of accesses and volume of transfer data allowed to a server per unit time for each user. Denies unreasonably vast volumes of access.
- Separation of content distribution server and authentication server: Disables the connection from a general external machine to a server that distributes content. This measure will improve safety and at the same time focus on protecting an authentication server.
DDoS (Distributed Denial of Service)

**[Actions]**
- Servers
  - Access log administration for user access tracking and recording
- Clients
  - Firewall installation
  - Introduction of easy way to setup for home users
  - Installing ingress/egress filter by tie-up with ISP

**[What is DDoS?]**
Distributes an attack program to a number of terminals and executes a DoS attack to a server all at once. The administrators of the terminals that are contaminated with an attack program may not be aware of the state of the terminals. Normally, it is impossible to track down such an attacker.

**[DDoS characteristics for HIKARI Service]**
When end users are connected to a network for bi-directional video communication, there is a possibility of these machines being used as so-called “zombie” units for DDoS attack. Since the knowledge and consciousness levels of end users with regard to security are generally lower than that of the server administrators, the scale of the damage can be expected to be higher.

**[Measures]**
- It is common knowledge that there are no effective measures against DDoS attacks in general even for a present day network. Therefore, it is necessary to manage logs to enable trace and history management.
- The measures at the client’s end are important.
  - The traditional measures such as introduction of a firewall are quite effective. However, vulnerabilities remain if you take into account the C2C service between home users, which is being reviewed by HSAC.
  - A system that allows a home user to add a patch easily when a security hall is detected will be an effective measure. However, such a system may be misused and an attack program may be implanted instead.

The installation of ingress/egress filter for an ISP is an effective measure. To install the filter, cooperation with the ISP is necessary.
6-2. Protection Against Terminal Attack

Contents of Section 6-2

6-2-1. Possibility of Terminal Attack for Each Service
6-2-2. Preventing Terminal Attack
### 6-2-1. Possibility of Terminal Attack for Each Service

#### Possibility of terminal attack for each service

<table>
<thead>
<tr>
<th>Type of service</th>
<th>Terminal attack</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B2B2C</strong></td>
<td></td>
</tr>
<tr>
<td>Stored content</td>
<td></td>
</tr>
<tr>
<td>Download Type</td>
<td>Stored content downloading (music, video, file)</td>
</tr>
<tr>
<td>Streaming Type</td>
<td>Stored content streaming Delivery (Net JukeBox, VoD, VideoClip, CM)</td>
</tr>
<tr>
<td>Live content</td>
<td></td>
</tr>
<tr>
<td>Real-time Type</td>
<td>Live content Delivery (TV program relay, Live concert, remote monitoring)</td>
</tr>
<tr>
<td>Time-shift Type</td>
<td>Time-shift delivery of live content (Replay of television broadcasting)</td>
</tr>
<tr>
<td><strong>C2C</strong></td>
<td></td>
</tr>
<tr>
<td>Uni-Directional Video Communication</td>
<td>Remote monitoring, Remote nursing</td>
</tr>
<tr>
<td>Bi-Directional Video Communication</td>
<td>TV Phone, Net culture school (conversation base), Net culture school (lecture base)</td>
</tr>
<tr>
<td>Streaming Type</td>
<td>Net TV, Net game</td>
</tr>
</tbody>
</table>

| [1*] Sending Offensive material with large capacity or harmful application into a terminal |
| [2*] Interruption in video phone or net-live delivery |
| [3*] Peeping of monitoring image |
| [4*] Illegal acquisition of information from a terminal or illegal operation of a terminal |

[n*]: See Notes

A terminal at home is conventionally connected to a network using dialup. The risk of attack on a terminal will be increased by a bi-directional connection using a broadband access network. The following terminal attacks are possible for each service indicated in the table.

1. Sending a large volume of offensive material and harmful applications to a terminal
2. Interrupting to a video phone or net-live stream
3. Peeping of monitoring image
4. Unauthorized operation of a terminal or unauthorized acquisition of information from a terminal

The details of each attack and issues for measures are given in the following pages.

For the security measures of a terminal, measures are required against individual attacks as indicated in the above table. Also, the security policy for the entire system (including terminals) must be clearly defined and this policy must be seriously implemented.
If a large volume of video data, virus, or harmful application is sent to a terminal connected to a network for bi-directional video communication, an auxiliary storage may overload or the terminal operation degrades gradually.

To prevent such troubles, a method is required to identify whether the received data is requested by a user at the terminal or sent without authorization. When a file is automatically received at a terminal, IP filtering must be performed at the home gateway and the user side must block the packet data based on information such as the recipient IP address and sender IP address.

Internal IP address of a terminal must be concealed from the Internet to protect against unauthorized rewriting of an IP address; and also a measure[*1] must be provided against the source IP address spoofing interference.

(*1: For example a packet with an internal IP address is discarded at the home gateway.)

It is also necessary to record an access log at the home gateway against external port scanning. An anti-virus function is also necessary at the terminal side.
It is possible to receive an unauthorized connection or interruption while receiving a live video or bi-directional video communications at the terminal.

To prevent such attacks, proper measures are necessary. For example, an IP address of a participant’s terminal should be checked after communication is started by an application and any other data from a remote sender of IP should be denied. Or, IP address filtering should be performed at the home gateway and the user must block the packet data based on information such as recipient IP address and sender IP address.

The local private IP address of a terminal must be concealed from the Internet to protect against unauthorized rewriting of an IP address; and also a measure[*1] must be provided against the source IP address spoofing interference.

(*1: For example, a packet with an internal IP address is discarded at a home gateway.)
6-2-2. Preventing from Terminal Attack (3/4)

Implementation of video monitoring, such as security surveillance by a continuous online status, is expected for HIKARI Service. We believe that video information peeping will occur in such services.

To prevent such infiltration, a person attempting a connection must be verified. The IP of a sender must be checked at the home gateway of the terminal and connection should be denied if the sender’s IP cannot be identified. Also, data should be encrypted.

A private IP address used locally must be concealed from the Internet, and also measure[*1] against the source IP address spoofing interference must be provided as measures against peeping.

(*1: For example, a packet with an internal IP address is discarded at the home gateway.)
6-2-2. Preventing from Terminal Attack (4/4)

[4] Illegal acquisition of information from a terminal or illegal operation of a terminal

One example of unauthorized operation of a terminal and unauthorized acquisition of terminal internal information is theft of electronic money. This illegal act is not directly related to HIKARI Service. However, the chances of using an IC card and electronic money via a terminal can be expected to increase with more pay viewing and electronic transactions. Therefore, this risk will also increase.

The possible measures against theft of electronic money are: (a) a system that always performs user authentication using a physical device at the terminal even when it is not connected to a network (service operation by inserting an IC card into a PC) and (b) use of a board dedicated to protection against tampering. As a safeguard for access to electronic money, it is possible to carry out the service operation only when a user inserts a physical key each time he or she accesses the electronic money account.
7. Technological Issues for Copyright

Contents of Chapter 7

7-1. Copyright Protection-related Trends
7-2. Illegal Actions and Protection Measures Related to Content Distribution
7-3. Element Technologies and Combining These Technologies to Prevent Illegal Actions
7-4. Issues Derived from Cache/Mirror

One of the principal aims of this HIKARI Service Architecture Consortium is to promote distribution of information content by clarifying the HIKARI Service Platform architecture and its technological conditions. To achieve this objective, it is necessary for a content holder to feel secure about using HIKARI Service Platform. Therefore, copyright protection is an important issue. In this chapter, we have compiled the unauthorized actions related to copyright and the element technologies used for preventing such actions. We also extracted the issues for the mirror/cache.
7-1. Copyright Protection-related Trends

<table>
<thead>
<tr>
<th>Name</th>
<th>Trends</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Element technologies</strong></td>
<td></td>
</tr>
<tr>
<td>Electronic watermark</td>
<td>- Used for checking the validity of a content and embedding CCI. Examples of functions, characteristics, application examples, application fields, positioning in cIDf/CPTWG, and systems</td>
</tr>
<tr>
<td>Encryption</td>
<td>- Used for concealing (protecting) information, individual authentication, completeness of data, and prevention against rejection.</td>
</tr>
<tr>
<td>Others</td>
<td>Technologies used for privacy protection, access control, infringement detection, etc.</td>
</tr>
<tr>
<td><strong>Organizations for standardization and protection system</strong></td>
<td></td>
</tr>
<tr>
<td>SDMI</td>
<td>Copyright protection standards for music distribution</td>
</tr>
<tr>
<td>4C</td>
<td>CPRM • Performs content encryption/decryption using a SD card and DVD-RAM.</td>
</tr>
<tr>
<td>CPTWG</td>
<td>DTCP (SC) • Copyright protection of images on a bus such as IEEE1394. Mainly focuses on AKE, encryption, CCI, and system update.</td>
</tr>
<tr>
<td>Others</td>
<td>cIDf (content ID implementation, MPEG (planning a policy for MPEG-7 and MPEG-21)</td>
</tr>
</tbody>
</table>

First, we have studied the trends related to copyright protection. Next, we put the results of this study into two categories: a) element technology and b) organization for a standardization/protection system, and then summarized these categories in the above table.

- **Element technologies**
  - The “Electronic watermark” technology embeds access conditions/validity verification and copyright information in a content to disable unauthorized access or identify a person who makes an unauthorized copying attempt. The advantage of a watermark is that it does not require a special system or process except at the time of embedding information and detection, and removal of the watermark is difficult. The watermark is used for content that is of high value and are accessed and used over a long period of time. For example, a watermark is applied to movies and music, art images, e-books, and evidence photographs.
  - The “Encryption” technology mainly has two functions, “concealing” and “authentication”. The safety and reliability of a network society can be implemented by: a) protecting digitalized information using a “concealing” function and b) adding trust options, such as personal authentication and completeness of data, to the information.
  - Besides the above are technologies such as “privacy protection”, “access control”, and “infringement detection” that monitor a communication line to detect unauthorized access to a network.

- **Organizations for standardization and protection systems**
  - Secure Digital Music Initiative (SDMI) is a project jointly established by the Recording Industry Association of America (RIAA) and five major music industry corporations in U.S.A. SDMI has designed a standardized format with specifications for regulating illegal copies.
  - The 4C Group promotes Content Protection for the Recordable Media (CPRM) technology for copyright protection related to storage devices.
  - Copy Protection Technical Working Group (CPTWG) is an international organization that advocates a copy protection technology involving movie production companies and major electric appliance manufacturers. The copy protection technology promoted by the group for digital data is Digital Transmission Content Protection (DTCP).
  - Among others is the Moving Picture Experts Group (MPEG) which is working on: a) distribution of royalty and billing by assigning a content ID, b) digitalization of content in danger of unauthorized access, and c) studies into technologies and architecture required for distributing data in a network.
7-2. Illegal Actions and Protection Measures Related to Content Distribution (1/3)

The above figure shows a complete system underlining copyright protection for a B2B2C video delivering service (mentioned in Chapter 2). This figure clarifies the copyright problems and role of the players in each phase.

In the figure, the arrow indicates the flow of video (content), signal (processing such as authentication), and money for both distribution system players and copyright-related players. The portion with the shaded background is HIKARI Service Platform. We assume that illegal actions against copyright of three types of distribution services (download, stream, and live) are overwhelmingly higher after the content is delivered to the users.

In the user terminal environment at home, two or more PCs and home information machines are more likely to be connected by LAN, WAN, or the wireless network. It is also possible that content may be distributed from a home to the outside on physical media or over the Internet. The following measures must be provided against illegal copies and illegal distribution of content. The measures against illegal copies of content include stopping playback by a tamper resistant module in the Content Player. The measures against illegal distribution of content include tracing of illegal use and watermark detection by the illegal monitor department.

According to the concept of the super-distribution system, when a copy of specific content is handed over to the user, playback is enabled when this user officially acquires/purchases the license for the content.

We believe that the flow of the system will change in accordance with the billing and license acquiring method. For example, if an advertisement is embedded in the content, a user may be exempted from payment of the viewing charge or it may be reduced.

Note, however, that the billing method basically will not change because of the type of service (because the storage and stream content distribution services are most likely to be a pay-per-view or fixed rate system).

The numbers [1] to [10] and items (1) and (2) in the figure are described on the next page.
Comments on Previous Page’s Index number

<table>
<thead>
<tr>
<th>Index</th>
<th>Description</th>
<th>Example of an illegal action and defending against it.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Content registration/authentication</td>
<td>(1) Illegal copy&lt;br&gt;(2) Trace of Illegal use</td>
</tr>
<tr>
<td>2</td>
<td>User authentication and device authentication&lt;br&gt;Problem: Spoofing/User property alteration</td>
<td>Note: There exists the model that the users have to pay no charge or less charge by embedding the advertisement.</td>
</tr>
<tr>
<td>3</td>
<td>Delivery request by User</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>User’s License Permission</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Delivery of content to end users&lt;br&gt;Problem: Denial of received content/Request to resend</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Content storage&lt;br&gt;Problem: Alteration of content usage rights (ex. copy count, use (play) count, and all), Interpolation of content</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Content replay&lt;br&gt;Problem: Pick up original data and re-distribute.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Billing and payment&lt;br&gt;Problem: Alteration of charge/Denial of payment</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Content copy</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>License obtaining and content usage</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>All&lt;br&gt;Problem: DoS, Tapping by outsider</td>
<td></td>
</tr>
</tbody>
</table>

The numbers indicated in the figure on the previous page represents each phase in the content distribution flow. The phases in which an illegal action may occur are suggested with the associated problem above.

As a result of our studies, the illegal actions and methods of attacks on a server or at the user end seem to be the same in both HIKARI Service Platform and a band with a narrow bandwidth.

Note, however, that the chances of illegal actions at a mirror server, cache, or IDC will increase since a broadband network creates an easier environment for video distribution.
The illegal actions involving this phase are ‘spoofing’ and ‘user property alteration’. For example, a user authorized to view a desired content pretends to be a user who has authorization to access sports-related content and subsequently downloads the latter information. User/equipment authentication by means of certificates can be performed as a measure to this problem.

Similar spoofing as in [2] also occurs in phases [3] and [4].

The illegal actions involving this phase are ‘denial of received content’ and ‘request to resend’. To be more specific, a user who has received delivered content successfully wrongfully notifies a server that the reception was unsuccessful and requests the content to be resent. By saving the received content in another dedicated terminal, the user can acquire two copies of the content. The measure to this problem includes confirmation of content delivery and recording of a viewing log.

The illegal action involving this phase is ‘metadata alteration’. The information associated with a specific content (for example, number of viewings allowed and number of copies that can be made) is normally written in a header. If such information is rewritten, a user can view a content that is restricted to one viewing only, as many times as they want. Also, a user may be able to make a copy of content when copying is not allowed. The measures to this problem include: a) encryption of a header using encryption technology, b) confirmation of validity of a header by means of the digital signature technology, and c) embedding of metadata into a content using digital watermarks.

The illegal actions involving this phase are ‘illegal capturing of original content and re-delivering by unauthorized user’. When an authorized user views a desired content, the decrypted content is distributed to this user’s monitor. This illegal action takes place during the transmission process. To be more specific, an unauthorized user captures a decrypted content and re-delivers it. The measure to this problem includes using a tamper resistant decryption module to prevent content from being illegally captured during the delivery process.

The illegal action involving this phase is ‘alteration of charge’. A user tampers with the viewing log to alter his or her own viewing records, and either reduces or deletes the viewing charges even after the content have been actually viewed. The measure to this problem includes: a) method that denies access to a viewing log using an encrypted protocol, and b) method that disables tampering by providing tamper resistance for equipment that stores the viewing logs.

The illegal actions involving this phase are ‘copy without rights and secondary use’. For example, when a content data is a moving picture, a user might attempt to cut one scene from this movie and use it as a still picture. As the measure to this problem, an electronic watermark can be embedded in the content.

**General problems:** Illegal actions in general involve ‘tapping of information on a network’. For example, during user authentication or while content is being delivered to a user, another user illegally obtains information such as the content data and decryption keys distributed over a network. As the measure to this problem, data encryption is performed when a content or decryption key is distributed.

**Problems specific to a PC:** ‘Alteration of viewer conditions’ and ‘tapping the memory and busses’ are some of the illegal actions involving a PC. For example, the conditions for viewing the content is altered so that the user can save content that are not allowed to be stored. In the case of latter, an unauthorized user appropriates and saves content while the decrypted content data is distributed to a memory or bus. Two measures for these problems are: a) a method that always encrypts content data except when a user is viewing the content and b) a method that utilizes a dedicated board. Method b) encrypts all data input/output to/from this board, prevents the decrypted data from being distributed to the outside, and provides tampering resistance for the board.
7-3. Element Technologies and Combining These Technologies to Prevent Illegal Actions (1/2)

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Explanation</th>
<th>Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encryption</td>
<td>Content encryption</td>
<td>Allow only compliant equipment to replay content by sending decryption keys to them.</td>
</tr>
<tr>
<td>Authentication</td>
<td>Confirm identity and legality of users and their equipment</td>
<td>Send content and decryption keys to only compliant equipment.</td>
</tr>
<tr>
<td>Digital Watermarking (WM)</td>
<td>Embed copy-control and legality information into content</td>
<td>Detect illegal content and detect illegal content to prohibit its replay.</td>
</tr>
<tr>
<td></td>
<td>Embed copyright information into content</td>
<td>Search networks for illegal content and identify the person who distributed it.</td>
</tr>
<tr>
<td>Digital signature</td>
<td>Prevent modification of (meta) data</td>
<td>Prohibit modification of copyright information written as meta data in content header.</td>
</tr>
<tr>
<td>Tamper resistant module</td>
<td>Prevent invasion into and modification of equipment</td>
<td>Protect decrypted content and decryption keys and impose content use by compliant equipment.</td>
</tr>
</tbody>
</table>

‘Authentication’ is used to prevent illegal actions involving the operation of unauthorized equipment. The authentication process verifies whether the equipment in use complies with the required specifications. Authentication also is used to block content from being distributed to a user who has not paid the charges. ‘Encryption’ makes it possible to allow only a user who has authorized equipment to view the content by transmitting a decryption key only to such equipment or user. Also, various DRM apply a billing management method. When this billing management method is applied, encrypted content is distributed to a user in advance, and the decryption key is sent to the user in exchange of a bill or a purchase contract for the content.

‘Watermark’ prevents illegal action after an authorized user purchases a specific content. In this method, either copy control information is embedded in the content data to restrict copying or validation information is embedded in the data to detect unofficial content and stop its playback. Normally, the detailed copyright information or access information is added to the content using a header description instead of a watermark. The problem is tampering of such data. To prevent such illegal action, an encryption technology such as digital signature must be attached to prevent or detect tampering. The transfer of copyright information data using a header and its protection measures has been already developed for commercial use in various DRM.

As a final measure, a tamper resistant module is used to protect all preventive functions against illegal actions (mentioned above) from external attacks.
In the distribution system of a Content Provider, metadata such as access conditions is attached to the content in the form of a header. When metadata is added to the content, a digital signature may be used in order to prevent tampering of this metadata. Also, validity information and copyright information are embedded in the content using an electronic watermark. Next, authentication is performed for a user (who purchases a content) or relevant content player to verify the content distribution source/destination, and then the content is sent to the user after it is encrypted. The content player decrypts the content, verifies the validity based on the metadata defined in a header or electronic watermark, and then performs viewing and copying operations according to the access conditions. The access conditions are attached to the content in the form of metadata or watermark.

Sometimes, a key required for decrypting a content data is acquired from a License Service Provider. Billing management is performed by getting the user to accept a content purchase contract and settling the charges for the delivery of this key. These Content Player functions are protected using a tampering resistant module.

The above functions for preventing illegal actions can be violated by certain methods in some cases (for example, unscrambling of encrypted data). Also, in some cases, the abovementioned function against illegal copying is not fully implemented due to cost- or convenience-related reasons. (For example, a content player is implemented in PC software, which provides poor tampering resistance.) Furthermore, there are cases where a monitor output can be captured. As a result, illegal copies and pirated editions can be made and distributed.

The Illegal Copy Surveillance System can detect a pirated edition by means of an electronic watermark and identify an unauthorized user to stop illegal distribution. The Content Player can detect a pirated edition by verifying its validity and stop playback to prevent use of a pirated edition.
7-4. Issues Derived from Cache/Mirror (1/2)

1. The model without trans-coding → No re-encryption
- The cache/mirror servers need at least the equivalent security protection level for the original distribution servers even in the case that the distributed content is already encrypted and water-marked.
- Caching/mirroring may not be permitted for valuable content.

The objective of the HIKARI Service (or broadband system) architecture is the distribution of large volumes of data including moving pictures. To distribute these contents efficiently, a Distribution Provider is likely to use architecture consisting of mirrors and caches in multiple phases. It is presumed that all contents data at the Distribution Provider are encrypted for copyright protection.

First, we study a case where the type (download, unicast streaming, multicast streaming, live streaming, etc.) and format (MEPG-2, MEPG-4, etc.) of delivery are not converted in mirror/cache servers.

In such a case, the level of protection equal to that of a video content delivery server (source server) is good enough. For high-value content, however, a cache must be disabled in order to prevent data leaks from it. The mirror/cache servers are classified based on their use (storage period of mirror/cache) into the following three types.

(1) Temporary storage servers (cache server)
- The case that the CDN/cache servers preserve the distributed content temporarily without the Content Provider's plan.
- The possibility of stealing the cached data from the cache/mirror servers is smaller than that from original servers. Of course, if there are malicious applications on the cache servers, they may steal the cached content.
- If a program for malicious intent is installed in a cache, the data saved in this cache could be read by this program.

(2) Delivery/distribution servers
- The case that the CDN/mirror servers preserve the distributed content suitable to the Content Provider's plan.
- The lifetime of preservation is longer than that in the case (1).
- → The possibility of stealing the cached data is larger than that in the case (1).

(3) Network Storage servers for personal use
- Case where a user stores personal content in the network storage server.
- Only the user that has placed the personal content in the network storage has the qualification to access it.
- If the data stored in a network server is encrypted or watermarked, the risk factor of copyright infringement is basically the same as when this data is stored at home.
- Note, however, that the period over which data is stored in a public server and thus exposed to outsiders is quite long, and therefore the chances of this data being appropriated increases proportionally.

The objective of the HIKARI Service (or broadband system) architecture is the distribution of large volumes of data including moving pictures. To distribute these contents efficiently, a Distribution Provider is likely to use architecture consisting of mirrors and caches in multiple phases. It is presumed that all contents data at the Distribution Provider are encrypted for copyright protection.
2. When the delivery type and format conversion are performed:

→ Re-encryption is required

<table>
<thead>
<tr>
<th>Problem</th>
<th>On the trans-coding server, decrypted content exists temporarily.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solutions</td>
<td>1. Use the special hardware devices: The Content Provider prepares the special devices that trans-code the encrypted content without exposing decrypted content on the server.</td>
</tr>
<tr>
<td></td>
<td>2. Use the contract: Let the distribution and delivery providers agree the contract that they are responsible for all damages by leak of decrypted content.</td>
</tr>
<tr>
<td></td>
<td>3. Adopt the model without trans-coding: The Content Provider prepares all types of content which the delivery and distribution providers require.</td>
</tr>
</tbody>
</table>

On the other hand, data is cached sometimes while the delivery type is changed or data format is converted. Such an operation generally takes place when the Delivery Provider attaches a value-added service to the content. When the delivery type is changed or data format is converted, the data must be re-encrypted. Therefore, unprotected raw data is temporarily saved in a mirror/cache and measures against leaking must be provided.

The type of delivery to be used may be changed as follows.

1. Content data is delivered up to the point of the streaming distribution server in a network by a download-type delivery from the Distribution Provider. Subsequent to this streaming distribution server, the delivery type is changed to streaming in order to deliver the content to a user.

2. Edge service type. Broadcasts of all channels are stored at an edge server and the content is delivered to a user who downloads the data as and when required from the server.

3. Similar type of service that stores data delivered by streaming instead of broadcasting and downloads the data on-demand.

Furthermore, the following value-added service types that carry out format conversion in a mirror/cache server may be implemented.

1. A service which converts content data distributed in the MPEG-2 format into the MPEG-4 format for use with mobile terminals.

2. A service that provides HD quality data after converting it into the SD quality for test viewing. Similarly, there may be a Delivery Provider who embeds a CM into content data at the mirror/cache server.

As these examples show, when the delivery type is changed or the format is converted in a mirror/cache server, the encrypted content data in the mirror/cache server must be decrypted, converted, and then encrypted again. For such an operation, it is impossible not to have raw data temporarily saved in a cache server. Therefore, we have studied copyright protection methods to prevent data leak from this cache server. Three such methods are outlined below.

1. A Content Provider supplies special hardware (or black box) that prevents the raw data from Distribution Provider to each Delivery Provider from leaking to the outside, and requests the players to use such hardware.

2. A contract is drawn up to make the Delivery Provider liable for data leaks occurring from a mirror/cache.

   → To assure the viability of this method, an option is also available for embedding an electronic watermark that identifies each mirror/cache server. Note, however, that this option is technologically quite difficult to implement if the watermark has to be applied to multiple phases.

3. The content of all patterns (delivery models and formats) are created and encrypted by the Distribution Provider (source server). The Delivery Provider does not perform trans-coding (mirror/cache service).
8. Technical Issue on the HIKARI Service Terminal

Contents of Chapter 8

8-1. Required Terminal Functions for HIKARI Services
8-2. Functions and Configurations of HIKARI Service Terminals
8-3. Technical Issues for Functions and Configurations of HIKARI Service Terminal
8-4. Conclusion

Section 8-1 describes the functions required for HIKARI Service (broadband net-live and bidirectional communication services).

Section 8-2 describes the category of terminals to be used in a practical situation, and explains how the service functions are implemented for each terminal type.

Section 8-3 describes the technical issues for implementing HIKARI service.
8-1. Required Terminal Functions for HIKARI Services

Contents of Section 8-1

8-1-1. Terminal Functions Required for HIKARI Service
8-1-2. Terminal Functions Required for Broadband Net-live Service
8-1-3. Terminal Functions Required for Bi-directional Video Communication Service
8-1-1. Terminal Functions Required for HIKARI Service

This chapter clarifies the technological issues related to functions and configurations of HIKARI Service terminals.

The functions reviewed in this chapter include the service selection function, display function for service selection, and handling of main signals such as video and audio signals. The billing, authentication, security functions, and copyright management function are excluded from the studies in this chapter. Billing, authentication, and security functions are very important in implementing HIKARI service, and the HIKARI Service terminal must provide the application software and required hardware for this purpose as a prerequisite. This chapter, however, focuses on HIKARI Service specific functions.

Similarly, the general Internet connection functions and functions related to the interface with a home gateway (HGW) between HIKARI Service Network and a home terminal are excluded from the studies in this chapter. HGW stands for a device that implements the functions defined by Digital Home-Network Forum in Japan. Basically, HGW provides the functions used to terminate a service between Wide Area Network (WAN) and home network and to connect and manage the terminal in a home.

This chapter studies terminal function and configuration from the perspective of implementing the two types of the characteristic HIKARI Services (broadband net-live broadcasting and bi-directional communication): (a) playback/display function of video/audio stream, (b) video and audio input function, (c) browsing function for service menu selection, and (d) content storage function.

Functions reviewed in this report:

- Menu display and service selection function
- Audio/video playback and display function
- Communication function (with content server)
  - See Chapter 9 for transmission and distribution control protocols.
- Content storage function and special playback function

Other functions:
- Billing, authentication functions
- Security functions such as secure communications
- Copyright management function
- Internet connection function (ISP connection, IP address acquisition)
- Clock setting
- Network QoS function
- Connectivity with home gateway (HGW)
8-1-2. Terminal Functions Required for Broadband Net-live Service

Basic functions
• Service selection, content selection
  - Multi-angle video selection function
• Content viewing function
  - Basic video control functions (playback, stop, pause)

Extended function 1
• Storage function (simple storage and playback)

Extended function 2
• Featured service functions provided by a service provider
  (Example: Interactive video control)
• Special playback functions
  (Example: Smooth reconstruction of multi-angle video, time-shift playback function)

To enable viewing of broadband net-live content using a terminal provided with variety of functions/performance options, the terminal functions related to content viewing are classified.

The basic functions required for viewing a broadband net-live service must include those allowing a user to select a service/content, playback video and audio, and select and view multi-angle video, regardless of the terminal resolution.

We presume that the functions classified in “Extended function 1” will allow a variety of time-shift operations on a terminal equipped with a storage function.

We expect that the functions classified in “Extended function 2” will provide more advanced video reconfiguration processing and display capacity. Extended function 2 enables a special playback function and viewing control of a video and audio link with a service provider.
Basic functions

- Selection and login of communication service (such as remote learning)
- Bi-directional communication function
  - Reception/display of video and audio of a speaker (such as a teacher/lecturer)
  - Transmission of participant video/audio and interactive commands

Extended function 1

- Download and storage of data associated with the service (such as teaching materials)
- Multi-stream video reception, selection, and display

Extended function 2

- Communication function in accordance with a service
  (Example: Interactive communication between student users and multiple viewers and precise localization of audio and video)
- Display of high-resolution document images and switching of video and high-resolution image

The functions of terminals related to content viewing are classified to enable viewing of bi-directional video communication content on a terminal provided with a variety of functions/performance options.

The basic functions for viewing a bi-directional communication service must include those that allow a user to select a service/content and perform bi-directional video communication with a teacher, regardless of the terminal resolution.

The functions classified in “Extended function 1” enable download and storage of teaching materials and reception, selection, and display of multi-stream video content.

The functions classified in “Extended function 2” provide more advanced N:N simultaneous interaction and high-precision content video display.
8-2. Functions and Configurations of HIKARI Service Terminals

Contents of Section 8-2

8-2-1. Categories of HIKARI Service Terminals
8-2-2. Functions and Configurations of Broadband Net-live Service Terminals
8-2-3. Functions and Configurations of Bi-directional Communication Service Terminals
8-2-1. Categories of HIKARI Service Terminals — Overview

The basic functions and configuration of HIKARI Service terminals are classified as follows.

- **Basic concepts**
  Terminals are classified in accordance with their service functions and video/audio playback capacity. Service functions are classified into three types, basic function and extended functions 1 and 2, as defined in Section 8-1.

  We classified the terminals based on the following video/audio playback capacities: (a) the lowest quality level of video/audio playback should be equivalent to VGA video quality and telephone audio quality; (b) the next level is TV and stereo quality; and (c) the highest quality level enables viewing of HD large screen with surround sound. A terminal with the highest quality level allows a user to enjoy the rich services provided only by HIKARI Service.

  On the other hand, a PC continues to serve as a general-purpose terminal platform covering a wide range of functions and performances in the HIKARI Service age.

  We hypothetically set up three terminal categories: Type 1, Type 2, and Type 3. Type 1 supports basic functions only and offers a video/audio quality equal to VGA and telephone. Type 2 supports functions up to the scope of extended function 1 and offers a video/audio quality equal to SDTV and stereo. Type 3 supports functions up to the scope of extended function 2 and offers a HD large screen and surround sound.
8-2-1. Categories of HIKARI Service Terminals — Types 1 to 3

1) Type 1 (HIKARI Service operation on an inexpensive and simple dedicated terminal)

<table>
<thead>
<tr>
<th>Simple dedicated terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Provides functions to receive the basic service</td>
</tr>
<tr>
<td>• Limited video display capability</td>
</tr>
<tr>
<td>• Limited web browsing functions</td>
</tr>
</tbody>
</table>

2) Type 2 (HIKARI Service operation by a combination of legacy terminals)

<table>
<thead>
<tr>
<th>Existing video terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Provides functions to receive the basic service</td>
</tr>
<tr>
<td>• Supports some of the extended function services (Extended function 1)</td>
</tr>
<tr>
<td>• Provides a Java execution environment for applications</td>
</tr>
<tr>
<td>• Limited web browsing functions</td>
</tr>
</tbody>
</table>

3) Type 3 (Full capacity for rich and high grade services provided only by HIKARI Service)

<table>
<thead>
<tr>
<th>HIKARI Service dedicated terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Provides all basic functions and extended functions and handle the highest quality video of HIKARI Service.</td>
</tr>
<tr>
<td>• Supports broadband net-live service functions that allow a user to enjoy flexible visual perspectives such as multi-angle video.</td>
</tr>
<tr>
<td>• Provides realistic, large-screen video for the bi-directional communication service, with precise audio localization.</td>
</tr>
</tbody>
</table>

The ranks of the terminal types are defined below.

Type 1 terminals are used in cases where HIKARI Service is received on a simple dedicated terminal. In this case, the terminal is not necessarily used exclusively for HIKARI Service. A PDA, for example, is included in this category. This terminal can receive only the basic functions in accordance with the conditions for functions/performance.

Type 2 terminals are used in cases where HIKARI Service is received by combination of legacy multimedia equipment and a terminal equipped with communication functions. Examples of terminals falling under this category are the combination of a set top box and TV or the combination of a game machine and TV. Although there are limitations in some of the functions, this type can handle a service equivalent to conventional video.

Type 3 terminals are used for implementing services that take full advantage of HIKARI Service, and are HIKARI Service dedicated terminals.
8-2-2. Functions and Configurations of Broadband Net-live Service Terminal

The functions required for a broadband net-live service terminal are described below. Obviously, by implementing these functions, all terminals should support “content distribution using VoD”, which is a typical HIKARI Service.

1) Menu display and select function
   This function indicates mainly the browser functions required for a terminal.

   **Type 1**: Provides basic browser functions.
   **Type 2**: Enables a Java execution environment in addition to the basic browser functions.
   **Type 3**: The latest browser functions and a Java execution environment are essential.

2) Audio/visual playback display function
   **Type 1**: Provides a VGA resolution display and telephone quality audio. Advanced display functions such as simultaneous multi-angle display are not supported.
   **Type 2**: Provides a TV resolution display that can handle high-quality video images such as 480P. Type 2 should offer stereo sound.
   **Type 3**: Provides a HD resolution display that can handle HD video images. Type 3 should support surround sound. Advanced display options such as simultaneous multi-angle display should be supported also.

3) Storage function
   The storage function provides support for a downloading service and is therefore treated as an optional function.

4) Communication function
   The communication functions are not very different between terminal types. All terminals must have the communication functions for receiving broadband net-live broadcasting services.

5) Other functions
   Other functions include functions such as PPPoE, DHCP, and NTP. Types 2 and 3 terminals, in particular, are equipped with a Java execution environment and can implement extended functions such as an original player software supplied by a service provider.
The functions required for a bi-directional communication service terminal are described below.

1) Menu display and select function

This function is same as that of the broadband net-live broadcasting service.

2) Audio/visual playback display function

**Type 1:** Offers the display capacity of MPEG-4 only, and telephone quality audio.

**Type 2:** Offers the display capacity of MPEG-4 only, but provides stereo sound.

**Type 3:** Offers a display capacity equivalent to HD resolution and can display MPEG-2 video images. Type 3 also supports surround sound. Advanced display capacity such as simultaneous display of multiple users is also provided.

3) Input/output function

Dedicated terminals such as Types 1 and 3 are equipped with a built-in input/output function. Type 2 terminals are equipped with either an external or built-in input/output function.

4) Storage function

The storage function provides a textbook downloading service and is therefore treated as an optional function.

5) Communication function

The communication functions are not very different between the terminal types, and the terminals must have the communication functions required for bi-directional video communication service.

6) Other functions

Types 2 and 3 terminals equipped with the Java execution environment can implement extended functions such as video and textbook collaboration.
Contents of Section 8-3

8-3-1. Issues for Broadband Net-live Service
8-3-2. Issues for Bi-directional Video Communication Service
(1) Multi-angle video select function
   • Simultaneously displays two or more multi-angle video images or switches these images.
   • Allows simultaneous display of the main screen and sub-screens.

(2) Storage function (See Chapter 5 for the storage function at the server side.)
   • Stores a live broadcasting stream in a terminal and allows a user to view it at any convenient time.
   • Enables a time-shift playback such as fast-forward, rewind, and jump.

(1) Multi-angle video select function
   • Focuses targets on multi-angle playback as a multi-stream application service.
   • Multiplexes multi-angle video images using multi-stream technology.
   • The video/audio between screens must be synchronized.

(2) Storage function
   • The inevitable requirement of storing data on a terminal must be studied since some services store data in a content server or edge server.
   • When data is stored in a terminal, technologies such as copyright management, restricted number of playback operations, and a date/time limit for playback are required.
8-3-1. Issues for Broadband Net-live Service — Multi-angle (1/3)

Issues for studies on the multi-angle video selection function (pre-requisition)

<table>
<thead>
<tr>
<th>Main content</th>
<th>Sub-content</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 Mbps (MPEG-2 equivalent)</td>
<td>1 Mbps (MPEG-4 equivalent)</td>
</tr>
</tbody>
</table>

Content can be switched between:
- Main content
- Main content + Multiple sub-content
- Multiple sub-content.

- This figure shows an example of a multi-angle video select function.
- This figure shows an example of the main screen and sub-screens of all angles being displayed simultaneously.
- The main screen demands DVD quality. Therefore, 6 Mbps (MPEG-2) is used.
- Since VHS quality is sufficient for the sub-screens, 1 Mbps (MPEG-4) is used.
8-3-1. Issues for Broadband Net-live Service — Multi-angle (2/3)

**Issues for study on multi-angle video select function**

(Issue 1) Multi-stream configuration
- A comparison between the method for multiplexing the main/sub-screens into a single stream and the method for dividing them into multiple streams must be performed.
- The study must take into account the terminal/server processing capacity (future trend) and synchronization method between streams.

(Issue 2) Delivering and receiving the information related to the main screen and associated sub-screens
- A standardized procedure must be provided so that the server can notify a terminal of information related to the main screen and its associated sub-screens.

(Issue 3) Synchronization between streams (for time-shift function)
- Since synchronization of the angles between different streams is not guaranteed, there must be a system of periodical synchronization to prevent an out-of-sync state.

(Issue 1) The stream configuration must be determined in accordance with the terminal processing capacity.
- Video/audio synchronization can be easily performed in the case of a single stream. However, the required processing ability may exceed the ability of a terminal if there are a large number of processing streams.

(Issue 2) Delivering and receiving the information related to the main screen and its associated sub-screens
- Regulations must be established to define URI and PID of a main screen in the MPEG-2 TS descript of a sub-screen so that the sub-screen can be clicked to use it as a main screen.

(Issue 3) Synchronization between streams
- A gradual out-of-sync may occur in the case of MPEG-2 TS over UDP, when multiple sub-screens or a main screen are sent in different streams.
- A process using the MPEG-2 time stamp for synchronization is required.
- When RTP is used, the RTP time stamp function can be used.
8-3-1. Issues for Broadband Net-live Service — Multi-angle (3/3)

Example of protocol sequence

- This is an example of the sequence used in a method that multiplexes a content into two streams. The one is for the main screen and the other is for the several sub-screens.
- The terminal and server operations are not heavy-loaded in this method. The means for synchronization between screens are necessary, however. Therefore, we have proposed a protocol based on the assumption that PID of a main screen is defined in MPEG-2 TS of a sub-screen.
8-3-1. Issues for Broadband Net-live Service — Storage Function

1) Implementation by means of a storage function in a cache server in CDN

We are assuming that the content of a broadband net-live service is basically distributed from a network when time-shift playback is performed. If the network does not provide the time-shift function, it can be implemented by storing the content at a terminal.

See the time-shift function in Chapter 5 for the server storage function.
An example of a HIKARI Service terminal configuration with a storage function for implementing the time-shift function is shown in the above figure.

As shown in this figure, there are three data stream flows for implementing the time-shift function.

1) Stream receive/playback from Ether
2) Saving on storage equipment (HDD) from Ether
3) Playback from HDD

Furthermore, display graphics such as menus are implemented by a control instruction from the CPU to a graphic controller.

Since various data shares a bus, the performance of the bus must be reviewed carefully.
8-3-1. Issues for Broadband Net-live Service — Terminal Storage Function (2/3)

**Issues for studies on HDD**

(Issue 1) The characteristics (write retry and defective cluster processing) of HDD may affect the video playback service.

(Issue 2) A file system for a simultaneous write and read operation to/from HDD is an issue to be resolved.

These issues can be resolved by installing external HDD that can provide real-time operation via IEEE1394 or USB.

Since HDD with advanced functions is unlikely to be installed in a terminal, the following HDD issues must be reviewed.

1) The characteristics (write retry and defective cluster processing) of HDD may affect the video playback service.
   - Write retry: If writing to HDD is unsuccessful, writing will be retried for several seconds. These retries will cause a buffer under-run and the stream data could be lost.
   - Problem of defective clusters: When HDD has a defective cluster, another location is allocated as a substitute cluster. Therefore, the reading speed from HDD will become inconsistent.

2) This issue is related to issue 1. A file system for simultaneous write and read operation to/from an HDD is an issue to be resolved.

These issues can be resolved by installing external HDD that can provide real-time operation via IEEE1394 or USB.
Other issues to be studied

(Issue 1) Some type of encryption system is required for copyright protection when data is stored on HDD, and a decoding function for the encryption must be implemented. However, if the same encryption system as that of the received stream can be applied, a common decoder can be used for decoding.

(Issue 2) Since a user can freely view the content by storing the data in HDD, it is necessary to realize a system for implementing DRM (restricted number of playbacks and date/time limit for a playback) and a billing process for the DRM.

The following issues must also be reviewed for installing HDD in a terminal.

1) Some type of encryption system is required for copyright protection when data is stored on HDD, and a decoding function for the encryption must be implemented. However, if the same encryption system as that of the received stream can be applied, a common decoder can be used for decoding.

2) Since a user can freely view the content by storing the data in HDD, it is necessary to realize a system for implementing DRM (restricted number of playbacks and date/time limit for a playback) and a billing process for the DRM.
1) Multi-point Control Unit (MCU) connection type
   • Enables small-scale to large-scale conference
   • In addition to small-scale to medium-scale N:N and 1:N conferences, the flexibility of a conference setting is high. 1:N large-scale lecture/curriculum can be realized.
   • Enables the connection of terminals with different capabilities.
   • Enables a large-scale conference by means of cascade connections.
   • The terminal side requires a narrow bandwidth network. The MCU side requires a wide bandwidth network.

2) P2P type
   • Suited for a conference with a comparatively small number of participants.
   • The basic service is N:N bi-directional communications such as video images of all conference participants.
   • The communication bandwidth for (N-1) times terminals is required.

There are two methods for setting up an N:N conference
1) connect terminals via the Multi-point Control Unit (MCU).
2) connect terminals mutually by P2P.

The P2P method does not require the high-function MCU and can set up a “conference” very easily. This method, however, is unsuitable for a large-scale conference.

On the other hand, the MCU connection method can be used to configure a conference with a small group to large group of people by means of cascade connections. It can also be used to connect terminals with different capabilities using the transcoding function. Therefore, this method offers high flexibility for different types of a conference.

We focus our study on N:N conferencing using MCU in this report.
8-3-2. Issues for Bi-directional Video Communication Service (2/5)

(1) MCU and N:N conference support – Functions required for MCU –

- Multiple protocol and format support
  - Communication control protocol (SIP, H.323, etc.)
  - Audio (G.711, G.722, G.723.1, G.728, G.729, AAC, etc.)
  - Video (MPEG-2, MPEG-4, etc.)
- Emphasis on a speaker: Continuous presence, voice-activated video selection
- Audio transmission/reception: AGC, mixer setting
- Multi-rate support
- Audio/video transcoding function
- Cascade connection function
- Link to directory server/service such as Gate Keeper

To provide N:N conferencing for Types 1 to 3 and PC terminals, the following functions are required:

(a) transcoding function for connecting terminals with multiple audio/video formats and various communication bandwidths,
(b) video merge function for highlighting a speaker in continuous presence,
(c) cascade connection function for configuring a large-scale conference by providing interfaces between multiple MCU.

When the high-quality MPEG-2/MPEG-4 video transmission is used, the delay amount of the audio transmission for lip synchronization increases. The video format may be selected according to the situation. For example, high-quality video is required for a remote class/lecture in a 1:N conference. On the other hand, some N:N conferences have greater emphasis on the audio exchanges and may not require high-quality video.
8-3-2. Issues for Bi-directional Video Communication Service (3/5)

(1) MCU and N:N conference support – Video images from MCU –

1) Video transmission method

- **“Composite video images”**
  - MCU transmits data as a single video stream.
  - Note: Enabled on Type 1 to 3 terminals and PC since there is only one receiving stream.

- **“Individual video images”**
  - MCU transmits data as multiple video streams.
  - Note: Enabled on Type 3 terminals and PC since there are multiple receiving streams. Also, the communication bandwidth required in the downstream is N times (or N−1 times).

2) Emphasis of a speaker

- **Voice Activate**
  - Displays an image of the person speaking only.

- **Continuous Presence**
  - Displays conference participants together with a speaker using “composite video images”.
  - Note: There is also a system that switches a display by fixing the display position of all or some of the conference participants (for a large number of conference participants) or by changing with “Voice Activate”.

There are two methods to send video from an MCU.

1) MCU merges the video received from each terminal and sends it out as a single video stream.
2) MCU sends the video received from each terminal to each terminal (transcoding is performed as required).

This method is free from quality degradation and processing delays associated with the video merging by MCU. The multi-streaming used in this method, however, requires a substantial increase of the communication bandwidth.

When a “composite video” is sent from an MCU, all types of terminals including Types 1 to 3 and PC can receive the video. Only Type 3 and PC terminals that can receive multi-streams can be connected for “individual video”.

There are three methods to emphasize a person speaking.

1) Voice Activate method.
2) Continuous Presence method.
3) a combination of these two methods.

These method are implemented in MCU and terminals require only the capability to receive a single stream.
(2) Issues – Asymmetrical communication –

1) “Teacher” and “Student” screens

“Teacher” screen
Displays “Student” video image.

“Student” screens
For example, “Teacher” screen of video image is displayed in large size and the other “Student” images are displayed in a smaller size in Mr. A’s screen.

2) Conversation in “Another room”

Communication as a whole
All participants can hear the speech.

Conversation in “Another room”
Mr. A, teacher has a 1:1 conversation with Ms. B in “Another room”.

We consider an example of an N:N conference where a single teacher lectures multiple students in remote locations.

An N:N conference for distance learning may be implemented in the following way. The images of all or some of the students attending the distance learning are displayed on the teacher’s terminal. On each student’s terminal, the image of the teacher, who will obviously speak most of the time, is displayed in a window larger than those used for images of other students.

It is possible to connect the Types 1 to 3 terminals by creating a composite video at an MCU. If the MCU distributes multi-stream video, only Types 2 or 3 can be available.

Generally, all participants listen to the same speech in an N:N conference. It is also possible for a participant to communicate with other specified/selected participants by creating “another room”. The “another room” function will complicate processing of the composite video at the MCU. Therefore, it is more realistic to use this function by the multi-streams method at MCU and to use Type 2 or higher level terminals.
8-3-2. Issues for Bi-directional Video Communication Service (5/5)

(2) Issues – Implementation of service AP(Application) –

1) Example of service Application

   (1) Web page sharing
   (2) Whiteboard
   (3) Chat (text base)
   (4) Application sharing (including sharing of documents)
   (5) AV content playback sharing

2) Example of Application execution

There are many examples of combining the service applications such as document video sharing with video/audio talk function with a bi-directional communication service.

Service AP provides: (1) web page sharing, (2) whiteboard, (3) chat (text base), (4) application sharing (including sharing of documents), and (5) AV content playback sharing.

There are two methods to implement Service AP.

(A) starting a program at the teacher and student terminals and sending the status of the teacher terminal to the student terminals as an event information

   ((1), (3), (4), (5));

(B) sending the screen image of the program execution on the teacher terminal to the student terminals as video information ((2), (4)).

Method (A) requires less information transmission between terminals. However, a video display timing control needs to be devised for each terminal. System (B), on the other hand, does not require service AP at the student terminals and therefore these terminals can have a low processing capability. It must be noted that student terminals may not be able to receive sufficient video quality due to the capability of a teacher’s terminal and the quality of the communication line.
Contents of Section 8-4

8-4-1. Summary
8-4-2. Future Issues
8-4-1. Matters Revealed in This Section

We have defined the requirements of a terminal that can be used for HIKARI Service. We also have clarified the central technical issues to be solved by clarifying the functions and configurations of HIKARI Service terminal.

- Terminal classification based on the expected service level (Types 1 to 3).
- Classifying the terminal functions and configurations on the basis of the broadband net-live service and bi-directional communication service.
- Extracting issues for terminal functions including the roles shared with HIKARI Service Platform.
8-4-2. Future Issues

Future issues are summarized below.

- Clarification of requirements and functions for a variety of HIKARI Services.
- Studies on terminal functions and configurations considering easiness of realization and popularization. (Considerations for terminal cost and sharing with other services.)
- Clarification of terminal functions and configurations including higher application layers such as billing and copyright which are necessary for the total implementation of a service.
- Clarification of roles of terminals shared with the HIKARI Service Platform, home network, and home gateway (HGW).
9. Technical Issues on Protocols and Content Formats

Contents of Chapter 9

9-1. Target HIKARI Services and Protocols
9-2. Protocols for Menu Display
9-3. Audio/Video Stream Formats
9-4. Media Data Transfer Protocols
9-5. Stream Delivery Control Protocol
9-6. Communication Connection Protocols
9-7. Conclusion of Protocols and Formats
9-1. Target HIKARI Services and Protocols

Contents of Section 9-1

9-1-1. Protocols for Live/VoD Services
9-1-2. Protocols for Bi-directional Video Communication Services
9-1-1. Protocols for Live/VoD Services (1/3)

Example of stream delivery using the unicast system

We have selected the live/VoD and bi-directional communication services of HIKARI Service as examples in this section.

The following three systems are used for the live/VoD service.

Sequence of stream delivery using the unicast system (VoD)
(1) The user terminal/browser accesses a portal server and receives a menu.
(2) The user terminal/browser starts up the video player.
(3) The stream delivery is controlled using RTSP between the user/video player and distribution server.
(4) The distribution server uses RTP or UDP to send a stream to the user/video player.

The stream is sent from distribution server to user terminal using 1:1 IP unicast.

The menu display data transfer protocols, stream delivery control protocols, and media data transfer protocols are explained in Sections 9-2, 9-5, and 9-4, respectively.

The page description and script languages that are transferred using menu display data transfer protocols are explained in detail in Section 9-2.

The audio/video formats of data transferred by media data transfer protocols are explained in detail in Section 9-3.
9-1-1. Protocols for Live/VoD Services (2/3)

Example of stream delivery using the IP multicast system

Sequence of stream delivery by IP multicast system

1. The user terminal/browser accesses a portal server and receives a menu.
2. The user terminal/browser starts up the video player.
3. The distribution server uses RTP or UDP to send a stream to the user/video player according to the IP multicast address and port number (and distribution source IP address) notified by the portal server.

The stream is sent from distribution server to user terminals using 1:N IP multicast.

The menu display data transfer protocols and media data transfer protocols are explained in Sections 9-2 and 9-4, respectively.

The page description and script languages that are transferred using menu display data transfer protocols are explained in detail in Section 9-2.

The audio/video formats of data transferred by media data transfer protocols are explained in detail in Section 9-3.
Example of stream delivery using a splitter server

When the target network is not capable of IP multicast delivery

1. The user terminal/browser accesses a portal server and receives a menu.
2. The user terminal/browser starts up the video player.
3. Stream delivery is controlled using RTSP between the user/video playback player and splitter server.
4. The splitter server uses RTP or UDP to send a stream to the user/video player.

Multicast (or unicast) is used to send the stream between a distribution server and splitter server, and 1:1 IP unicast is used between the splitter server and user terminal.

The menu display data transfer protocols, stream delivery control protocols, and media data transfer protocols are explained in Sections 9-2, 9-5, and 9-4, respectively.

The page description and script languages that are transferred using menu display data transfer protocols are explained in detail in Section 9-2. The audio/video formats of data transferred by media data transfer protocols are explained in detail in Section 9-3.
Another service we selected as an example of HIKARI Service is bi-directional video communications.

The Session Initiation Protocol (SIP: RFC 2543) or H.323 are used as control protocols (communication connection protocol) for bi-directional video communications. However, RTP is used as the transfer protocol for media data.

This document does not cover directory access protocols.

The communication connection protocols and media data transfer protocols are explained in detail in Sections 9-6 and 9-4, respectively.

The audio/video formats of data transferred by media data transfer protocols are explained in details in Section 9-3.
9-2. Protocols for Menu Display

Contents of Section 9-2

9-2-1. Menu Display Data Transfer Protocols
9-2-2. Page Description Language
9-2-3. Script Language
9-2-4. Security on Menu Display
9-2-5. Starting Method of Video Player
9-2-6. Conclusion on Menu Display
9-2-1. Menu Display Data Transfer Protocols (1/5)

The HyperText Transfer Protocol (HTTP) is used as the menu display data transfer protocol. Generally, HTTP Basic authentication is used to restrict access to menus. The Secure Socket Layer (SSL), however, is used for data that requires high security (see issues discussed later in this section and Section 9-2-4).

- **HTTP/1.0**
  - The basic specification of HTTP is simple and stateless. The server only responds to a request from a user. The following specifications, however, have been provided in addition to this basic specification. As a result, HTTP has surpassed its original purpose of “transferring of hyper text”, and can be used for various types of data transfers including audio and video. This improvement is one of the factors that have accelerated the widespread use of the web.
    - Expandable request method
    - Expandable response code (status code)
    - Expandable header

- **HTTP/1.1**
  - The specifications for improving the web performance and problems of HTTP/1.0 have been corrected and added. The main specifications that were added to HTTP/1.1 are listed below.
    - Persistent Connections
    - Content Negotiation
    - Caching in HTTP
    - Virtual Hosts

Besides having the extra functions added to HTTP/1.0, the HTTP/1.1 specifications have been more rigidly defined in order to avoid communication troubles caused by implementation ambiguities in the previous version. HTTP/1.1-compliant servers and user terminals must be able to interpret a request and response sent from HTTP/1.0-compliant user terminals and servers.

Our analysis regarding the implementation conditions of HTTP/1.1 for commercial applications (PC, integration) at the present time confirms that it has been adequately put to practical use. Therefore, it is recommended that a terminal provide support for this version. We believe, however, that HTTP/1.0 support should be also provided. For example, a dedicated terminal for various HIKARI Service applications requires HTTP/1.0 for a light load implementation.

See the following specifications for details about each HTTP version.

- **HTTP/1.0** : RFC 1945
- **HTTP/1.1** : RFC 2616

Basic authentication is defined in RFC 1945 and RFC 2617 (HTTP Authentication) together with Digest authentication, which is a more secure authentication system. For the study, we will discuss a case that demands higher security measures against disclosure.
9-2-1. Menu Display Data Transfer Protocols (2/5)

The basic sequence of menu display data transfer using HTTP is shown in the figure below.

Outline of menu display data transfer using HTTP

(1) Connection request
A TCP connection is established when a user (terminal) specifies a host name and port number and requests a connection to a menu server (such as a service portal), and the server accepts this request.

(2) Connection response
The server returns a response to the user.

(3) Menu request
The user specifies a menu in a HTTP request message to be sent to the server. The GET method is used to specify the URL of the required menu. The server can be notified of additional information (user information, ...) using a header field. A host header that is used to specify a virtual host is compulsory for a HTTP/1.1 request.

(4) Menu transmission
The server locates the menu requested by a user based on an URL specified in the HTTP request message, and sends it as a HTTP response message. In some cases, a menu is dynamically created using CGI or ASP.

(5) Disconnection
A user who receives the server response for the request closes the TCP connection established between user and server. Note, however, that if a continuous connection is established, multiple HTTP requests/responses can be exchanged in a single TCP connection.
9-2-1. Menu Display Data Transfer Protocols (3/5)

The menu display data transfer sequence including basic authentication is shown in the figure below.

The sequence figure above skips the TCP/IP establishment and disconnection phases. Two sets of requests and responses are included in the above sequence. This example does not care whether these requests/responses are carried out in the same or separate TCP/IP sessions.

(1) Menu request
A user sends a HTTP request message to a server and requests a menu.

(2) Response and authentication request
The server specifies the following items in a HTTP response message to be returned to the user, to notify the user about the need for authentication.
   Status code 401 (Unauthorized)
   Header field WWW-Authenticate:Basic realm="Menu URL"

(3) Menu re-request
The user sends a new request message to the server and transmits the user authentication information using the following header field included in this request message.
   Header field Authorization:Basic ID:Password
   “ID:Password” portion is encoded using BASE64.

(4) Menu transmission
The server will verify the user authentication information sent from the user after BASE64 decoding. When the user is authorized for access, the server will send the menu data according to the request.
9-2-1. Menu Display Data Transfer Protocols (4/5)

Study issues for the application of HTTP

- Compatibility between HTTP/1.0 and HTTP/1.1
  A request method, status code, and header field have added to HTTP/1.1 to expand its functions. Conversely, some of the HTTP/1.0 functions have been discarded. Basically, compatibility between these versions is required so that HTTP/1.1 servers and user terminals can interpret messages sent from HTTP/1.0 servers/user terminals. However, there are specification-related issues in HTTP/1.0 that should be eliminated before defining this protocol.

- Should cookies be supported?
  This technology has been developed to add a status maintenance function to HTTP (implemented without status codes until now). Cookies are mainly used for the following applications and should be supported.
  - Session management for server applications using CGI
    Although the function for setting and retrieving user information is the same, a cookie is used temporarily during a series of communications (normally called ‘session’) with the user and does not have to be saved permanently on storage media.
  - Setting and retrieving of user information
    The server sets the user information including expiration of service for a user terminal. This user information is retrieved when the user connects to the network on subsequent occasions. The user must retain the cookie on storage media for the period of time specified by the server.

One of the problems of HTTP/1.0 is the support for continuous connection. The HTTP/1.0 specification for continuous connection is in an experimental phase and is not supported by many hosts. For example, suppose there is a server that does not support continuous connection and is responding to a continuous connection request message sent from a user. This continuous connection between the user and server is made via the HTTP/1.0 proxy server. The proxy expects the connection to be disconnected after a one-time request and response exchange. On the other hand, the server and user terminal do not disconnect the line, and both connections will be unsuccessful.

The issues related to a cookie are outside of HTTP specifications. The cookie, however, should be included among essential user functions in order to support the various types of services provided by a server. At the implementation level, if a cookie is used just for a server application using CGI (as mentioned above), the only requirement is whether temporary processing can be performed. On the other hand, if the user information must be stored for a specified period, a function for retaining cookies sent from multiple servers is required. Therefore, the implementation level required by a terminal must be reviewed.
9-2-1. Menu Display Data Transfer Protocols (5/5)

Study issues (continued from previous page)

- About security
  When the Basic authentication is used on HTTP, the user authentication information (user ID, password) is encoded in base64. However, base64 itself is defined in RFC 2045 (MIME), which enables a third party to decode the data easily. Thus, there is a safety-related problem unless the Basic authentication is used on a communication channel encrypted by SSL. This problem also applies to cases where: a) a user’s privacy information is exchanged between the user terminal and server using a cookie and b) the menu data customized for each user’s preference is sent to the user from the server. Therefore, it is necessary to define that various data transfers related to menu displays must be carried out after establishing a secure communication channel between the user and server using encryption such as SSL. (Details about SSL are given in Section 9-2-4.)
We recommend using HyperText Markup Language (HTML) for menu page description.

- **HTML3.2**
  HTML3.2 is an older version that contains many tags and attributes related to the appearance of a document. HTML3.2 cannot be excluded because of the large percentage of application integrated with this version at the present time and also because of its implementation on browsers designed for dedicated terminals.

- **HTML4.0 (4.01) + Cascading Style Sheets (CSS)**
  This version isolates the tags and attributes of document appearance from HTML and specifies them in a style sheet instead. W3C recommends this HTML version. There are several types of style sheets. However, CSS is the standard for HTML. In this report, we have selected CSS1 and CSS2 for our applications.

- **DHTML**
  The DHTML version of HTML corresponds to the “HTML4.0 + CSS” application indicated above. DHTML enables dynamic display of a page by controlling the HTML tags or style sheet using a script language.

See the following recommendations on the sites of the original designers (W3C) for further information about the above specifications.

- HTML3.2 : [http://www.w3.org/TR/REC-html32](http://www.w3.org/TR/REC-html32)
- HTML4.0 (4.01) : [http://www.w3.org/TR/html4](http://www.w3.org/TR/html4)
- Cascading Style Sheets, level1 (CSS1): [http://www.w3.org/TR/REC-CSS1](http://www.w3.org/TR/REC-CSS1)
- Cascading Style Sheets, level2 (CSS2): [http://www.w3.org/TR/REC-CSS2](http://www.w3.org/TR/REC-CSS2)
- DOM Level2 : [http://www.w3.org/2001/05/level-2-src.zip](http://www.w3.org/2001/05/level-2-src.zip)

HTML4.0 (4.01) + CSS has support of browsers for built-in systems in addition to the major PC browsers. It is the most recommendable language in the current circumstances because it provides mutual connectivity and user service application environment.
The following study issues concern the page description language.

- **Regulations of CSS**
  Full CSS1 support should be provided. The extent of CSS2 support must be reviewed.

- **About DHTML support**
  The regulations of DHTML must be reviewed. A system based on the W3C Document Object Model (DOM) seems appropriate.

CSS1 regulates the basic concept and various properties used mainly to specify an HTML document display style on a PC screen. Many commercial browsers have CSS1 support. The disadvantage of CSS1 is the specification limits in its document layout definition.

CSS2 is an extended version of CSS1 that basically provides backward compatibility for the latter. The newly added main specifications of CSS2 are listed below.

- Considerations for indication on various media (paper, audio synthesizer, TV, text terminal, ...)
- Application of table style sheet
- Application of positioning schemes

The scope of the CSS2 regulations should be reviewed extensively. There is no commercial browser that provides full CSS2 support at this time.

It seems that support for “Positioning”, which is essential for screen animation using DHTML, is appropriate for the current state of the web and the effects on menu display.

DOM Level2 with CSS support seems appropriate for compliance with DHTML regulations regarding DOM. Due to its wide scope, however, the extent of compliance with the regulations must be reviewed.
9-2-3. Script Language (1/2)

We recommend JavaScript as the script language for page description.

- **JavaScript1.1**
  Suitable for a terminal that does not require DHTML support. The object addition and event support functions of Version 1.1 are far more advanced than Version 1.0. Version 1.1 is good enough for providing interactive capability to a terminal.

- **JavaScript1.3**
  The core portion (Core JavaScript) complies with ECMA-262 Edition 2 (ECMAScript Edition 2). Whereas it does not offer too many enhancements on Version 1.2, this version does provide Unicode support. Although the implementation load is lighter than that of Version 1.5, Version 1.3 has various compatibility problems including support for DHTML.

- **JavaScript1.5**
  The core portion (Core JavaScript) complies with ECMA-262 Edition 3 (ECMAScript Edition 3). Version 1.5 provides additional functions such as exception processing during operation and offers highly reliable programming facilities. Furthermore, Version 1.5 supports the W3C Document Object Model (DOM) specifications for browser extended portion and provides considerations for DHTML compatibility.

See the ECMA site (http://www.ecma.ch) and the following URLs for ECMA-262 Specifications.


See the Netscape (developer of JavaScript) site for JavaScript Specifications. (URL below)

The following issues concern script language.

- **Significance of compliance with ECMA-262**
  ECMA-262 allows the addition of original types and objects. Vendor proprietary components in implemented objects that become associated with HTML elements often cause compatibility problems.

- **Relationship with DHTML**
  The portion related to object specification, which is associated with HTML elements and CSS. The script language probably should be defined based on W3C Document Object Model (DOM) specifications instead of JavaScript (ECMAScript).

The implementation of “JavaScript” as the script language must be reviewed. In this report, JavaScript is a user script language. Here, JavaScript is divided into a “core portion” and “extended portion for browser”.

The core portion complies with JavaScript Version 1.2 or later (ECMA-262) as mentioned in the previous page. The specifications of this portion can be easily stipulated. It is also a good idea to specify this portion as “ECMAScript Edition x”.

The extended portion for browser, on the other hand, has a close association with the object model of an HTML document and should be stipulated as DOM specifications. However, even if this portion, which is used for association with a script language, complies with the same DOM specifications, it still possible that it may contain a specific ‘jargon’ for each vendor implementation. Therefore, if the extended portion for browser has to be strictly stipulated, it must include the specifications of the user script language (JavaScript) as implementation standards.
Secure Socket Layer (SSL) is used when high security is required for data exchange between a user terminal and server on the HTTP base menu display. It is recommendable to use SSL Version 2.0 or Version 3.0 with a 128-bit common key.

SSL is a protocol located in the transport layer and session layer (of the OSI reference model) used for adding the function of exchanging information by encrypting data. SSL can be used together with a protocol of application layer such as HTTP. SSL is ranked as follows in the TCP/IP protocol group.

<table>
<thead>
<tr>
<th>HTTP</th>
<th>LDAP</th>
<th>IMAP</th>
<th>SSL</th>
<th>TCP/IP</th>
</tr>
</thead>
</table>

The following three functions can be implemented by SSL.

1. **Server authentication**
2. **User authentication**
3. **Encryption of session**

A secure communication is required between server and user terminal on the HIKARI Service Platform for transmission and authentication of personal information. We recommend using Secure Socket Layer (SSL) for communications on HIKARI Service in this case. The widespread use of SSL on the Internet has made it an essential technology. SSL was designed and developed by Netscape and at present the current standard is version 3.0. A public key encryption (RSA) is used together with the common key encryption in SSL. The common encryption key lengths are 40, 56, and 128 bits. However, we recommend that a user terminal provide support for a 128-bit encryption key because 40- or 56-bit keys are no longer safe, considering the high capacity of the latest computers.

SSL uses both public key encryption and common key encryption for data being exchanged. SSL is widely used with HTTP since it is also necessary to encrypt the data to be exchanged on HTTP. The following three functions can be implemented by SSL: (1) server authentication, (2) user authentication, and (3) session encryption. To implement (1) and (3), the pair of a common key and public key and their certificates are required at the server end. To implement (2), the pair of these keys and their certificates are required at the terminal end also. At the present time, however, most of the terminals do not support (2).

Visit the following sites for further information.

Details of SSL protocol

1) Hello Request (Server → Client)
   A server sends “Client Hello” request to a terminal.

2) Client Hello (Client → Server)
   This message is sent to change the encryption parameters when the “Hello Request” is received at initial connection to a server. The SSL version, random number, session ID, and supported encryption and data compression methods are sent to a server.

3) Server Hello (Server → Client)
   Server response to “Client Hello” request. This message specifies the encryption and data compression methods to be used among those indicated in the “Client Hello” message. Includes the server’s SSL version, random number, session ID, and selected encryption and data compression methods.

4) Server Certificate Message (Server → Client)
   The certificates of a server are sent to a terminal. This message is defined in a list that includes all certificate chains up to the Route Certificate Authority. The certificates comply with X.509.

5) Server Key Exchange (Server → Client)
   This message is sent from a server to a terminal when a server does not have certificates, or it has certificates that can be used only for signatures.

6) Certificate Request (Server → Client)
   Request for terminal certificates from a server.

7) Server Hello Done (Server → Client)
   Notifies that all messages are sent from server to terminal.

8) Client Certificate (Client → Server)
   When terminal authentication is performed, a terminal sends its own certificates to a server.

9) Client Key Exchange (Client → Server)
   A terminal sends data, which serves as the basis for a master secret key to be used for generating an encryption key during a session. In the case of RSA, data is encrypted using the public key of a server.

10) Certificate Verify (Client → Server)
    Sends the data used for terminal authentication by a server. Retrieves a hash value of the existing data and encrypts the data using the terminal secret key. The server will decode the data using the terminal public key and compare the data with the hash value retrieved in the same manner to perform certificate verification.

11, 12) Finished (Client → Server, Server → Client)
    A terminal notifies the server that it was successfully authenticated and vice versa.
Three possible starting methods are indicated below.

(1) Method using HTML “<a>” tag
   Directly specifies a connection address after “a href=”.
   Example: `<a href="rtsp://hostname:10000/content.mpg">Content</a>
   A general-purpose browser such as IE cannot interpret URI. Therefore the versaility of this method in general is low.

(2) Method using metafile
   Specifies a link address of the “<a>” tag in a metafile and starts an application by identifying the MIME-Type and extension.
   Example: MIME-Type : application/mpeg2-stream
            File description : rtsp://hostname:10000/content.mpg
   The description of a file is expandable.
9-2-5. Starting Method of Video Player (2/2)

(3) Method of drawing on a page using a plug-in function

- Method using the "<embed>" tag
  
  Example: `<embed src="rtsp://hostname:10000/content.mpg"
  type="application/mpeg2-stream"
  width="320" height="200" />

  Note, however, that, the "<embed>" tag is not defined in HTML 4.0.

- Method using "<object>" tag
  
  Example: `<object data="rtsp://hostname:10000/content.mpg"
  type="application/mpeg2-stream"
  width="320" height="200">

  Content cannot be played back.
  Access <a href="http://info.server.com/info.html"> here </a>
  for further information.
  </object>

  Note, however, that some browsers do not provide support for the
  "<object>" tag.

A description using a meta file (method (2)) appears to be most appropriate. However, the description format in metafile must be reviewed.

Three methods for starting the video player at the user end are given above. General-purpose browsers do not support Method (1) at the present time and thus it has little practical use.

Method (2) offers expandability and has been used for starting the Windows Media Player and RealPlayer. This method seems to be the most appropriate. However, the description format of a meta file must be discussed in the future.

Method (3) can maintain best consistency for menu description among these methods. However, the "<embed>" tag is not defined in HTML 4.0, and many browsers do not provide support for the "<object>" tag, which has succeeded the "<embed>" tag. Therefore, the implementation of this method must be reviewed in the future with consideration for trends.
9-2-6. Conclusion on Menu Display

- HTTP should be used as the transfer protocol for menu display data.
  - HTTP/1.0 and HTTP/1.1 may be used for this purpose. HTTP/1.1 is recommended, however.
  - The Basic authentication of HTTP is used to restrict access to a menu.

- HTML should be used as the language for menu page description.
  - HTML3.2, HTML4.0 (+CSS), and DHTML may be used for this purpose. HTML4.0 is recommended, however.
  - The W3C DOM Specifications must be defined for DHTML compatibility, which must be reviewed further.

- JavaScript should be used as the script language for page description.
  - JavaScript1.1, JavaScript1.3, and JavaScript1.5 may be used for this purpose.
  - JavaScript1.3 is recommended for a terminal that does not require DHTML support.
  - JavaScript1.5 is recommended for a terminal that requires DHTML support.

- SSL should be used to ensure security during a data transfer.
  - SSL2.0 and SSL3.0 may be used. A 128-bit common key or higher is recommended.
  - Issues such as the authentication system and issuer of certificates must be reviewed further.
9-3. Audio/Video Stream Formats

Contents of Section 9-3

9-3-1. Outline of Audio/Video Stream Formats
9-3-2. Outline of Video Formats
9-3-3. MPEG-2 Video Formats
9-3-4. MPEG-4 Video Formats
9-3-5. Outline of Audio Formats
9-3-6. Outline of Multiplex Formats and MPEG-2 Multiplex Formats
9-3-7. MPEG-4 Multiplex Formats
9-3-1. Outline of Audio/Video Stream Formats

Main audio/video stream formats

- ISO/IEC Standards
  MPEG-1, MPEG-2, MPEG-4, etc.
- ITU-T Standards
  H.263, etc.
- De facto standards
  Real Audio/Video, Windows Media Technology, QuickTime, etc.

Audio/video stream formats

- ISO/IEC Standards
  MPEG-1 and MPEG-2 implement the standards for VideoCD and broadcasting, respectively.
  The MPEG-4 standards target low reliability networks at lower bit rates.
  Developers are currently making an effort to provide ultra high-quality (Studio Profile) and
  streaming with broadband network support (Advanced Simple Profile, Fine Granularity Scalable
  Profile) for MPEG-4.
- ITU-T Standards
  These standards are basically a subset of the ISO/IEC standards.
  H.263 is widely used for picture phone and video conferences at low transmission speeds.
  (Subset of the MPEG-4 Simple Profile)
  There is no stream format that can use broadband at the present time.
- The de facto standards implemented by several companies
  Real Audio/Video, Windows Media Technology, and QuickTime are widely used for content
  delivery via the Internet.
9-3-2. Outline of Video Formats

**MPEG-2/MPEG-4 Video Formats**

- **MPEG-2 Video**
  - Used for broadcasting and high quality (DVD quality). A full lineup of software and hardware is available.
  - Main Profile is used for the most part.
  - Compliance with bandwidth (1.5 M to 80 Mbps) assumed for FTTH.
  - MPEG-2 Video is more vulnerable to out-of-sync and packet loss compared with MPEG-4.

- **MPEG4 Visual (natural video content)**
  - Wider range of models (including animation) available than MPEG-2.
  - MPEG4, however, is used only for natural video content. It is not used for other content.
  - Among these formats, the most extensive lineup of software and hardware is available for Simple Profile (up to 384 k).
  - The hardware and software that support up to the subset (up to 2 M) of Core Profile are commercially available.
  - Includes various information for compensating out-of-sync and packet loss.

Video formats

MPEG-2 Video and MPEG-4 Visual are the main standards. Note, however, that the software and hardware lineups are available for only portions of the standards. (A terminal can be configured at comparatively low cost for these portions.)

MPEG-2 Video is suitable for DVD quality.

There are high-quality support standards for MPEG-4. However, the standards ready for commercial use are of VTR quality at the present time.

It is known that MPEG-4 is relatively secure against out-of-sync and packet loss. Technically speaking, “MPEG-4 is provided with an extensive information base for recovering out-of-sync and packet loss” only. The amount of data that actually can be recovered depends on the decoder.
### 9-3-3. MPEG-2 Video Formats

<table>
<thead>
<tr>
<th>Profile and level of MPEG-2 video</th>
<th>Simple 4:2:0</th>
<th>Main 4:2:0</th>
<th>SNR Scalable 4:2:0</th>
<th>Spatial Scalable 4:2:0</th>
<th>High 4:2:2</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td>MP@HL</td>
<td></td>
<td>HP@HL</td>
</tr>
<tr>
<td>1920 x 1080 x 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1920 x 1152 x 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-1440</td>
<td></td>
<td></td>
<td>MP@H14</td>
<td></td>
<td>HP@H14</td>
</tr>
<tr>
<td>1440 x 1080 x 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1440 x 1152 x 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main</td>
<td>SP@ML</td>
<td>MP@ML</td>
<td>SNP@ML</td>
<td></td>
<td>HP@ML</td>
</tr>
<tr>
<td>720 x 480 x 29.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>720 x 576 x 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>MP@LL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>352 x 288 x 29.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Video formats that are actually used are shown in the portion. There are a significant number of software/hardware available for these portions.
- SNR Scalable and Spatial Scalable are systems that can provide multiple bandwidth support, and are aimed at simultaneously distributing stream data to broad bands and narrow bands. There are no actual implementation examples, however.

### MPEG-2 Video

A significant number of MPEG-2 Video software/hardware is available for Profile/Level, which is used in digital broadcasting.

If this portion is used for video content delivery (live, VoD), the parts can be shared with a digital TV and thus a terminal can be manufactured at low cost.

The MPEG-2 formats used for bi-directional video content require large-scale buffering, which makes it difficult to minimize delays.
9-3-4. MPEG-4 Video Formats

- **MPEG-4 Visual Profile**
  - MPEG-1 and MPEG-2 are compression formats of natural video.
  - In addition to natural video, MPEG-4 can handle animation, texture, separating from background, merging, and coding by setting them as multiple objects.
  - In practice, however, MPEG-4 is used for natural video and it normally handles one object, same as MPEG-2.
  - Widespread application of Profile (such as Main Profile), which is used commercially (DVD, etc.) for MPEG-2, may not be possible.
  - Only the Simple Profile (up to 384 k) and subset of Core Profile (video of up to 2 M using a compression method similar to Simple Profile) will probably be put to commercial use at present.

The main Profiles associated with the natural video of MPEG-4 Visual are listed below.

- **Simple Visual Profile (Standardized by MPEG-4 Version 1)**
  Basic coding of natural video. Coding is performed at a specified bit rate. The bandwidth is from 64 k to 384 k. Most of the current commercially available MPEG-4 equipment provides support only for this Profile.

- **Simple Scalable Visual Profile (Standardized by MPEG-4 Version 1)**
  Scalable time and space coding that supports two bit rates. Up to 256 k.

- **Core Visual Profile (Standardized by MPEG-4 Version 1)**
  Coding for irregular shapes and time scalability have been added to Simple Visual Profile. Up to 2 M.

- **Main Visual Profile (Standardized by MPEG-4 Version 1)**
  Targets DVD quality, but the function puts too much load. Use of this Profile is unlikely even in the future.

- **Advanced RealTime Simple Profile (Standardized by MPEG-4 Version 2)**
  Reduces delay for picture phones, video conference, and monitoring. This Profile is not used at the present time, but it has the potential for future use (up to 2 M).

- **Advanced Coding Efficiency Profile (Standardized by MPEG-4 Version 2)**
  High compression rate coding. The function is even heavier than Main Visual Profile. (384 k to 38.4 M)

- **Advanced Simple Profile/Fine Granular Scalability Profile (standards in planning phase)**
  More efficient than Simple Visual Profile. Lighter (presumably) than Advanced Coding Efficiency (ACE). Emphasizes streaming via the Internet. Fine Granular Scalability (FGS) Profile (128 k to 8 M) provides scalability that takes delivery to multiple bandwidths into account.

- **Simple Studio Profile/Core Studio Profile (standards in planning phase)**
  High-quality video that can be edited. It is difficult to ascertain whether the use of these Profiles will be widespread in the future (180 M to 1.8 G).
9-3-5. Outline of Audio Formats

- MPEG-2/MPEG-4 Audio Formats
  - MPEG-2 Audio
    - Audio implemented by MPEG-1. (MPEG-1 Audio Layer I, II, III)
    - Used for broadcasting and high quality (AAC: Advanced Audio Coding).
  - MPEG-4 Audio
    - In addition to the natural sound of MPEG-1 and MPEG-2, this system provides natural audio of low bit rates (2 k to 64 k) and advanced functions such as MIDI, Text-to-Speech, and placement of a 3D sound source.
    - In practice, however, MPEG-4 is used only for natural audio. The mainstream MPEG-4 Audio applies the same audio format as MPEG-2.

MPEG-2 Audio

MPEG-1 Audio Layers I, II, and III and AAC have been adopted by MPEG-2 Standards. AAC can handle higher sound quality sound, and so has been added to the standards. MPEG-1 Audio Layers I and II and AAC are also used for digital broadcast. Some AAC have surround sound (5.1 ch sound).

MPEG-4 Audio

The low bit rate sound is used for commercial purposes, such as Global System for Mobile Communications (GSM) and Adaptive Differential Pulse Code Modulation (ADPCM). MPEG-4 Audio is mainly used for IP telephones and cellular phones at the present time. ADPCM and Pulse Code Modulation (PCM) enable a lower delay. The highest quality level generally used by AAC sound is stereo and CD.
9-3-6. Outline of Multiplex Formats and MPEG-2 Multiplex Formats

- **Multiplex formats**
  Formats used for bundling video, audio, and other data. An issue for multiplex formats is a system that synchronizes multiple streams.

- **MPEG-2 multiplex formats**
  - MPEG-2 Program Stream (PS)
    Format suitable for data storage.
  - MPEG-2 Transport Stream (TS)
    Transmission format.

Since the amount of jitters in a transmission via the Internet is higher than that of a broadcast, the synchronization signal of TS cannot be used (data is played back by fixing the frame rate at the playback side).

Due to this reason, there is very little merit in using TS (transmission can also be performed by PS). However, TS is widely used at the present time.

MPEG-2 multiplex formats

MPEG-2 PS and MPEG-2 TS are available. The former is mainly used for storage media and the latter for transmission.

Although there is not much difference between these formats when it comes to transmission over the Internet, MPEG-2 TS is used more often.

The synchronization signal of MPEG-2 TS cannot be used for synchronizing because of the fluctuations in network transmission.

Therefore, the playback frame rate must be fixed to the terminal clock. It is possible that out-of-clock will occur for a long playback. However, this should not be a big problem.
9-3-7. MPEG-4 Multiplex Formats

MPEG-4 multiplex formats

• MPEG-2 TS
  MPEG-2 TS was originally designed for MPEG-2. However, this format can be applied for MPEG-4.

• MPEG-4 Systems (MP4)
  Multiplex format designed for MPEG-4. As a format, it provides special playback support for a downloading operation. It cannot be used for a long period of streaming.

• RFC 3016/3GPP TS26.234
  This is the most popular format for Internet streaming at the present time. It can synchronize with RTP and perform special playback by RTSP. Small header overhead. Extended functions are required for filing data. This format has been stipulated for next-generation cellular phones. It is included as a part of the MPEG-4 Systems Standards.

RFC 3016 is a very popular format. The system is under review for terrestrial digital broadcasting. If the MPEG-4 multiplex format is adopted for terrestrial digital broadcasting, the use of MP4 may become widespread (the MP4 format has been used for storage data). The RFC 3016 format has a proven track record in bi-directional video communication services.

MPEG-4 multiplex formats

MPEG-4 Systems (MP4) are a natural choice as far as the standards is concerned. MP4 does not have much of a proven track record yet.

In practice, only some portions of the MPEG-4 audio/video formats are used. Therefore, the use of MP4 is not necessary. MP4 uses a complicated system in order to support a variety of MPEG-4 audio/video other than natural audio/video.

Natural audio/video can be multiplexed even by the MPEG-2 multiplexing system. The standards called RFC 3016 used by the Internet Engineering Task Force (IETF) have a proven track record as the MPEG-4 multiplex format.
9-4. Media Data Transfer Protocols

Contents of Section 9-4

9-4-1. What is Media Data Transfer Protocol?
9-4-2. RTP
9-4-3. When RTP is Not Used
9-4-1. What is Media Data Transfer Protocol?

Unidirectional delivery from server to terminal and bi-directional video communication between terminals are two types of media data communications used. These communications consist of the following protocols.

- **Control protocols**
  - Protocols used for controlling start, stop, and special playback of a media data transfer.
  - A protocol such as RTSP is mainly used for video delivery. SIP and H.323 are mainly used for bi-directional video communications.

- **Media data transfer protocols**
  - Protocols used for actual media data transfer.
  - UDP or UDP with RTP are mainly used.

In this section, media data transfer protocols are described.

See Sections 9-5 and 9-6 for control protocols.
9-4-2. RTP (1/6)

Outline of RTP

- This protocol has been stipulated in RFC 1889 and RFC 1890.
- Transmission protocol between server and user terminal for transmitting media data that demands critical real-time such as interactive multimedia or video delivery services.
- In addition to a UDP header, an RTP header (see the next page) is added to handle the following issues that could be a problem when real-time media is transmitted.
  - Network fluctuation
  - Media synchronization
  - Packet loss
- RTP Control Protocol (RTCP) is used for controlling an RTP stream. (RTCP uses a port that is different from the RTP port, and has a message format that is unlike the RTP format. However, these two ports are operated in a closely coordinated manner.)
9-4-2. RTP (2/6)

Outline of RTP header

The definitions of RTP header elements are summarized below.

- **RTP version:**
  Indicates the RTP version.

- **Padding flag:**
  A flag used for indicating whether padding (dummy data to make up the data length) is added at the end of a packet.

- **Extension header flag:**
  A flag used for indicating whether an extension header is added at the end of the RTP header.

- **Number of CSRC:**
  Indicates the number of CSRC identifiers.

- **Marker bit:**
  The usage of marker is defined for each payload type. This bit is used to indicate the border of an image frame for video.

- **Payload type:**
  Indicates the type of coding used for the transmission data. (See 9-4-2 3/6 for details.)

- **Sequence No.:**
  Sequence number used for detecting a packet loss at the receiving end. A random number decides the initial value.

- **Time stamp:**
  Indicates the time when the first byte of a packet is sent. The payload type defines the clock reference frequency.

- **SSRC identifier:**
  Synchronization transmission source identifier. This identifier is used to identify a synchronization transmission source so that the same value is assigned to multiple streams which must be processed in a combination, such as the audio stream and video stream of the same user.

- **CSRC identifier:**
  Host identifier. This identifier is used to identify a host that prepared each stream element included in a packet. For example, when streams from multiple terminals are transmitted as a single stream by a mixing process on a mixer, an identifier of each source terminal is indicated.
Payload type
The key audio/video coding system for RTP is defined as payload type.
The following table indicates the payload type of the defined main
audio/video streams.

<table>
<thead>
<tr>
<th>Payload type</th>
<th>Coding system</th>
<th>Payload type</th>
<th>Coding system</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>G.711 µ-law (Audio)</td>
<td>33</td>
<td>MPEG-2 TS (Audio/video)</td>
</tr>
<tr>
<td>8</td>
<td>G.711 A-law (Audio)</td>
<td>34</td>
<td>H.263 (Video)</td>
</tr>
<tr>
<td>9</td>
<td>G.722 (Audio)</td>
<td>35-71</td>
<td>Undefined</td>
</tr>
<tr>
<td>14</td>
<td>MPEG-Audio (Audio)</td>
<td>72-76</td>
<td>Reserved</td>
</tr>
<tr>
<td>15</td>
<td>G.728 (Audio)</td>
<td>77-95</td>
<td>Undefined</td>
</tr>
<tr>
<td>31</td>
<td>H.261 (Video)</td>
<td>96-127</td>
<td>Dynamic assignment</td>
</tr>
<tr>
<td>32</td>
<td>MPEG-Visual (Video)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The payload header is defined on the next page.
## Payload header

Example: H.263 (RFC 2190: RTP Payload Format for H.263 Video Streams)

<table>
<thead>
<tr>
<th>IP header</th>
<th>UDP header</th>
<th>RTP header</th>
<th>H.263 Payload header</th>
<th>H.263 Media data</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>P</th>
<th>Mode Flag</th>
<th>S</th>
<th>Syntax-based Arithmetic Coding Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>PB-frames Mode</td>
<td>A</td>
<td>Advanced Prediction Option</td>
</tr>
<tr>
<td>SBIT</td>
<td>Start Bit Position</td>
<td>reserved</td>
<td></td>
</tr>
<tr>
<td>EBIT</td>
<td>End Bit Position</td>
<td>DBQ</td>
<td>Differential Parameters for the B Frame</td>
</tr>
<tr>
<td>SRC</td>
<td>Source Format</td>
<td>TRB</td>
<td>Temporal Reference for B</td>
</tr>
<tr>
<td>I</td>
<td>Intra-frame Encoded Data</td>
<td>TR</td>
<td>Temporal Reference</td>
</tr>
<tr>
<td>U</td>
<td>Unrestricted Motion Vector Option</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The rules are defined in RFC for a specific coding system. For example: a) when the minimum coding unit (macro block in the case of video data) is divided into packets, the data must not cross over two packets, and b) an RTP payload header must be implemented so that frame switching can be detected by a marker bit.

Besides the payload header stipulated for H.263, an RTP payload header for the following media formats is defined by IETF.

- RFC 2032 : RTP Payload Format for H.261 Video Streams
- RFC 2190 : RTP Payload Format for H.263 Video Streams
- RFC 2250 : RTP Payload Format for MPEG-1/MPEG-2 Video
- RFC 3016 : RTP Payload Format for MPEG-4 Audio/Visual Streams
- RFC 2198 : RTP Payload Format for Redundant Audio Data
- RFC 3119 : RTP Payload Format for MP3 Audio
- RFC 2658 : RTP Payload Format for PureVoice(tm) Audio
- RFC 3047 : RTP Payload Format for ITU-T Recommendation G.722.1
9-4-2. RTP (5/6)

■ RTCP
RTCP is a protocol used for exchanging control information between the sender and recipient who perform a media data transfer (streaming) using RTP, so that the media display/playback quality can be maintained.

■ Implemented functions
RTCP is used to notify information for resolving the following issues of media data transmission using RTP.
- Flow control
- Clock synchronization
- Comparison of playback time between media
- Recognition of information source

If the above control functions are implemented using RTCP information, the processing load increases. Therefore, the notified information may not actually be used.

RTCP functions are implemented by a combination of RTCP packets. The RTCP packet types are listed below.

Types of RTCP packets
- Sender Report (SR): An equipment that has sent a stream will use this packet to notify information related to the sent stream, and information related to a stream received from the other equipment.
- Receiver Report (RR): An equipment that has not sent a stream uses this packet to notify information related to a stream received from the other equipment.
- Source Description (SDES): This packet is used to notify the relationship between the SSRC/CSRC values of the RTP packet and user information such as a mail address.
- BYE: An equipment that disconnects communication will use this packet to notify the other equipment of the disconnection.
- Application (APP): This packet is used to notify control information inherent in an application that does not comply with RTCP.

An RTCP packet is transmitted using UDP. The port number to be used by RTCP is the RTP port number plus 1.
This figure shows an example of flow control. When network congestion occurs and data is lost, the video is distorted. In that case, the display quality at the receiving end will improve if the media data compression rate is increased and the transmission rate is decreased.

An RR packet notifies the sender side of the packet-discard rate to enable the above control. (The implementation of this function depends on the implementing side. Generally, the level of network congestion that causes packet loss changes significantly within a short period of time. Therefore, a significant effect is rarely achieved.)
9-4-3. When RTP is Not Used

Instead of using RTP, stream delivery for media data transfer is also enabled by writing the media data in UDP.

Example: MPEG-2 TS
One packet of MPEG-2 TS has a fixed length of 188 bytes. These packets can be sequentially transmitted by writing ‘N’ number of MPEG-2 TS packets.

This system is expected to achieve high network quality for HIKARI Service access network. However, this matter must be verified.

Why RTP is not used

- RTP is used together with RTCP. When the RTCP functions are implemented, the processing load on a user terminal becomes high. (There is no need to use RTCP data for control purposes when a substantial network quality is ensured, packet loss does not occur too often, and delays are rather consistent.)
- The RTP common header is not required when a stream containing a time stamp or multiplex information (such as MPEG) is processed.
- There are many cases where RTP is not used for streaming at a high bit rate. (This is because of the large overhead of the RTP header.)
9-5. Stream Delivery Control Protocol

Contents of Section 9-5

9-5-1. What is Stream Delivery Control Protocol?
9-5-2. Outline of RTSP
9-5-3. Needs for Common RTSP
9-5-1. What is Stream Delivery Control Protocol?

- What is the stream delivery control protocol?
  It is an application level protocol that performs actual stream controls such as start and pause of content delivery. This protocol implements stream delivery in a combined operation with the media transfer protocol (as discussed in Section 9-4).

- Types of control protocols
  The stream delivery control protocols, which are adopted by each vendor for their products at the present time, are roughly divided into the following two types.
  - RTSP (compliant with RFC 2326)
  - Vendors’ original protocols
    (Microsoft Windows Media Technology, etc.)

The stream delivery control protocol to be used for HIKARI Service must be able to process media data in various formats and have open-ended specifications. HSAC uses RTSP (compliant to RFC) as stream delivery control protocol. (See the following sections for information about RTSP.)
What is RealTime Streaming Protocol (RTSP)?

**History**
- RTSP was jointly designed by Real Networks, Netscape, and Columbia University.
- It is defined in RFC 2326.
- RTSP was promoted to RFC from the Internet draft in August, 1998, and is the Proposed Standard at present.

**Functions**
- Multimedia presentation control protocol between user terminal and server
- Application level protocol for controlling real-time data transmission. This protocol enables control for transmission on-demand.
- A request message is formed by the combination of a method and header.
- A target media data either can be generated in real-time or stored in advance (before delivery)

**Example of RTSP request message description**

```
PLAY rtsp://<IP Address>:<Port Number>/content.mpg RTSP/1.0
CSeq: 2
Session: 1234567890
Range: npt=25.00-
```

“PLAY” is called a method; and “CSeq”, “Session”, and “Range” are called headers.
9-5-2. Outline of RTSP (2/2)

### Basic sequence

<table>
<thead>
<tr>
<th>Client</th>
<th>Video Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESCRIBE</td>
<td>Request the information of the content. Acknowledged. Requested data will be sent.</td>
</tr>
<tr>
<td>SETUP</td>
<td>Requesting the content. Acknowledged. The setting has been completed.</td>
</tr>
<tr>
<td>PLAY</td>
<td>Please distribute the content. Acknowledged. The content will be distributed.</td>
</tr>
<tr>
<td>PLAY</td>
<td>Perform trick play (fast-forward, rewind). Acknowledged. Trick play will be performed.</td>
</tr>
<tr>
<td>ANOUNCE</td>
<td>delivery of the content until the end has been completed. Acknowledged.</td>
</tr>
<tr>
<td>TEARDOWN</td>
<td>Ending the viewing of this content. Acknowledged. End now.</td>
</tr>
</tbody>
</table>

- **DESCRIBE**
- **SETUP**
- **PLAY**
- **PLAY**
- **ANNOUNCE**
- **TEARDOWN**

Data Streaming
9-5-3. Needs for Common RTSP

The protocol stacks to be created vary according to the expandable description and difference in the interpretation of RFC 2326.

The details of RTSP implemented in the current product system of each vendor are different.

Mutual connectivity between different system equipment is very rare.

We have selected the common RTSP to establish a common platform because it ensures minimum connectivity.

Purpose:
- Designing a protocol stack that will become a reference model of RFC 2326.
- Making it easier to implement on each vendor’s product.
  (See the attached System Interface Requirements on HIKARI Service Network Architecture.)

This time, we have studied a recommended model of the streaming delivery control protocol, which is one of the common platforms for video delivery services.

We have designated minimum protocol stacks to establish mutual connectivity of the basic operation portions such as “starting playback”, “stop”, and “fast-forward”. See the attached System Interface Requirements on HIKARI Service Network Architecture for further information.
9-6-1. Outline of Communication Connection Protocols

The available main communication connection protocols are SIP and H.323.

1) SIP

The SIP standards were established by IETF. SIP Version 1 was proposed in February 1996. Version 1 was merged with the Simple Conference Invitation Protocol (SCIP), which was also proposed in February 1996. This merged version is called SIP Version 2 and was released in December 1996.

Although the standardization process is not complete (draft-ietf-sip-rfc2543bis is the standard at the present time), the details of this protocol have been completed up to the commercial stage.

2) H.323

H.323 standards were established by ITU-T. The regulations of video conference system using line switching are revised to the regulations applicable to “low-reliability network”. There are common regulations for both (such as audio/video coding regulations).

The studies in this document cover SIP because of its expandability and future potential. We have studied the communication connection protocol based on SIP for bi-directional video communication services.

We cannot determine at present whether H.323 or SIP will become the dominant protocol in the future, and therefore we should pay close attention to the market trends.

<table>
<thead>
<tr>
<th></th>
<th>SIP</th>
<th>H.323</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encode</td>
<td>ASCII text + CR/LF</td>
<td>ASN.1 PER</td>
</tr>
<tr>
<td>Field definition</td>
<td>Very moderate restrictions</td>
<td>Strict restrictions are required for the content of each field and the field length.</td>
</tr>
<tr>
<td>Test</td>
<td>Easy confirmation</td>
<td>A tool for confirming exchanged information is required.</td>
</tr>
<tr>
<td>Transport</td>
<td>Either TCP or UDP. (Disconnection during session is accepted.) The round trip time of UDP can be reduced.</td>
<td>TCP only (UDP cannot be used.)</td>
</tr>
<tr>
<td>Address</td>
<td>SIP URL (Same as H.323)</td>
<td>H.323 alias, e-mail address, Phone No, URL...</td>
</tr>
<tr>
<td>Complication</td>
<td>Simple (Five messages for setup)</td>
<td>Complicated (Eighteen messages for setup.)</td>
</tr>
<tr>
<td>Scalability</td>
<td>Stateless</td>
<td>Problem in Gatekeeper scalability. The statefullness restricts the number of simultaneous calls.</td>
</tr>
<tr>
<td>Targets of regulations</td>
<td>Signaling only</td>
<td>Umbrella (Manages various related regulations.)</td>
</tr>
<tr>
<td>Conference</td>
<td>Also implemented by multicast</td>
<td>An MCU is required (if there are four or more participants in the conference).</td>
</tr>
<tr>
<td>Market trends</td>
<td>On the way to wider applications (WindowsXP provide SIP support)</td>
<td>3-years ahead in standardization. There are several vendors.</td>
</tr>
<tr>
<td>Expandability</td>
<td>Flexible expansion is enabled (local expansions may be used).</td>
<td>Expansion is possible but there is little flexibility.</td>
</tr>
</tbody>
</table>
9-6-2. Outline of SIP

Basic sequence of SIP (SIP protocol core)

SIP is used for initializing communications.

The above sequence illustrates minimum message exchanges. The protocol required for such minimum message exchanges is defined first by RFC as the SIP protocol core.

The above figure shows the SIP message exchanges between Terminal A and Terminal B. These messages are exchanged when Terminal A (and Terminal B) have the knowledge of the location of Terminal B (and Terminal A).

If the location of a terminal is unknown (for example, when a remote communicating party moves, or in the case of a wide-area network), direct message exchanges between terminals cannot be performed.

In that case, the exchange will involve either a server (proxy server) that relays the SIP message or a server (redirect server) that instructs the transfer destination address, in order to identify the location of the remote party.

Even if there is a proxy server for relaying the SIP message, audio/video data transfer using RTP is not always carried out via proxy server.
9-6-3. Expansion of SIP Protocol Core (1/6)

Addition of reliable delivery
No confirmation for ‘Ringing’ and ‘Queuing’.
→ This feature is supported by an expansion protocol.

Core protocol sequence

| (1) INVITE |
| (2) 180 Ringing |
| (3) 200 OK |
| (4) ACK |

Expansion protocol sequence

| (1) INVITE |
| (2) 180 Ringing |
| (3) 180 Ringing |
| (4) PRACK |
| (5) 200 OK |
| (6) ACK |

Issues for SIP core protocol

No message received confirmation for a kind of backup response (180 Ringing, 182 One in The Queue, etc.) except “OK” for “INVITE” is available and there is no means to confirm whether a message was actually transmitted to a remote terminal.

To resolve the above issue, the addition of Provisional Response ACK (PRACK) has been proposed (Draft-ietf-sip-100rel).

“Ringing” will be transmitted repeatedly until PRACK is received to ensure that “Ringing” has reached the remote terminal.

No concern is necessary for the minimum operation on a terminal that does not support PRACK.
Support for resource reservation

When it takes too long to set the prerequisites (such as resource reservation) of communication → This feature is supported by an expansion protocol.

Expansion of the SIP protocol core

Timeout occurs when it takes too long to set the prerequisites (such as required resources reservation including allocation of bandwidth) of communication.

To resolve the above problem, the preCondition MET (COMET) method was added. COMET notifies a remote terminal of the completion of resource reservation. This feature enables a sender to wait until the execution conditions of the sequence following “Ringing” are ready.
9-6-3. Expansions to SIP Protocol Core (3/6)

- Information notification during communications
  A separate session is required for information notification during communications
  → This feature can be supported by the expansion protocol without establishing a new session.

Issues for SIP protocol core

Conventionally, a separate session is required for notifications during communications. In the case of a simple information transmission, however, it is better to provide the notification within the frame of the SIP protocol.

To provide support for the above feature, the INFO method was added. (RFC 2976)
9-6-3. Expansion of SIP Protocol Core (4/6)

Asynchronous event notification

No asynchronous event notification (Repetition of polling is required.)
→ This feature is supported by the expansion protocol.

Issues for SIP protocol core

No asynchronous event notification is available. One of the solutions for this issue was confirmation by polling. If a large number of events need to be confirmed, it is difficult to use the polling method. Also processing becomes complicated.

Due to the reasons above, the SUBSCRIBE and NOTIFY methods were added for asynchronous event notification. (draft-ietf-sip-events)

The events required for notification are sent to a remote terminal using SUBSCRIBE, and the notification is performed using NOTIFY if the event actually occurs. Normally, the change of status of a remote terminal and the change of the number of participants are notified as an event.
9-6-3. Expansion of SIP Protocol Core (5/6)

Transfer processing

Transfer processing during a call is not defined.
→ This feature is supported by the expansion protocol.

Issues for the SIP protocol core

No standards have been defined in regard to the procedures for realizing the transfer function that is frequently used for telephones. A conventional framework can be used to establish a new session and disconnect the old one. But, no standards are set up for the control over a message and therefore it was done by an original control sequence. This became an issue of mutual operation performance.

Since transfer processing is used frequently, it is recommendable to define the standards of the procedures for realizing the transfer function within the frame of SIP. For this reason, the REFER method was added. (draft-ietf-sip-cc-transfer)

Accompanying REFER, Refer-To: and Referred-By were added to the header.
9-6-3. Expansion of SIP Protocol Core (6/6)

Other expansion methods

• MESSAGE
  Sends an instant message.

• DO
  The commands themselves are defined in the Device Messaging Protocol (DMP).

The following expansion methods have been proposed.

• MESSAGE
  Sends an instant message. (draft-rosenberg-impp-im)

• DO
  Sends the commands for equipment control. (draft-moyer-sip-appliances-framework)
  H.248 uses the DO method for controlling a device that has a closely connected relationship.
  Unlike H.248, SIP uses the DO method for controlling a device that has a loosely connected relationship.
  The commands themselves are included in the message body and defined in the Device Messaging Protocol (DMP).
9-6-4. Issues for SIP

- Issues of SIP
  - Interoperability
    - SIP is a protocol related to the initialization of a session. It does not define how each session should be used in subsequent communications. Since it can be freely expanded, a problem could occur in mutual connectivity as in RTSP.
    - The significance of the SDP description and its mode of operation are not clearly defined.
  
  - Congeniality with Firewall/NAT
    - The data to be exchanged during communications should be encrypted in some form. The IP address and port number of the remote terminal are defined in the SDP description (encryption is recommended) included in the message body of SIP. Therefore, when the data passes through NAT, these values must be replaced. Therefore, there must be a system that is interlinked with NAT/Firewall and SIP Proxy at the application level and also provides an encryption function.

For better accuracy, the standards defined for H.323 are stricter than those of SIP. SIP has more potential in regard to support for the variety of media expected in the future. NAT-related problems are also an issue for H.323.
9-6-5. Session Description Protocol (SDP) and Its Issues

- **Session description**
  Specifies an address, bandwidth, expiration time, and media type that are defined in the message body of the INVITE method using Session Description Protocol (SDP: RFC 2327).

- **Issues for SDP**
  Originally, SDP was not designed for combined use with SIP. Although SDP is also used for RTSP, there are inappropriate or incomplete description formats for these protocols.

  SDP Next Generation (SDPng) has been proposed. (draft-ietf-mmusic-sdpng) SDPng can enable a richer session description which is also more suitable for negotiation.

SDP is also used for the session description of RTSP. It was originally designed for conference systems. Therefore, SDP is not suitable for both SIP and RTSP.

There are the following main problems.

- Insufficient session description capacity
- Insufficient expression capacity for negotiation

SDP Next Generation (SDPng) is being proposed to resolve these problems (no relationship with IPng). The description will be based on XML in SDPng.
9-7. Conclusion of Protocols and Formats

Contents of Section 9-7

9-7-1. Protocols and Formats for Live/VoD
9-7-2. Protocols and Formats for Bi-directional Video Communication
9-7-1. Protocols and Formats for Live/VoD (1/2)

Summary of audio/video formats

- The following formats are mainly applied at present because of the solid software/hardware support for their proven track record in digital broadcasting services.
  
<table>
<thead>
<tr>
<th>Format</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video format</td>
<td>MPEG-2 MP@ML, MP@H1440</td>
</tr>
<tr>
<td>Audio format</td>
<td>MPEG-1 Audio Layer I, II/AAC</td>
</tr>
<tr>
<td>Multiplex format</td>
<td>MPEG-2 Transport Stream</td>
</tr>
</tbody>
</table>

- The following formats have future potential.
  
<table>
<thead>
<tr>
<th>Format</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video format</td>
<td>MPEG-4 Advanced Simple Profile, Fine Granular Scalability Profile</td>
</tr>
<tr>
<td>Audio format</td>
<td>AAC</td>
</tr>
<tr>
<td>Multiplex format</td>
<td>MPEG-4 Systems</td>
</tr>
</tbody>
</table>

The live and VoD audio/video formats are summarized as above.

For video formats, we expect to see MP@HL support in MPEG-2. (MP@HL is HDTV standards that are not used for actual broadcasting at the present time.)

AAC is required for a surround sound of 5.1ch, for example.

It is difficult to estimate the timing of practical MPEG-4 application using compression methods other than Simple Profile. Advanced Simple Profile is suitable for a fixed bandwidth and Fine Granular Scalability Profile is suitable for delivery to multiple bandwidths.

The timing of the widespread use of formats relies on a format's application to next-generation cellular phones and terrestrial digital broadcasts, and the substantial PC software support.
Conclusion on media data transfer protocols
- We recommend the use of MPEG-2 TS over UDP in the case where media is MPEG-2 Transport Stream.
- Besides the above, we generally recommend using RTP as the media data transfer protocol.

Conclusion of stream delivery control protocols
- RTSP is used as a stream delivery control protocol.
- There is a wide scope of interpretation for the RTSP standards and this raises the issue of mutual operation performance. The attached System Interface Requirements on HIKARI Service Network Architecture provides detailed definitions of the regulations concerning minimum mutual operation performance.

We can also expect to see instances where MPEG-2 PS is used as audio/video formats. If there are such instances, a further review of the media data transfer protocol must be carried out.

Complies with ISMA if MPEG-4 is used.
Summary of audio/video formats
- The following formats are mainly applied at present because of the solid software/hardware support they have.
  Video format : MPEG-4 Simple Profile@Level3 (384 k),
                Core Profile@Level2subset (2 M)
  Audio format : ADPCM/PCM/MPEG-1 Layer I, II/AAC
  Multiplexed format: RFC 3016

- The following formats have future potential.
  Video format : MPEG-4 Advanced RealTime Simple Profile
  Audio format : AAC-Low Delay
  Multiplexed format: MPEG-4 Systems (considerations for a low delay factor)
                    (Relies on the conditions of application to third generation cellular
                     phones and solid PC software support.)

The audio/video formats used for bi-directional communication are summarized as above.
Core Profile (2M), which uses the same compression algorithm as Simple Profile, shows the highest promise regarding high quality.
The ADPCM/PCM audio formats support telephone sound quality. MPEG-1 Audio Layer I, II/AAC supports broadcast quality.
The AAC has a long delay, but this seems to be within the acceptable range compared with the video delay.

Although a low delay factor has been taken into consideration for standardization of the Advanced RealTime Simple Profile of MPEG-4, DVD quality is not likely to be implemented.
The AAC Low Delay format has been designed to implement low delay by AAC. The level of reduction in sound quality must be evaluated.
The multiplex system may be similar RFC 3016, which has been designed for low delay and low load by the MPEG-4 Systems format.
Conclusion on media data transfer protocols
- We recommend using RTP as the media data transfer protocol. The payload format is stipulated in RFC 3016 when MPEG-4 Video/Audio is used.
- The media data transfer protocols to be applied when the MPEG-4 Systems format is used for multiplexing must be reviewed further.

Conclusion of communication connection protocols
- SIP and H.323 are protocols with future potential.
- SIP is more expandable and flexible than H.323. However, SIP might have problems regarding mutual operation performance.
- The detailed specifications of SIP will take shape by the end of 2002. The products that support SIP standards have already started coming out in the market. We should pay attention to the future market trends.
Chapter 5

- TVAF: Specification Series: S-3 On: Metadata (Normative), SP003v1.1, Aug. 2001
- TVAF: Specification Series: S-4 On: Content Referencing (Normative), SP004v11, April 2001
- IETF RFC2326: Real Time Streaming Protocol (RTSP)
- IETF RFC1890: RTP Profile for Audio and Video Conferences with Minimal Control
- IETF RFC1112: Host Extensions for IP Multicasting
- IETF RFC2710: Multicast Listener Discovery (MLD) for IPv6
- IETF RFC2117: Protocol Independent Multicast-Sparse Mode (PIM-SM)
- IETF RFC1075: Distance Vector Multicast Routing Protocol
- IETF RFC1584: Multicast Extensions to OSPF
- IETF RFC2189: Core Based Trees (CBT version 2) Multicast Routing
- IETF RFC1633: Integrated Services in the Internet Architecture: an Overview
- IETF RFC2206: RSVP Management Information Base using SMiv2
- IETF RFC2207: RSVP Extensions for IPSEC Data Flows
- IETF RFC2208: Resource ReSerVation Protocol (RSVP) Version 1 Applicability Statement Some Guidelines on Deployment
- IETF RFC2209: Resource ReSerVation Protocol (RSVP) -- Version 1 Message Processing Rules
- IETF RFC2475: An Architecture for Differentiated Services
Chapter 5

Chapter 6
Chapter 9

- Peter Kent, John Kent: Official Netscape JavaScript, BNN
- ISO/IEC 13818-1, "Information technology - Generic coding of moving pictures and associated audio information - Part1:Systems"
- ISO/IEC 13818-1, "Information technology - Generic coding of moving pictures and associated audio information - Part2:Video"
- ISO/IEC 13818-1, "Information technology - Generic coding of moving pictures and associated audio information - Part3:Audio"
- ISO/IEC 13818-1, "Information technology - Generic coding of moving pictures and associated audio information - Part4:Conformance"
- ISO/IEC 13818-7, "Information technology – MPEG-2 Advanced Audio Coding(AAC)"
Chapter 9