Unified Multimedia Transmission Protocol for Multimedia Conference System

Jun Sato, Hideki Mori
Dept of Information and Computer Science
Toyo University
2100 Kujirai, Kawagoe, Saitama
Japan, 350-0815
jun@sb.cs.toyo.ac.jp, mori@mo.cs.toyo.ac.jp

Michiaki Katsumoto
Communication Research laboratory
Ministry of Posts and Telecommunications
4-2-1 Nukui-kita, Koganei, Tokyo
Japan 184-8795
katumoto@crl.go.jp

Koji Hashimoto, Yoshitaka Shibata
Faculty of Software and Information Science
Iwate Prefectural University
152-52 Sugo, Takizawa, Iwate
Japan 020-0173
{hashi, shibata}@iwate-pu.ac.jp

Abstract

In order to realize general multimedia information network applications on the same system such as combination of video-on-demand and multimedia conference, it is required to support variety of presentation types, such as stored multimedia presentation, captured realtime presentation(live presentation), fine grained and scene synchronizations, etc. Since, these presentation technologies have developed separately, it is very difficult to realize the more practical application such as multimedia conference system which integrates the both stored and live presentations at the same time. Therefore, unified protocol for these presentations provided in the practical multimedia application is required. In this paper, we propose the unified multimedia transmission protocol and the layered protocol architecture to realize the multimedia applications. In addition, we design new presentation control functions which are organized independent controllers: including service controller(SC), presentation controller(PC) and media controller(MC) based on the proposed hierarchical protocol architecture to support variety of presentation types. As one of possible application example based on our proposed architecture, we design a multimedia conference system and its flow of the conference, which provide both stored and live presentations at the same time.

1. Introduction

The realization of multimedia information network applications such as video-on-demand(VoD), multimedia conference and electronic museum, are being expected as the development of the high speed network and high speed compoter technologies. However, it is required to support the variety of presentation types at the same time on the multimedia applications. For an example, VoD system provides the stored media presentation which is consisted of both audio/video streams stored and transmitted by taking fine grained media synchronization from the distributed DB servers to user station. On the other hand, a multimedia conference system must provide not only live audio/video presentation which is generated from user station in realtime, but also stored media presentation as the reference data on the conference at the same time. Furthermore, the electronic museum system must provide high quality presentation which is integrated by different type of media streams such as audio, video, text and image based on the presentation scenario. In addition, the system must also perform hypermedia functions to navigate from the current presentation to another presentations. Historically, these presentation technologies has been separately developed for each application, and attention was not paid for unifying these presentations even though such unified presentation method is very important particulary when practical multimedia application is taken into consideration. This requires a new multimedia transmission protocols and functions which unify variety of the presentation type.

Furthermore, in order to realize realtime communication using the multimedia conference system, media synchronization[1], media transform, media flow and rate controls[2], and End-to-End QoS guarantee[3][4] functions are also required. So far, we have already investigated the layered protocol architecture named Packet Audio/Video System(PAVS), and designed and implemented audio/video synchronization[6], dynamic rate control[7] and QoS guarantee control[8] methods based on proposed protocol architecture. However, our PAVS architecture supported just only the stored audio/video presentation, but did not support variety of presentation types such as both stored and live presentations.

In this paper, we reorganize the previous PAVS architecture into a unified multimedia transmission protocol architecture which provides the variety of presentation types uniformly and integrates these presentations. In this protocol architecture, presentation control functions including three new controllers: **service controller(SC), presentation controller(PC)** and **media controller(MC)** are introduced. The SC controls multiple presentations concurrently in multimedia application. The PC controls multiple media streams, which consist of one presentation. The MC performs the sophisticated media transmission functions such as synchronization, media transform and media flow control functions. In another words, several MCs are managed by one PC to carry one presentation, and several PCs are managed by one SC to unify the several presentations based on proposed protocol architecture. Thus, using protocol architecture function, more flexible and more general presentations including live and stored media presentations, fine grained and scene synchronization, and scenario based and interactive presentations are concurrently provided. As a concrete example of application based on our proposed protocol architecture, we design the multimedia conference system which provides both stored and live presentations and its flow of the conference to verify the functionality and performance of our suggested protocol architecture.

In section 2 the layered protocol architecture based on the unified multimedia transmission protocol are proposed, and the presentation control functions organized by three new controllers based on the protocol architecture are explained in section 3. In section 4, 5 and 6, three new controllers are explained respectively. Furthermore, multimedia conference system as one of actual example of application based on our new protocol architecture is designed in section 7.

2. Unified Multimedia Transmission Protocol Architecture

The layered architecture of our proposed protocol is shown in the Figure 1.

In this architecture, we have introduced the three new function layers: synchronization, media transform and media flow control layers between application and transport layers in the OSI reference model, to correspond to provide flexible presentation capability. The synchronization layer makes various media synchronization[6] including the inter/intra media synchronizations between the continuous as well as discrete media. The media transform layer performs media format conversion, media compression/decompression. And the media flow control layer performs packet flow control and packet loss control according to the load condition of the used computers and networks[7]. Besides of those functions abobe, each layer has the QoS guarantee functions considering resource management, QoS mapping and admission control to maintain the End-to-End QoS. The detail of these functions is explained in the paper [4][8].

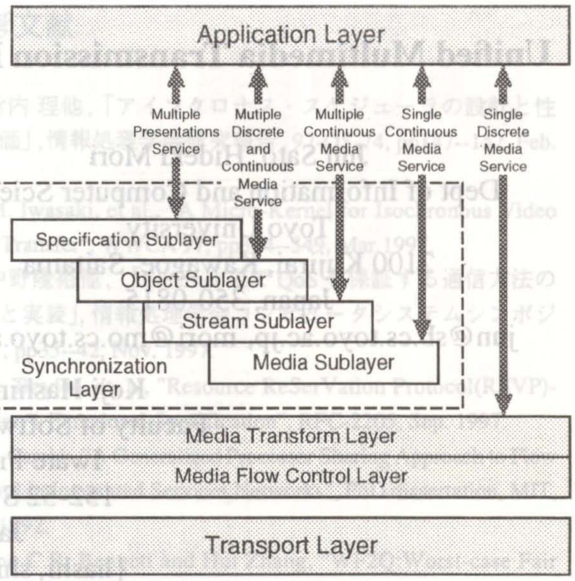

Figure 1. Protocol Architecture

Furthermore, we divide the synchronization layer into four sublayers to flexibly support variety of service types as follows[1]:

Media Layer: Intra-media synchronization within a single continuous media stream such as audio or video is performed.

Stream Layer: Lip(fine grained) synchronization between the related audio and video is performed.

Object Layer: Scene synchronization which integrates the different type of media stream such as audio, video, image and text is performed based on the presentation scenario.

Specification Layer: Multiple presentations are controlled to realize sophistcal multimedia application. For example, the stored presentation and the live presentation are integrated into a presentation window for realtime communication while hypermedia functions in onother presentation window is used interactively.

In addition, we define the protocol data units which are delivered between each layer independently to execute the functions required in each layer, as shown in Figure 2.

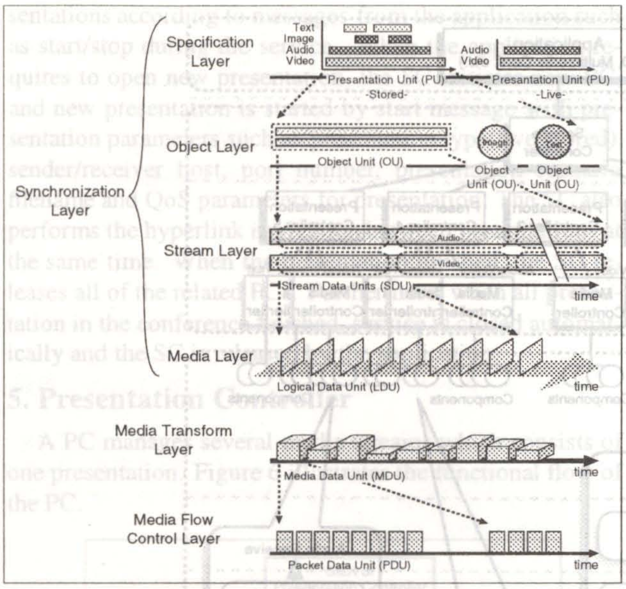

Figure 2. Protocol Data Unit

Packet Data Unit(PDU): The data unit which is delivered between media flow control layer and transport layer. It is equivalent to one packet transmitted on the network.

Media Data Unit(MDU): The data unit which composes the media data, and is handled under the media transform layer. It is equivalent to a compressed audio/video frame.

Logical Data Unit(LDU): The data unit which also composes the media data, and is handled over the media transform layer. It is equivalent to a decompressed audio/video frame.

Stream Data Unit(SDU): All of LDUs in a single media, between the synchronization point at which lip synchronization between audio and video is taken.

Object Unit(OU): The whole of single media such as audio, video, text and image.

Presentation Unit(PU): It is equivalent to one presentation, which is organizaed by multiple object units(OU).

3. Presentation Control Functions

In this section, presentation method in the multimedia conference is discussed to design the presentation control functions of our proposed protocol architecture to support variety of presentation types, and describe the functional module which are consisted by three new controllers, **service controller(SC)**, **presentation controller(PC)** and **media controller(MC)**.

3.1. Presentations in the Multimedia Conference

Multiple presentations are provided at the same time in the most of multimedia services. For a typical example, it is simply assumed that a multimedia conference is held between the three users, A, B and C, as shown in Figure 3.

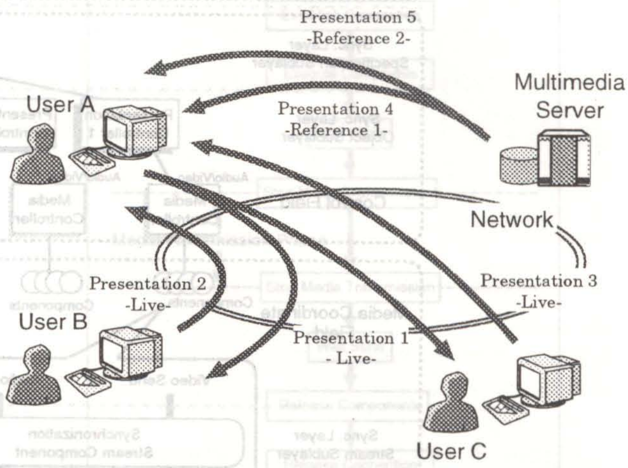

Figure 3. Presentations in the Multimedia Conference

Here, the following presentations which is handled by user A during the conference is assumed as follow: 1) Live audio/video presentation provided by user A to user B and C, using multicast technique. 2) Live audio/video presentation provided by user B. 3) Live audio/video presentation provided by user C. 4) Stored multimedia presentation provided by as the reference data used for the conference. 5) Stored multimedia presentation including the detail information of presentation 4). The presentations 4) and 5) are related using hyperlinking. The Figure 4 indicated functional module configuration when these five presentation are used at the same time.

3.2. Functional Module

We classify these functional modules illustrated by Figure 4 into the three fields: application field, control field and media coordinate field.

Application Field: The application field is located at the application layer on the protocol architecture shown in the Figure 1. The functions needed for multimedia conference application are existed in this field. For an example, in the multimedia conference system, QoS negotiation between user stations and floor control functions for conference are performed in this application field. Furthermore, the application manages SC which exists in the lower layer and informs the presentation control message such as start/stop and user's QoS requirement messages to the SC.

Control Field: The control field is located between the specification layer and object layer on the protocol architecture, and includes the three controllers: the SC which manages multiple presentations concurrently for service, the PC which manages several media streams

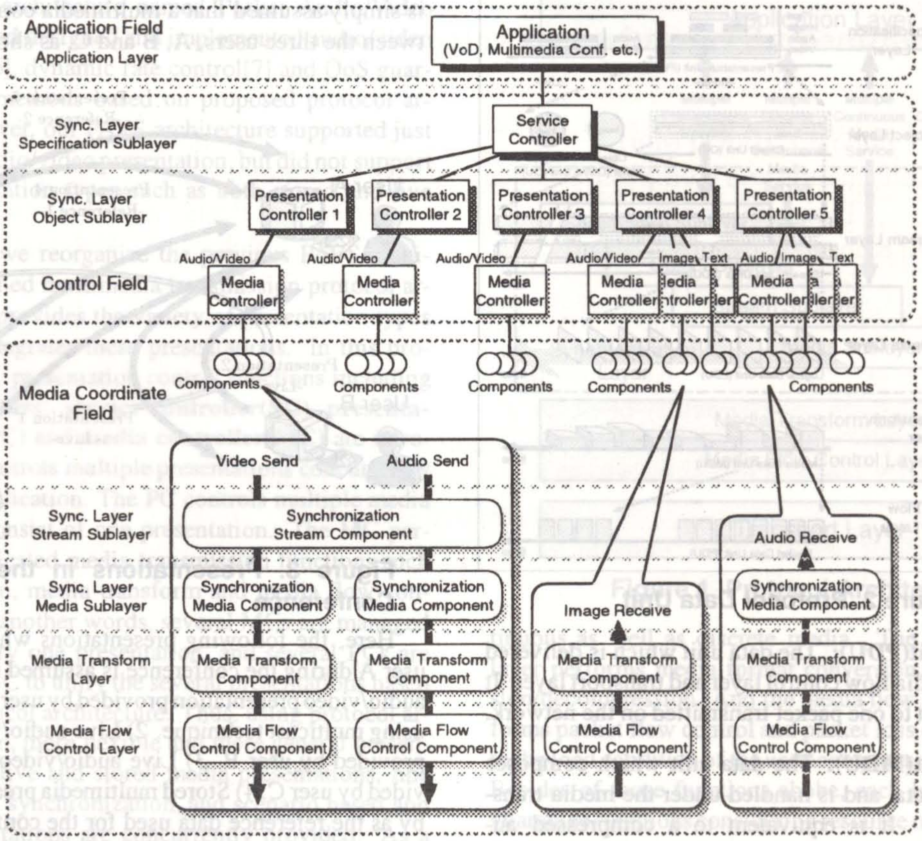

Figure 4. Functional Module

consisting one presentation, and the MC which manages the functions to execute single media transmission. A SC informs the presentation control message such as start/stop and QoS requirement messages to the PCs depending on the requirement from application. The SC also controls context between each presentation. For an example, navigation between these presentations based on the hyperlinking is performed. A PC handles one presentation based on the interactive operation by user such as presentation start, stop, rewind and forwarding and manages multiple MCs to realize the presentation. A MC manages multiple components for synchronization, media transform, media flow control to realize multimedia transmission. The detail functions of each controller will be explained in section 4, 5 and 6.

Media Coordinate Field: The media coordinate field has several components such as lip(inter-media) synchronization component, intra-media synchronization component, media transform component, media flow component and device component. These components are selected by the MC according to user's QoS requirements and the presentation type.

4. Service Controller

A SC manages multiple presentations in parallel. Figure 5 illustrates the functional flow of the SC.

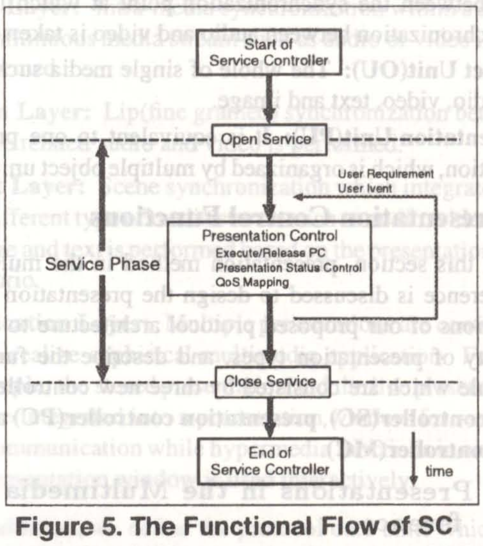

Figure 5. The Functional Flow of SC

When a service is initiated by user, a SC is immediatly executed by the application. The SC controls multiple presentations according to messages from the application such as start/stop during the service. When the application requires to open new presentation, the SC initiates new PC, and new presentation is started by start message with presentation parameters such as presentation type(live/stored), sender/receiver host, port number, presentation scenario filename and QoS parameters for presentation. The SC also performs the hyperlink navigation between presentations at the same time. When the presentation finished, the SC releases all of the related PCs. Furthermore, when all presentation in the conference finished, service is closed automatically and the SC is released by the application.

5. Presentation Controller

A PC manages several media streams which consists of one presentation. Figure 6 illustrates the functional flow of the PC.

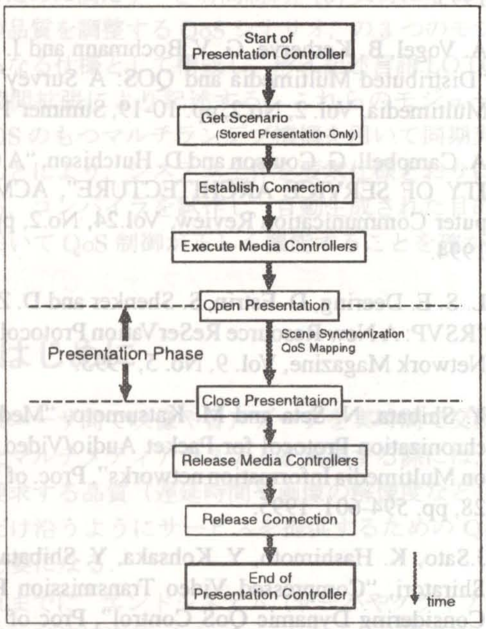

Figure 6. The Functional Flow of PC

When a PC is initiated by the SC, the PC executes multiple MC according to media streams required for the presentation. The presentation is started as soon as preparation of each MC is completed. During the presentation, the PC controls each MC to execute scene synchronization. Finally, when all of the media streams transmission are completed the PC closes the presentation and releases each MC.

6. Media Controller

A MC manages several components for media transmissions. Figure 7 illustrates functional flow of the MC.

When a MC is executed by PC, the MC selects multiple components such as media synchronization, media transform and media flow control components according to user's

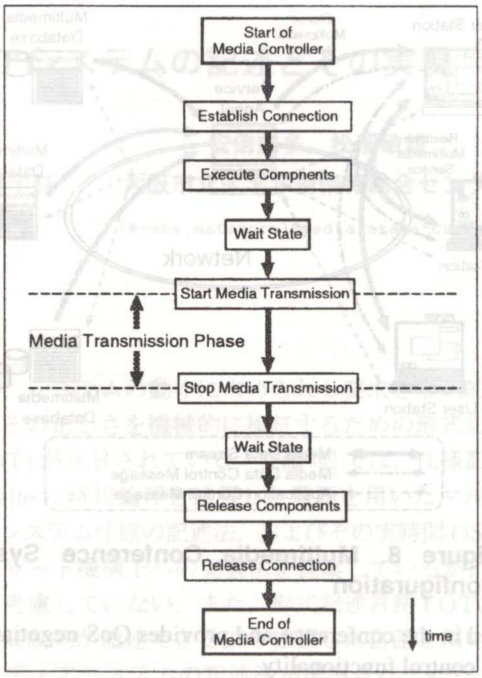

Figure 7. The Functional Flow of MC

QoS requirements and presentation type and establishs connections between these components to transmit the media data. After the preparation of each component is completed, the MC waits until media transmission message comes from PC to start media transmission. When the message comes from PC, the MC executes components to actualy transmit media data actually. After the media transmission finished, the MC waits until the message comes from PC which manages the MC, to release the components and has connections.

7. Multimedia Conference System

In this section, we describe the multimedia conference system as a concrete example of application based on our suggested protocol and message flow of the multimedia conference system.

In this multimedia conference system, while the live audio/video presentations are provided between users, stored media multimedia presentation are also provided as the reference data simulatanously. Each media stream used for the stored media presentation are distributed in the databases on the network. In order to realize these multimedia conference system, we introduce following structure model consisted of **user stations(US), service agents(SA)** and **multimedia data bases(MDB)** as shown in Figure 8.

The US provides the presentations to user and captures the live audio/video and transmits then to all of the other users. The MDB store several media data as reference data used for the conference, and each media data are managed by a SA. The SA also manages the user information partic-

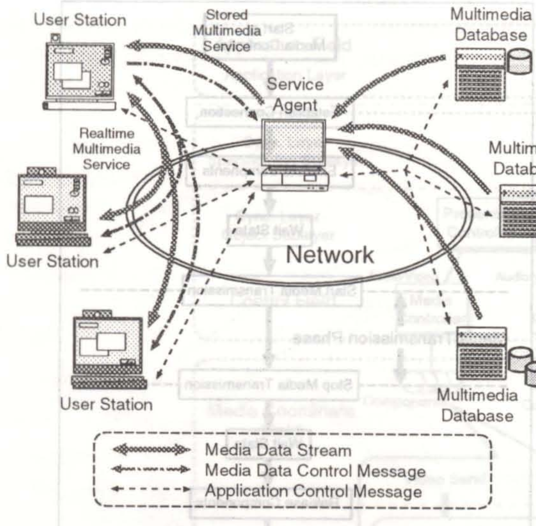

Figure 8. Multimedia Conference System Configuration

ipated in the conference and provides QoS negotiation and floor control functionality.

8. Conclusions

In this paper, we proposed a unified multimedia transmission protocol architecture and the presentation control functions based on three controllers, SC, PC and MC on the proposed protocol architecture to support variety of presentation types such as both stored and live presentation types. Furthermore, we designed the multimedia conference system as one of actual example of application based on our unified protocol architecture. In the multimedia conference system, both live and stored presentations and several presentations are supported and integrated on the user station at the same time.

Currently, we are prototyping the multimedia conference system in the Communication Research Laboratory(CRL) in the Ministry of Posts and Telecommunications where live audio/video presentation is alredy realized to be able to hold in multimedia conference within the lab. In the prototype system, we implement the USs, SA and MDBs on several Silicon Graphics Workstations by C language, and a number of controllers and components are realized in parallel using POSIX Thread technology.

In the near future, we will also realize the stored presentation as the reference data in addition to the live audio/video presentation and the QoS guarantee functions based on the RSVP[5] protocol. More practical conference will be held among three locations, CRL(Tokyo), Toyo University(Saitama) and Iwate Prefectural University(Iwate) in Japan.

References

[1] G. Blakowski, R. Steinmetz, "A Media Synchronization Survey, Reference Model, Specification, and Case Studies", IEEE JSAC, Vol.14, No.1, pp.5-35, January 1996.

[2] H. Kanakia, P. P. Mishra and A. Reibman, "An Adaptive Congestion Control Scheme for Real-Time Packet Video Transport", ACM SIGCOMM'93, pp. 20-31, 1993.

[3] A. Vogel, B. Kerherve, G. V. Bochmann and J. Gecsei, "Distributed Multimedia and QOS: A Survey", IEEE Multimedia, Vol. 2, No.2, pp. 10-19, Summer 1995.

[4] A. Campbell, G. Coulson and D. Hutchison, "A QUALITY OF SERVICE ARCHITECTURE", ACM Computer Communication Review, Vol.24, No.2, pp. 6-27, 1994.

[5] L. S. E. Deering, D. Estrin, S. Shenker and D. Zappala, "RSVP: A New Resource ReSerVation Protocol", IEEE Network Magazine, Vol. 9, No. 5, 1993.

[6] Y. Shibata, N. Seta and M. Katsumoto, "Media Synchronization Protocol for Packet Audio/Video System on Multimedia Information networks", Proc. of HICSS-28, pp. 594-601, 1995.

[7] J.Sato, K. Hashimoto, Y. Kohsaka, Y. Shibata and N. Shiratori, "Compressed Video Transmission Protocol Considering Dynamic QoS Control", Proc of ICPP98 Workshop, pp. 95-104, 1998.

[8] K. Hashimoto and Y. Shibata, "Performance Evaluation of End-to-End QoS Using Prototyped VOD System", Proc. of ICOIN-12, pp.175-178, January 1998.