Adapting H.323 Terminals in a Service-Oriented Collaboration System

Wenjun Wu[1] Hasan Bulut[1] Ahmet Uyar[2] Geoffrey Fox[1]
[1]Community Grid Computing Laboratory, Indiana University
[2]Department of Electrical Engineering and Computer Science, Syracuse University

Abstract

Global Multimedia Collaboration System (Global-MMCS), a scalable, robust and service-oriented collaboration system, can bridge H.323, SIP, Access Grid clients and 2.5G/3G cellular phones in audiovisual collaborations. The H.323 gateway in Global-MMCS enables H.323 terminals to interact with other clients, and provides them the complete H.323 conference control services.

Keywords

1. Introduction

Collaboration and videoconferencing systems have become a very important application in the Internet. There are various solutions to such multimedia communication applications, among which H.323 [1], SIP [2] and Access Grid [3] are well-known. It will bring substantial benefits to Internet users if we can build an integrated collaboration environment, which combines these systems into a single easy-to-use, intuitive environment.

H.323 defines a set of tightly integrated protocols (H.225 [4], H.245 [5], and H.243 [6]) for call signaling and session management. Most H.323 conferencing systems depend upon a centralized or cascaded MCUs for media delivery, processing and session management. In contrast, Access Grid relies upon Internet multicast to distribute media data and uses a much more loosely coupled control mechanisms.

XGSP (XML based General Session Protocol) [7] is a common, interoperable framework based on Web services technology for creating and controlling multipoint collaborations. XGSP uses a unified, scalable, robust overlay network to support audiovisual and data group communication over heterogeneous networking environments. XGSP offers a distributed, flexible conference management mechanism for integration of various collaboration communities. Moreover, XGSP specifies a common audiovisual signaling protocol for interactions between different audiovisual collaboration endpoints. Just like the text messages in SIP, XML is used in XGSP to describe the common signaling protocol because it makes the protocol easier to be understood and to interact with other Web based components.

Based on this framework, we have developed a prototype system called Global-MMCS (Global Multimedia Collaboration System) [8] to support interoperable collaborations, including videoconferencing, instant messaging and streaming. In terms of multipoint audiovisual collaboration, Global-MMCS can be regarded as a scalable and service-oriented virtual MCU (Multipoint Control Unit) for bridging H.323 terminals, Access Grid clients and 2.5G/3G cellular phones capable of image uploading in the same audiovisual sessions.

Typical solution of centralized H.323 MCU is not scalable as overlay based and service oriented approach. Global-MMCS separates the MCU services into separate service modules: MC (Multipoint Controller) and MP (Multipoint Processor). And MP is further divided into two major tasks: media distribution and media processing such as mixing and transcoding. Global-MMCS uses NaradaBrokering [9] as media distribution overlay. NaradaBrokering is an event brokering middleware, which supports publish-subscribe messaging model with a dynamic collection of brokers. We also provide a general and scalable service platform (Media Servers) that can execute various media processing. Therefore Global-MMCS MP can provide rich and flexible media processing services, communicating with H.323 terminals and Access Grid with both unicast and multicast RTP communication.

Furthermore, H.323 MCUs only handle one or two protocols (H.323, SIP), but this virtual MCU supports H.323 terminals, Access Grid and wireless clients by flexible adapting and filtering mechanisms. The Global-MMCS MC implements the XGSP conference control framework to provide the audiovisual session management for H.323 terminals and Access Grid separately. As a key component in this virtual MC, the H.323 Gateway offers H.323-style control services to H.323 terminal by bridging conference control service between H.323 and XGSP.

2. H.323 Gateway Services

Figure 1 shows the architecture of the Global-MMCS and the services of the H.323 Gateway. The NaradaBrokering overlay network provides the group communication services to all the components. For

the transportation of RTP packets over the NaradaBrokering network, RTP packets have to be encapsulated into special NaradaBrokering events named "RTP event" by *RTPLink*s in the edge brokers before they can be routed in the overlay network. In order to connect Access Grid audiovisual sessions, Global-MMCS creates a multicast RTPLink for each active Access Grid RTP session. For H.323 terminals, Global-MMCS also creates unicast *RTPLink*s for audio and video H.245 RTP channels.

Unlike Access Grid clients (VIC & RAT) which are capable of receiving, mixing and rendering multiple audio and video streams, some H.323 terminals like video phone can only handle a single video/audio stream. Therefore H.323 terminals need mixing services provided by MediaServers. A MediaServer acts as a factory for media service providers, creating media service instances and allowing the instances to run in a service container. All service providers need to implement the interface required by the container. Currently, we implemented three types of media service providers: AudioMixerServer, VideoMixerServer, and ImageGrabberServer. ImageGrabberServer can transcode video streams into JPEG images and distribute these video snapshots to clients so that they can have a visual stream list for video switching. More service providers can be added by following the guidelines and implementing the relevant interfaces. Each MediaServer is independent of other MediaServers and new ones can be added to cope with the increase of media processing overhead.
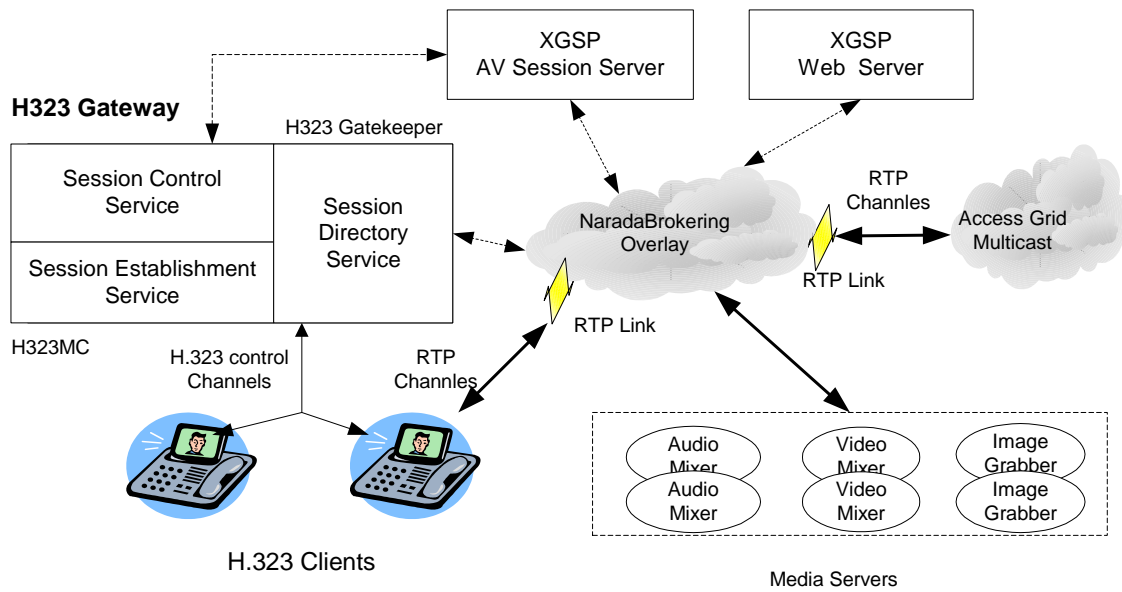

Figure 1. H.323 Gateway Services

The XGSP AV Session Server implements the XGSP audiovisual session services including: join and leave XGSP AV session, video switching and floor control as well as the management of media processing services. Working with the XGSP AV Session Server, two functional components in the H.323 Gateway: H.323 Gatekeeper and H.323 MC provide the services including session directory service, session connection service and session control service to bridge H.323 terminals into XGSP audiovisual sessions.

2.1 Session Establishment Service

The session establishment service of H.323 Gateway enables H.323 terminals to join or leave XGSP audiovisual sessions. XGSP framework specifies the procedure about how to create and destroy public and private XGSP audiovisual sessions. If a H.323 terminal wants to join an activated XGSP session, it has to make a H.225 call to the H.323 gateway with the right *session ID*. The session directory service shall forward this call to the registered H.323 MC. H.323 MC processes the call and translates it into the XGSP AV session join procedure.

Three pieces of information are needed for establishing a call between two endpoints, namely the signaling destination address, local & remote media capabilities, and local & remote media transport addresses at which the endpoint can receive the media packets. H.323 spreads them over different stages of

the call setup, while XGSP puts them into three different join session XML messages. Figure 2 illustrates the translation between the H.323 call and the XGSP procedure.

H323 Terminal   H323 Gateway

H.225 Call Setup

H225.Setup
JoinAVSession

H225.Connect
JoinAVSession OK
with the RTPLinks <IP Addr, Port>
& capability description

H.245 Capability Exchange

Terminal Capability Set
ACK
Terminal Capability Set
ACK

Open Audio and Video Logic Channels

OpenLogicChannel (Video)
ACK
ACK with video RTPLink
<IP Addr, Port>
OpenLogicChannel (Video)
ACK
JoinAVSession (Video)

OpenLogicChannel (Audio)
ACK
ACK with Audio RTPLink
<IP Addr, Port>
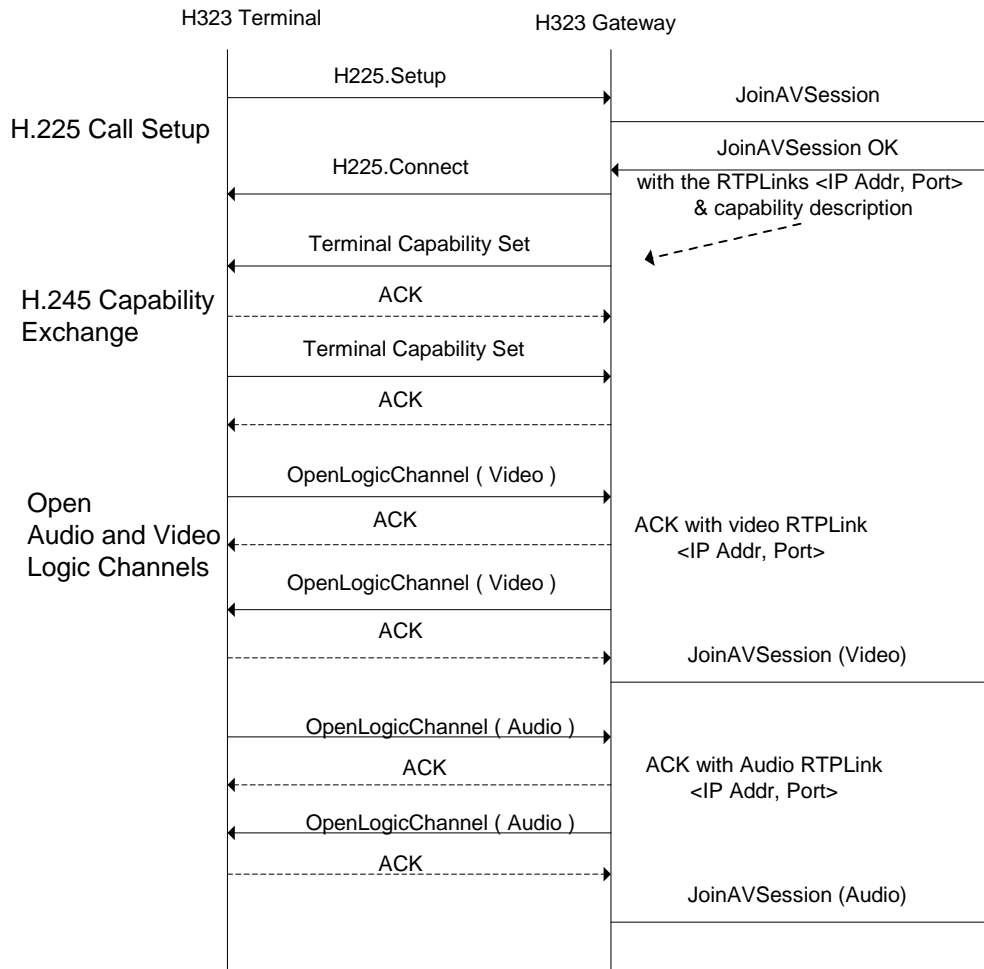OpenLogicChannel (Audio)
ACK
JoinAVSession (Audio)

Figure 2. The translation of H.323 call into XGSP procedure

This join session procedure has three important steps: H.225 call setup, H.245 capability exchange and audiovisual logical channel creation.

(1)H.225 Call Setup

A H.323 terminal sends H.225 setup message with the signaling destination address that represents *SESSION ID* for this meeting. H.323 gateway parses the H.225 setup message, and sends "*JoinAVSession*" with the *SESSION ID.*   This "Join-AV-Session" informs the XGSP AV Session Server that a H.323 terminal wants to join the session. XGSP AV Session Server determines whether the request of the H.323 terminal can be allowed and fill the information of audiovisual RTPLinks of the broker in the "Join-AV-Session-OK" message. If the H.323 gateway gets the *"OK"* response, it sends H.225 *CONNECT* to complete the H.225 procedure. And it keeps the RTP channel information < IP Address, port > for the phase of OpenLogicChannel.

(2)H.245 Capability Exchange

Each H.323 terminal has its H.245 media capability description. The establishment of a H.323 conference involves the procedure of common media capability negotiation, which requires the MC to send a capability set to the endpoints in the conference. We use SDPng [10] for the capability description of Global-MMCS media services, which can be mapped into H.245 media capability description. SDPng uses XML to describe the session media information and is quite appropriate for the capability exchange between XGSP AV Session Server and H.323 terminals.

Access Grid clients don't use capability negotiation because they can handle streams in different codecs dynamically. However a H.323 terminal can't support this because it depends upon H.245 media negotiation procedure to get the codec agreement. Therefore Global-MMCS has to provide some transcoder services between Access Grid clients and H.323 terminals.

(3)Open Audio and Video Logic Channels

Following the exchange of capabilities and master-slave determination, the procedures of H.245 shall be used to open logical channels for the audiovisual streams. After the H.323 Gateway gets the *OpenLogicChannelACK*, it retrieves the transport address of the H.323 terminal and sends "*JoinAVSession (video)*" and "*JoinAVSession (audio)*" with the address information to the Session Server. The Session Server creates transmitting RTPLinks in the NaradaBrokering brokers for the incoming H.323 terminal when it gets the "*JoinAVSession*" request from the H.323 gateway. After that, the H.323 Gateway waits for the *JoinSession.OK* message to get the transport address of the receiving RTPLinks < IP Address, UDP Port Number > and includes the information into *OpenLogicalChannelACK* to notify the requesting H.323 terminal.

2.2 Session Directory Service

Collaboration users need to know how many active sessions are available and associated information. The conference announcement can be implemented either in the session protocol or through the out-of-band communication. XGSP framework divides the conference advertisement information into two levels: one is the collaborative conference calendar, which describes high-level meta-data about the organization of the conference including meeting time, meeting topic and so on. The other is the detailed information for audiovisual clients to join the conference, for example the session identification in the system and transport addresses associated with the session. The high-level conference calendar is actually implemented as a web-services running in the XGSP web server. Each active entry in the calendar has a pointer to the detailed session description. H.323 users can firstly visit this conference calendar and then start their H.323 calls using the detailed information.

The session directory service provided by the H.323 Gatekeeper keeps the registration of H.323 terminals and the alias name of active sessions. Since there are more than one audiovisual sessions in Global-MMCS for H.323 terminals, the H.323 Gatekeeper must keep the session alias list for the calling transport address of the H.323 MC. In addition, new AV sessions can be created and activated at any time by the conference participants using the XGSP protocol. H.323 gateway listens to these events through the NaradaBrokering and updates its session alias list. When a H.323 terminal wants to join a XGSP AV session, it will call the H.323 MC with the session alias name. The H.323 Gatekeeper shall translate this alias into the calling address and route the conference call to the H.323 MC.

The gatekeeper can also be used to enforce the admission control for H.323 terminals. For example, we can only allow a H.323 terminal to join a XGSP session if the owner of the terminal has logged into Global-MMCS and get authenticated. Each time when a H.323 terminal makes a registration at the Gatekeeper, it should generate a unique client ID for this terminal in Global-MMCS. The format of the client ID is h323:user@host, in which user identifies the owner of the terminal and host is the DNS name of the H.323 terminal. When this terminal tries to join the XGSP AV session after the registration, the gatekeeper can check whether the user of this terminal has been activated and give the Call Accepted or Call reject response.

2.3 Session Control Service

Audiovisual session control basically involves the following tasks: creating/destroying sessions, allowing endpoints to join/leave session and make audiovisual selection, managing audiovisual service elements and so on. In our previous research work on XGSP [11], we implemented it in two function components named audiovisual session management and floor control. XGSP audiovisual session management defines the basic control procedure, while audiovisual floor control defines the roles in the audiovisual session management and how to enforce the resource management based on the roles.

Our approach is to use point-to-point H.245 control channels between H.323 terminals and the H.323 Gateway. On the top of H.245 control channels, our XGSP service allows H.323 users to vote for meeting

chair, request floors, and make audiovisual mixing and switching. The session control service is implemented by the H.323 MC in the Gateway. Figure 3 shows the components involved in XGSP session control for H.323 terminals.
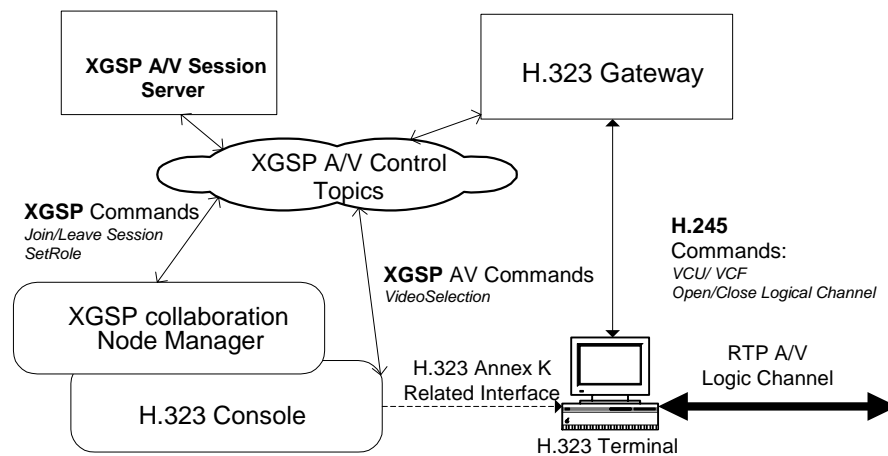

Figure 3. XGSP H.323 Session Control Architecture

The XGSP Collaboration Node Manager is the user interface for the XGSP conference management service. Every Node Manager can create collaboration endpoint instances, and invoke the methods of starting, stopping and role setting provided by the instances. Particularly, it can launch a H.323 console, a XGSP audiovisual control client for H.323 users. The XGSP session control service involves the communication among the Node Managers, the H.323 consoles and the XGSP AV Session Server. The H.323 Gateway intercepts the XGSP messages and translates the procedure into H.245 control procedure.

H.323 Annex K allows third party control of an H.323 call based on a HTTP channel for user interaction. Vendors of H.323 terminals have to define a custom interface between H.323 terminals and web browsers to support H.323 Annex K. Our H.323 consoles can use this interface to communicate with H.323 terminals. When a user tries to join a XGSP AV session using a H.323 terminal, he need to launch a H.323 console from his XGSP Node Manager at first. Through the Annex K related interface, the H.323 console asks the H.323 terminal of the user to start a conference call to H.323 gateway. The user can also use the H.323 console to ask the H.323 terminal to end the conference call and leave the session.

Video switching enables a H.323 terminal with the capability of single video rendering to select many streams in the XGSP session. The XGSP AV Session Server keeps a video stream list for the activated session. And the H.323 console can copy the list and allow the user to make the video selection. Upon the selection, the H.323 console sends a "*XGSP-Video-Selection*" command to the Session Server to subscribe a new video stream and unsubscribe the old one. The H.323 Gateway also receives the command and follows the procedure of H.243 to perform video switching for the connected H.323 terminal. This procedure involves two H.245 commands: *videoFreezePicture* (VCF) and *videoFastUpdatePicture* (VCU). VCF commands the video decoder to complete updating the current video frame and subsequently display the frozen picture until receipt of the freeze-picture release control signal. And VCU commands the far-end video encoder to enter fast-update mode at its earliest opportunity and send an I-frame to the video decoder. After a H.323 user makes a video selection, the H.323 Gateway shall send a VCF command to the H.323 terminal for decoder freezing. Furthermore, when a video stream from a H.323 terminal is selected, the H.323 Gateway will send a VCU to that terminal for fast-update.

Each user in a XGSP AV session may have different roles. For an audio session, a user can be either a speaker or listener who isn't allowed to send audio stream. For a video session, a user can be either a sender or viewer who isn't allowed to send video stream. The XGSP conference chair has the power of setting different roles to users in a public AV session. The XGSP conference chair sends "*SetRole*" messages to other users through its XGSP Node Manager. The destination Node Manager parses the message and notifies the local H.323 console of the action of role setting. The H.323 console indicates the new role to the destination user. At the same time, the H.323 Gateway also gets the message and starts a "Request Open/Close Logical Channel" H.245 procedure to open or close the audiovisual logical channels.

2.4 Distributed H.323 Gateways

It's well known that both firewalls and NATs are serious obstacles to the deployment of Internet conferencing. The packet filtering policies in firewalls are usually static on IP/transport layer whereas H.323 signaling messages convey dynamic IP addresses and UDP port numbers in their H.245 payloads. Therefore firewalls will block the messages of H.245 channels so that H.323 call signaling procedure and audiovisual data communication will fail. Similar issues affect NAT devices. The H.323 terminals behind a NAT use the inside IP address in H.245 messages which also lead to the failure of building H.323 calls.

As an overlay networking on Internet, NaradaBrokering can provide communication channels through the firewall and NAT boundaries. On top of the NaradaBrokering overlay network, a distributed H.323 Gateway architecture is introduced to address the issue of firewalls and NATs. For each administration domain behind a firewall or NAT, we setup a local H.323 Gateway and a NaradaBrokering node.
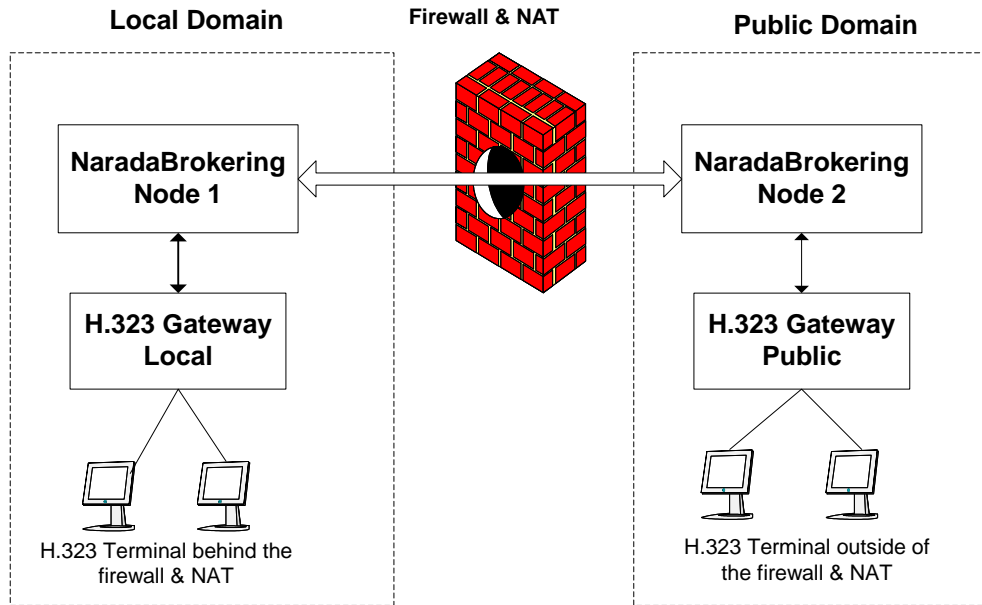

Figure 4. Distributed H.323 Gateways for different administration zones

This local NaradaBrokering broker builds a channel across the firewall or NAT with the outside public broker. This channel carries all the traffic of Global-MMCS between the local administration domain and the public domain. The local H.323 Gateway uses the XGSP protocol to communicate the XGSP AV Session Servers. We can have one or multiple XGSP AV Session Servers. One of them can be placed inside the local administration domain if necessary. H.323 terminals in this domain can still use H.225 and H.245 protocols to connect to the local H.323 Gateway. And their RTP/RTCP data can be routed by the local NaradaBrokering node to the whole NaradaBrokering overlay.

Since each H.323 Gateway only keeps the state from H.323 terminals in its own administration zone, there is no state consistency issue for H.323 Gateways. They just need to keep their own replica of some global information from XGSP conference servers. This distributed H.323 architecture not only solves the problems of firewall but also improves the scalability of H.323 services because each H.323 Gateway solely deals with limited workload generated by local H.323 terminals.

3. Implementation and Related Work

The H.323 Gateway is implemented on the base of protocol stack provided by OpenH323 [12]. The H.323 Gateway has two servers: H.323 Gatekeeper and H.323 MC. We modified GNU Open Gatekeeper [13] to support the Session Directory Service. Our H.323 Gatekeeper can get the list of logged users in the system and the ownership list from web server, and perform the security checking based on the information. We also created three functional components for the H.323 MC, including the Session Event Listener, Session Connection Service Handler and Session Control Handler.

The Session Event Listener is responsible for the registrations of XGSP audiovisual sessions, intercepting the session events in the XGSP control topic and exchanging RAS messages with H.323 Gatekeeper. The Session Connection Service Handler working with the H.225 handler in the OpenH323

library, implements Session Establishment Service. Instead of creating real H245 logical channels between the H.323 MC and H.323 terminals, it builds external RTP channels that are RTPLinks between H.323 endpoints and the edge brokers of the NaradaBrokering. The Session Control Service Handler supports video switching and floor management.

During the development of H.323 MC, we find that the video-switching for H.323 terminals between video streams from Access Grid is more than supporting H.243 control procedure. There is a mismatch between the video decoder of H.323 terminal and the encoder of Access Grid VIC. The H.261 encoder of VIC only generates I-frame with motional macro blocks, and doesn't set the "Freeze Picture Release" bit in H.261 bitstream [14].

When a H.323 terminal for example Polycom ViaVideo, makes a switch from one VIC H.261 stream to the other, its H.261 decoder cannot generate the new video image for quite a long time because it has been frozen. In order to address this issue, we introduce a packet filter which can parse the H.261 bitstream and set the "Freeze Picture Release" to "1". The packet filter runs as an embedded component inside the RTPLinks of the NaradaBrokering and only servers H.323 terminals. In addition, H.323 Gateway reopens the video logical channel when video switch happens. In this way, the initiation of the new logic channel will refresh state of the video decoder so that the H.323 terminal can display the new video picture faster.

When the VIC display the video from a H.323 terminal, the problem of video refresh also arises because the H.261 encoder of the H.323 terminal doesn't send I-frame all the time. Since the VIC can't send "VCU" to request the fast-update of the remote encoder, we have to create some motion in front of H.323 devices to refresh the video. Now the user just need to make video selection via our own audiovisual applet and the H.323 Gateway will send VCU commands to the target H.323 terminal.

We also introduced an audio transcoding in the audio mixing service. The audio client of Access Grid (RAT) sends LINEAR audio data by default whereas LINEAR is not supported by H.323 terminals. The audio mixing service decodes LINEAR audio from Access Grid and mixes them into ULAW audio which can be handled by any H.323 terminals.

VRVS [15] is a videoconferencing system over a reflector infrastructure, which can provide some services like our H.323 Gateway. We believe that our H.323 Gateway provides better services than VRVS, especially in the services of session directory and session control. More importantly, VRVS is a close project, which has no idea of the open "virtual MCU" like Global-MMCS.

4. Conclusion

This paper focuses on how to adapt H.323 terminals in Global-MMCS, an integrated, scalable and service-oriented virtual MCU. The H.323 gateway helps to enables H.323 terminals to connect this MCU by bridging conference control service between H.323 and XGSP. Because the whole system is based on the open and service-oriented architecture, it is easy for further extension. In the next step, we will try to use SOAP communication service in NaradaBrokering for XGSP web-services such as the negotiation of stream transportation. Such a hybrid media and web-services framework allows other developers to add new services on top of Global-MMCS. We believe that Global-MMCS as a pure software solution provides an opportunity to create an integrated, scalable and easy-to-use audiovisual collaboration environment in a lower price than customized hardware solution.

5. Reference

1. Recommendation H.323, Packet-base multimedia communications systems, ITU, Sep. 1999.
2. J.Rosenberg , H.Schulzrinne, et al. "SIP: Session Initiation Protocol", RFC 3261,
   Internet Engineering Task Force, http://www.ietf.org/rfc/rfc3261.txt.
3. R Stevens, , M. PapkaE, T. Disz, et al Access Grid Project, http://www.accessgrid.org.
4. Recommendation H.225, "Calling Signaling Protocols and Media Stream Packetization for Packet-based Multimedia Communication Systems", ITU, Feb, 2000.
5. Recommendation H.245, "Control Protocols for Multimedia Communication", ITU, Feb 2000.
6. Recommendation H.243, "Terminal for Low Bit-rate Multimedia Communication", ITU, Feb 2000.
7. W. Wu, G. C. Fox, H. Bulut, A. Uyar, H. Altay, "Design and Implementation of A Collaboration Web-services system", Journal of Neural, Parallel & Scientific Computations, Volume 12, 2004.
8. Global Multimedia Collaboration System (Global-MMCS), http://www.globalmmcs.org

9. G.C.Fox, S. Pallickara, "The Narada Event Brokering System: Overview and Extensions", proceedings of the 2002 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'02)

10. K. Ott. Bormann, IETF Draft SDPng: "Session Description and Capability Negotiation", http://www.dmn.tzi.org/ietf/mmusic/sdp-ng/

11. W. Wu, H. Bulut, A. Uyar, G. C. Fox, "A Web-Services based Conference Control Framework for Heterogenous A/V collaboration", 7th IASTED International Conference on Internet and Multimedia Systems and Applications, August 13-15, 2003, Honolulu, Hawaii.

12. Open H.323 Project, http://www.openh323.org

13. GNU Gatekeeper Project, http://www.gnugk.org

14. Recommendation H.261, Video Codec for audiovisual services at p*64 kbit/s, ITU.1993.

15. H.Newman, P. Galvez, G. Denis, VRVS, http:// www.vrvs.org