Design and Practice of the Interworking System of the Heterogeneous-VoIP

CHIH-CHANG SHEN, TUNG-MING KOO[1],BO-LIN KUO
Department of Information Management
National Yunlin University of Science and Technology
123 University Road, Section 3, Douliou, Yunlin 64002
TAIWAN, R.O.C.
g9220804@yuntech.edu.tw ; koo@yuntech.edu.tw

Abstract: - At present, the VoIP is a fashion application of Internet. However, because they cannot communicate each other among various protocols of VoIP, it is necessary to install the VoIP-application for that specific group if these users having different VoIP protocols wish to communicate. For interworking between different VoIP groups, the two major players of the instant message, Yahoo Messenger and MSN Messenger, have provided an integration project. Thus, the interworking of the heterogeneous-VoIP has become a trend. Besides, the VoIP users can be classified into two groups, SIP and Skype. Therefore, to communicate both groups of users, SIP and Skype, a connection system of the heterogeneous-VoIP with a man-in-the-middle mechanism is provided and practiced in this study.

Keyword: P2P, VoIP, SKYPE, SIP, Heterogeneity, Man-in-the-Middle

1 Introduction

The VoIP technique is to transmit the voice service on the basis of IP network after compressing the voice signals to data packets, i.e., the service of voice can be forwarded via an open Internet. Since the cost of transmitting the instant voice data on Internet is lower than the conventional public switched telephone network, PSTN, lots of vendors such as Skype, Google Talk, and MSN Messenger of Microsoft, successively join this VoIP market.

The application of the VoIP is based on the connection of protocols. Now, the main protocols of VoIP are the H.323 as well as SIP (Session Initiation Protocol). H.323 is a standard developed by the International Telecommunications Union (ITU) and it contains a set of protocols for VoIP, videoconference, and message-share. SIP is newer and simpler protocol designed for VoIP. It is a replacement program of H.323 developed by the Internet Engineering Task Force (IETF). At present, many types of equipments at central offices and terminals follow the protocols of H.323 and SIP. Because SIP is simpler and more flexible as well as comprehensible than H.323 for general people and it seemingly has prospects, the major protocol discussed in this study is SIP.

Skype is a VoIP application by the P2P technique. Because of free of charge and excellent quality of communication, there are more and more users joining the Skype group. The facts of 200-thousand downloading person-time every day, 140 million users around the world, and 4-million simultaneous online users show prospects of Skype. In addition to a shareware, the business strategy of Skype has expanded to value-added service such as communication with the general urban telephones and instant chat messages to the cellular phones. To accelerate its development, Skype has released the API to users for joining their originalities into Skype and diversifying applications of Skype.

1.1 Causes and purposes of this study

Skype, the famous VoIP application, owns enormous users within two years. Because of these stable customers, Skype cooperates with communication industries of different countries and portal websites to introduce value-added service such as Skype-Out as well as Skype-In that can connect with the conventional telephones. However, because Skype uses closed protocol, only the authorized cooperation vendors can develop value-added service owned by Skype. Besides, Skype does not support the open protocols, SIP. Thus, users installing Skype cannot communicate with the SIP users.

SIP is open standard protocol. Many famous companies such as Cisco and Microsoft support the SIP protocol. Even though the voice quality of SIP

[1]Author to whom correspondence should be addressed

is not as good as Skype, it still has enough level. Because SIP needs the proxy, location, and Registrar servers to aid routing of messages and locating of income calls for connecting the mutual connection, the load of servers is very heavy. However, with progressive information technology, the speed of hardware is getting swift so that the issues of servers with heavy load have not been the bottleneck of developing SIP. Meanwhile, since many vendors devote to research SIP and expand its capabilities, the prospects of SIP are still prosperous.

By the above introductions, we have known that both heterogeneous-VoIP protocols, Skype and SIP, own enormous users. If both groups of users can be connected and integrated, i.e., users from Skype and SIP can communicate each other, the benefits between both parties will be maximized. Based on this consideration, the applications gateway bridging both heterogeneous-VoIP networks will be designed according to the concept of man-in-the-middle and users from Skype and SIP can exchange messages and contact each other. Consequently, a framework of interworking for the heterogeneous-VoIP can be established that the communication can be built between various Heterogeneous-VoIP networks.

1.2 Restrictions in studies

Because of confidential protection for the protocol of Skype, the buildup of connection, application of the Skype account, voice, and encryption as well as unencryption of the instant chat messages must be indirectly processed via Skype Phone. After communication with the Skype Phone through the Skype API, the unencrypted voice or instant messages can be saved into the virtual sound card and then forwarded to the SIP Phone via the communication applications developed by the SIP API. For instant messages, after unencrypted instant messages are acquired from the Skype Phone through Skype API, they can be transmitted to the SIP Phone via SIMPLE protocol.

2 Related technique

The man-in-the-middle will be utilized to exercise the research and practice of instant communication of VoIP/IM(Instant Message) for Skype and SIP. For the relevant techniques and theorems, we will discuss four major types, SIP (Session Initiation Protocol), SIMPLE (SIP for Instant Messaging and Presence Leveraging Extension), RTP (Real-time Transport Protocol) as well as RTCP (Real-time Transport Control Protocol), and Skype. We devote to integrate the above techniques into the systematic architecture of this study for building a complete man-in-the-middle mechanism that fulfills the communication between Skype and SIP.

2.1 SIP

The Session Initiation Protocol (SIP), mainly applied on VoIP, is a protocol standard developed by the Internet Engineering Task Force (IETF). SIP protocol allow building a multimedia-streaming meeting among users and this multimedia streaming contains voice, video, or any medium based on Internet. Its application scope is very extensive, for example, Internet game, peer-to-peer file sharing and applications of instant message [2][5][6].

By descriptions of RFC 3261, SIP is based on ASCII and belongs to the application layer in OSI. It is a signal connection to build, maintain, and terminate peer-to-peer and multi-peer communication. Meanwhile, through Gateway and PSTN, it can build a connection to integrate the whole communication network.

In the SIP environment, there are two major parts, SIP servers and SIP User Agents. And the SIP Servers contain four major elements, Registrar Server, Location Server, Proxy Server, and Redirect Server.

2.2 SIMPLE

SIMPLE (SIP Instant Messaging Present Leveraging Extension) is a protocol based on SIP to set up a SIP Extension method, Message, and a process transmitting the instant messages on SIP. Thus, the advantage of SIMPLE is to inherit the SIP protocol and it is not necessary to set up other new protocol [1].

2.3 RTP and RTCP

When users on both sides connect via SIP, the method of voice compress, format of packet, and transmission protocol will be decided. At present, the transmission of instant voice data is finished through the Real-Time Transport Protocol (RTP). The real-time is a characteristic of RTP. But UDP lacks overhead for the packet transmission so that the transmission speed of UDP is faster than TCP. Therefore, UDP is selected as a protocol of packet transmission in the Transport Layer by RTP. The major definitions about RTP are described in RFC 3550. Besides, RTP also provides another protocol, RTCP, for control. This RTCP will provide a control to monitor quality of all streaming from new users[4][7].

2.4 Skype

Skype, the first VoIP application of P2P developed by KaZaa in 2003, supports voice messages and instant chat messages. In essential, Skype, which is very close to applications like MSN and Yahoo IM, can support voice telephone, instant chat message, videoconference, and crony list. However, the protocols and the application technologies executed on the lower layers are totally different [3][8].

Skype, like the file sharing application KaZaa, is operated according to a concept of overlay P2P network. The overlay network contains two major nodes, ordinary node and super node. The ordinary node, an application of Skype, can build the voice messages and transmit the instant chat messages. The super node is an endpoint that the ordinary node enters the Skype network. Any node with an open URL, enough CPU speed, memory capacity, and network bandwidth is qualified to become a super node. An ordinary node has to be connected with a super node for connecting the Skype network and registered in the login server for the identity verification in future logins.

3 Methods in studies

In this study, a connection mechanism of the heterogeneous-VoIP is constructed under the man-in-the-middle architecture. Through this mechanism, both groups with different protocols can mutually communicate in accordance with the man-in-the-middle mechanism. The connection architecture of the heterogeneous-VoIP consists of four modules, Skype Packet forwarding module (Skype-PF), SIP Packet forwarding module (SIP-PF), Voice Capturing module (VoCP), and User Mapping module (UM). If both users, on the Skype-end and SIP-end respectively, wish to communicate each other, they should manually key in the opposite side's URL in the heterogeneous-VoIP system. Then the connection and voice packet forwarding can be built through the "SIP-PF" and the "Skype-PF". Finally, users on both sides of SIP and Skype can directly communicate via these modules.

The following is an introduction of process by two different situations, SIP-to-Skype and Skype-to-SIP:

SIP-to-Skype:

When a SIP user wants to communicate with a Skype user, the process of operation is:

a. SIP Phone dials the SIP URI corresponding to the Skype account and builds a connection with the "SIP-PF". Then the user manually keys in the Skype account of the income user.
b. The Skype Packet forwarding module will inquire the Skype account corresponding to the SIP URI in the UM and dial the Skype account of the income user for a connection with the Skype phone.
c. After finishing the communication, the "Skype-PF" will save the voice messages forwarded by the Skype phone to the "VoCP".
d. The "Skype-PF" will capture the voice messages from the Skype phone in the "VoCP" and then forward to the SIP Phone.
e. When the "Skype-PF" receives the voice messages from the SIP Phone, it will save these messages into the "VoCP".
f. After the "Skype-PF" captures the voice messages from the SIP Phone in the "VoCP", these messages will be forwarded to the Skype Phone.
i. Thus, after the "SIP-PF" and the "Skype-PF" connect to the SIP Phone and the Skype Phone respectively, the previous steps (c)~(f) will be repeated to ensure the mutual communication between the SIP Phone and the Skype Phone.
g. If any party of the SIP Phone or the Skype Phone wishes to disconnect the conversation, the "SIP-PF" and the "Skype-PF" will terminate connection with the SIP Phone and the Skype Phone.

Skype-to-SIP:

When a Skype user wishes to communicate with a SIP user, the process is as mentioned above. But when the user changes his caller into SIP Phone and dials with Skype account, the flow will be dotted line in Fig. 1

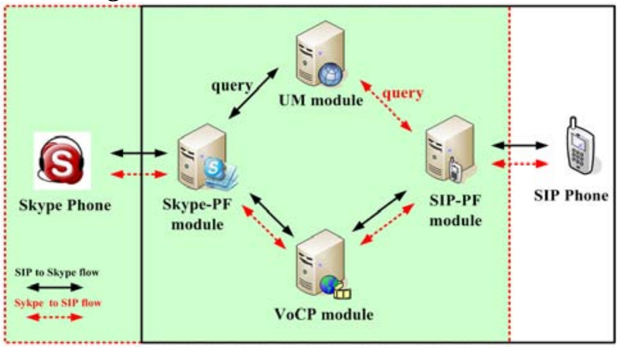

Fig.1, the process of call flow

3.1 Skype Packet forwarding module(Skype-PF)

The major functions of this module is to build a connection with the Skype Phone, receive the voice data from the Skype Phone, forward the voice data from the SIP phone to the Skype Phone, receive the instant chat messages from the Skype Phone and forward them to the SIP-PF, and forward the instant chat messages from the SIP Phone to the Skype Phone. To achieve the previous functions, the Skype API is employed in this study and the Skype-PF can indirectly build a connection with the Skype Phone for transmission of voice messages as well as chat messages by using the Skype API.

3.2 SIP packet forwarding module(SIP-PF)
This module is responsible to forward the Call Setup messages, voice packets, and instant message packets. In this process of operation, we can observe that the SIP-PF will receive the messages from the SIP UA and appropriately process them and then corresponds to the Skype-PF to make a communication between Skype AP and SIP AP. The communication between the SIP-PF and the SIP AP can be divided into two categories, Request (INVITE, CANCEL, BYE, MESSAGE, and INFO, etc.) and Response (Status code) and they can be adequately processed by category difference. For instance, when the SIP-PF receives the INVITE Message from the SIP AP, the SIP-PF will drive the Skype-PF to connect with the Skype AP. After AP of the Skype end rings, the SIP-PF will feedback 180 Ringing Message to AP of the SIP end and then wait answering from AP of the Skype end. Then the SIP-PF will re-feedback 200 OK Message to AP of the SIP end. Finally, the SIP-PF as well as the Skype-PF will finish connections with the SIP Phone and the Skype Phone respectively.

3.3 Voice Capturing module(VoCP)
The functions of this module are to save voice data from the SIP Phone as well as forward to the Skype Phone and save the voice data fro the Skype Phone as well as forward to the SIP Phone. Usually, there are two channels for voice output and input at each VoIP device. Normally, the input device is the microphone and the output device is the speaker or earphone. The Voice Capturing module(VoCP) will generate two virtual sound cards, one card is responsible to process sound output of the SIP Phone and forward this processed voice data to the Skype Phone as a part of input of the Skype Phone; and another one will process the output of voice data from the Skype Phone and forward them to the SIP Phone as a sound input of the SIP Phone. Thus, the "VoCP" is only responsible to forward the voice data. For capturing as well as forwarding of the voice data from the Skype Phone as well as the SIP Phone, the "Skype-PF" and the "SIP-PF" will process them. The architecture of operation is shown in the following figure.

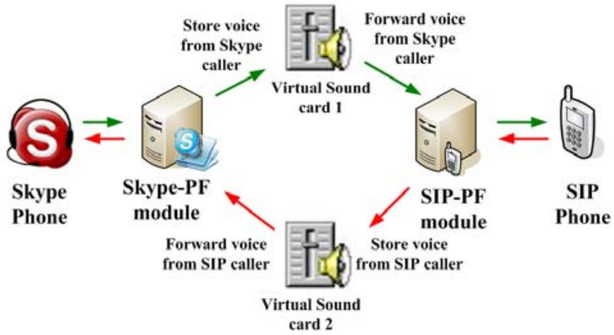

Fig. 2, process of operation of the VoCP

3.4 User mapping module(UM)
The major functions of this module are to design a mapping relationship of accounts between the "SIP-PF" as well as the "Skype-PF" and the SIP Phone as well as the Skype Phone and generate accounts of SIP as well as Skype. To construct the communication capability between SIP and Skype, the first thing is to generate a communication method for both. After the Skype account and the SIP URI have generated, the program can build connections and the mapping relationships between the accounts and the SIP Phone as well as the Skype Phone via the "SIP-PF" as well as the "Skype-PF". If there is one Skype user named Alice wishes to communicate with a SIP user, sip:Bob@10.20.3.4, the program will generate accounts mapping to the Skype user and the SIP user respectively. We suppose that the mapping account to Skype is sky66 and another mapping account to SIP is sip:sk66@10.20.3.6. The connection between the account sip:sk66@10.20.3.6 and the SIP user sip:Bob@10.20.3.4 can be constructed via the "SIP-PF". And the "Skype-PF" can use this account skypesip666 to build a connection to the user Alice. The architecture figure is shown below:

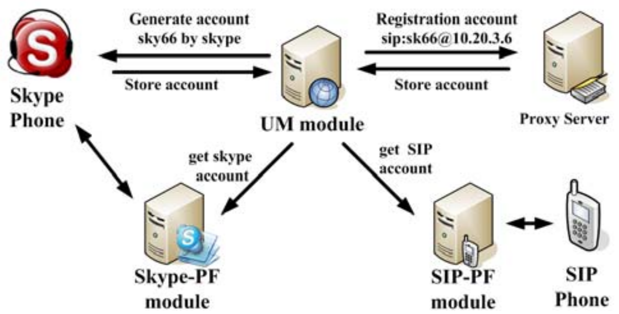

Fig. 3, process of operation of the UM

3.5 Interworking framework of the heterogeneous-VoIP

In this study, the feasibility of VoIP interworking between heterogeneous platforms is verified using the communication applications of SIP and Skype. The study focuses on the one-to-one conversation and the transmission of instant messages. After this model has been proved feasible, we expect that there are more VoIP communication applications included to become a general framework.

Because most VoIP communication applications do not completely open source codes, the APIs released from the VoIP vendors are the only sources for development. Thus, the general framework is restricted to the released APIs and some functions like voice packets and message packets forwarded in call setup cannot be directly processed. The only tools for communication via indirect control are the API forwarding control message released from vendors and VoIP applications.

Under the interworking framework of the heterogeneous-VoIP, the key point of operation is how to forward the voice data of VoIP to another heterogeneous-VoIP. To solve this problem, the virtual sound card is designed for instant capturing and data saving. Assume that VoIP A and VoIP B belong to different protocols. If both parties wish to talk, VoIP A is assigned to save the processed voice data into the input end of the virtual sound card and then VoIP B retrieves them at output end of the virtual sound card for data forwarding. Similarly, VoIP B also saves the processed voice data into the virtual sound card and VoIP A reads them from the virtual sound card for data forwarding. Therefore, if VoIP A and VoIP B can talk each other, they need two virtual sound cards, one card is for voice data output from VoIP A as well as voice data input from VoIP B and another one is for voice data input from VoIP A as well as voice data output from VoIP B. The process of operation is shown in the following figure.

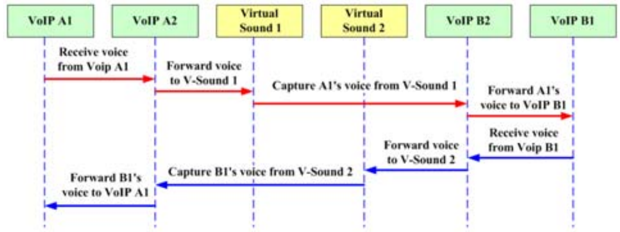

Fig. 4, process of operation under the general framework of the heterogeneous-VoIP

Until now, the feasibility of communication among various VoIP systems such as Skype, SIP Phone, MSN Messenger, and Google Talk has been verified. The method is to build a communication connection between homogeneous-VoIP systems, then save the voice data into the virtual sound cards, and finally forward them to the heterogeneous-VoIP systems. For instance, the first step for the communication between MSN Messenger and Google Talk is to build a connection by using two MSN Messenger users, MSN 1 and MSN 2. Then two Google Talk users, Google Talk 1 and Google Talk 2, have to build another connection each other. At this moment, MSN2 receives voice data from MSN 1 and save them to the virtual sound card 1. Google Talk 2 captures voice data from the virtual sound card 1 and then forwards them to Google Talk 1. Thus, the virtual sound card 1 is output of MSN2's voice and input of Google Talk 2's voice. Similarly, the virtual sound card can be input of MSN2's voice, i.e., output of Google Talk 2's voice. The architecture of this general framework is shown below:

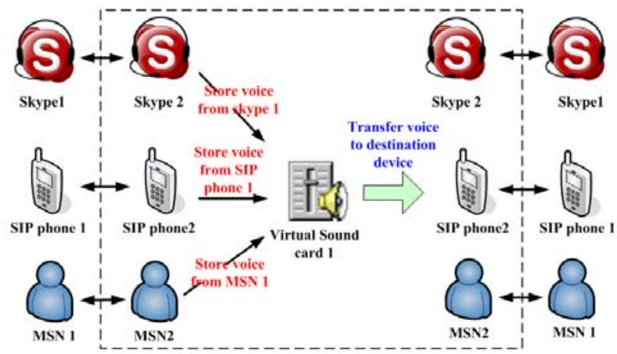

Fig. 5, the architecture of this general framework

4 Effect assessment of this system

The effect assessment of this system is to measure the difference of voice quality in various network environments. The voice quality of systems can be objectively measured from packet loss, jitter, and packet latency.

In measuring voice quality of VoIP, the voice quality will be classified into three levels by the loss rate of packet, voice jitter, and voice latency. The standards are listed below:

Table 1, measurement index of voice quality

Voice quality	Time of packet latency	Loss rate of packet	Voice jitter
Good	< = 150 ms	0%	1 ms
Fair	150 – 400 ms	1%	20 ms
Poor	>= 400 ms	5%	60 ms

The experiment in this study is designed to measure voice quality of VoIP by restriction of uploading bandwidth. Firstly, the communication between pure SIP users is measured and secondly the measurement condition is changed to a connection via the heterogeneous-VoIP. Then the voice quality of the above different conditions can be compared. Since the voice quality of SIP is not as good as Skype, the restriction from SIP will lead to degradation of voice quality when a Skype user builds a connection via the interworking application of VoIP. Meanwhile, it is necessary to add some specific hardware equipments for measurement of voice quality of Skype. Thus, the comparison of voice quality is made between SIP and the interworking application of heterogeneous-VoIP.

Table 2, result of experiment

	SIP to SIP	Skype to SIP		
Bandwidth for uploading	10KBps	10KBps	20KBps	32KBps
Jitter	1.49 ms	47.78 ms	23.96 ms	0.98 ms
Packet latency	25.22 ms	628.87 ms	341.6 ms	65.27 ms
Loss rate of packet	0%	4%	1.9%	0%

In the above table, when the uploading bandwidth from SIP to SIP is 10kbps, the voice quality is the best. Compared Skype with SIP at the same uploading bandwidth, the voice quality will be substantially decreased when interworking is built between Skype and SIP via interworking applications of the heterogeneous-VoIP. In fact, it will spend some time for data processing at forwarding voice data. Moreover, the bandwidth will be occupied when the “SIP-PF” and “Skype-PF” forwards voice data to SIP Phone and Skype Phone respectively. Thus, the voice quality is hard to accept if the uploading bandwidth is 10 kbps. When the uploading bandwidth is changed to 32kbps, the testing data at jitter, packet delay, and loss rate of packet can reach a good level. Thus, for an acceptable voice quality, the uploading bandwidth for a communication between Skype and SIP via interworking applications of the heterogeneous-VoIP must be at least 32kbps.

5 Conclusions and suggestion

In this study, the interworking system of the heterogeneous-VoIP is built on the client-end rather than on the server-end. By this method, the communication connection and capturing/ forwarding of voice data will be processed via users’ CPU. Furthermore, the format of communication connection is one-to-one so that the overloading on the server-end for processing multi-users will not be considered.

Besides, the practice in this study only focuses on the one-to-one conversation and instant messages and it still lacks functions like multiparty call, instant video forwarding, crony-list sharing, and voice chat room. Therefore, any researcher who is interested in this field can continuously enhance the functions of interworking platform of the heterogeneous-VoIP by relevant RFC of SIP and Skype API.

ACKNOWLEDGEMENT

This work was supported in part by National Science Council under the Grants NSC 96-2221-E-224-026.

Reference:

[1] A. Johnston, S. Donovan, R. Sparks, C. Cunningham, K. Summers, *Session Initiation Protocol (SIP) Basic Call Flow Examples* , RFC 3665, IETF, Dec., 2003.

[2] Campbell, Ed., J. Rosenberg, H. Schulzrinne, C. Huitema, D. Gurle, *Session Initiation Protocol (SIP) Extension for Instant Messaging*, IETF , Dec., 2002.

[3] Dennis Bergström, *An ananlysis of Skype VoIP application for use in a corporation environment*, Dennis, Oct., 2004.

[4] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, *RTP: A Transport Protocol for Real-Time Applications*, RFC 3550, IETF, Jul., 2003.

[5] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, *SIP: Session Initiation Protocol*, RFC 3261, IETF, Jun., 2002.

[6] J. Rosenberg, H. Schulzrinne, *Reliability of Provisional Responses in the Session Initiation Protocol (SIP)*, RFC 3262, IETF, Jun., 2002.

[7] M. Handley, V. Jacobson, *SDP: Session Description Protocol*, RFC 2327, IETF, Apr., 1998.

[8] Salman A. Baset, Henning Schulzrinne, *An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol*, Department of Computer Science Columbia University, Sep., 2004.