International Symposium on Information Theory and its Applications, ISITA2006
Seoul, Korea, October 29–November 1, 2006

A proposal of Voice over IP Passing through Firewall and its Evaluation

Masashi ITO[†], Akira WATANABE[‡]

Graduate School of Science and Technology, Meijo University
1-501 Shiogamaguchi, Tenpakuku, Aichi, 468-8502 Japan
E-mail: [†]m0641502@ccmailg.meijo-u.ac.jp [‡]wtnbakr@ccmfs.meijo-u.ac.jp

Abstract

In recent years, IP telephone has achieved remarkable progress on the Internet due to ”low price ”, ”continuous connections”, and ”high speed communication rate”. However, it is not easy to use IP telephone over firewall and NAT because of their restrictions of communications. We have proposed the system called SoFW (SIP over Firewall) that solves the problem. In this paper, detailed functions and its implementation method of SoFW are described.

1. INTRODUCTION

Due to the spread of broad band communications and development of backbone networks among ISPs, the transmission capacity of the network has been considerably increased. Therefore, the quality assurance of Voice over IP (VoIP, hereinafter) becomes a level of practical use, and it has become popular among enterprise networks and home networks.

However, in case of enterprise networks, Firewall (FW, hereinafter) [1] that is located between the enterprise network and the Internet, prevent the communications of VoIP between the terminals. It is expected that expansion of the VoIP is further promoted if it can safely pass through FW.

There is a protocol referred to as SIP (Session Initiation Protocol) [2] which is standardized by IETF (Internet Engineering Task Force) owing to easy implementation and expandability, is being paid attention that it can be used for various kinds of multimedia services. SIP has been employed to most of VoIP systems provided by ISP. A SIP system consists of user agents and a SIP server, and provides functions of registering user locations for the SIP server, and of relaying dial messages based on its location. However, in the SIP system, it is needed that IP address of callee terminal, or IP address of the SIP server to which the callee terminal belongs to can be identified by caller when dial starts. For that reason, dialing can not be performed in the environment in which NAT (Network Address Translator) [3] exists between the communication terminals. Further, in most cases, FW limits the communications to the applications such as mail and Web server access from an inside of the enterprise network to the Internet, and blocks other communications. If VoIP is to be introduced in a network under such limitations, a security policy of the enterprise network must be changed, and degradation of security accompanied thereby possibly occurs. So, changing the security plicy is troublesome and very difficult.

Some systems have been proposed, in which VoIP can pass through the barrier of FW and NAT. They are, for example, HCAP [4], Skype [5], and SoftEther [6]. SoftEther enables any applications to path through FW and NAT, not limited to VoIP.

HCAP and Skype provide an HTTP tunnel between terminals in an enterprise network and a relay server on the Internet. Special applications are used for dialing, and voice streams are relayed with packets embedded in HTTP GET and POST messages. Thus, VoIP can pass through FW and NAT if the environment is capable of accessing web site on the Internet. However, there are problems that special functions are required for the terminals. In case of SoftEther, software referred to as Virtual LAN Card is implemented in a PC on a private address side, and software referred to as Virtual HUB is implemented in a PC on a global address side. Virtual IP address and MAC address are allocated in Virtual LAN Card. Virtual LAN Card and Virtual HUB construct a virtual Ethernet by embedding Ethernet frames in a protocol capable that can pass through FW and NAT, such as HTTPS, SSH or the like. Terminals connected to the virtual Ethernet can freely communicate across FW and NAT. In order to construct a VoIP system on the virtual Ethernet, a SIP server and VoIP terminals are to be connected to the virtual Ethernet. However, in this system, it raises problems that a network originally protected by FW is exposed to danger, and further, an integrated control

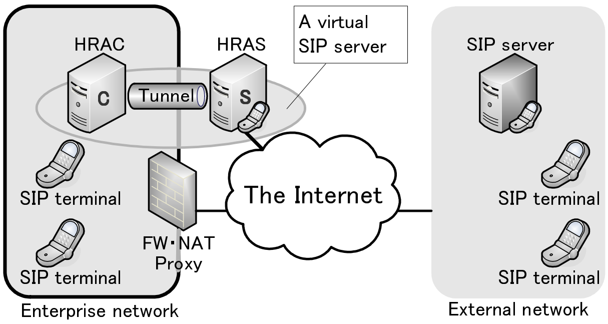


Figure 1: Configuration of SoFW.

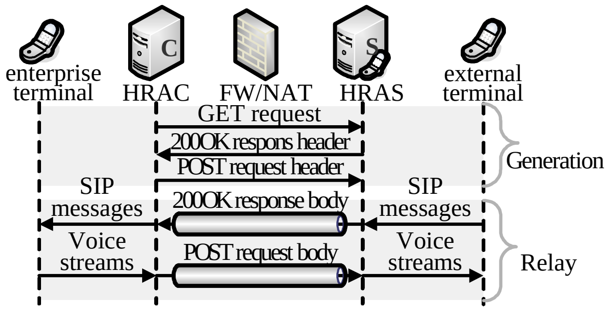


Figure 2: Sequence from generation of an HTTP tunnel to termination of a call.

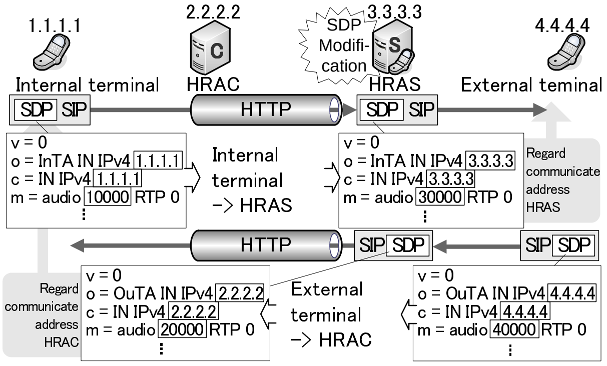


Figure 3: Procedure of changing SDP.

of IP addresses in the virtual Ethernet is needed.

Thus, we have been proposed the system called SoFW (SIP over Firewall) to solve the problems. In SoFW, two types of relay agents are placed inside and outside of FW/NAT, one by one, and all messages of SIP and voice streams from terminals are passed through in HTTP tunnel made by the relay agents. Since SoFW realizes passages over FW and NAT only by adding relay agents, it does not affect existing systems. This becomes very effective in case that people in the company are already using the SIP based VoIP system. In this paper, a principle of SoFW, its implementation, and evaluation are described.

2. A principle of SoFW

2.1. Outline

Fig. 1 shows the configuration of SoFW. In SoFW, HRAC (Half Relay Agent Client) is placed in an enterprise network and HRAS (Half Relay Agent Server) is placed on the Internet. Prior to the telephone communication, an HTTP tunnel is generated between HRAC and HRAS, and the two devices are functioned as a virtual SIP server having interfaces of a global and a private IP address. Voice streams are also relayed by the HTTP tunnel.

2.2. HTTP tunnel

Fig. 2 shows sequence of a generation of HTTP tunnel. HRAC establishes two TCP connections for a GET request and a POST request, defined by HTTP. When HRAS receives a GET request, it returns a header part of 200OK response. When the process is completed, HRAC and HRAS wait for SIP messages from end terminals. HRAC embeds a receiving SIP message in a body part of a POST request and transmits it to HRAS. HRAS embeds a receiving message in a body part of a 200OK response, and transmits it to HRAC. And then, if HRAC and HRAS recieve voice streams, HRAC embeds a receiving voice stream in a body part of a POST request and transmits it to HRAS. HRAS embeds a receiving voice stream in a body part of a 200OK response, and transmits it to HRAC.

2.3. Voice stream guidance

SoFW relays not only SIP messages but also voice streams to the HTTP tunnels between HRAC and HRAS. However, in normal SIP specifications, voice streams are directly exchanged between end terminals. In SoFW, to guide the voice streams to the HTTP tunnel, when SIP messages reach HRAS in dialing phase, HRAS changes type values described in SDP [7], body part of SIP messages.

Fig. 3 shows the procedure of changing SDP contents. Various kinds of information required for a voice communication is described in SDP as type values. The type values include IP addresses, port numbers and codec type which will be used for a voice communi-

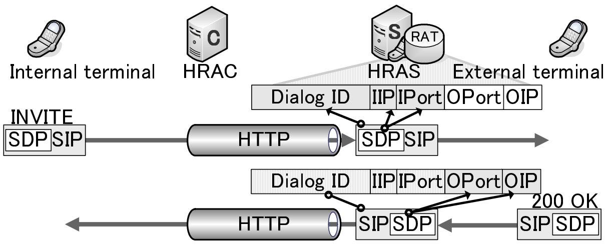


Figure 4: Generation of RAT.

cation by the terminals. In HRAS, an IP address of caller in SDP transmitted from an enterprise terminal is changed into the IP address of HRAS, and an IP address of callee in SDP transmitted from an external terminal is changed into the IP address of HRAC, respectively. The enterprise terminal, that receives the changed SDP, recognizes that the correspondent node is HRAC, and the external terminal recognizes that the correspondent node is HRAS, and thus, the voice streams are guided to the HTTP tunnel.

2.4. Determination of a routing path

As described in 2.3, end terminals send voice streams toward HRAC or HRAS, thus HRAC and HRAS have to determine the right path of voice streams to the end terminals. In SoFW, RAT (Relay Agent Table) specific to SoFW, is generated in HRAS from the information of SIP header and SDP contents during dialing operations. The paths of voice streams are determined with reference to RAT during voice communications.

Fig. 4 shows a flow of RAT generation when the dialing is started from an enterprise terminal. Dialog ID is formed by To, From, Call-ID which are obtained from SIP headers, and it identifies the communication. Others are obtained from SDP contents, and IIP and IPort show an IP address and a port number of an enterprise terminal, and OIP and OPort show an IP address and a port number of an external terminal. SDP is contained in INVITE message which is a start message of a caller and in 200OK which is the response of INVITE. When HRAS receives INVITE, it writes down the dialog ID, IIP and IPort in a RAT record. Next, when HRAS receives 200OK it retrieves the same communication of the RAT record from the dialog ID in the message, and adds OIP and OPort in the RAT record.

When dialing operation is finished, voice communications start, and the path of the voice streams is determined by RAT in HRAS.

Fig. 5 shows process flows of voice streams. When

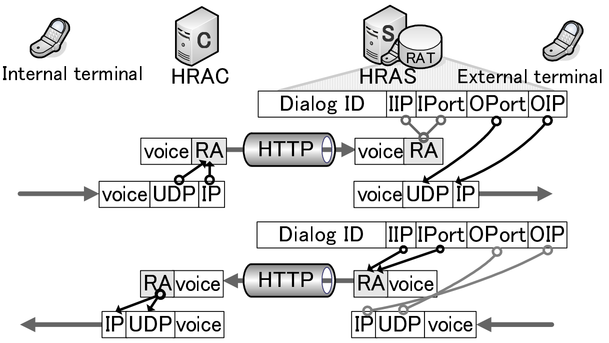


Figure 5: Processing flow of voice streams.

HRAC receives voice streams from an enterprise terminal, an IP address and a port number of the enterprise terminal are added to the voice data as an RA header, and the packet is relayed to HRAS. HRAS retrieves the corresponding RAT record from the information in the RA header, and changes the destination of the voice stream to the external terminal indicated in RAT and transmits the voice stream. When HRAS receives voice streams from an external terminal, HRAS retrieves the corresponding RAT record from a source IP address and a port number. The IP address and the port number of the enterprise terminal are added to the voice data as an RA header, and the packet is relayed to HRAC. HRAC changes the destination address of the voice streams to the enterprise terminal indicated in the RA header, and transmits the voice streams. When HRAS receives a BYE message which is a request for disconnection, a corresponding RAT record is retrieved from the dialog ID, and contents of the record are deleted.

3. Implementation method

HRAC and HRAS have been implemented as applications on FedoraCore30 (linux2.6.9), and the function of HRAS has been realized by a cooperation with SER [8] which is free software of SIP server. On the portion of dialing process in HRAS, other functions than SER are referred to as SIP Relay Server module. SOCKET is used for a connection between SIP Relay Server and SER. To achieve that, a simple modification is made to SER. Before SIP messages completing a series of processing by SER is about to leave for the SOCKET, a judgement function whether the message is addressed to an external terminal or to an enterprise terminal is added.

4. Evaluation results

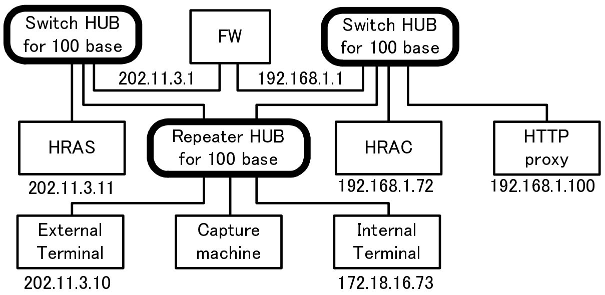


Figure 6: Structure of the experimental system.

Table 1: Specifications of the evalutation system.

device	specification	
HRAS /HRAC	CPU	Intel Pentium4 2.8GHz
	Memory	512MB
	NIC	Broadcom Tigon3 100BASE-TX
FW/NAT /Proxy	CPU	Intel Pentium3 600MHz
	Memory	256MB
	NIC	Global: Silicon Integrated System crop 100BASE-TX
		Privete: ADMtek FNW-9803-T 10/100BASE-TX
External terminal	CPU	Intel Pentium4 3.4GHz
	Memory	1GB
	NIC	Broadcom NetXtreme57xx 100BASE-TX
Internal terminal	CPU	Intel PentiumM 1.80GHz
	Memory	512MB
	NIC	Realtek RTL8139/810x 100BASE-TX

We have measured the end-to-end delay that contains packet processing time of HRAS, HRAC, FW, NAT and HTTP Proxy server.

Fig. 6 shows a structure of the experimental system, and Table 1 shows specification of devices. The security policy that passes through only HTTP access from inside to outside and TCP statefull inspection is set up in FW. SIP terminal uses X-Lite with G.711 codec. Total packet numbers are 100,000.

Table 2 shows end-to-end delay of the experimental system. Average of the delay is less than 1msec, and maximum delay is about 50-75msec.

Table 2: Measurement results of delay time.

	Outbound	Inbound
average(msec)	0.9535	0.9735
max(msec)	51.987	73.792

Table 3: Variance of delay time.

delay (msec)	-0.7	0.7 -1.3	1.3 -1.9	1.9-
Outbound(%)	0.41	98.31	1.27	0.01
Inbound(%)	1.09	98.02	0.78	0.01

Table 3 shows variance of delay time. Although maximum delays in Table 2 are fairly large, however, almost all delays are sufficiently small. That is seen from Table 3 which says that the number of delay time over 1.9msec is only 0.01%. It is said that end-to-end delay within 400msec is enough for IP telephone. So, it has become clear that SoFW can keep sufficiently high performance.

5. Conclusion

In this paper we have described the realization method of SoFW and its evaluation. From the results, it is shown that SoFW has sufficiently high performance. Hereafter, we will evaluate the performance of SoFW under the existence of back traffic, and in the case that plural telephone terminals exist in the system.

References

[1] Freed, N.: Behavior of and Requirements for Internet Firewalls, ,RFC 2979 (2000).

[2] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M. and Schooler, E.: SIP: Session Initiation Protocol, RFC3261 (2002).

[3] Egevang, K., and Francis, P.: The IP Network Address Translator (NAT), RFC 1631 (1994).

[4] Shinji Kunai.: Allround Internet Telephony Protocol -HTTP-based Conference Application Protocol ,IPSJ-JNK4403007

[5] Skype.
http://www.skype.com/home.html

[6] SoftEhter.
http://www.softether.com/jp/

[7] Handley, M. and Jacobson, V.: SDP:Session Description Protocol, RFC2327(1998).

[8] SER.
http://www.iptel.org/ser/

A proposal of Voice over IP Passing through Firewall and its Evaluation

The 2006 International Symposium on Information Theory and its Applications

October 29 – November 1, 2006

COEX, Seoul, Korea

Masashi Ito and Akira Watanabe

Graduate School of Science and Technology, Meijo Univ, Japan

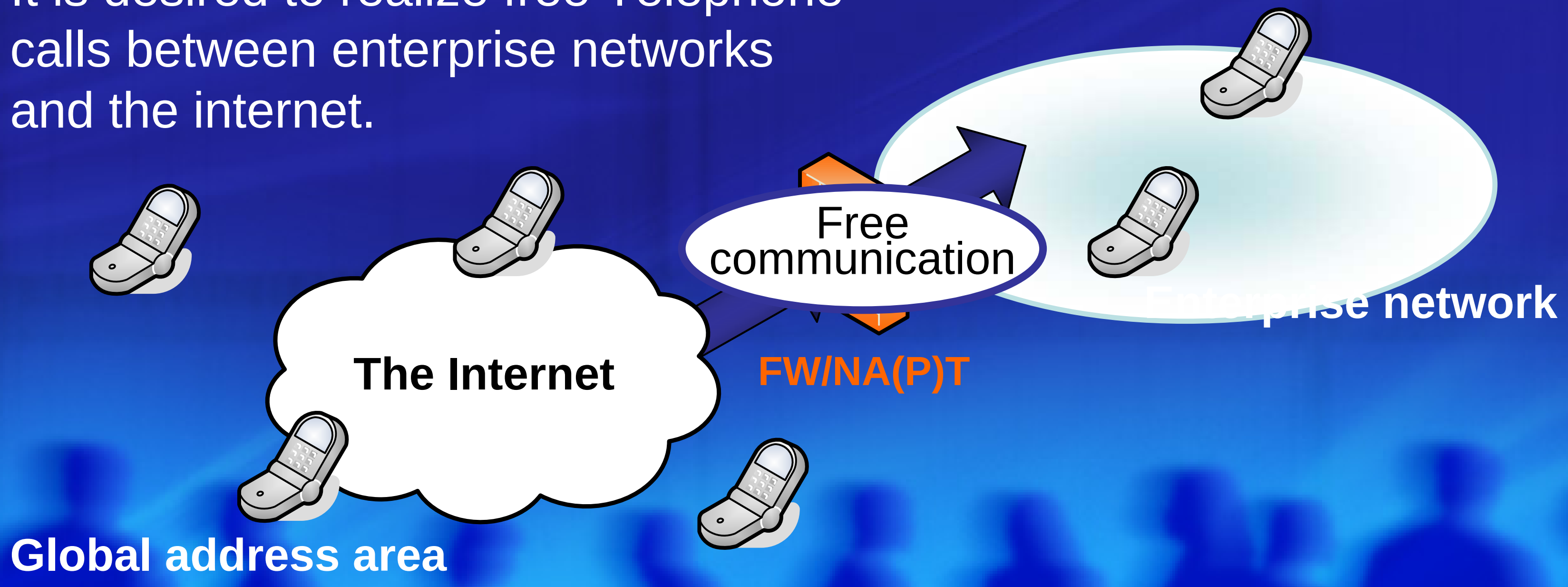
Background

■ Spread of broad band communications.
■ Quality of VoIP become a level of practice use.

⇨ IP Telephony has become popular in enterprise networks and the internet.

To improve convenience of IP telephony further…

It is desired to realize free Telephone calls between enterprise networks and the internet.

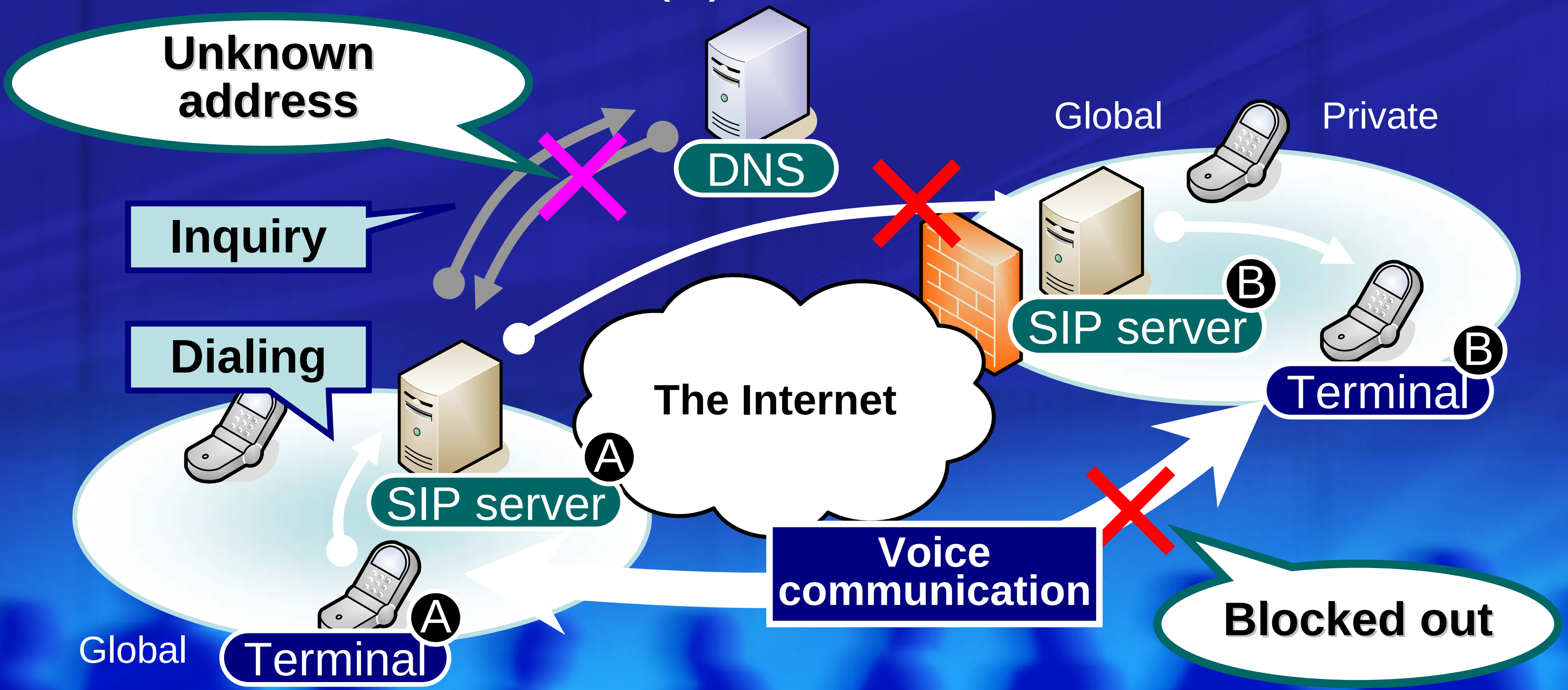


Related studies

SIP（Session Initiation Protocol）

SIP has been paid attention as a session initiation protocol, with its easy implementation and expandability.

If there exist FW／NA(P)T …



Related studies

- Firewall remodeling method

  Specific Firewall opens proper port numbers and IP addresses, and closes them dynamically.

  UPnP, VoIPsecureGW, Connect-VPnP, IP - NAT, etc...

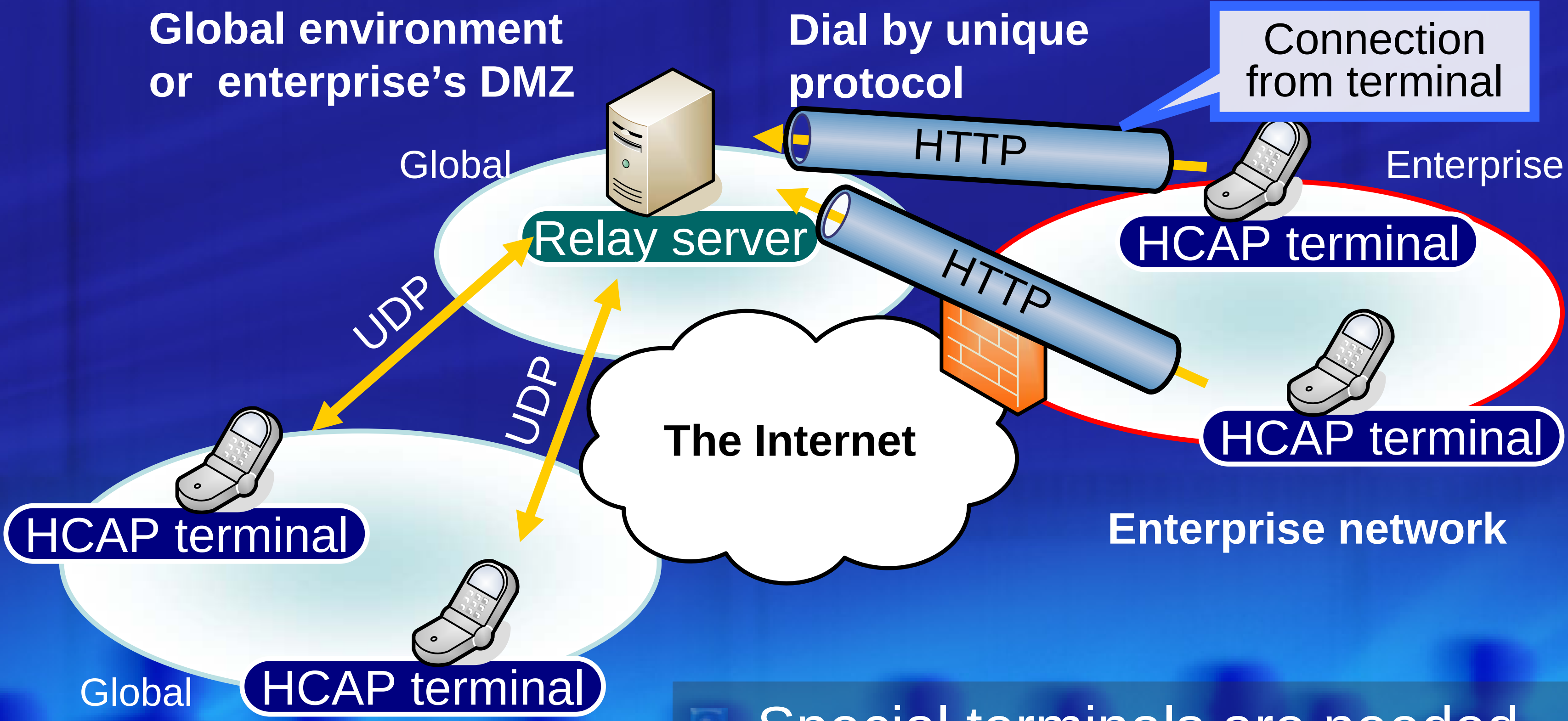
↓

It is needed to change security policy

- Terminal remodeling method
  - Skype
  - HCAP

Related and early studies HCAP

Terminals and a relay server generate HTTP tunnels between them at the initial phase, and all dial and voice streams are relayed in the tunnel.



- Special terminals are needed

Principle of SoFW

SoFW（SIP over Fire Wall）

Purpose of the system

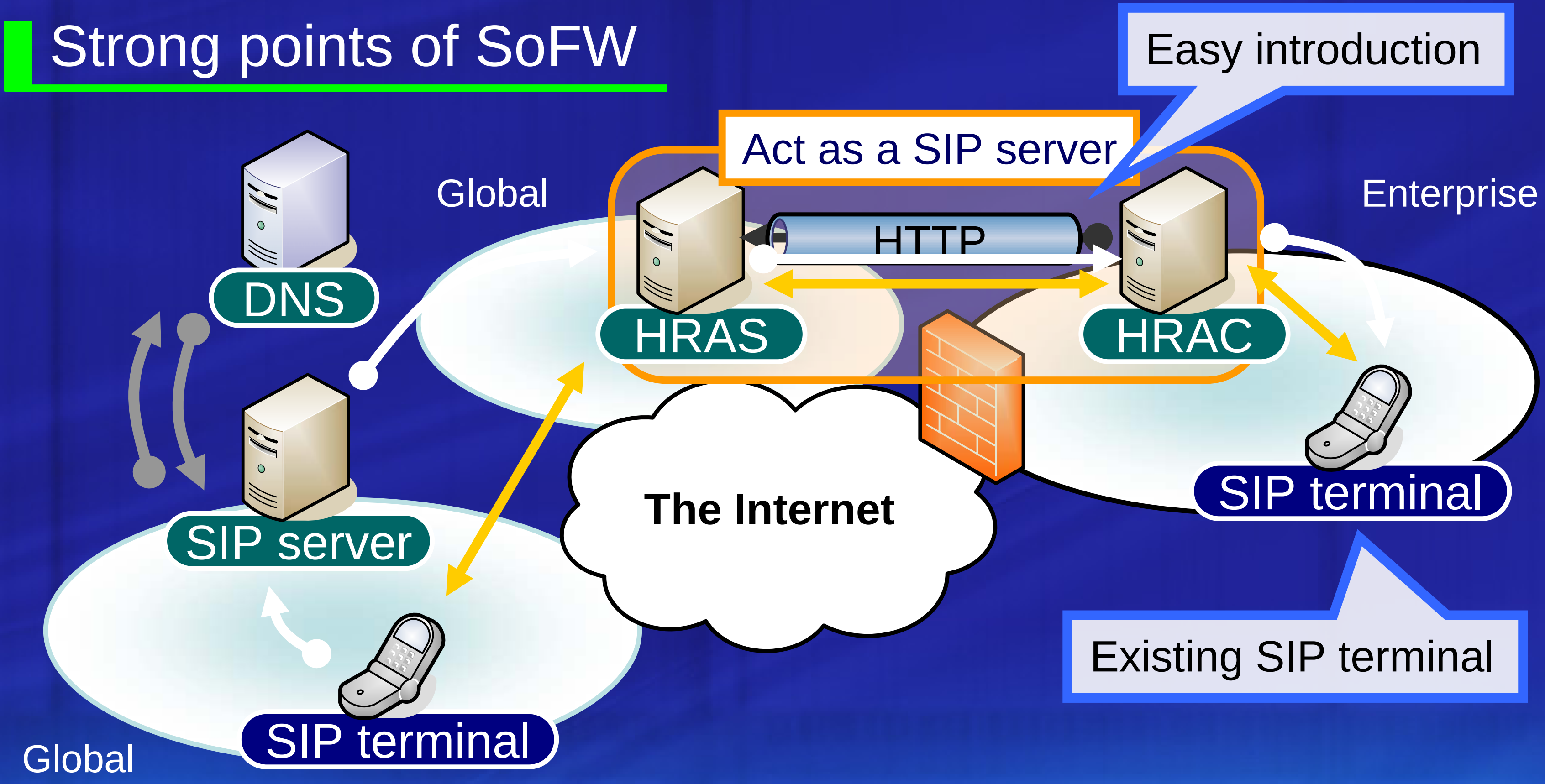
- Easy introduction
- Existing SIP terminals

SoFW relays SIP and Voice streams using two devices which set in an enterprise network and a global network respectively.

Principle of SoFW

System configuration

Strong points of SoFW



HRAS（Half Relay Agent Server）is set up in global

HRAC（Half Relay Agent Client）is set up in enterprise

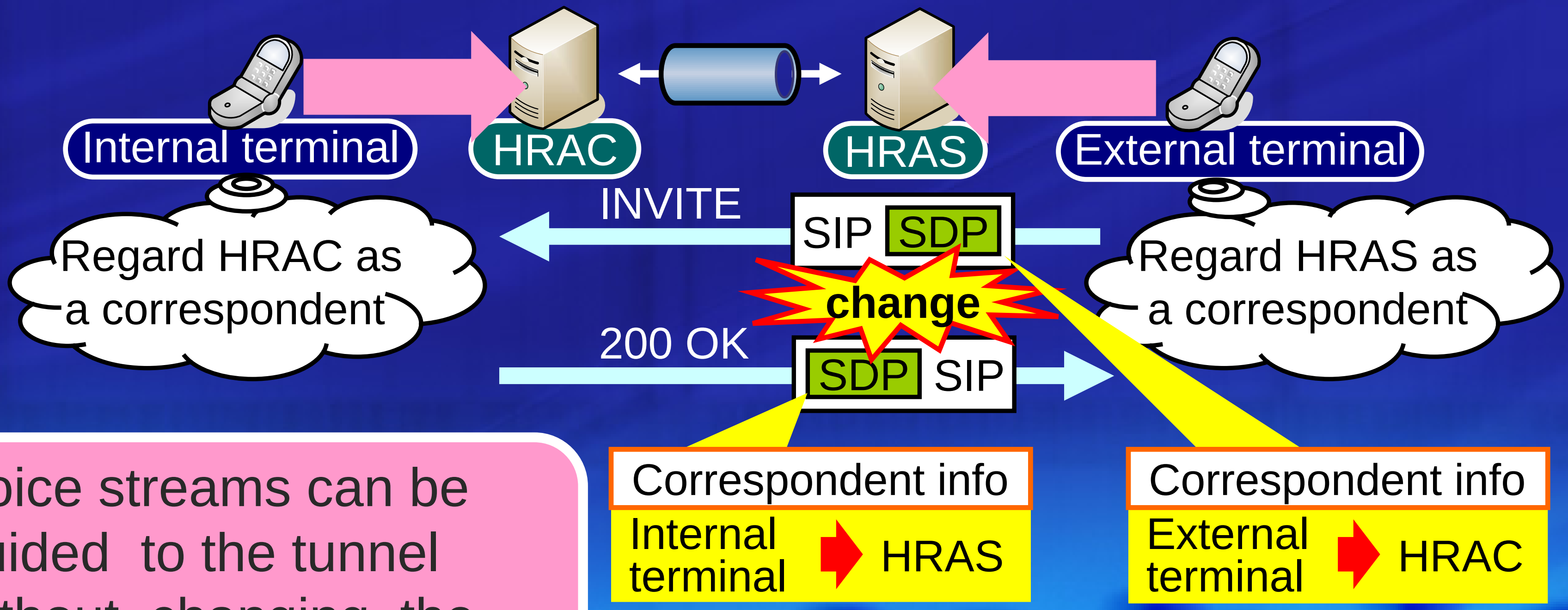
Principle of SoFW

Voice stream guidance

In normal SIP specification,
voice streams are directly exchanged between end terminal.

It is needed to guide the voice streams into the HTTP tunnel.

Change the content of SIP （the dialing phase）



Correspondent info

Internal terminal ➡ HRAS

Correspondent info

External terminal ➡ HRAC

Voice streams can be guided to the tunnel without changing the functions of SIP terminals

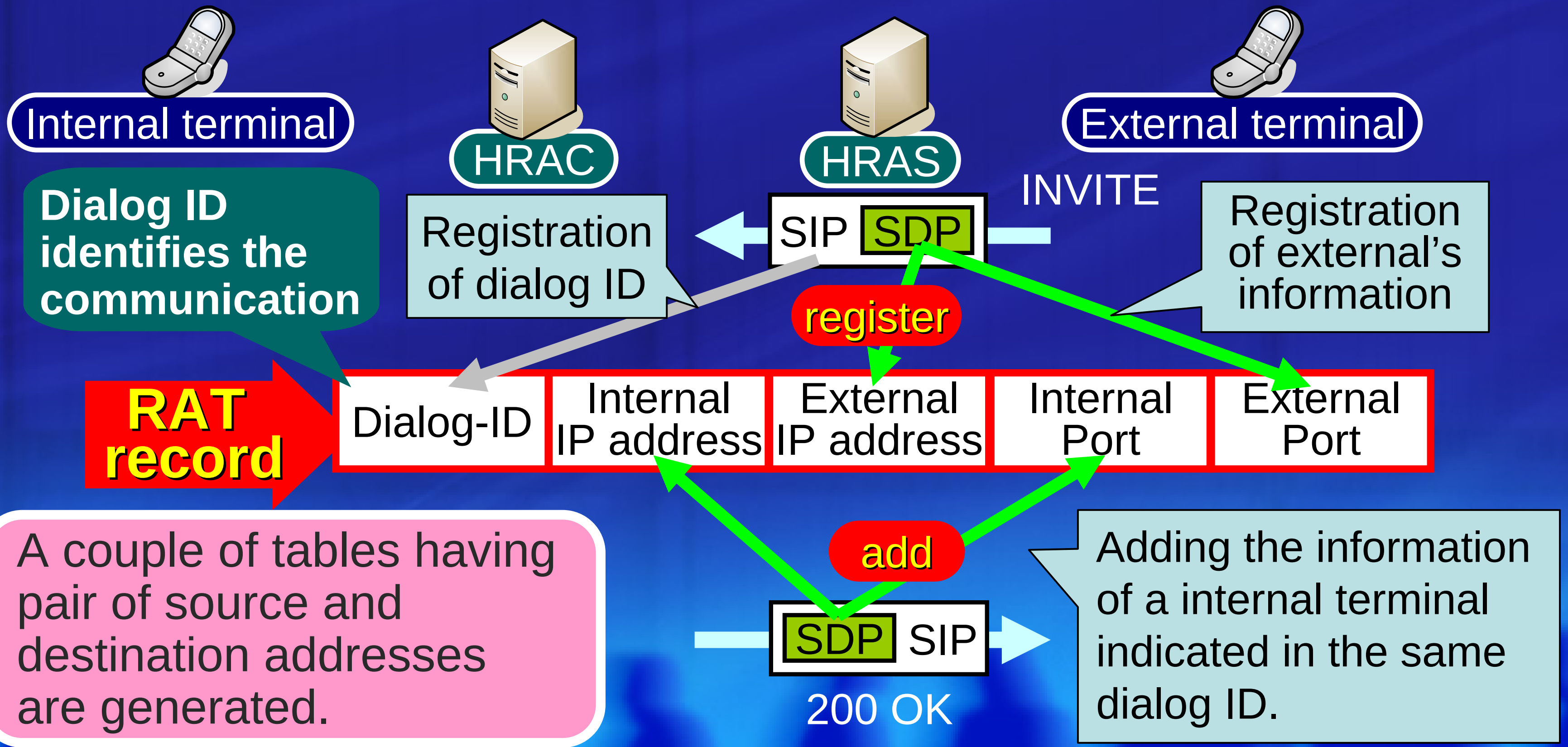
SDP: Session information for voice communication.

Principle of SoFW

IP address change function

HRAS determines the routing path of voice streams referring **RAT (Relay Agent Table)**.

Generation of RAT （the dialing phase）



Dialog ID identifies the communication

Registration of dialog ID

Registration of external's information

Dialog-ID	Internal IP address	External IP address	Internal Port	External Port

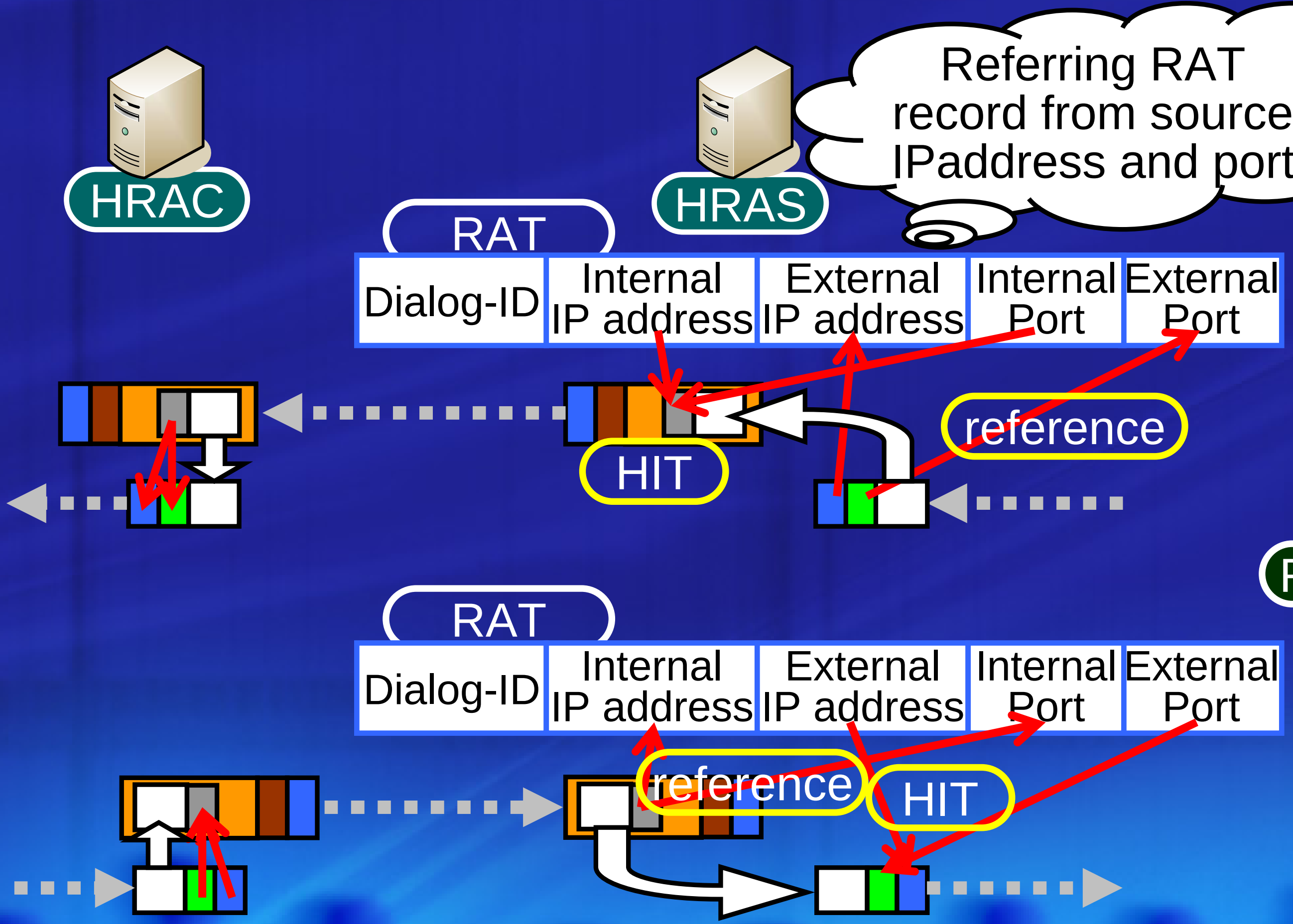
A couple of tables having pair of source and destination addresses are generated.

Adding the information of a internal terminal indicated in the same dialog ID.

Principle of SoFW

IP address changing

Reference RAT （voice stream phase）



Referring RAT record from source IPaddress and port

RAT

Dialog-ID	Internal IP address	External IP address	Internal Port	External Port

RAT

Dialog-ID	Internal IP address	External IP address	Internal Port	External Port

- 🟦 IP
- 🟩 UDP
- 🟫 TCP
- 🟧 HTTP
- ⬜ RA header
- ⬜ Voice data

RA(Relay Agent)header

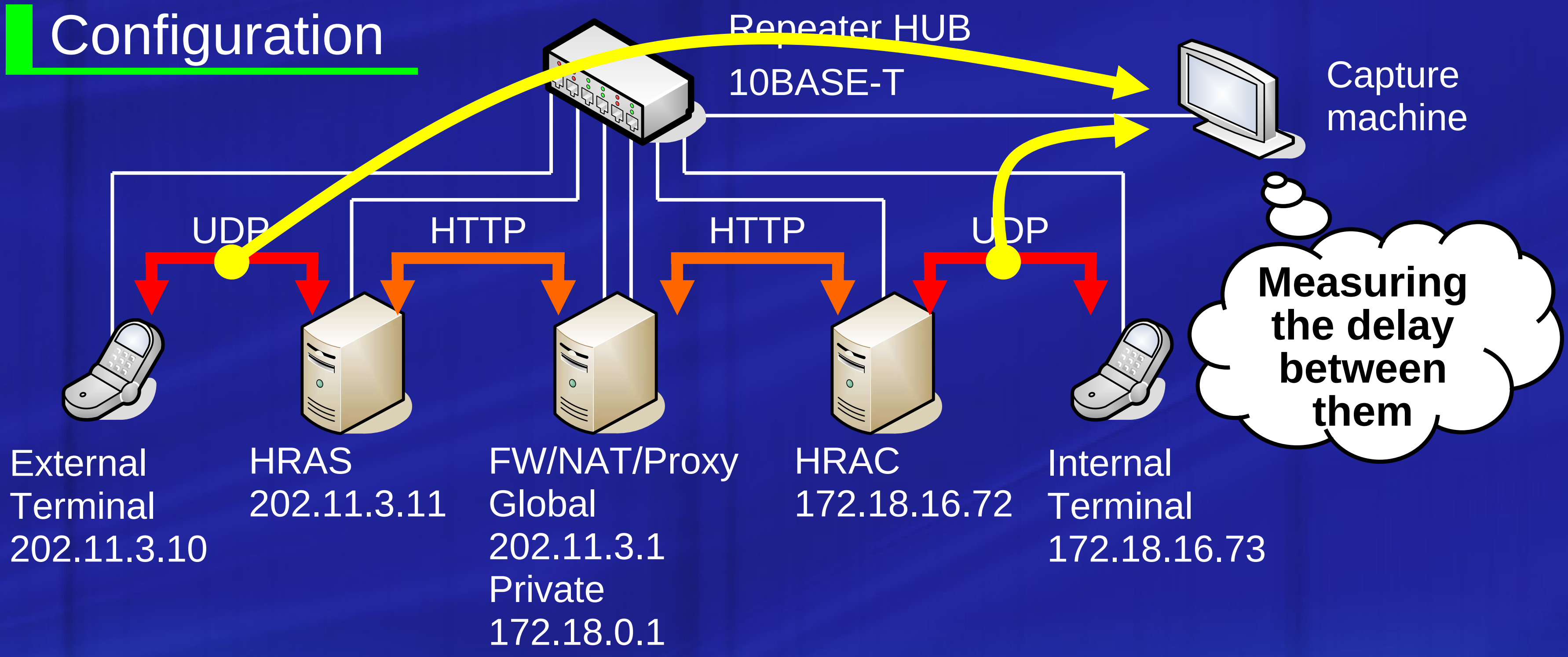
In tunnel it added IP address and port number to save information of terminal

Implementation

- SoFW is implemented in an application layer system of Linux, Fedora core3.0.
- SIP server function in HRAS is realized with SER(SIP Express Router) which is free SIP software.
- Multithread is used for parallel processing.

Evaluation

Configuration



We have measured delay time of voice communications.

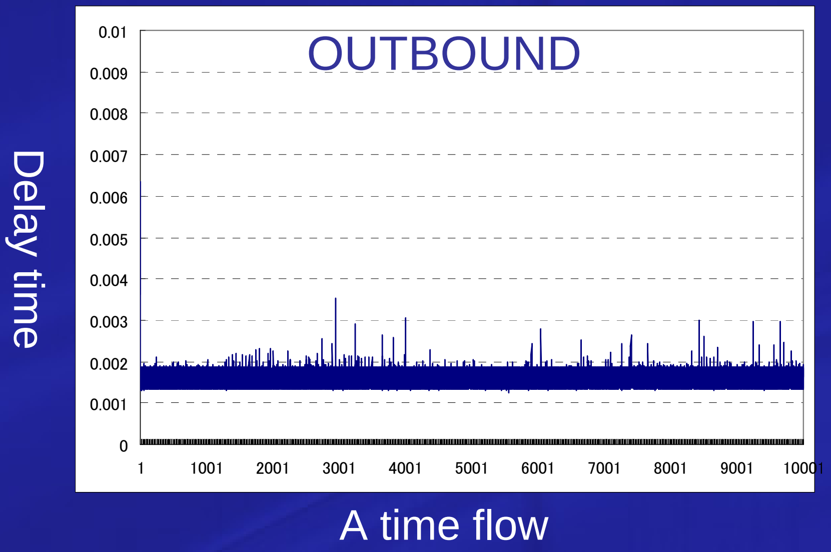
- No other traffic.
- X-Lite for Windows as SIP terminal.
- G.711 as Voice codec.

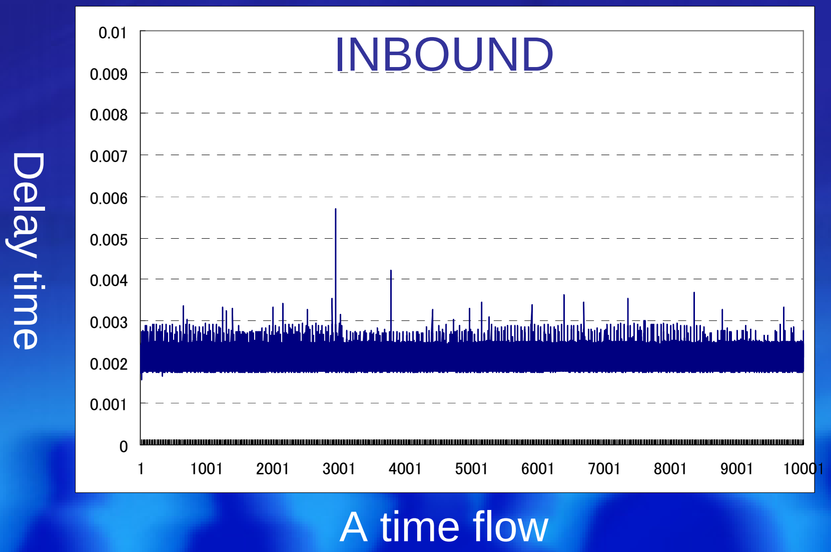
Evaluation

Results of evaluation

Direction of voice stream	Added delay of SoFW
Outbound	1.6 msec
Inbound	2.1 msec

Number of sample packet: 10000



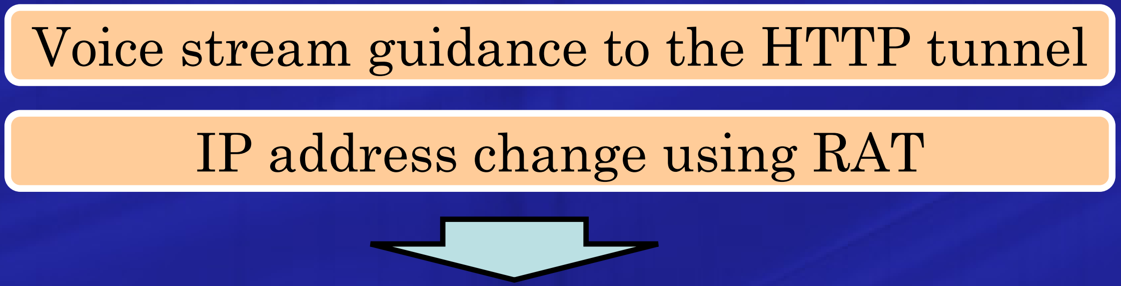


Telephone permissible range is 200msec.

Added delay of SoFW is sufficiently time small.

Summary and future plans

- Conclusion
  - Proposal of IP telephone system passing through FireWall.
  - Explanation of HRAC・HRAS.



Voice stream guidance to the HTTP tunnel

IP address change using RAT

    - Easy introduction
    - Existing SIP terminal
  - Delay time of SoFW is sufficiently small.
- Hereafter
  - Evaluation of throughput degradation with TCP.
  - Evaluation in the case when there exist a number of terminals.