A Realization method of Voice over IP System Passing Through Firewall and its Implementation

Masashi Ito and Akira Watanabe

Graduate School of Science and Technology, Meijo University, Japan
m0432004@ccmailg.meijo-u.ac.jp , wtnbakr@ccmfs.meijo-u.ac.jp

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

1 Introduction

Due to spread of broad band communications and development of backbone networks between 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) are located between the enterprise network and the Internet, prevent the communications of VoIP between the terminals [1]. It is expected that expansion of the VoIP is further promoted if it can safely pass through FW.

There is a protocol based on the existing telephone specification referred to as H.323 [2] which is standardized in an early stage by ITU-T (International Telecommunication Union Telecommunication) as the session initiation protocol of VoIP. However, now, SIP (Session Initiation Protocol) [3] [4] 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 multi-media services. SIP has been employed to most of VoIP systems provided by ISP [5] [6]. 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) [7] 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.

Some systems have been proposed, in which VoIP can pass through the barrier of FW and NAT. They are, for example, HCAP [8], Skype [9], and SoftEther [10]. 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, and there is wasteful traffic caused by constant connections of HTTP flows on FW. 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 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, outline of SoFW and its realization method are described in 2, implementation method is shown in 3, and conclusion is described in 4.

2 Outline of SoFW and its Implementation

2.1 Outline of SoFW

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.

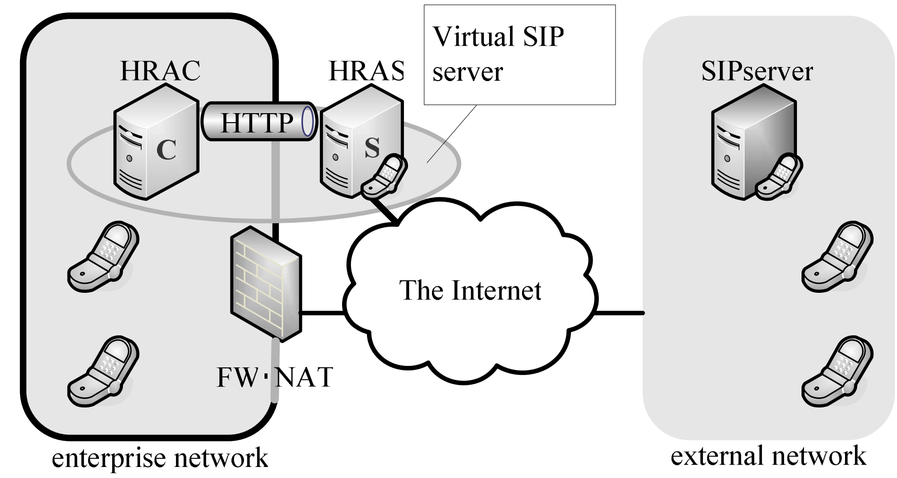


Fig.1. Structure of SoFW

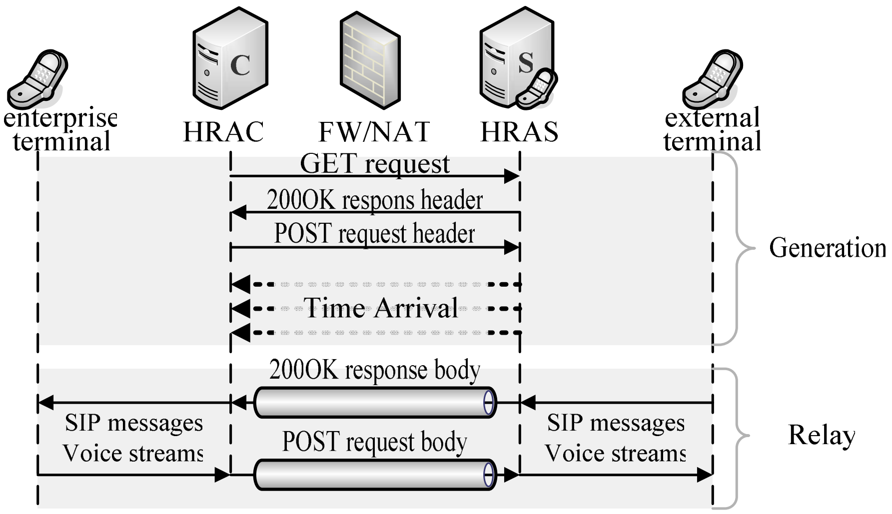


Fig.2. Sequence of a generation of HTTP tunnel

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

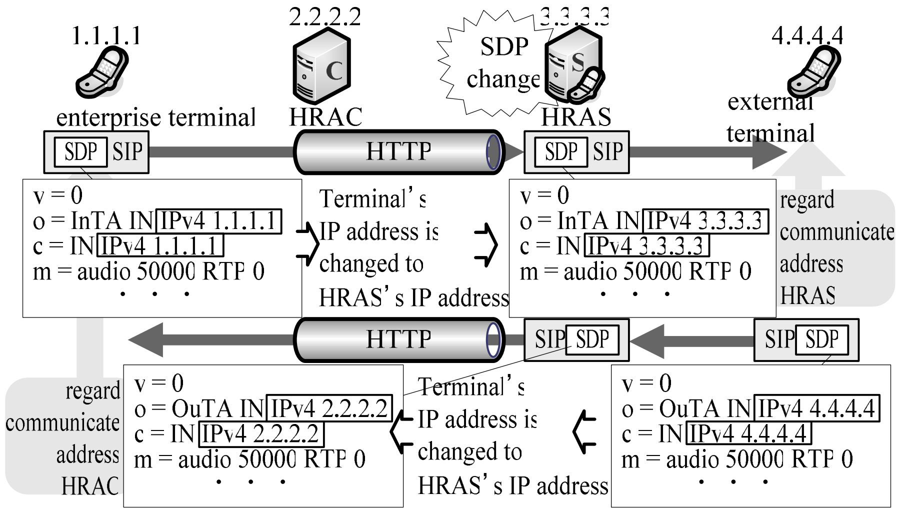


Fig.3. Procedure of changing SDP contents

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. HRAS sends to HRAC messages referred to as Time Arrival at every predetermined time during waiting time in order to keep the TCP connections alive.

2.3 Voice stream guidance by the change of SDP contents

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 [11], 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 communication 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.

Table.1. Contents of RAT

Contents	Explanation
To	Information of destination terminal
From	Information of source terminal
Call-ID	Session descriptor
IIP	IP address of an enterprise terminal
IPort	Port number of an enterprise terminal
OIP	IP address of an external terminal
OPort	Port number of an external terminal

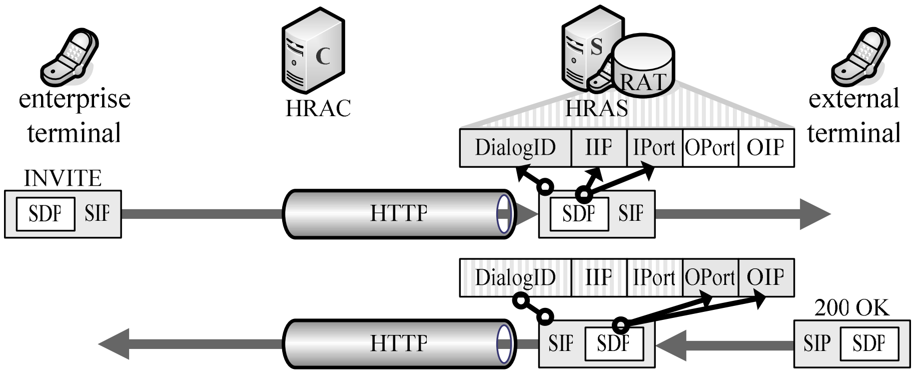


Fig.4. A flow of RAT generation

2.4 Determination of a routing path of voice streams using RAT(Relay Agent Table)

As described in 2.2, 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. Table. 1 shows contents of RAT. To, From, and Call-ID are obtained from SIP headers and they form a dialog ID which 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. Fig. 4 shows a flow of RAT generation when the dialing is started from an enterprise 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.

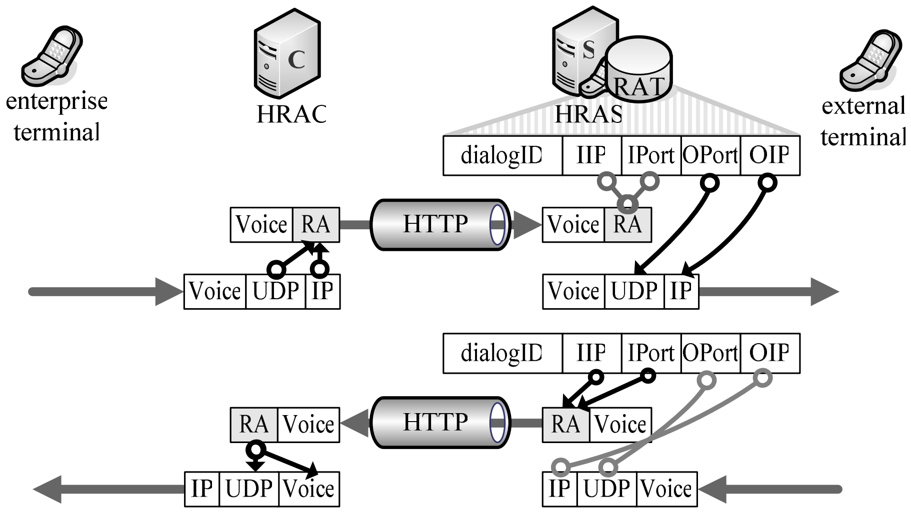


Fig.5. Process flows of voice streams

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 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 external 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 is 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 [12] which is free software of SIP server. Fig. 6 shows the function of HRAS and HRAC, and its data flow. Table 2 indicates functions of HRAC and HRAS. On the portion of dialing process in HRAS, other functions than SER are referred to as SIP Relay Server module. In dialing phase, SIP messages dialed from an external terminal to an enterprise terminal are processed in HRAS by SER at first, and then processed

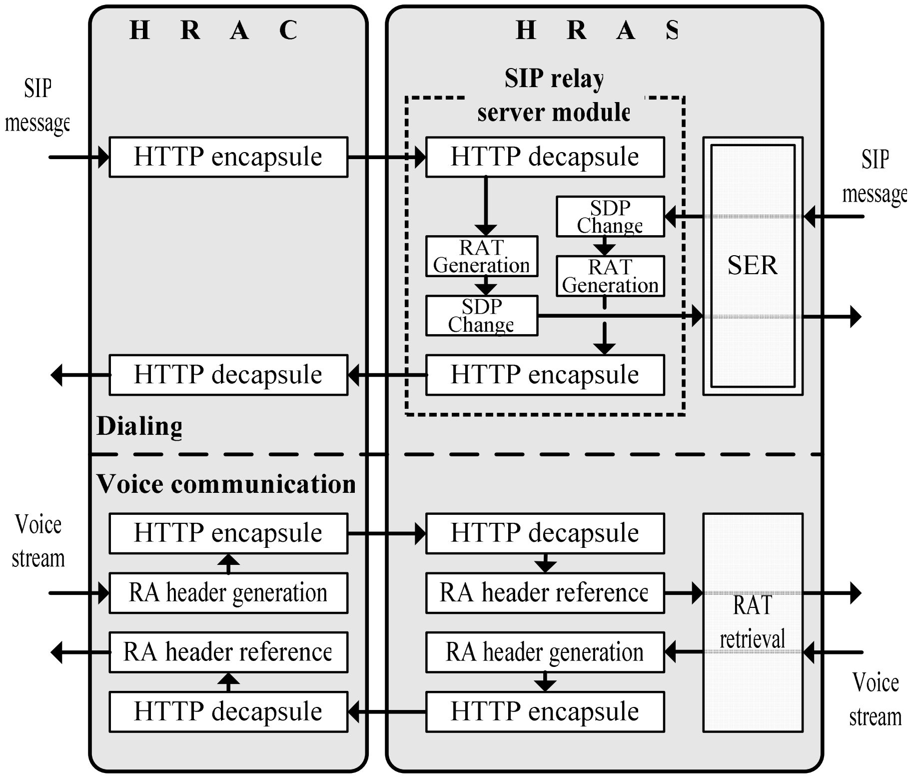


Fig.6. Function and data flow

Table.2. Function of HRAC and HRAS

	Achievement method	Function
HRAS	HTTP tunnel	HTTP encapsule HTTP decapsule
	Voice stream guidance	SDP change
	Determination of a routing path of voice streams	RAT generation RAT retrieval RA header reference RA header generation
	Cooperation with SIP server	SER
HRAC	HTTP tunnel	HTTP encapsule HTTP decapsule
	Determination of a routing path of voice stream	RA header reference RA header generation

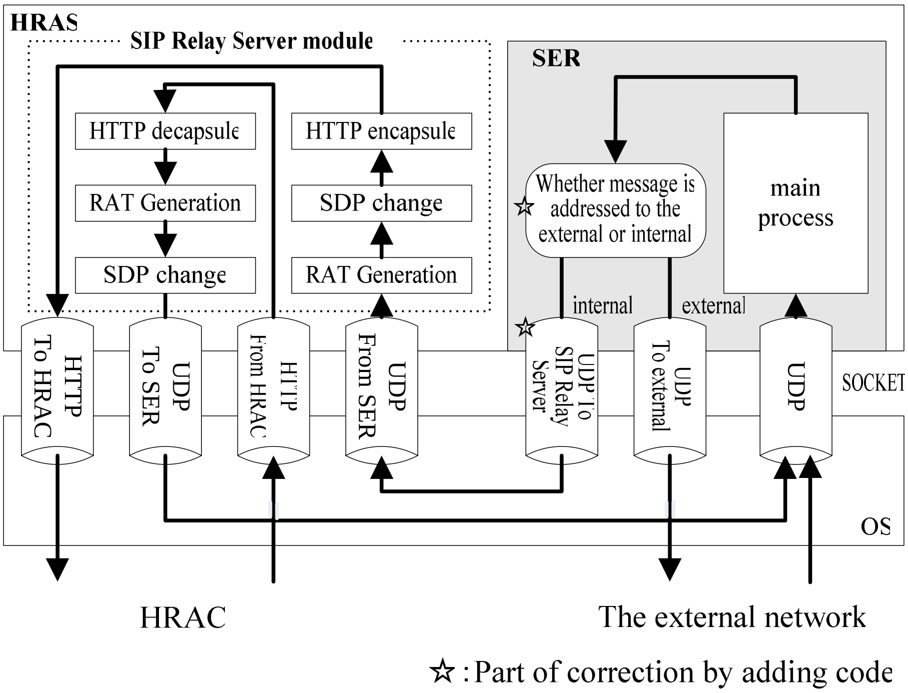


Fig.7. The connection of SER and SIP relay server modules

by SIP Relay Server, in which SDP change, RAT generation, and HTTP encapsule, and relayed to HRAC. In HRAC, the SIP messages are transmitted to an enterprise terminal after processed by HTTP decapsule. SIP messages dialed from an enterprise terminal to an external terminal are relayed to HRAS after processed by HTTP encapsule in HRAC, and processed in HRAS by SIP Relay Server at first, and then SER, and transmitted to an external terminal. In voice communication phase, voice streams transmitted from the external terminal to the enterprise terminal are processed in HRAS in a order of RAT retrieval, RA header generation, and HTTP encapsule, and relayed to HRAC. In HRAC, voice streams are processed in a order of HTTP decapsule and RA header reference, and transmitted to the enterprise terminal. Voice streams transmitted from the enterprise terminal to the external terminal are processed in HRAC in a order of RA header generation and HTTP encapsule, and relayed to HRAS. In HRAS, voice streams are processed in a order of HTTP decapsule, RA header reference, and RAT retrieval, and transmitted to the external terminal. SOCKET is used for a connection between SIP Relay Server and SER, and its connection method is shown in Fig. 7. Fig. 7 shows the boundary portion between SER and SIP Relay Server in Fig. 6 in detail. SIP messages transmitted to HRAS from an external terminal reach SOCKET generated by SER, then after processed by SER, reach SOCKET generated by SIP Relay Server, and then relayed to HRAC. On the other hand, SIP messages relayed to HRAS through HRAC reach SOCKET generated by SIP Relay Server, and after processed by SIP Relay Server, they reach SOCKET generated by SER. Then after processed by SER, they are transmitted to an external terminal. In case of replying a response like SIP registration message transmitted from an enterprise terminal, a reply message (200OK) is made by SER and relayed back to HRAC through SOCKET generated by SIP Relay Server again. To achieve the above-described flow, a simple modification is made to SER. As shown in Fig. 7, 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 Conclusion

In this paper we have described the realization method of SoFW and its implementation. We are now implementing the system, and are going to evaluate delay of voice communications near future. This time, we have limited the application of SIP to voice communications, however, SIP have been paid many attentions to various kinds of applications. Studies for applications other than IP telephone should be considered hereafter.

References

1. N.Freed : Behavior of and Requirements for Internet Firewalls, IETF RFC 2979 (200.10)
2. H.323, Packe Based Multimedia Communications Systems, ITU-T Recommendation, 1998
3. Handley, M., Schulzrinne, H., Schooler, E. and Rosenberg, J.: SIP : Session Initiation Protocol, RFC2543(1999)
4. J. Rosenberg,et all “SIP: Session Initiation Protocol” IETF RFC3261(2002.6)
5. Petri Koskelainen, Henning Schulzrinnc, Xiaotao Wu : VoIP : A SIP-based conference control framework, ACM press 53-61 (2002.5)
6. Stcfan Bcrgcr, Honning Schulzrinnc, Stylianos Sidiroglou, Xiaotao Wu : Conferencing : Ubiquitous computing using SIP, ACM press 82-89(2003.6)
7. K. Egevang, P. Francis : The IP Network Address Translator (NAT), IETF RFC 1631 (1994.5)
8. Shinji Kunai.: Allround Internet Telephony Protocol -HTTP-based Conference Application Protocol ,IPSJ –JNK4403007
9. Skype.: “http://www.skype.com/home.html”
10. SoftEhter.: ” http://www.softether.com/jp/”
11. Handley, M. and Jacobson, V.: SDP : Session Description Protocol, IEETF RFC2327(1998)
12. SER : “http://www.iptel.org/ser/”

A Realization method of Voice over IP System Passing Through Firewall and its Implementation

The International Conference of Information Network 2006

Jan, 16-19, 2006

Sendai, Japan

Masashi Ito and Akira Watanabe

Graduate School of Science and Technology, Meijo Univ, Japan

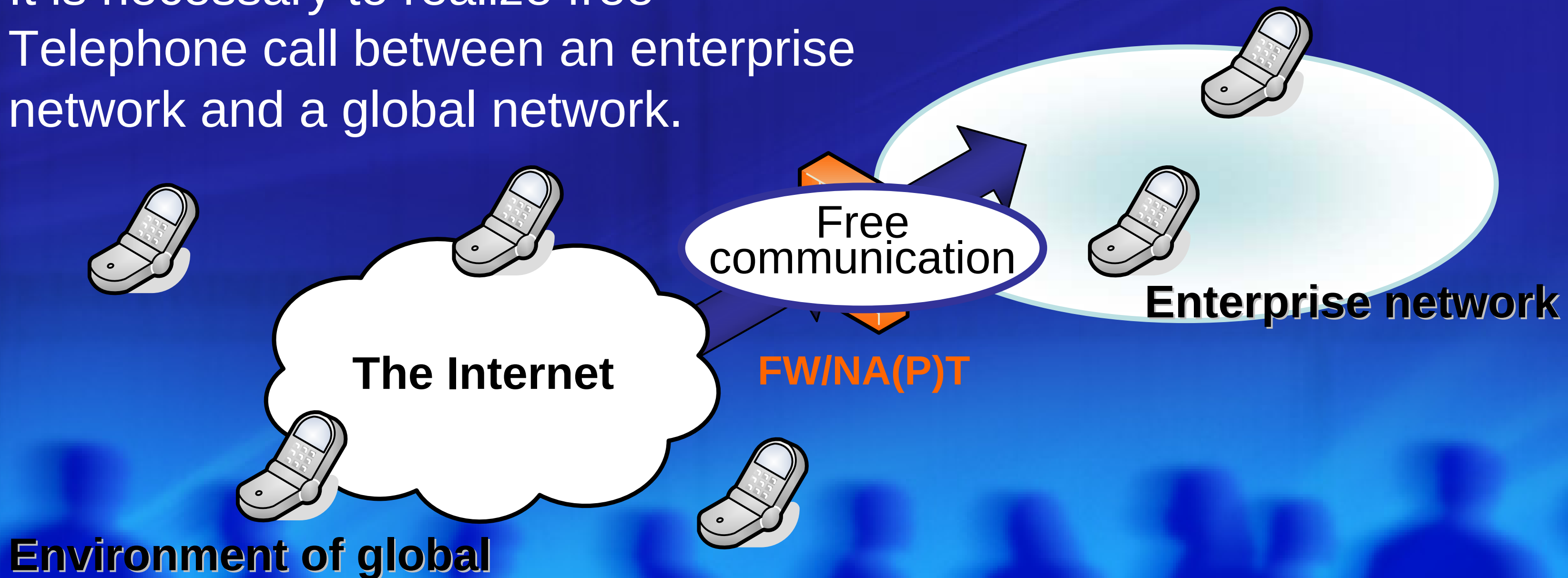
Background

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

→ IP Telephony has became popular among enterprise networks and home networks

To improve the convenience of IP telephony further…

It is necessary to realize free Telephone call between an enterprise network and a global network.

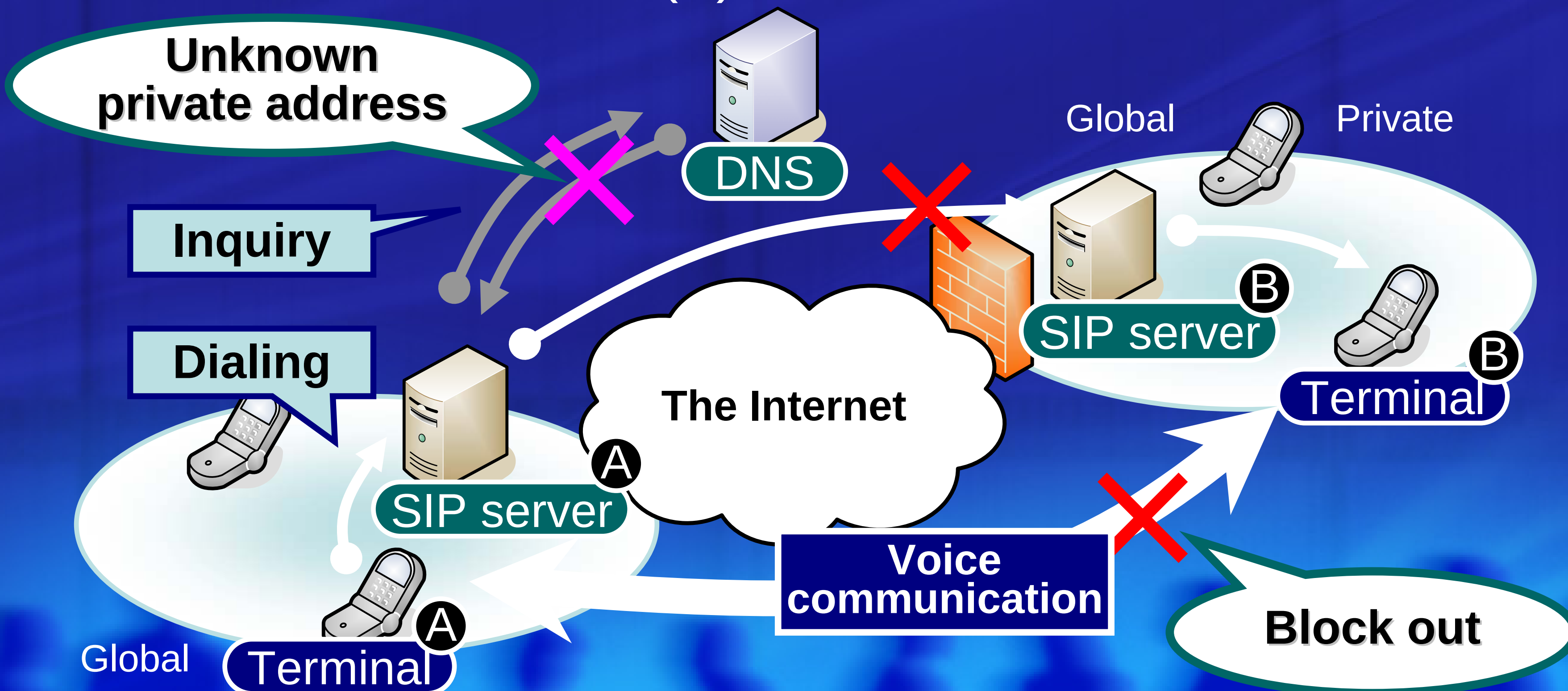


Related studies

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

- Functions are implemented in FireWall itself

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

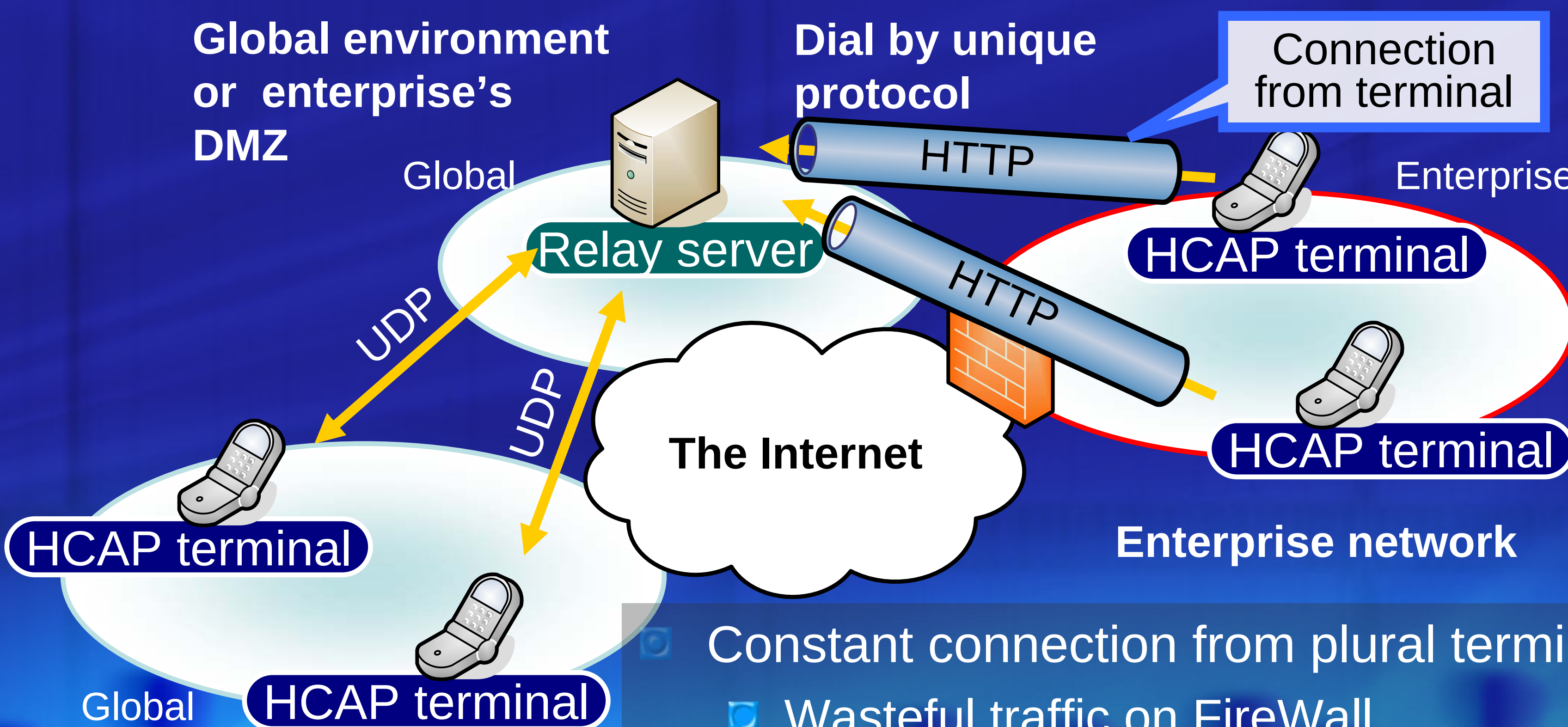
  UPnP, VoIPsecureGW, Connect-VPnP, IP - NAT, etc...

It is needed to change security policy

- Functions are implemented in terminals and servers
  - Skype
  - HCAP
  - SoftEther（protocol free）

Related and early studies HCAP

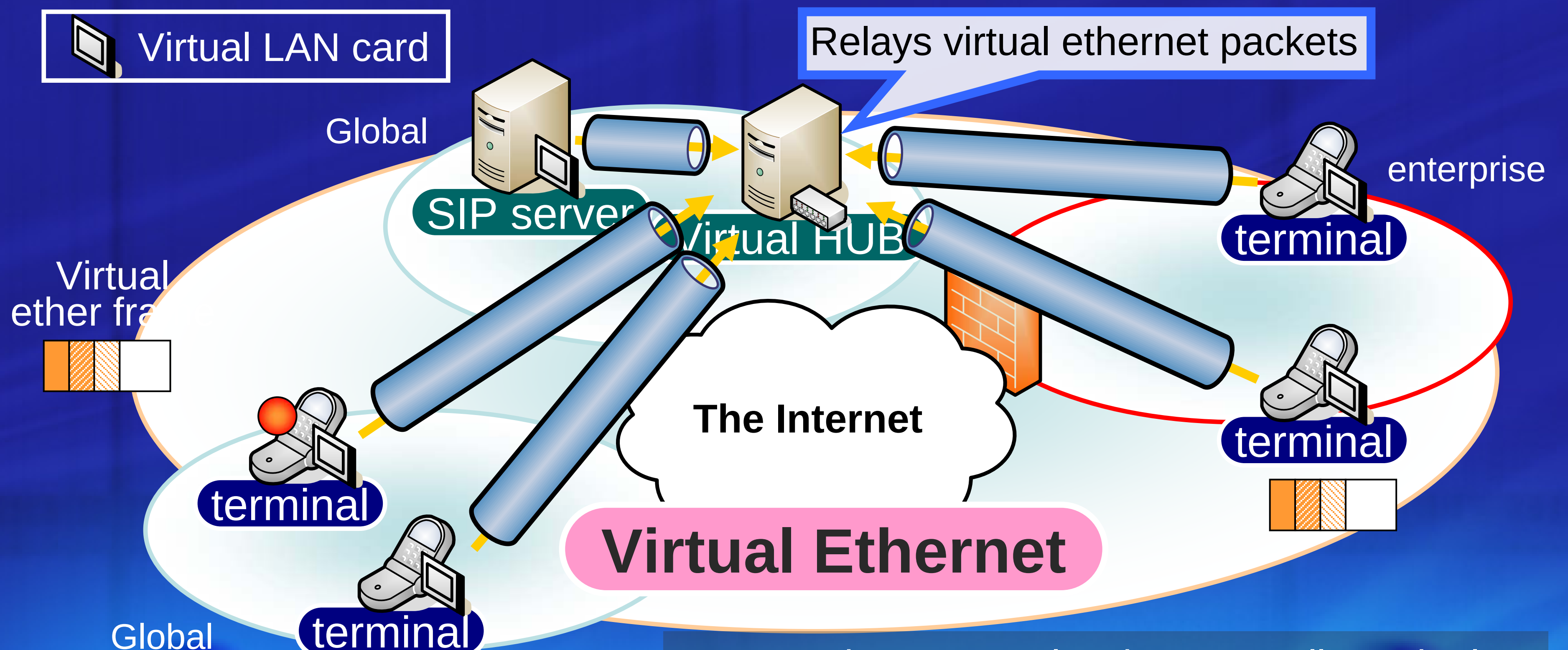
HCAP terminals and a HTTP relay server generate HTTP tunnels between them at the initial phase, and the tunnels relay all dial and voice streams.



- Constant connection from plural terminal
  - Wasteful traffic on FireWall
  - Special terminals are needed

Related and early studies

This system generates a virtual network, among terminals having the function of virtual LAN card, and a relay server having the function of virtual HUB.



Virtual MAC and IP address are allocated in terminals

- Security are very low because all terminals are located in the same LAN.
- Integrated control of IP addresses is needed.

Principle of SoFW

SoFW（SIP over Fire Wall）

Purpose of the system

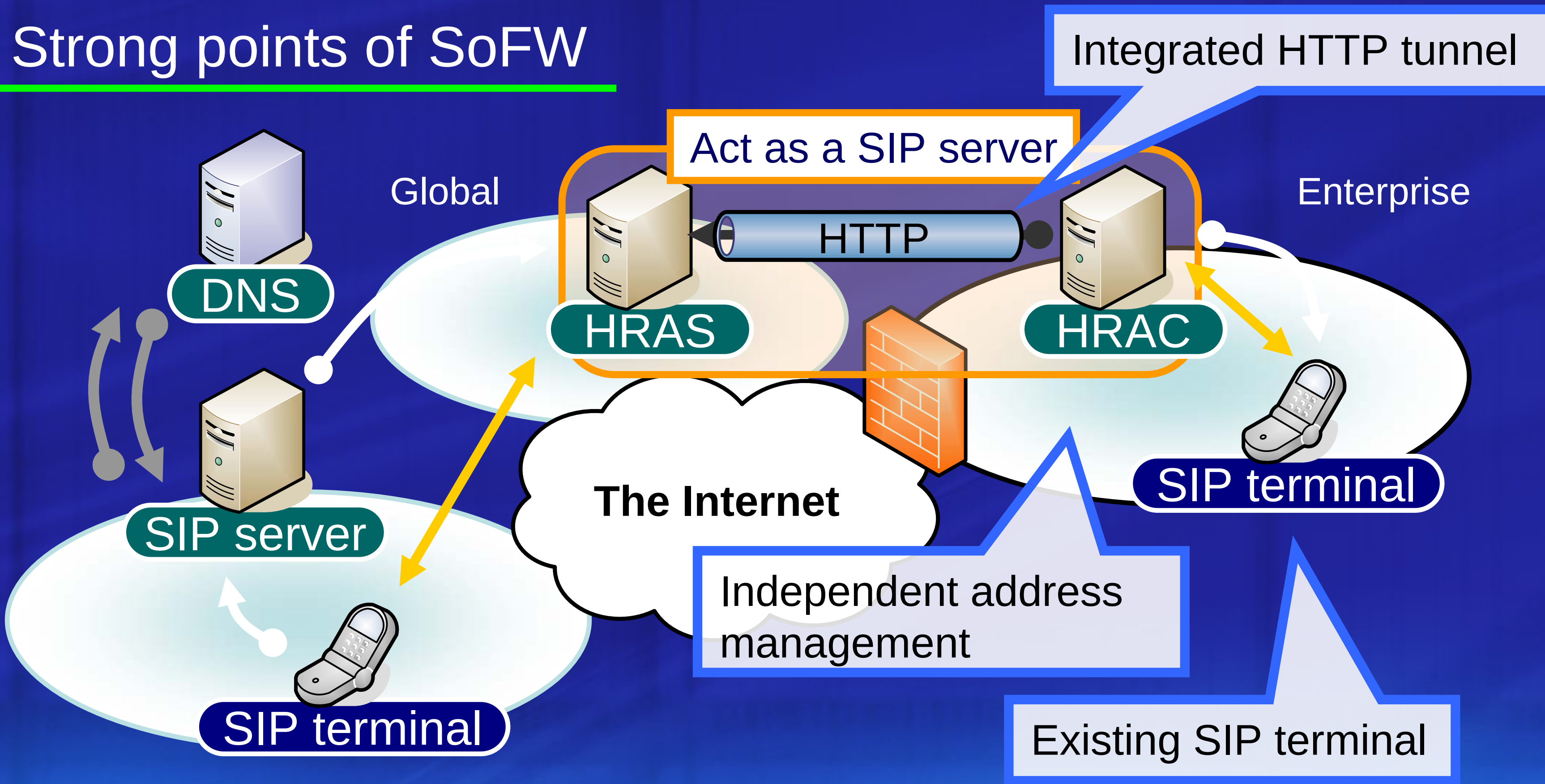
- Easy introduction
- No wasteful traffic on FireWall
- Separation of different types of address area
- Secure network

It relay SIP and Voice streams are relayed using two devices which are set up in an enterprise network and a global network.

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

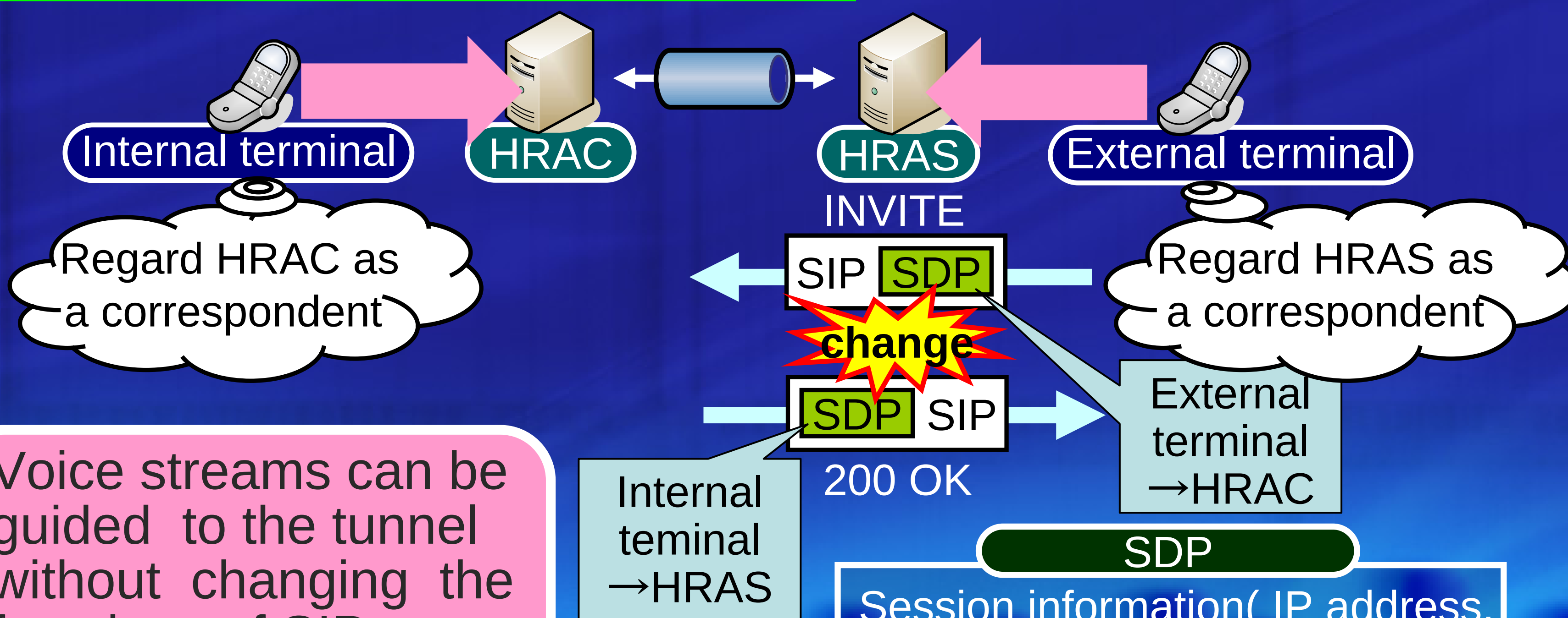
Principle of SoFW

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 connection information (In dialing phase)



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

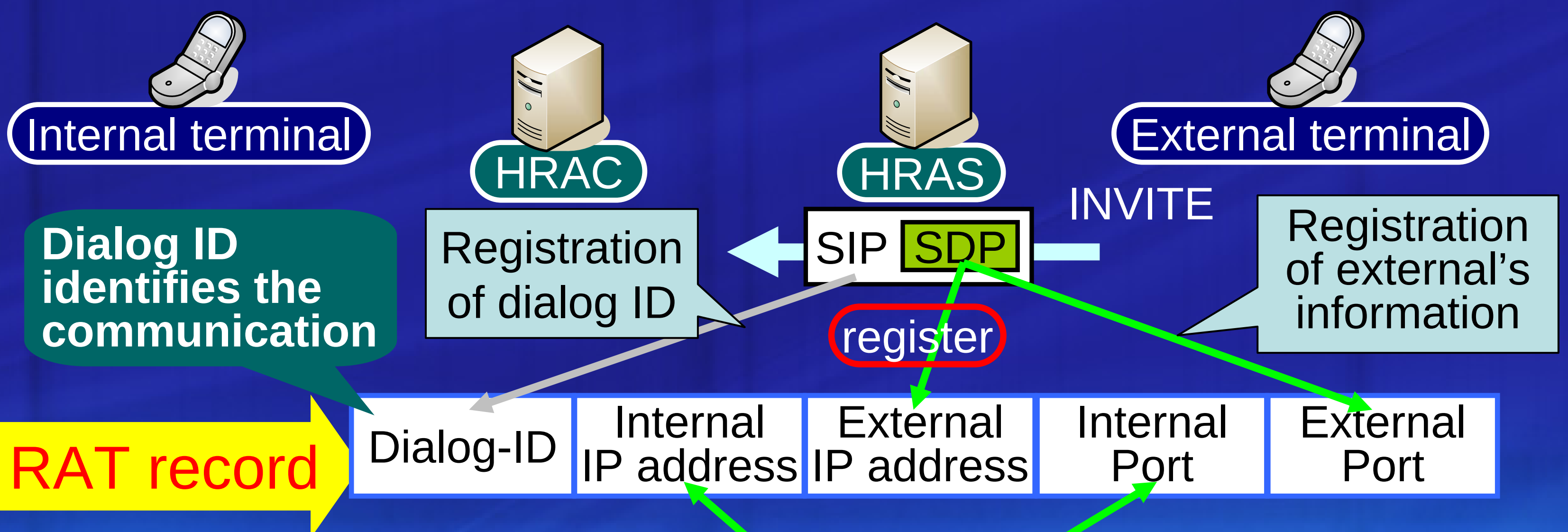
SDP
Session information(IP address, port number etc.) for voice communication.

Principle of SoFW

Relay agent table

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

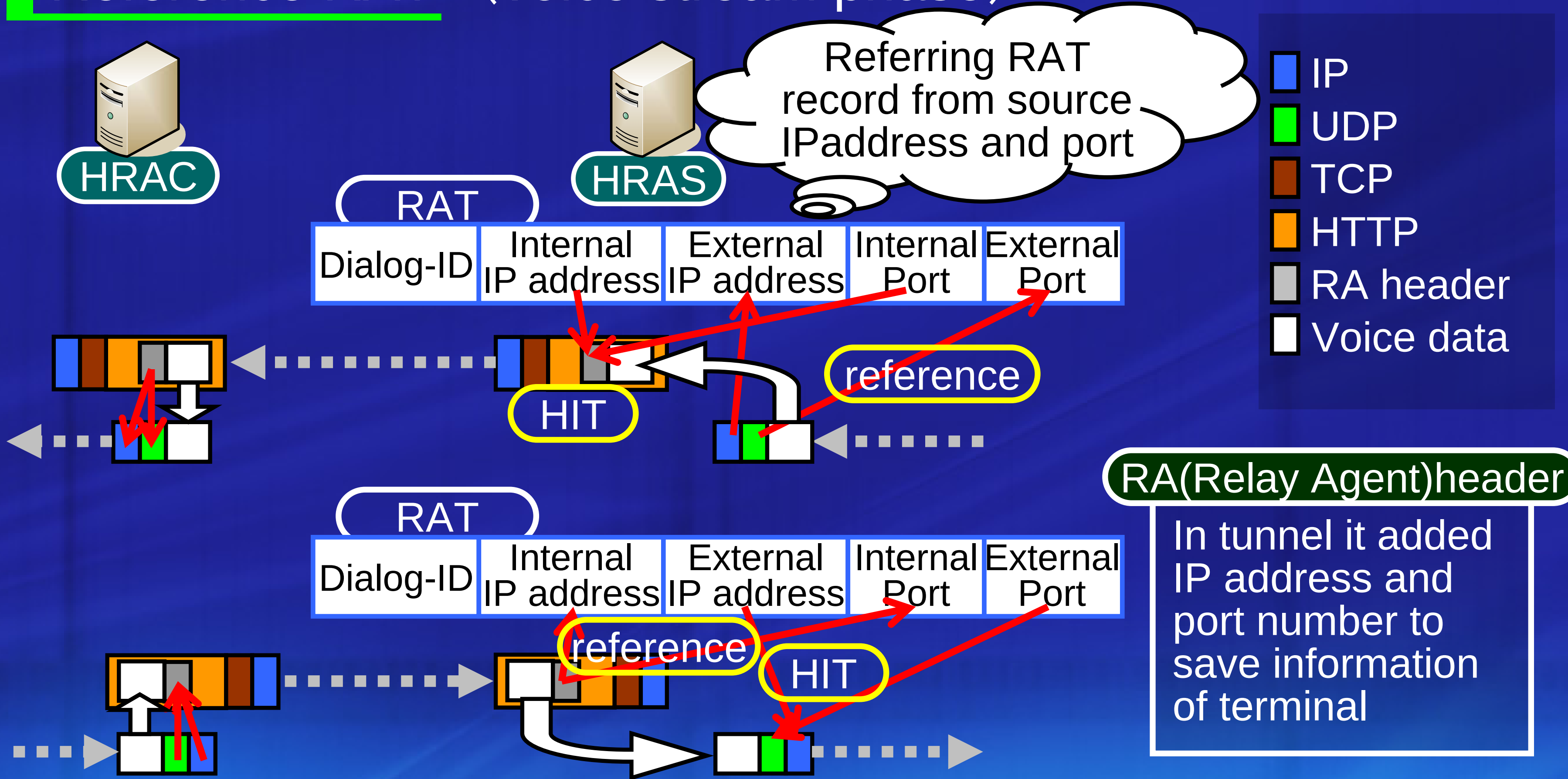
Generation of RAT （in dial phase）



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

Principle of SoFW

Reference RAT （voice stream phase）



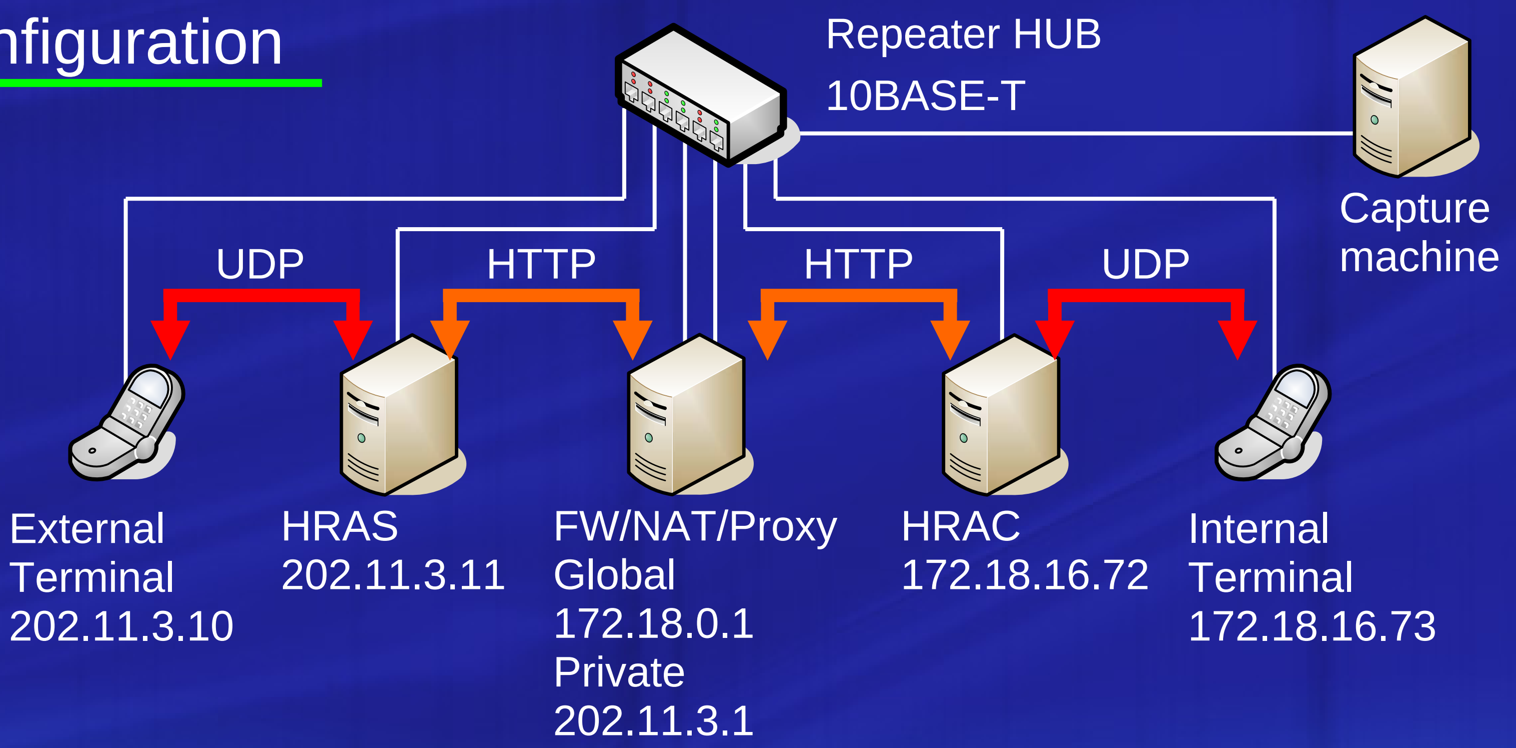
By using RAT, SoFW can keep independence of different address areas.

Implementation

- SoFW is implemented as an application of Fedora core3.0 (linux 2.6.9)
- SIP server function in HRAS is realized with SER(SIP Express Router) which is free SIP server
- In HRAS, main processes are implemented.
- In HRAC, as minimum process is implemented.
- Multithread is used for parallel processing to speed up memory reference.

Evaluation

Configuration



- We measured throughput of the processing time of SoFW.
- No traffic.
- X-Lite for Windows as SIP terminal.
- G.711 is used as Voice codec.

Evaluation

A result of experimentation

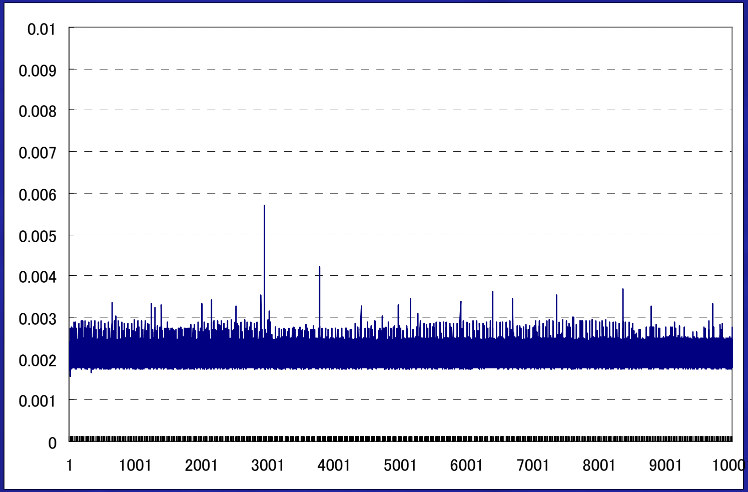
Direction of voice stream	Added delay of SoFW
Outbound	1.641msec
Inbound	2.087msec

Number of sample packet: 10000

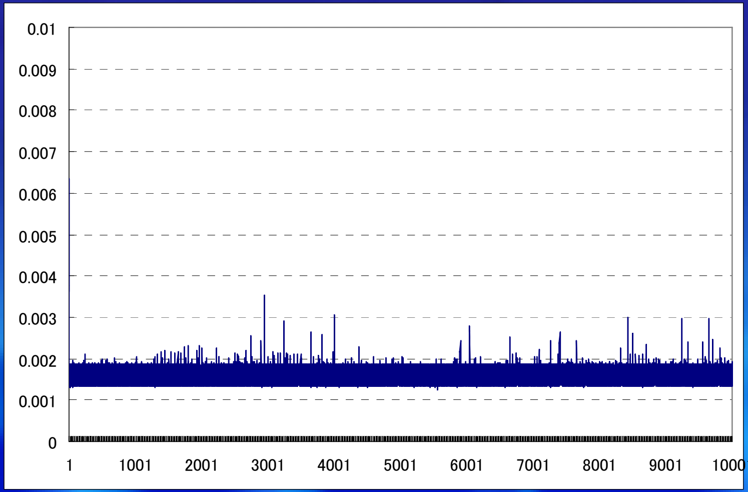
Permissible range 200msec.

Added delay of SoFW is sufficiently small.

INBOUND



OUTBOUND



Summary and future plans

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

Guidance voice stream to tunnel

Determine routing path by using RAT

- Easy introduction
- No wasteful traffic on FireWall
- Separation of different types of address areas
- Secure network

  - Added delay of SoFW is sufficiently
- Hereafter
  - Evaluation of throughput degradation with TCP.
  - Evaluation on in the case when there exists a number of terminals.