

**AN ANALYSIS OF VOICE OVER INTERNET PROTOCOL (VOIP)
AND ITS SECURITY IMPLEMENTATION**

By

MUHAMMAD RIZAL BIN AZMI

2378

FINAL REPORT

**Submitted to the Electrical & Electronics Engineering Programme
in Partial Fulfillment of the Requirements
for the Degree
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)**

Universiti Teknologi Petronas

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

Copyright 2006

by

Muhammad Rizal bin Azmi, 2378

CERTIFICATION OF APPROVAL

**AN ANALYSIS OF VOICE OVER INTERNET PROTOCOL (VOIP)
AND ITS SECURITY IMPLEMENTATION**

by

MUHAMMAD RIZAL BIN AZMI

2378

A project dissertation submitted to the
Electrical & Electronics Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfillment of the requirement for the
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

Approved:




Mr. Azlan Awang
Project Supervisor

**UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK**

June 2006

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



MUHAMMAD RIZAL BIN AZMI

ABSTRACT

Voice over Internet Protocol (VoIP) has been in existence for a number of years but only quite recently has it developed into mass adoption. As VoIP technology penetrates worldwide telecommunications markets, the advancements achieved in performance, cost reduction, and feature support make VoIP a convincing proposition for service providers, equipment manufacturers, and end users. Since the introduction of mass-market VoIP services over broadband Internet in 2004, security and safeguarding are becoming a more important obligation in VoIP solutions. The purpose of this final year project is to study and analyze VoIP and implement the security aspect using Secure Real-time Transport Protocol (SRTP) end-to-end media encryption in the Universiti Teknologi PETRONAS (UTP) laboratory. Extensive research, evaluation of case studies, literature reviews, network analysis, as well as testing and experimentation are the methods employed in achieving a secure and reliable VoIP network. With the given time frame and adequate resources, the study and analysis of VoIP and implementation of SRTP should prove to be very successful.

ACKNOWLEDGEMENTS

First and foremost, I would like to praise God Almighty who have helped and guided me through the tough times in completing my two semester Final Year Project (FYP) in Universiti Teknologi PETRONAS (UTP). Without His blessing, I may not get to where I stand today.

My utmost gratitude goes to my FYP supervisor, Mr. Azlan Awang, lecturer in UTP Electrical and Electronics Engineering faculty who has been excellent in providing the information and guidance to me regarding data networks, communications, as well as other areas throughout the whole course. His leadership, supervision, and guidance have led to the success of this project. In addition, thanks to the FYP committee for providing me direction and guidance throughout the entire project.

I would also like to express my appreciation to Mr. Anang Hudaya, lecturer in UTP Information Technology faculty for his counsel in the network security area. Not forgetting Mr. Ruslan Idris, lab technician of UTP Data Communications Lab 02 who has been very helpful in assisting me in lab-related works as well as network setup and configuration. Many thanks also to Bobby Ryxler and Megat Azim Jamil, UTP graduate students who have assisted me in my research throughout the project.

In this little moment, I would like to convey my deepest gratitude to my mother, Mrs. Rohana Yusuf, for her full encouragement. Also, thanks to the most important person to me, Ernadia Os'hara Omar for her support and understanding.

Last but not least, thanks to all my friends; Irman Johan, Khairil Zakwan Chung, Khalid Yashaiya, M Fadzril, M Nizar, M Taufiq, Nyleonald Taba, Suzianna Termidzi, Wan Hirman and others; who have contributed their part and effort as well as their acquaintance in making this project a huge achievement.

TABLE OF CONTENTS

CERTIFICATION	i
ABSTRACT	ii
ACKNOWLEDGEMENT	iii
CHAPTER 1:	INTRODUCTION	1
	1.1 Background of Study	1
	1.2 Problem Statement	2
	1.3 Objectives and Scope of Study	2
CHAPTER 2:	LITERATURE REVIEW/THEORY	4
	2.1 VoIP Fundamentals	4
	2.2 VoIP Protocol Stack	7
	2.3 VoIP Security Overview	9
	2.4 VoIP Components	9
	2.5 QoS Issues in VoIP Security	11
	2.6 Secure Real-time Transport Protocol	13
CHAPTER 3:	PROJECT WORK	15
	3.1 Methodology	15
	3.2 Hardware	16
	3.3 Router Configuration	17
	3.4 Software	22
CHAPTER 4:	RESULTS AND DISCUSSION	32
	4.1 VoIP Test Call Results	32
	4.2 H.323 Protocol Discussion	34

4.3	SIP Protocol Discussion	.	.	37
4.4	H.323 and SIP Comparison	.	.	39
4.5	PSTN to VoIP Transition	.	.	41
4.6	SRTP Testing Results	.	.	43
4.5	SRTP Discussion	.	.	45
CHAPTER 5:	CONCLUSION AND RECOMMENDATION			47
5.1	Conclusion	.	.	47
5.2	Recommendation	.	.	48
REFERENCES	.	.	.	49
APPENDICES	.	.	.	50

LIST OF FIGURES

- Figure 2.1 Basic VoIP Architecture (Call Originating)
- Figure 2.2 VoIP protocols in the TCP/IP protocol stack
- Figure 2.3 Secure Real-time Transport Protocol (SRTP) header
- Figure 3.1 VoIP Hardware Connection in Data Communications Lab
- Figure 3.2 Cisco 3725 Multiservice Access Router ports
- Figure 3.3 Example of basic VoIP connection showing respective dial-peers and call legs
- Figure 3.4 MiniSIP User Agent (UA) Graphical User Interface (GUI)
- Figure 3.5 MiniSIP account settings menu
- Figure 3.6 MiniSIP security settings menu
- Figure 3.7 Snom 360 User Agent (UA) Graphical User Interface (GUI)
- Figure 3.8 Snom 360 web interface menu – Login Information
- Figure 3.9 Snom 360 web interface menu – RTP Line Settings
- Figure 3.10 CommuniGate Pro Server web interface menu – User Accounts
- Figure 3.11 Ethereal capture for an SIP call with RTP stream analysis
- Figure 3.12 Cain & Abel VoIP Sniffer window
- Figure 4.1 H.323 Protocol Stack
- Figure 4.2 Basic H.323 call setup and teardown between two terminals
- Figure 4.3 SIP protocol in the IETF multimedia conferencing architecture
- Figure 4.4 Basic SIP call setup and teardown between two endpoints
- Figure 4.5 Ethereal RTP stream analysis (SRTP-disabled)
- Figure 4.6 Ethereal RTP stream analysis (SRTP-enabled)



LIST OF TABLES

- Table 2.1 Encrypted calls channel impact per DSP using various codecs on Cisco analog voice modules
- Table 3.1 Comparison of various CODEC characteristics
- Table 4.1 Comparison between H.323, SIP, and MGCP/H.248/Megaco protocols
- Table 4.2 Parameters of SRTP-disabled VoIP call from user 110 to user 111
- Table 4.3 Parameters of SRTP-enabled VoIP call from user 110 to user 111
- Table 4.4 Comparison of VoIP parameters and recommended upper limits



CHAPTER 1

INTRODUCTION

1.1 BACKGROUND OF STUDY

The aspiration to carry real-time voice in addition to data over packet-switched networks started in the 1970s. The Advanced Research Projects Agency funded a project in the year 1975 to look at the feasibility of integrated voice and data packet networks. Integrated Services Digital Network (ISDN) research also began in the same era but only received some attention in the early 1980s. In the mid-1990s, the first commercial Voice over Internet Protocol (VoIP) client software produced by VocalTel Communications Ltd (a United States based company) was released. Mass-market VoIP services over broadband Internet began sprouting in 2004 and now more organizations look into VoIP as a fundamental shift away from circuit-switched telephony.

In the VoIP security area, Voice over IP Security Alliance (VoIPSA) was formed in February 2005 from the collaboration of VoIP and Information Security vendors, providers, and thought leaders. Its mission is to promote the current state of VoIP security research, VoIP security education and awareness, and free VoIP testing methodologies and tools. Through the formation of this group, resources could be gathered more efficiently and discussions could be done in a more organized and centralized manner. Currently, top VoIP vendors like Cisco and Avaya are gaining ground in securing VoIP by innovating maximum security networks in their VoIP systems.

1.2 PROBLEM STATEMENT

Voice over IP routes voice transmission over any Internet Protocol (IP) network such as the Internet or a Local Area Network (LAN). Analog telephony requires per-call costs but VoIP can eliminate this once the IP network backbone is in place. The IP networks backbone can carry simultaneous data such as video, data, and voice, which allows more information to travel, thus increasing flexibility. VoIP routes voice traffic over Packet-Switched (PS) networks, as opposed to analog telephony which transmits over the Circuit-Switched (CS) network. This factor makes it more susceptible to security threats, vulnerabilities, and attacks as it is easier to tap into the PS network than the CS network. Protecting the VoIP network would ensure the integrity, confidentiality, and availability of the system while protecting the privacy of end-users. An in-depth research on VoIP security would minimize or eliminate the risks of threats in the network. Hence, a study in VoIP security is highly recommended.

1.3 OBJECTIVES AND SCOPE OF STUDY

The objectives and scope of study are as follows:

- ***Study the fundamentals of VoIP networks***

A stable and flexible VoIP system would reduce or eliminate per-call costs of traditional circuit-switched telephony when calls are initiated. It would also create a highly-flexible and upgradeable telephone system which is built 'future-proof'. As VoIP is bound to replace analog telephony, research on the fundamentals of VoIP is required.

- ***Study the configuration of switches and routers for VoIP***

An understanding of the configuration of switches and routers is essential in enabling and optimizing VoIP. This includes knowledge of the originating and destination addresses, routing addresses, codec and quality-of-service selection. Properly configured switches and routers would create a robust and resilient VoIP network.

- ***Protecting the privacy and confidentiality of VoIP users***

Voice conversations can contain crucial information such as confidential company information and credit card numbers. Traditional analog telephony running over a CS network is deemed almost 100% secure in protecting user conversations. As VoIP runs over data networks, it is susceptible to packet capturing and eavesdropping. Thus, implementation of SRTP would ensure the privacy and confidentiality of users.

- ***Implementing a security mechanism without affecting VoIP performance***

Securing VoIP is basically done either in the application/transport layer using Secure Real-time Transport Protocol (SRTP) or the network layer using Internet Protocol Security (IPSec). The advantage of SRTP over IPSec is that it has minimal or no significant impact on the bandwidth of VoIP traffic. IPSec also has the added disadvantage of complicated Network Address Translation (NAT) and firewall traversal as well as compulsory Operating System (OS) and User Agent (UA) interaction.

- ***Practice project management skills***

The final year project is also a chance to practice project management skills. This is important and crucial as it determines the success or failure of a certain project. Essential aspects include time management, communication skills, and presentation.

CHAPTER 2

LITERATURE REVIEW/THEORY

2.1 VOIP FUNDAMENTALS

VoIP is fundamentally the routing of voice calls over any IP network including LANs and the Internet. It differs from traditional phones which uses circuit-switched lines by transmitting voice over a packet-switched network. VoIP transforms analog phone signals into digital signals and then transmits them to the receiver via an IP network. The following diagram depicts a basic VoIP architecture on the origin side using three different types of dialers:

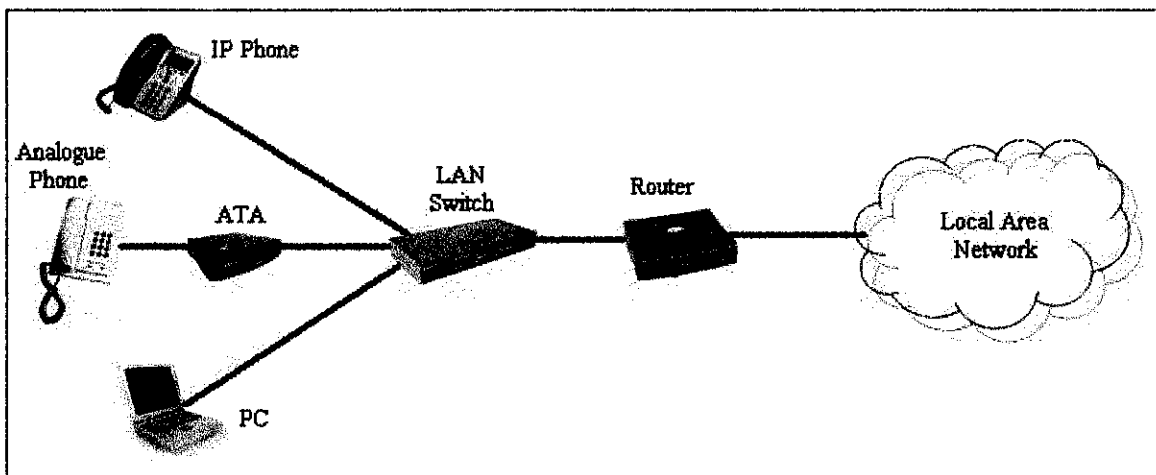


Figure 2.1: Basic VoIP Architecture (Call Originating)

Generally, there are three types of equipment which are used to make a VoIP call. They can co-exist and be utilized on the same VoIP network. The available options for a VoIP-enabled phone are:

- IP Phone
- Analog Phone with Analog Telephone Adaptor (ATA)
- Software Phone via a Personal Computer (PC)

IP Phone

An IP Phone is basically a VoIP-ready telephone. It can be plugged in directly to an RJ-45 port of a switch without additional hardware or software and usually has an on-screen display integrated into it for number displays, Calling Line Identification Presentation (CLIP), and other purposes. Some of the more expensive IP phones come with a built-in display for video calls. There are also wireless IP phones which can be utilized for Voice over Wireless LAN (VoWLAN).

Analog Phone with Analog Telephone Adaptor (ATA)

An Analog Telephone Adaptor (ATA) is required if legacy analog phones are used to make VoIP calls. This equipment acts as a medium between analog telephones and the IP network, allowing it to be operational for VoIP. It usually supports one to two voice ports independently and has an internal Ethernet switch which allows for a direct connection to the Ethernet LAN or a PC via an RJ-45 interface. By using ATAs, the analog telephones can still be kept for utilization over an IP network, thus saving purchase costs.

Software Phone via a Personal Computer (PC)

A software phone is an IP phone in software form. It is installed in a PC and functions the same as an IP phone. Software phones require a headset or a microphone and speakers connected to the soundcard of the PC to function. Most of them are available for free downloads on the Internet such as SJphone, Adore Softphone, and others. Microsoft NetMeeting, which comes bundled with Microsoft Windows XP operating systems, can also be used as a software phone. The advantages of software phones are customizable user interfaces, integration of address books, speed dials, and other features.

Users need one of the mentioned VoIP-enabled phones connected to the IP network to make and receive VoIP calls. To establish a connection, a form of signaling is required which is understood by both clients such as Signaling System #7 (SS7). A suitable naming and addressing scheme is also required in addition to standards and protocols. Suitable protocols include H.323 and the current leading Session Initiation Protocol (SIP). This is vital for interconnectivity between different VoIP vendors and products.

H.323 was initially created by the International Telecommunications Union, Telecommunications Standardization Sector (ITU-T) to provide audio-visual communication sessions on any packet network but has evolved to address the needs of VoIP networks. SIP was developed by the Internet Engineering Task Force (IETF) for interactive user sessions that involve multimedia elements. It is currently the leading protocol for VoIP and gradually replacing H.323 in this role.

After a connection is established, the format of the media to be transmitted is decided. The data transmitted can be purely-audio or include video and/or other data together. In Public Switched Telephone Network (PSTN) systems, it is always assumed that it would be an audio only 3.1 kHz voice call. For VoIP, the amount of network bandwidth has to be determined to suit either a 'best-effort' or 'quality-of-service' capability.

The IP packets would then be sent via a 'media stream'. The voice signals would be digitally sampled by the calling equipment (e.g; IP Phone, PC soundcard) and encoded into a format using a codec before transmitting over the IP network and received using the opposite process.

2.2 VOIP PROTOCOL STACK

Voice over IP is sent through the network via a set of protocols in the TCP/IP five-layer model. The most important layers concerning VoIP are the top three; Application, Transport, and Internetwork.

In the application layer, there are three protocols involved when sending VoIP packets. The first is called the Network Time Protocol (NTP). This protocol enables timing, which helps ensure that the signals are transmitted and received within the proper time frame to assure quality. The next protocol is called the Real-time Transport Protocol (RTP). The purpose is to ensure that the data streams can be reconstructed accurately. Jitter can occur, where the individual packets of the data stream experience variation in delay times. To overcome this, data is buffered at the receiving end of the link so that it can be played out at a constant rate. The third protocol, RTP Control Protocol (RTCP), provides feedback on the quality of the transmission link while RTP transports the digitized samples of real time information. However, these protocols do not reduce overall delay time or have any control over Quality-of-Service (QoS).

Next, it is sent to the transport layer where the User Datagram Protocol (UDP) protocol is used. Transmission Control Protocol (TCP) is much slower than UDP, but it provides guaranteed delivery of data packets. Because voice is a real-time application, it is more important that the voice packets are sent as quickly as possible, hence UDP is utilized for VoIP. A connectionless protocol, UDP routes the datagrams to its correct destination port but does not perform any sequencing or ensuring data reliability.

In the Internetworking layer, the Internet Protocol (IP) is used. At this layer, IP addressing is added to the packets. Every VoIP phone or computer acting as a VoIP phone gets a unique IP address that routes delivery of VoIP packets to and from the caller and receiver during the life of the call. Like UDP, it is also a connectionless protocol.

There are also other protocols involved in the transmission of voice packets over an IP network. To manage QoS, the Resource Reservation Protocol (RSVP) is utilized. It is a signaling protocol that requests a certain amount of bandwidth and latency in every network hop that supports it.

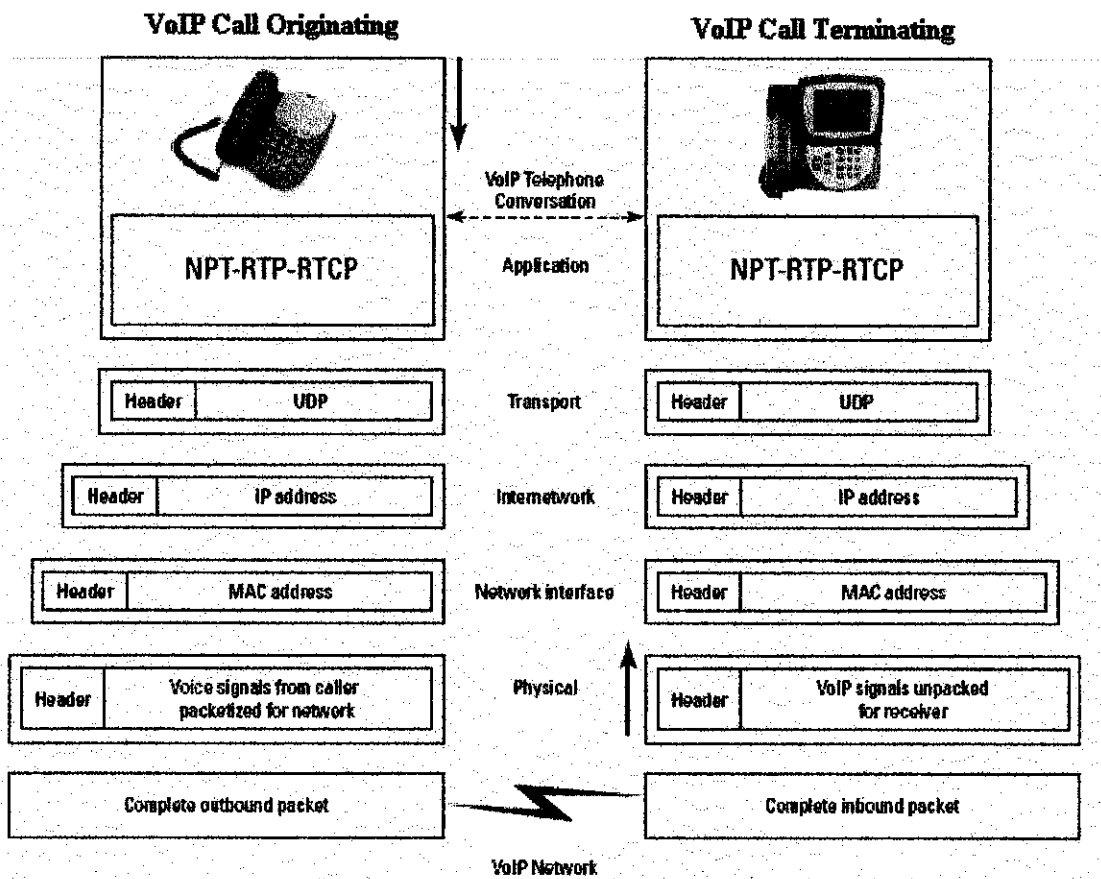


Figure 2.2: VoIP protocols in the TCP/IP protocol stack

2.3 VOIP SECURITY OVERVIEW

Although digitized voice travels in packets just like any other data, existing network architectures and tools cannot simply be used without implementing changes. With the introduction of VoIP in data networks, the need for security is compounded because two invaluable assets must now be protected which is data and voice. As an example, when placing an order for merchandise over the phone, callers will read their credit card number to the receiver on the other end. Without a secure line, the transmitted data is not protected and would be exposed to information theft and subsequently credit card fraud. This would have a significant impact on transactions over the IP networks and has to be addressed before it goes out of hand. The current IP network architecture does not provide the same physical wire security as traditional analog telephony which is not as easily intruded and attacked.

2.4 VOIP COMPONENTS

There are several components of VoIP which has to be addressed in order to form a secure network. This includes:

- ***End-user equipments***

This is the point of access for VoIP users to establish and receive calls. Connections to the network may be in the form of physical cables or wireless. It is usually not protected with firewall devices and is therefore exposed to vulnerabilities. Softphones are more defenseless compared to IP phones as it runs on PC Operating Systems which have a higher risk of threats.

- ***Network components***

This includes the cabling, routers, switches, and firewalls in the IP network. Securing the components give the added weight of balancing the performance of VoIP while transmitting voice safely and securely. The Quality-of-Service (QoS), latency, jitter, and packet loss have to be managed in accordance to network security.

- ***Call processors***

Functions of the call processors include phone number to IP translation, call setup, call monitoring, user authorization, and signal coordination. Call processors are usually software that runs on an OS. It is subject to the vulnerabilities of the OS, the call processor application, as well as other applications running on the OS.

- ***Gateways***

This can be further classified as Signaling Gateways (SG), Media Gateways (MG), and Media Gateway Controllers (MGC). In general, they handle call origination and detection as well as analog to digital conversion. SGs manage the signal traffic between an IP network and a circuit switched network, while MGs manage media signals between them. MGCs manage traffic between the SGs and MGs. Vulnerabilities can exist between the internal IP network and the circuit switched network and IP attacks on the gateway itself.

- ***Protocols***

The two most common VoIP protocol used; H.323 and Session Initiation Protocol (SIP), both suffer security complications of their own. Both use random ports causing the need of firewalls to be setup to allow only H.323 and SIP traffic to flow through them. Network Address Translation (NAT) is a problem for both protocols because the IP and port on the IP header do not match those in the messages.

2.5 QOS ISSUES IN VOIP SECURITY

QoS is fundamental to the operation of a VoIP network. The implementation of various security measures can degrade QoS. These complications range from delaying or blocking of call setups by firewalls to encryption-produced latency and delay variation also known as jitter. As VoIP is a time-critical application, most of the security measures used in data networks is not applicable to VoIP. Latency, jitter, packet loss, and bandwidth congestion are the areas affected by security in VoIP.

2.5.1 Latency

Latency in VoIP refers to the time it takes for a voice transmission to go from its source to its destination. This has to be ideally as low as possible to ensure high voice quality. There is however, a practical low boundary limit for voice applications. A delay of 150 ms and below is not noticeable to the human ear. For the range of 150 – 250 ms, it is considered acceptable although it is detectable. PSTN networks latency are in this range. Delays above 400 ms are considered unacceptable for even general network purposes. Adding security measures need to be checked to ensure the latency limits are not exceeded.

2.5.2 Jitter

Jitter refers to non-uniform packet delays. It is often caused by low bandwidth conditions in VoIP and can be exceptionally detrimental to the overall QoS. Jitter can cause packets to arrive and be processed out of sequence. Due to the fact that RTP, on the transport layer, is based on UDP, packets out of order are not reassembled at the protocol level. Jitter however, can also be controlled throughout the VoIP network by using routers, firewalls, and other network elements that support QoS. It has to be ensured that delays in packet deliveries caused by security devices are kept uniform throughout the traffic stream to avoid jitter.

2.5.3 Packet Loss

Packet loss in VoIP would cause the voice conversation to be missing and silent during the fault. Packet loss can result from excessive latency, where a group of packets arrive late and must be discarded in favour of newer ones. It can also be the result of jitter, that is, when a packet arrives after its surrounding packets have been discarded from the buffer, making the received packet useless. There are also cases where a packet is not delivered at all in VoIP due to its dependence on UDP, which does not guarantee packet delivery during transport.

2.5.4 Bandwidth Congestion

As in data networks, bandwidth congestion can cause packet loss and a multitude of other QoS problems. Thus, proper bandwidth reservation and allocation is essential to VoIP quality. Congestion of the network causes packets to be queued, which in turn contributes to the latency of the VoIP system. Low bandwidth can also contribute to jitter, since packets will be delivered in spurts when an opening is detected in the traffic. The effective bandwidth, which is defined as the percentage of bandwidth carrying actual data with regard to the total bandwidth used is equally important in managing QoS. Introducing security mechanisms, such as encryption, reduces the effective bandwidth creating a decreased throughput and increased latency in the system.

2.6 SECURE REAL-TIME TRANSPORT PROTOCOL (SRTP)

SRTP is a security profile for RTP that adds confidentiality, message authentication, and replay protection to the protocol. It is a proposed standard in the IETF Audio-Video Transport Working Group, and is defined in RFC 3711. SRTP is ideal for protecting VoIP traffic because it can be used in conjunction with header compression and has no effect on IP QoS. These facts provide significant advantages, especially for voice traffic using low-bitrate voice codecs such as G.729 and iLBC (internet Low Bitrate Codec). The following diagram shows the SRTP header.

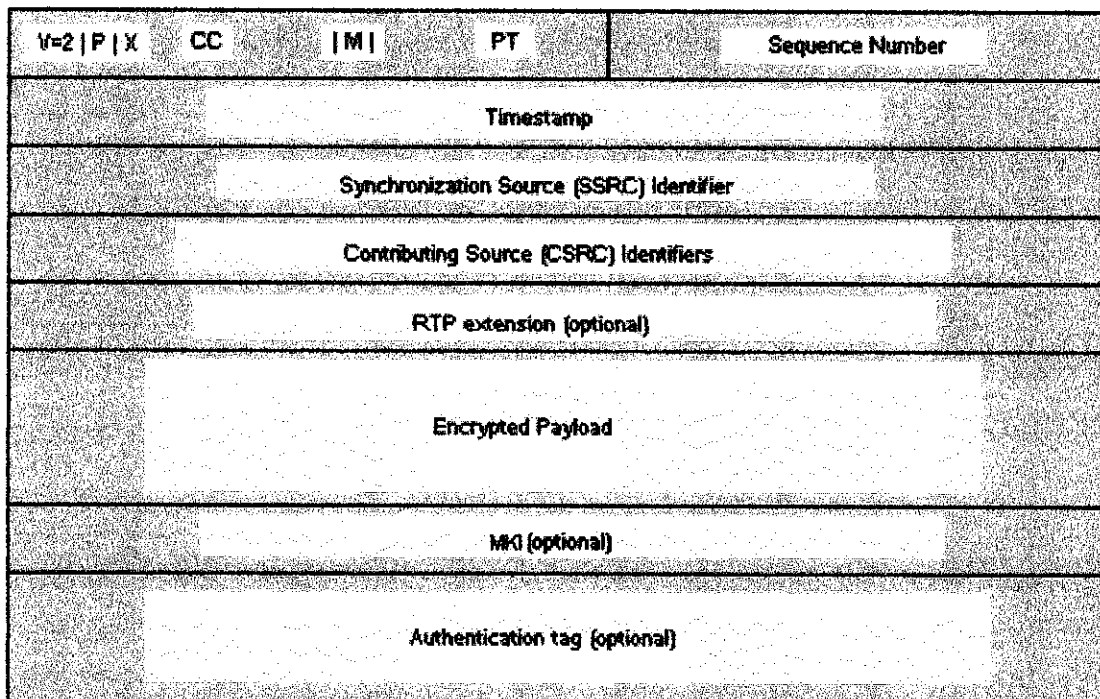


Figure 2.3: Secure Real-time Transport Protocol (SRTP) header

Although SRTP is used, all provided features, such as encryption and authentication, are optional and can be separately enabled or disabled. SRTP works by encrypting the voice conversation, rendering it meaningless to internal or external eavesdroppers who could have gained access to the voice domain. It also supports the Advanced Encryption Standard (AES) encryption algorithm.

Due to the fact that it is designed specifically for streaming real time data, media encryption using SRTP is more bandwidth efficient than IPsec. SRTP encrypts only the payload of a voice packet without adding additional encryption headers. Because of this, an SRTP-encrypted voice packet is almost indistinguishable from an RTP voice packet, allowing features like QoS and compressed-RTP to be supported without any additional development or packet manipulation.

SRTP media encryption is performed on Digital Signal Processing (DSP) modules available on Analog Voice Network Modules and not on the Central Processing Unit (CPU) for hardware phones. This enables efficient scalability as increasing the number of voice gateway interfaces with DSPs, or increasing the number of DSPs integrated on the routers will increase the number of DSPs available for secure calls. No additional call setup delays are introduced with encrypted calls, as the key exchange is completed as part of the normal call setup, and no extra messages are introduced.

Codecs	Regular Voice Call/DSP	Encrypted Voice Call/DSP
G.711	16 calls	10 calls
G.729a	8 calls	8 calls
G.729	6 calls	6 calls

Table 2.1: Encrypted calls channel impact per DSP using various codecs on Cisco analog voice modules

From table 1 shown previously, it is observed that there is minimal impact when using the G.711 codec and no channel capacity impact for encrypted calls utilizing the G.729 and G.729a codecs. This shows the advantage of performing SRTP media encryption in the DSP instead of the router CPU. While implementing SRTP, its sister protocol, Secure Real-time Transport Protocol (SRTCP) is always used in conjunction. It provides the same security-related features as the ones in SRTP including master key identifier, encryption, and authentication.

CHAPTER 3

PROJECT WORK

3.1 METHODOLOGY

A few methods are utilized to achieve the objectives of the project. This includes all the stages of development. They are as follows:

- ***Preliminary research and literature review***
This is the first stages of research and involves reviewing reference books, online tutorials, journals, and other materials which encompass the basics of VoIP.
- ***Problem analysis and data gathering***
The initial problem statement is analyzed and investigated. Data and information is gathered as much as possible.
- ***Tools and equipment identification***
The various tools and equipment required for the project is identified. This is elaborated in the next part of this report.
- ***Testing and experimentation***
The next stage of the project is the testing and experimentation with the tools obtained. This includes configuring the equipment and familiarizing with it.
- ***Presentation and demonstration***
The final part of the project in semester one is the presentation and demonstration of VoIP to examiners.

3.2 HARDWARE

A working demonstration is available for testing and research purposes at the Data Communications lab of the IT/IS department in Academic Building 02. The hardware consists of:

- Analog Telephones (2 units)
- Desktop PCs (Compaq Evo D510 – Multiple units)
- Switches (Cisco 3550 Catalyst – 3 units)
- Routers (Cisco 3725 – 3 units)
- Wireless Access Point (1 unit)

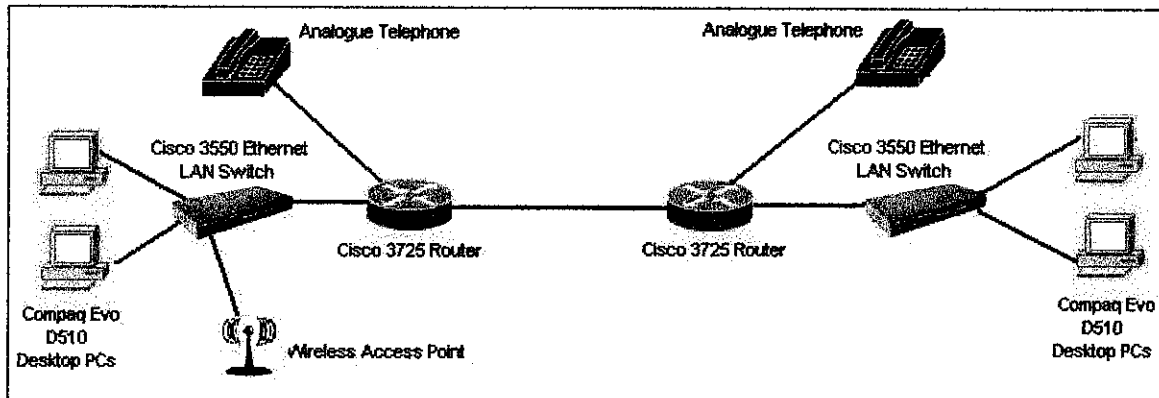


Figure 3.1: VoIP Hardware Connection in Data Communications Lab

The above network supports the H.323 and SIP protocol and utilizes IPv4. As this is an isolated network, the configurations of the router are very basic with the absence of firewalls and Network Address Translations (NATs). The software used to dial from the PC is Microsoft NetMeeting, which comes packaged with Microsoft Windows XP operating systems and SJ Phone, which is downloadable for free. The analog telephones are plugged in straight into the Foreign Exchange Station (FXS) port of the voice network module in the Cisco 3725 router.

Compaq Evo desktop PCs and the Wireless Access Point are connected to the Cisco 3550 Catalyst switches which are then connected to the router. Both routers are interconnected together. The Cisco 3725 Multiservice Access Router's configuration is setup according to the Cisco manual.

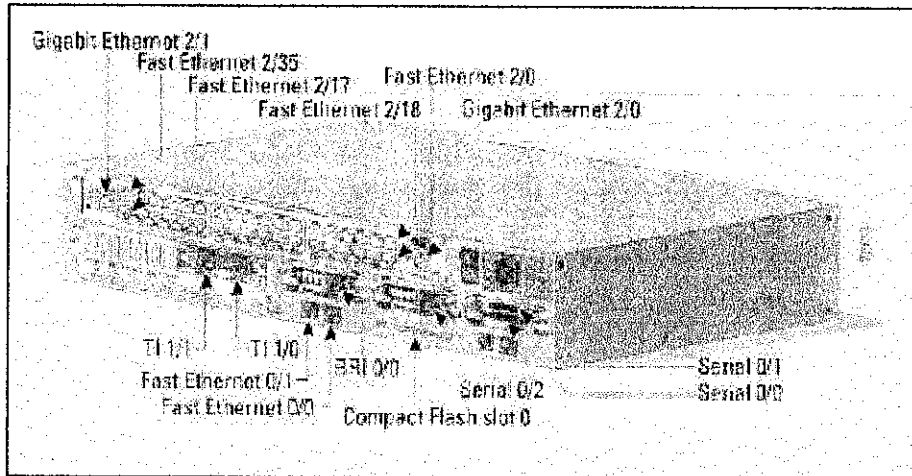


Figure 3.2: Cisco 3725 Multiservice Access Router ports

3.3 ROUTER CONFIGURATION

3.3.1 Configuring Dial-Peers

In VoIP, a *dial-peer* is referred to as an addressable call endpoint. A *dial-peer* configuration has to be defined for each call leg. A call leg is a logical connection between the router and either a telephony endpoint over a bearer channel or another endpoint using a session protocol. There are two different types of dial-peers, namely:

- POTS – A dial-peer describing the characteristics of a traditional telephony network connection. POTS peers point to a particular voice port on a voice network device.
- VoIP – A dial-peer describing the characteristics of a packet network connection. In the case of Voice over IP, this is an IP network. VoIP peers point to specific VoIP devices.

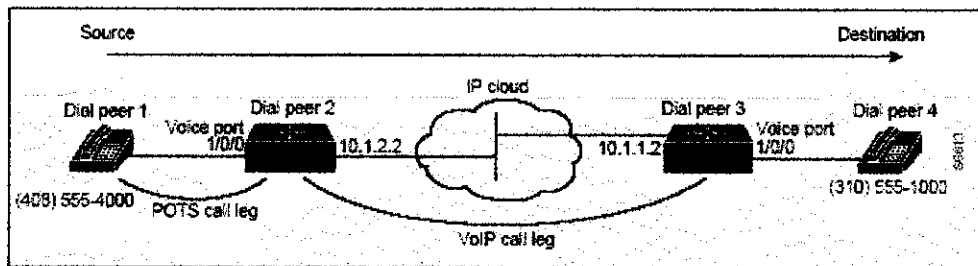


Figure 3.3: Example of basic VoIP connection showing respective dial-peers & call legs

Based on the figure above, the following commands are to be input to router 10.1.2.2 to configure the respective dial-peers:

```
dial-peer voice 1 pots                ! defines the type and number of dial-peer
destination-pattern 1408555....      ! defines the telephone number of the port
port 1/0/0                            ! specifies the port number

dial-peer voice 2 voip                ! defines the type and number of dial-peer
destination-pattern 1310555....      ! defines the telephone number of the port
session target ipv4:10.1.1.2         ! specifies the destination IP address
```

For router 10.1.1.2, the following commands are to be input to configure the respective dial peers:

```
dial-peer voice 4 pots                ! defines the type and number of dial-peer
destination-pattern 1310555....      ! defines the telephone number of the port
port 1/0/0                            ! specifies the port number

dial-peer voice 3 voip                ! defines the type and number of dial-peer
destination-pattern 1408555....      ! defines the telephone number of the port
session target ipv4:10.1.2.2         ! specifies the destination IP address
```

The *voice* command indicates that the dial peers are to be used for voice services. The subsequent number indicates a tag that uniquely identifies the dial peer. At the end of the first line, the type of dial-peer is defined either as *pots* or *voip*. The *destination-pattern* defines the destination telephone number associated with the VoIP dial peer. If the dial-peer is in *pots* mode, the number indicates the telephone number of the telephone connected to the voice port. If it is in *voip* mode, the number points to the telephone number of the router it connects to.

The *port* command associates the POTS dial peer with a specific voice port. The numbers indicate slot-number/subunit-number/port for the respective voice module. A *session target* command specifies the destination IP address. In this case, it is using IPv4 as the IP protocol version.

3.3.2 Configuring CODEC

CODEC is the abbreviation for Coder-Decoder. It determines how much bandwidth the voice session uses. A CODEC is used to transform analog signals into digital bit streams and vice versa. In the case of VoIP, it specifies the voice coder rate of speech for a dial-peer. Some CODECs require more processing power than others and each have different complexities, bit rate, Mean Opinion Score (MOS), and compression delay. Generally, there are four types of CODECs:

- G.711 - Describes the 64-kbps PCM voice coding technique.
- G.729 - Describes CELP compression where voice is coded into 8-kbps streams. There are two variations of this standard (G.729 and G.729 Annex A).
- G.723 - Describes a compression technique that can be used for compressing speech or audio signal components at very low bit rate as part of the H.324 family of standards. This CODEC has two bit rates associated with it: 5.3 kbps and 6.3 kbps.
- G.726 - Describes ADPCM coding at 40, 32, 24, and 16 kbps.

By default, Cisco VoIP uses the G.729 CODEC which is adequate for most networks. If voice quality is of high importance and high bandwidth is available, G.711 is usually utilized. The following shows the basic command for configuring the G.711 CODEC on router 10.1.2.2:

```
dial-peer voice 2 voip           ! defines the type and number of dial-peer
destination-pattern 1408555.... ! defines the telephone number of the port
codec g711alaw                 ! specifies the codec type
session target ipv4:10.1.1.22  ! specifies the destination IP address
```

Compression Method	Bit Rate (kbps)	MOS Score	Compression Delay (ms)
G.711 PCM	64	4.1	0.75
G.726 ADPCM	32	3.85	1
G.728 LD-CELP	16	3.61	3 to 5
G.729 CS-ACELP	8	3.92	10
G.729 x 2 Encodings	8	3.27	10
G.729 x 3 Encodings	8	2.68	10
G.729a CS-ACELP	8	3.7	10
G.723.1 MP-MLQ	6.3	3.9	30
G.723.1 ACELP	5.3	3.65	30

Table 3.1: Comparison of various CODEC characteristics

3.3.3 Configuring IP Precedence:

Under normal network conditions, both data and voice are given equal priority in network traffic and bandwidth consideration. To give real-time voice traffic a higher priority than others, there are basically two methods employed in the configuration of VoIP dial peers. The first is called IP Precedence. Shown below is the basic command for configuring IP Precedence on router 10.1.2.2:

```
dial-peer voice 2 voip           ! defines the type and number of dial-peer
ip precedence 5                 ! specifies the IP precedence class
```

The number at the end of *ip precedence* command identifies the classes for IP flow. A scale of 1 to 5 is utilized with 5 being the highest. If the networks receiving these packets have been configured to recognize precedence bits, the packets will be given priority over packets with a lower configured precedence value. A value of 6 & 7 is also applicable but is used for network routing and updates.

3.3.4 Configuring Resource Reservation Protocol (RSVP)

Another method for optimizing VoIP is by utilizing the Resource Reservation Protocol or RSVP. The RSVP protocol is used to facilitate Quality of Service (QoS) in VoIP. The example below shows configuration for *controlled load* quality of service on router 10.1.2.2:

```
dial-peer voice 2 voip           ! defines the type and number of dial-peer
destination-pattern 5551234     ! defines the telephone number of the port
req-qos controlled-load        ! specifies the type of QoS
session target ipv4:10.2.2.1    ! specifies the destination IP address
```

Every time a connection is made through a specific VoIP dial peer with RSVP configured, an RSVP reservation request is made between the local router, all intermediate routers in the path, and the final destination router. RSVP works only one-way. The VoIP dial peer on both ends of the connection has to be configured for RSVP to work. There are three types of QoS configurable for VoIP:

- Best Effort – RSVP makes no bandwidth reservation
- Controlled Load - RSVP guarantees a single level of preferential service, presumed to correlate to a delay boundary. Admission or capacity control is used to ensure that preferential service is received even when the bandwidth is overloaded.
- Guaranteed Delay - RSVP reserves bandwidth and guarantees a minimum bit rate and preferential queuing if the bandwidth reserved is not exceeded.

3.4 SOFTWARE

Several softwares are utilized to test and analyze SRTP using the SIP protocol. All the softwares used are open-source or distributed under free licensing for educational and research purposes. These include:

- SIP User Agents (MiniSIP, Snom 360)
- SIP Proxy & Registrar Server (CommuniGate Pro Server)
- Network Protocol Analyzer (OpenXtra Ethereal)
- Packet Sniffer (Cain & Abel)

3.4.1 SIP User Agent (UA)

An SIP User Agent or UA is an SIP based softphone that can be used for VoIP calls. It is installed in a desktop PC with a soundcard and requires a headset or a speaker and microphone combo. Currently, there are two SIP UAs that support SRTP and runs natively on Windows XP.

- MiniSIP User Agent

MiniSIP is distributed under the GNU General Public License (GPL). It runs on multiple platforms including Linux, Linux on Embedded Platforms (Familiar, Maemo), Windows XP/2000, and Windows Mobile 2003 SE. A binary version exists for Windows XP/2000 although still in its development phase. The user interface includes a graphical user interface (GUI) and a command prompt window for monitoring and debugging purposes. The developers of MiniSIP confirm that the UA runs stable on Linux and has some minor bugs on Windows. Frequent program crashes have occurred during testing and configuration but the UA is generally moderately stable for Windows. The following diagram shows the GUI of the MiniSIP UA.

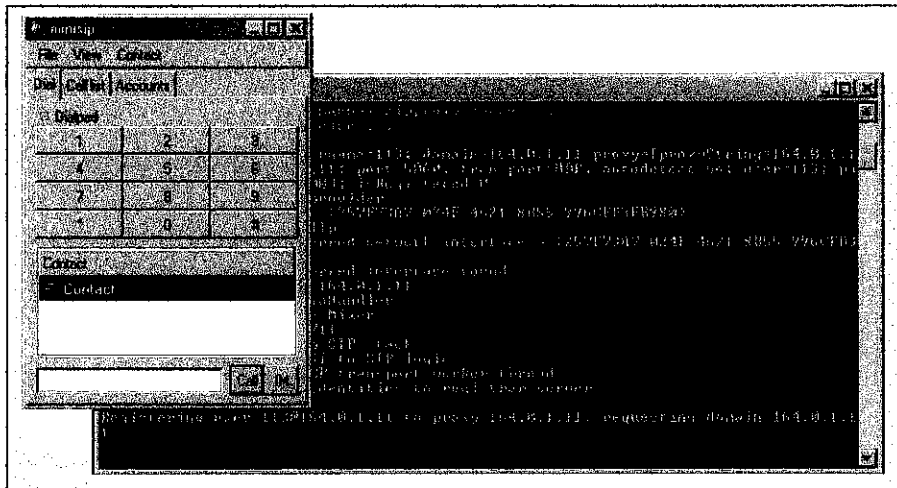


Figure 3.4: MiniSIP User Agent (UA) Graphical User Interface (GUI)

There are two external libraries required to run the MiniSIP softphone which are:

- **OpenSSL (version 0.9.7 or higher)** – A toolkit implementing the Secure Socket Layer (SSL) and Transport Layer Security (TLS) protocols as well as a general purpose cryptography library.
- **Glademmm:**
 - **GTK** – A multi-platform toolkit for creating graphical user interfaces (GUIs) originally for the X Window System.
 - **gtkmm** – The official C++ interface for the GTK GUI library

The features of MiniSIP include the ability to configure a key exchange for applying Secure Real-time Transport Protocol (SRTP). There are two selectable key exchanges which are the shared key and a Diffie-Hellman (D-H) key. Of the two, the shared key is less complicated to implement as the D-H key requires certificate setting and configuring which are more complex.

The account settings menu provides configuration options for user registration as well as authentication in order to establish connection. The fields include account name, SIP Uniform Resource Identifier (URI), SIP proxy and port, transport method, registration username & password, and registration time. The security settings menu provides options for enabling secured calls, key exchange types, certificate settings, as well as a pre-shared key option. The following are standard account settings for UAs used in testing:

- Account: 110
- SIP Uniform Resource Identifier (URI): 110@10.20.20.15
- SIP Proxy: 10.20.20.15
- SIP Registrar: 10.20.20.15
- Network Port: 5060
- Transport Method: UDP

Although not all UAs have the same configuration menu, the basic settings are similar. The following diagrams show the account settings and security settings menu for user account 110 with SRTP enabled using shared key exchange on MiniSIP:

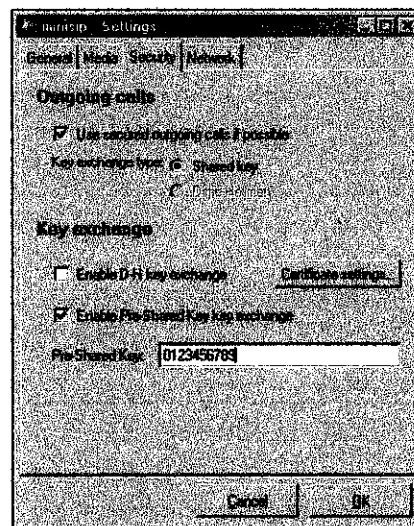
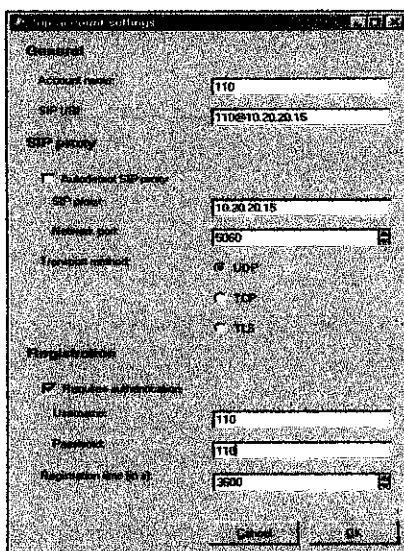


Figure 3.5: MiniSIP account settings menu Figure 3.6: MiniSIP security settings menu

- Snom 360 User Agent

The Snom 360 User Agent is a softphone version of the Snom 360 IP phone. It has most of the features available in the IP phone and provides free licensing for private and educational use. The Snom 360 works on Windows operating platforms and utilizes the Session Initiation Protocol (SIP) signalling protocol. Although it contains largely the functions of a real IP phone, only codecs G.711u, G.711a, and GSM are supported. Some of the features include:

- Web-Interface Style Configuration Menu
- 47 keys, 13 Simulated LEDs
- 12 Programmable Function Keys
- Multiple Language Support
- Secure RTP Encryption
- Advanced Call Control Features

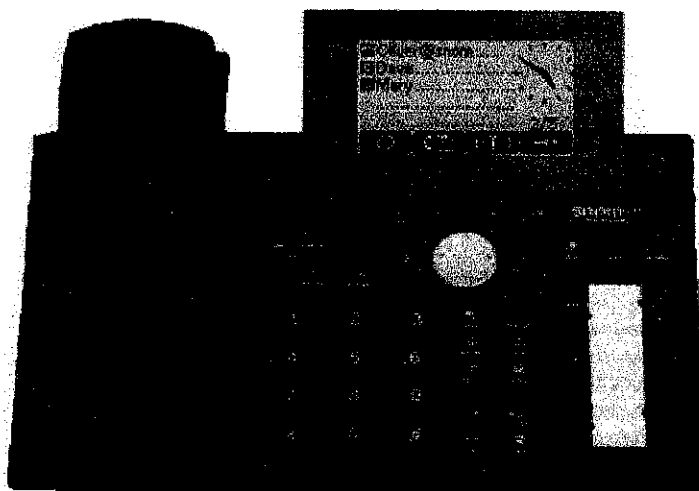


Figure 3.7: Snom 360 User Agent (UA) Graphical User Interface (GUI)

The Snom 360 UA utilizes ‘sdescriptions’ as the key-exchange protocol. It describes a way of establishing security parameters for SRTP with and SDP attribute a=crypto. This attribute is not actually a key management protocol but only conveys a set of parameters for SRTP. It is used to signal and negotiate cryptographic parameters for media streams in general, and SRTP in particular.

The configuration menu of Snom 360 UA is accessible via a web interface. The account settings menu is pretty similar to that of the MiniSIP UA. Necessary fields include line activation, account, password, SIP registrar, SIP outbound proxy, and authentication username. The security settings are configured in the RTP line settings menu and SRTP has the selectable option of ON or OFF. The following diagrams show the login information for user *111* and RTP options:

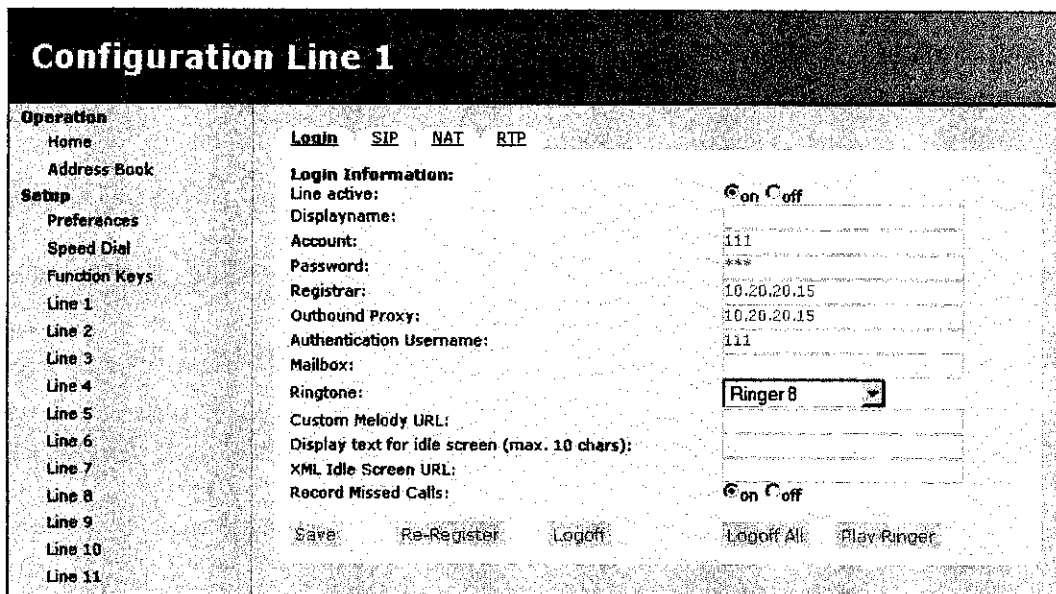


Figure 3.8: Snom 360 web interface menu – Login Information

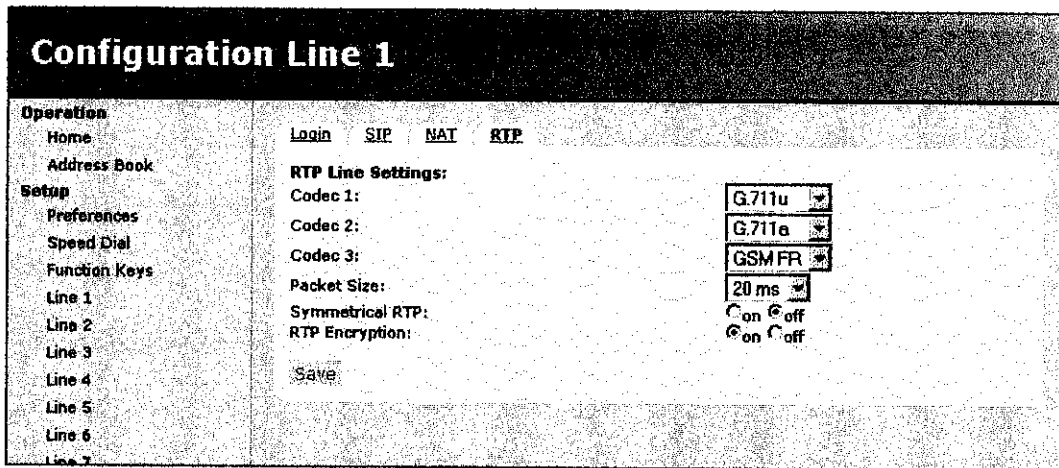


Figure 3.9: Snom 360 web interface menu – RTP Line Settings

3.4.2 SIP Proxy and Registrar Server

SIP calls require a register server to register and authenticate user and a proxy server to connect users and their UAs. Previously, Brekeke OnDO SIP Server was utilized to act as the SIP proxy and registrar server. The server managed to register users but no calls could be established between UAs. It was later found that the server is unable to support SRTP media encryption. Thus, a different SIP server has to be used in order to implement and utilize SRTP for testing and analysis purposes.

- CommuniGate Pro Server

The CommuniGate Pro Server is an Internet communications application server. It provides an integrated platform for "store-and-forward" (E-mail, Calendaring) and Real-Time (VoIP, Video, Instant Messaging, White Boards) communications. This includes the SIP registrar and proxy which registers users & UAs and establishes call connectivity. Most importantly, it provides support for SRTP as well as other security mechanisms like TLS and S/MIME.

Upon installation, the CommuniGate Pro Server exists as a service in Windows XP and could be manually started and stopped from the Control Panel (Administration Tools > Services > CommuniGate Pro Messaging Server). The configuration menu is accessible from a web interface by entering <http://127.0.0.1:8010> at a web browser in the PC installed with the server. User accounts have to be created and defined first before it can be registered. User authentication is done via the SIP server to ensure that communicating users are who they claim to be. The server also includes a log and an SIP trace to monitor the registration and real-time call process. The following diagram shows the accounts registered to the CommuniGate Pro SIP server.

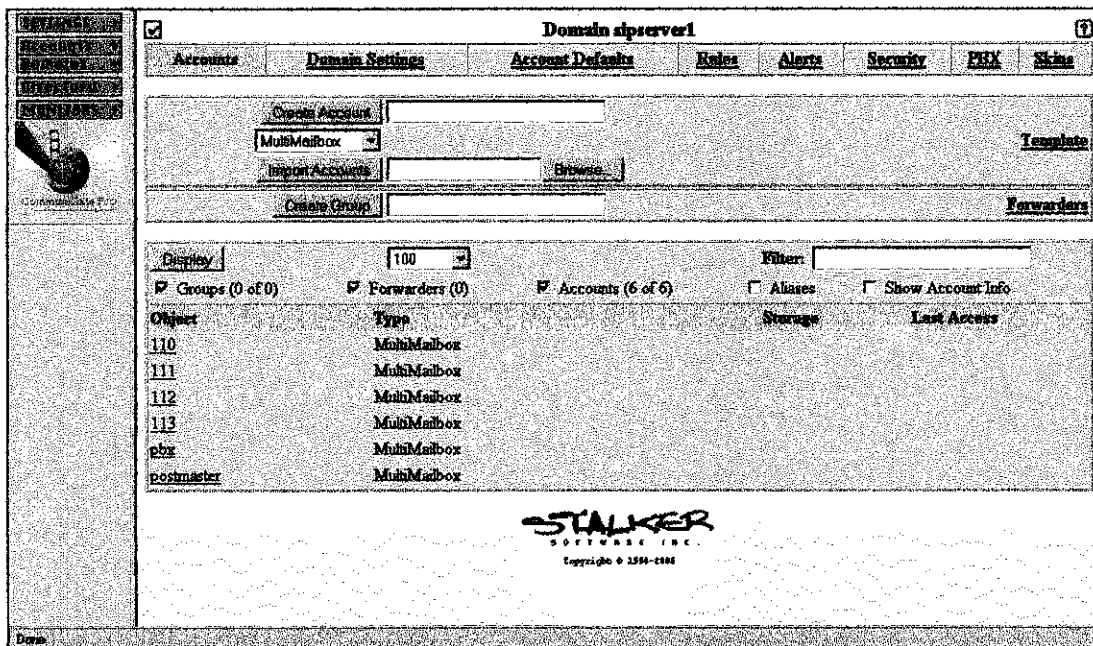


Figure 3.10: CommuniGate Pro Server web interface menu – User Accounts

3.4.3 Network Protocol Analyzer – Ethereal

Ethereal is a protocol analyzer used for network troubleshooting and analysis. A network packet analyzer works by capturing network packets data and displays them in detail. It is able to display encapsulation and single fields and also interpret their meanings. Although Ethereal is still technically beta software, it has a comprehensive feature set and is suitable for production use. It supports multiple platforms including Windows, UNIX, and Linux. Ethereal is an open source software released under the GNU General Public License.

OpenXtra Ethereal

For this project, the network protocol analyzer used to capture VoIP traffic is the OpenXtra Ethereal tracer. It is used to capture VoIP traffic and analyze the data to study the impact of SRTP on VoIP performance. It supports capturing VoIP traffic protocols such as Real-time Transport Protocol (RTP), User Datagram Protocol (UDP), Internet Protocol (IP), and Resource Reservation Protocol (RSVP). The two most common VoIP protocol, H.323 and Session Initiation Protocol (SIP), can also be captured on Ethereal. The following are the steps taken for capturing network protocols:

1. Firstly, the Network Interface Card (NIC) used in the computer is selected.
(Capture > Options > Interface)
2. Next, the Capture Filter is set to display only the relevant protocol to be captured which is UDP.
(Capture > Options > Capture Filter > UDP Only)
3. Ticking the checkbox to 'Update list of packets in real time' and 'Automatic scrolling in live capture' allows viewing the trace in real time.
(Capture > Options > Display Options)
4. Clicking on the 'Start' button would then begin the capture process.
(Capture > Start)

5. After several minutes of capturing, the 'Stop' button is depressed to halt the tracing process.

(Capture > Stop)

6. To begin another live capture process, steps 4 and 5 are repeated.

At the end of the capture, the protocol is then analyzed by clicking on the packet to reveal encapsulated data. As the protocol displayed is UDP, the data is first decoded as RTP (*Analyze > Decode As... > RTP*). The RTP protocol is then displayed in the window. Statistics of the RTP stream is analyzed by displaying the captured data table and graph (*Statistics > RTP > Stream Analysis*). The following diagram shows an Ethereal capture with the stream analysis present.

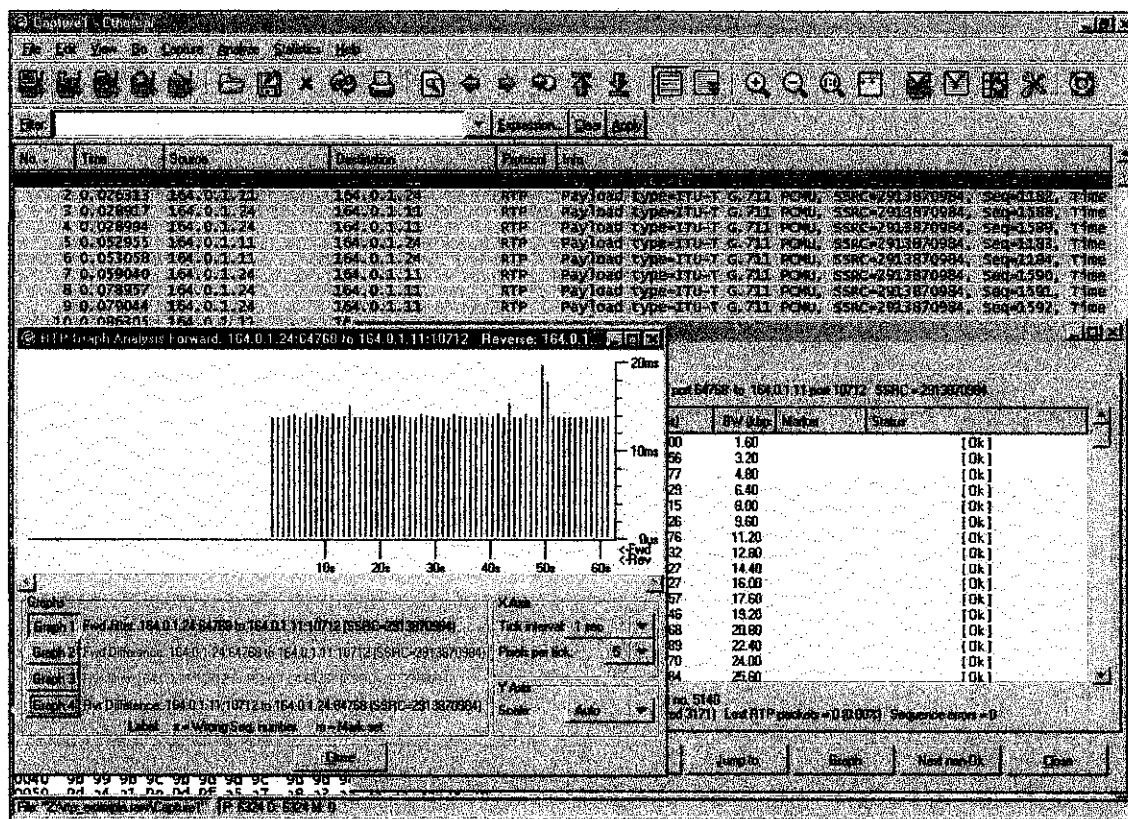


Figure 3.11: Ethereal capture for an SIP call with RTP stream analysis

3.4.4 Packet Sniffer – Cain & Abel v2.8.6

Packet sniffing is the process of listening to a network device for packets via software. Cain & Abel is initially developed as a password recovery tool for Microsoft Operating Systems. It also acts as a packet sniffer which enables capturing and recording VoIP conversations which are not secure over the network. This software is provided with free licensing for any usage. It is utilized to test the security of VoIP calls which are not encrypted at the application layer (RTP), in comparison to the encrypted VoIP protocol (SRTP).

The packet sniffer is installed at endpoints within the VoIP network. By simply selecting the appropriate network interface card, the VoIP conversation can be recorded and saved as a wave file in the PC for playback. This shows the vulnerability of a VoIP conversation without media encryption for security. Hence, applying SRTP encryption is crucial in ensuring the integrity and privacy of VoIP without sacrificing its call quality. The following diagram shows the Cain & Abel VoIP Sniffer window.

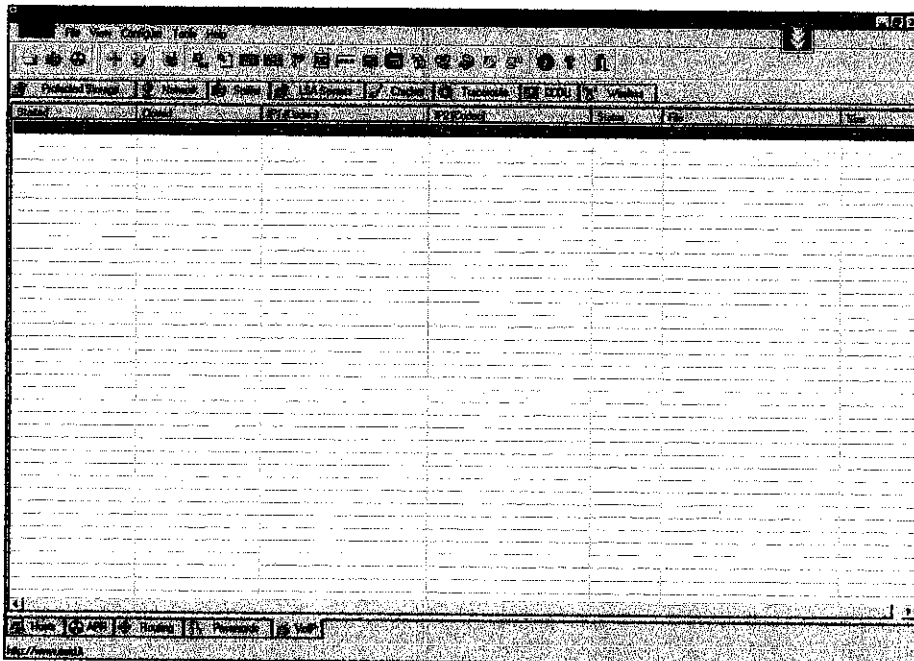


Figure 3.12: Cain & Abel VoIP Sniffer window

CHAPTER 4

RESULTS AND DISCUSSION

4.1 VOIP TEST CALL RESULTS

After all the equipment was setup according to the earlier mentioned configuration, the VoIP was put to test in the Data Communications Lab in Building 02. Mr. Ruslan Idris, the lab technician, was available on site to provide help and assistance. There are four types of calls tested in the lab which are:

- ***Phone to Phone***

Two analog phones are connected to two different routers containing two different voice modules. Both routers are interconnected to link both phones, allowing access for placing VoIP calls. Calls are made by dialing the destination telephone number 555-1000 and 555-4000, respectively. The calls were a success and no significant delay or jitter was heard during conversation.

Overall Result: OK

- ***Phone to PC***

The PC is connected to an Ethernet LAN switch which is connected to the routers. Initially, Microsoft NetMeeting was used during testing as the PC software phone, but was replaced with SJ Phone in the later stages as it had more features such as switching between SIP and H.323 protocols. NetMeeting only supports the H.323 protocol. The software phone has to be opened in order to receive calls and the speaker turned on to listen for ringing tones. The destination telephone number, 555-1234 and 555-8888 respectively, are entered in the user interface of the software phone to initiate calls. Calls received by both softwares and using both protocols were a success and no significant delay or jitter was heard during conversation.

Overall Result: OK

▪ ***PC to PC***

Both PCs are connected to two separate Ethernet LAN switches which are then connected to two separate routers. Both routers are interconnected to allow VoIP calls to pass through. The destination telephone number, 555-1234 and 555-8888 respectively, are entered in the user interface of the software phone to initiate calls. A microphone and speaker set is attached to the soundcard of both PCs to allow voice conversations to be heard and made. Calls received by the software phones were a success and no significant delay or jitter was heard during conversation.

Overall Result: OK

▪ ***PC to Phone***

The fourth and final type of call tested was from PC to the analog phone. The same configuration was employed where the PC was connected to the Ethernet LAN switch and the analog phone connected to the voice port in the router. Both routers are interconnected to allow access to and from both points. Calls are made by dialing the destination telephone number 555-1000 and 555-4000, respectively. The calls were a success and no significant delay or jitter was heard during conversation.

Overall Result: OK

4.2 H.323 PROTOCOL DISCUSSION

The VoIP demonstration at the lab was conducted using the H.323 protocol. Generally, H.323 implementation includes four logical entities or components: terminal, gateways, gatekeepers, and multipoint control units, of which the last three are optional. H.323 is an umbrella specification for the many different protocols that comprise the overall protocol stack. The protocols in the H.323 protocol suite include functions for:

- Call control & signaling (H.225.0, H.225.0/RAS, H.245)
- Audio processing (G.711, G.722, G.723.1, G.728, G.729)
- Video processing (H.261, H.263)
- Data conferencing (T.120)
- Media transportation (RTP, RTCP)
- Security (H.235)
- Supplementary services (H.450.1 – H.450.9)

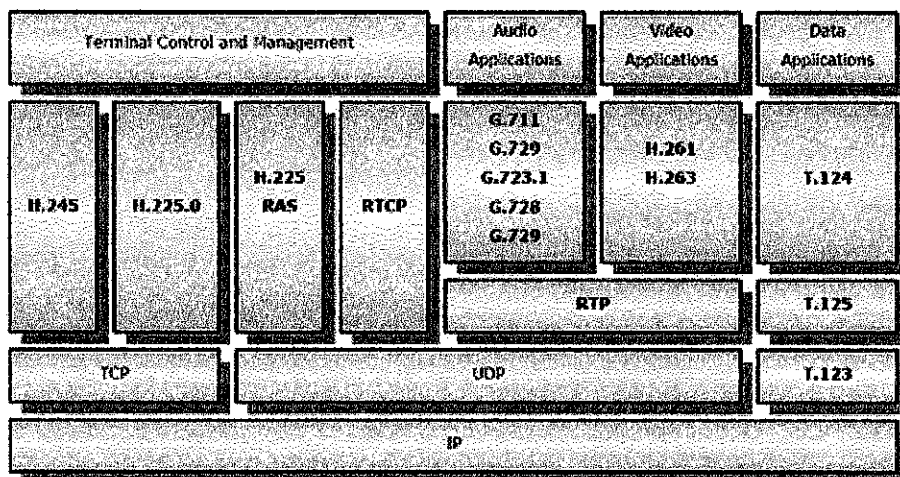


Figure 4.1: H.323 Protocol Stack

The steps below illustrate a basic call scenario between two H.323 terminals (Endpoint A and Endpoint B) without using a gatekeeper:

1. Endpoint A (the calling party) connects to endpoint B (the called party) and sends a **Setup** message (as specified in H.225.0) including the type of call (e.g., audio only), the called and the calling party number and address.
2. Endpoint B responds with an **Alerting** message. This message must be received by endpoint A before its setup time expires.
3. When the user at endpoint B picks up the call, a **Connect** message is sent to endpoint A.
4. Both terminals send their capabilities (media types, codec choices and multiplex information) in a **TerminalCapabilitySet** message.
5. Each terminal replies with a **TerminalCapabilitySetAck** message. In case the remote endpoint's capabilities are rejected, the message will be **TerminalCapabilitySetReject** and the terminals continue to send these messages until reaching a capability set supported by both endpoints.
6. Each terminal sends an H.245 **OpenLogicalChannel** message to open a logical channel at the remote endpoint, in order to set up the voice channels over which the media stream will be exchanged.
7. When ready to receive data, each terminal sends an **OpenLogicalChannelAck** to the remote endpoint, specifying the port number at which the remote endpoint should send the RTP data and the port number the RTCP data should be sent by the remote endpoint.
8. The endpoints exchange information in RTP packets. RTCP packets are sent during this exchange to monitor the quality of the data transmission.
9. When endpoint A hangs up, it must send an H.245 **CloseLogicalChannel** message for each channel opened with endpoint B.

10. Endpoint B replies with a **CloseLogicalChannelAck** message.
11. Endpoint A sends an H.245 **EndSessionCommand** and closes the channel after receiving the same message from endpoint B.
12. Both terminals send an H.225.0 **ReleaseComplete** message over the call signaling channel, which closes the channel and ends the call.

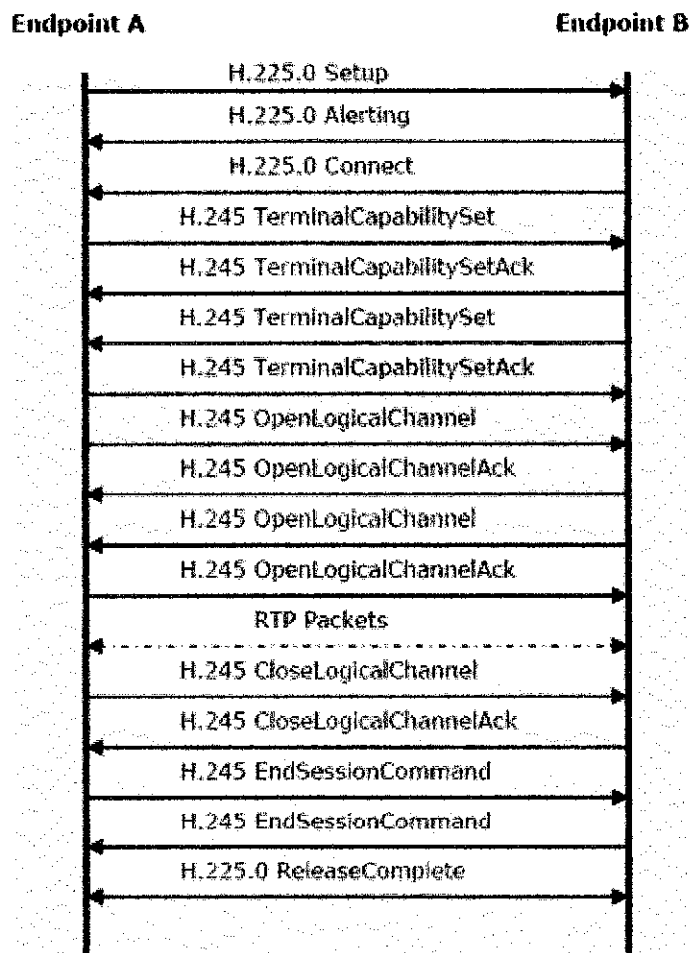


Figure 4.2: Basic H.323 call setup and teardown between two terminals

4.3 SIP PROTOCOL DISCUSSION

The SIP protocol was developed later than the H.323 protocol. It is an Internet Engineering Task Force (IETF) standard protocol for establishing, manipulating, and tearing down an interactive user session that involves multimedia elements such as video, instant messaging, and in this case real-time audio. SIP is a request-response protocol that works in the Application layer of the Open Systems Interconnection (OSI) communications model, and provides capabilities to:

- Determine the location of the target end point
- Determine the media capabilities of the target end point
- Determine the availability of the target end point
- Establish a session between the originating and target end point
- Handle the transfer and termination of calls

conference setup and discovery				conference course control		audio/ video	shared applications
SDP				RSVP	distribution control	RTP/ RTCP	reliable multicast
SAP	SIP	HTTP	SMTP				
UDP	TCP			UDP			
IP and IP multicast							
integrated services forwarding							

Figure 4.3: SIP protocol in the IETF multimedia conferencing architecture

There are two main components in the SIP architecture, namely:

- The SIP User Agent (UA) – This is the endpoint component which can be hardware or software utilizing SIP such as an IP-phone. It consists of two components; User Agent Client (UAC) and User Agent Server (UAS).
- The SIP Network Server – This component handles signaling associated with multiple calls providing name resolution and user location. Consists of three components; SIP Register Server, SIP Proxy Server, and SIP Redirect Server.

The following messages are utilized in a typical SIP call setup and tear down scenario:

- *INVITE* - Initiates the call session. The session description is included in the message body. Re-INVITE is used to change session state
- *ACK* - Confirms session establishment and can be used only with INVITE
- *BYE* - Terminates the call session
- *CANCEL* - Cancels a pending INVITE
- *OPTIONS* - Capabilities inquiry
- *REGISTER* - Binds a permanent address to current location and may convey user data

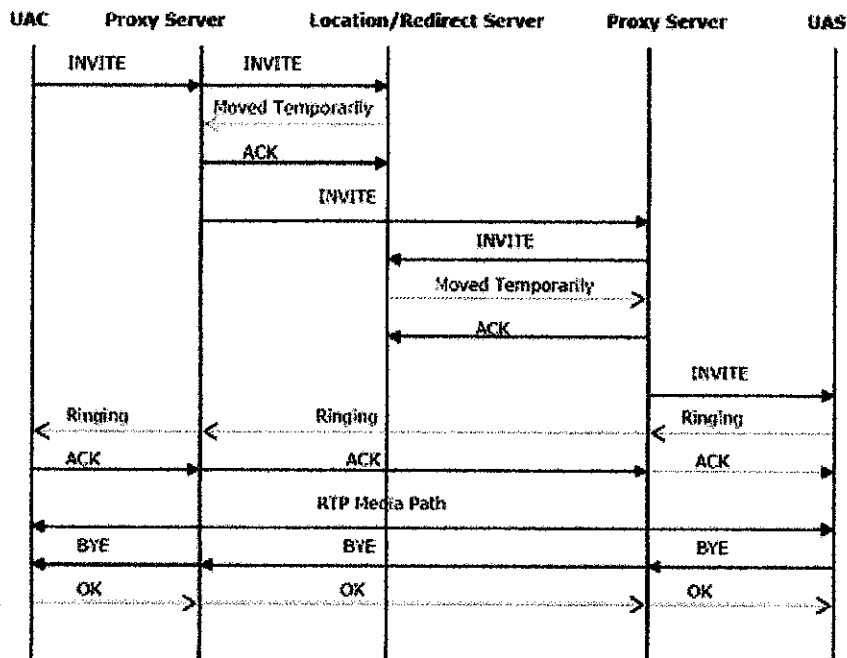


Figure 4.4: Basic SIP call setup and teardown between two endpoints

4.4 H.323 AND SIP COMPARISON

It can be concluded that H.323 and SIP are two of the best protocols to be utilized for VoIP. Although both protocols support VoIP, there are advantages and disadvantages between the two. There are also other protocols such as MGCP and H.248/MEGACO but are less accepted as they are mostly used for centralized architecture only. Distributed architecture is usually associated with SIP and H.323. They allow network intelligence to be distributed between endpoints and call-control devices. This creates a more flexible network and support more features.

From the architectural point of view, SIP is modular because it covers basic call signaling, user location, and registration. Other features are in other separate orthogonal protocols such as SDP for capabilities exchange, URLs for addressing, DNS for service location and TRIP for call routing. H.323 covers almost every service, such as capability exchange, conference control, basic signaling, QoS, registration, service discovery, and others. It specifies almost everything including the codec for the media and how packets are carried in the RTP.

In contrast to SIP, H.323 has a few disadvantages. It consumes more call setup time than SIP. H.323 requires about 12 packets for call setup while SIP requires only about 4 packets. It also necessitates TCP and UDP protocols during the call setup while SIP only requires UDP. In general, implementation of H.323 is much more complex and time-consuming compared to SIP. On the other hand, SIP has minimal capability exchange compared to H.323, although it is adequate for IP telephony. Since SIP is relatively new, most applications and hardware are currently utilizing the H.323 protocol.

	H.323	SIP	MGCP/H.248/Megaco
Standards body	ITU	IETF	MGCP/Megaco—IETF; H.248—ITU
Architecture	Distributed	Distributed	Centralized
Current version	H.323v4	RFC2543-bis07	MGCP 1.0, Megaco, H.248
Call control	Gatekeeper	Proxy/Redirect Server	Call agent/media gateway controller
Endpoints	Gateway, terminal	User agent	Media gateway
Signaling transport	Transmission Control Protocol (TCP) or User Datagram Protocol (UDP)	TCP or UDP	MGCP—UDP; Megaco/H.248—both
Multimedia capable	Yes	Yes	Yes
DTMF-relay transport	H.245 (signaling) or RFC 2833 (media)	RFC 2833 (media) or INFO (signaling)	Signaling or RFC 2833 (media)
Fax-relay transport	T.38	T.38	T.38
Supplemental services	Provided by endpoints or call control	Provided by endpoints or call control	Provided by call agent

Table 4.1: Comparison between H.323, SIP, and MGCP/H.248/Megaco protocols

Generally, H.323 and SIP as well as other protocols have their own advantages and disadvantages. Each addresses a different network requirement and need. Most equipment and applications today use H.323 as it was the first dominating protocol for VoIP although SIP is slowly paving its way to take over the role of H.323 as the lead protocol for SIP. Realistically, multiple VoIP protocols and architecture have been deployed and will co-exist in the future. Selection of the protocol or interconnected protocols is highly dependant on the usage of VoIP in the network. For the SRTP testing purposes, SIP is selected as the signaling protocol as most SRTP compliant softphones currently support SIP. In the IP telephony world, SIP is seen to be gaining ground over H.323 due to the fact that it was developed for the IP networks.

4.5 PSTN TO VOIP TRANSITION

Before VoIP can be implemented in a wide scale transition, such as in the UTP network, a few key issues have to be addressed. Besides the management's decision, other technical factors needed to be considered are:

- **Performance (QoS)**

Voice quality is very important in telephony. Delay, jitter, and packet loss can affect the overall performance of VoIP. This is typical for most network transmissions but for data, latency is not a big issue. For voice, a slight delay can cause losses in the conversation. For the human ear, a delay of 150 ms and below is not noticeable. For the range of 150 – 250 ms, it is considered acceptable although it is perceptible. This is valid for the traditional PSTN network. Delays above 400 ms are considered unacceptable for general network purposes. For efficient delivery of voice packets, the maximum jitter delay has to be 75 ms. Maximum packet loss tolerable is 3%.

- **Reliability**

The current circuit switched PSTN telephone system is considered to be very robust and reliable. Having a percentage of availability of 99.999%, it translates to about only 5 minutes of downtime per year. Data networks are less reliable due to complexity of protocols, vendors, operating systems, network systems, and others. The best private data networks are available 94% of the time on average. This translates to about 22 days in a year. For internet-based VoIP, only 61% reliability is recorded, meaning a downtime of 142 days annually.

▪ **Scalability**

Compared to PSTN equipment, networking equipment is designed to be readily scalable. Switches and routers can be connected to the existing network without much adjustment to the network. The advantage of interoperable standards such as H.323 and SIP means that upgrading and replacing in the network is easier. Multiple vendor equipments can exist without much hassle.

▪ **Service Life**

Typically, network equipments have a service life of 3 – 5 years. Servers have a service life of 4 years while switches and routers average on 5 years. However, there have been no studies on the service life of IP-PBXes and call managers. The approximate service life of a PBX in the PSTN system averages 7 – 10 years.

▪ **Electricity Consumption & Backup Power**

Both PSTN and VoIP require electrical power to operate. This varies according to the specification of the equipments used. The majority of networking equipment runs on AC power although there are some that utilizes DC power. Latest technologies enable power to be sent over Ethernet to power IP phones. This is done on unused wires of the cabling. Alternatively, IP phones can also be powered externally.

Analog phones can still operate even without electricity. However, IP phones require back up power from UPSes. Generally, it provides between 2 – 4 hours of backup power depending on the capacity of the UPS. Otherwise, generators can also be installed to power the whole building.

4.6 SRTP TESTING RESULTS

After all the equipment was setup according to the earlier mentioned configuration, SRTP in VoIP was put to test in the laboratory. Mr. Ruslan Idris, the lab technician, was available on site to provide help and assistance. Basically, two test scenarios were depicted to compare the security of VoIP between SRTP-disabled and SRTP-enabled conditions.

4.6.1 Test Scenario #1 (SRTP-Disabled)

The first test scenario depicts a standard SIP call without SRTP. After the accounts have been registered in the CommuniGate Pro Server, a call is initiated from user 110 (UA1) to user 111 (UA2) via their respective desktop PCs and UAs. OpenXtra Ethereal protocol analyzer is initiated to capture the RTP packet traces. Cain & Abel is started from a third PC to facilitate packet sniffing and recording the unsecured VoIP call.

Results:

PARAMETER	RESULT
Packet Sniffing	<i>Conversation audible</i>
Maximum Delay (latency)	<i>126.88 ms</i>
Maximum Jitter	<i>21.14 ms</i>
Packet Loss	<i>0 (0%)</i>
Maximum Bandwidth	<i>89.6 kbps</i>

Table 4.2: Parameters of SRTP-disabled VoIP call from user 110 to user 111

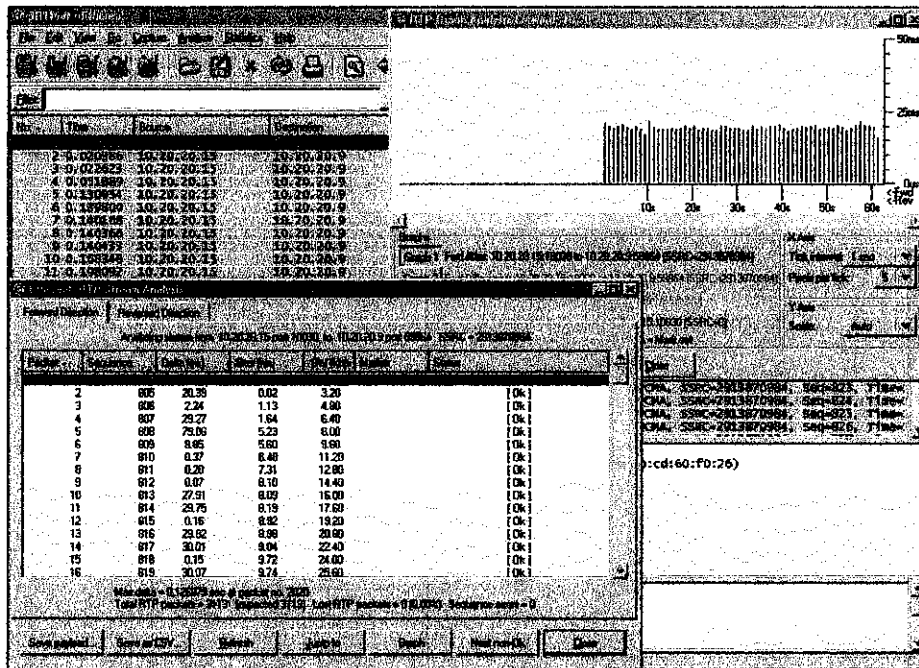


Figure 4.5: Ethereal RTP stream analysis (SRTP-disabled)

4.6.2 Test Scenario #2 (SRTP-Enabled)

The second test scenario depicts a standard SIP call with SRTP enabled. Similarly, a call is initiated from user 110 (UA1) to user 111 (UA2) via their respective desktop PCs and UAs. OpenXtra Ethereal protocol analyzer is initiated to capture the RTP packet traces. Cain & Abel is started from a third PC to facilitate packet sniffing and recording the secured VoIP call.

PARAMETER	RESULT
Packet Sniffing	<i>Only static noise audible</i>
Maximum Delay (latency)	<i>141.72 ms</i>
Maximum Jitter	<i>22.17 ms</i>
Packet Loss	<i>0 (0%)</i>
Maximum Bandwidth	<i>93.02 kbps</i>

Table 4.3: Parameters of SRTP-enabled VoIP call from user 110 to user 111

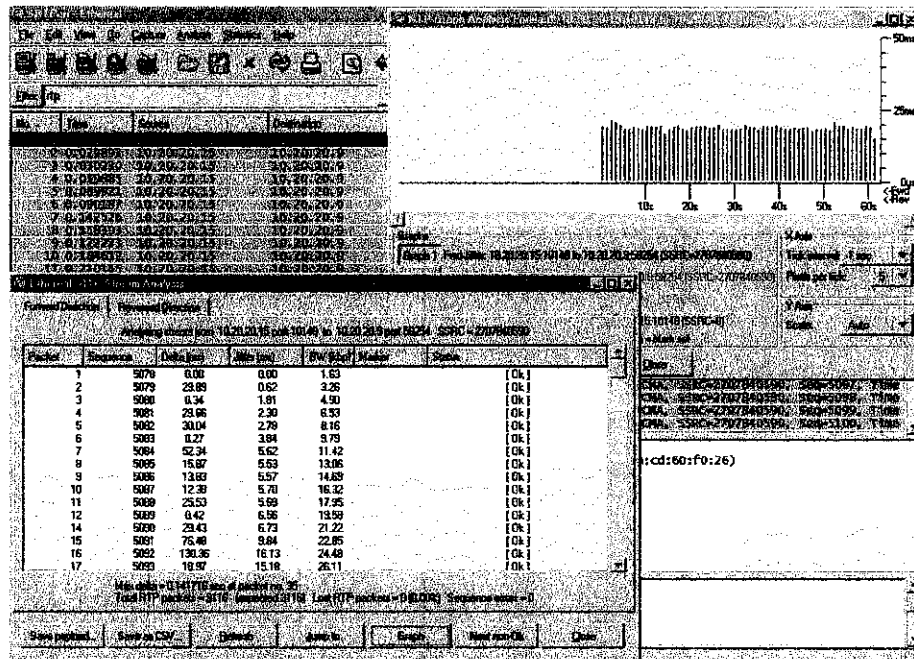


Figure 4.6: Ethereal RTP stream analysis (SRTP-enabled)

4.7 SRTP DISCUSSION

Based on the results and findings of the two test scenarios, it is found that SRTP proves its functionality by producing only static noise when capturing the VoIP conversation using Cain & Abel packet sniffer. In the SRTP-disabled state, the connection is exposed and vulnerable as the whole conversation is audible in the recording. Thus, SRTP media encryption is proved to be functional and usable for SIP calls.

In test scenario #1, the maximum delay or latency is recorded as *126.88 ms* while in test scenario #2, it is recorded as *141.72 ms*. This is in accordance with the ITU-T Recommendation G.114 that the upper bound for one-way voice traffic is *150 ms*. This also corresponds to the current latency bound for domestic Public Switched Telephone Network (PSTN) calls. This would equate to a similar call quality for VoIP and PSTN.

The jitter in test scenario #1 is found to be *21.14 ms* while in test scenario #2 is *22.17 ms*. The general rules is that jitter should be under *40 ms* for a high-quality voice call. Packet loss in both test scenarios is 0 (0%) and adheres to the <1% tolerance for acceptable-quality calls. The bandwidth increases only slightly from *89.6 kbps* to *93.02 kbps* from SRTP-disabled to SRTP-enabled test cases. This proves SRTPs low packet expansion and high bandwidth conservation.

PARAMETER	Test Scenario #1 (SRTP-Disabled)	Test Scenario #2 (SRTP-Enabled)	Upper limits for high-quality calls
Max. Delay	<i>126.88 ms</i>	<i>141.72 ms</i>	<i><150 ms</i>
Max. Jitter	<i>21.14 ms</i>	<i>22.17 ms</i>	<i><40 ms</i>
Packet Loss	<i>0%</i>	<i>0%</i>	<i><1%</i>

Table 4.4: Comparison of VoIP parameters and recommended upper limits

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

Studying Voice over IP and implementing SRTP for VoIP network security require some extensive research, studying, analysis, and testing. Although VoIP does exist for quite some time, the security aspects have yet to be studied widely and elaborately. Packet-switched telephony is said to pose a threat to circuit-switched telephony and would probably replace it altogether in the near future. Problems and setbacks of VoIP security have to be studied in length in order to sustain the growth and development of the system. SRTP provides an advantage over other security methods such as IPSec as it conserves bandwidth, has low packet expansion, and traverses NAT and firewalls better. The current technologies in VoIP have the potential to provide adequate protection to its users and providers while matching the security of its counterpart, the traditional PSTN. This draft report contains the finalized works and results of the project and serves as a key milestone to the overall project. With the given time frame and adequate resources, the study and analysis of VoIP as well as implementation of SRTP has proven to be very successful.

5.2 RECOMMENDATION

There are a few suggestions and recommendations for the continuation of the study of VoIP in general and VoIP security in particular. For campus-wide implementation in UTP, the IP networks have to be studied and the hardware and software required has to be identified. This requires the permission and consent of the IT and Media Services department to allow firewall and NAT traversal throughout the UTP network. The cost and benefits of VoIP implementation also have to be determined. The Quality of Service (QoS) in VoIP is to be experimented and tested further to identify which is the most suitable QoS to be utilized.

In the security aspect, further research is recommended in studying the SRTP key exchange protocols. As several key exchanges exist, a detailed study is required to determine the best protocol for implementation in end-user phones. A detailed identification of threats and vulnerabilities are also required to implement a full-proof and secure VoIP network. With additional budget, a better test scenario can be depicted using SRTP-supported hardware phones, firewalls, routers, and media gateways. This would provide a clearer picture on the implementation of SRTP in a real-world scenario.

REFERENCES

- Black, Uyles 1999, *Voice over IP*, New Jersey, Prentice Hall PTR.
- Campbell, Paul; Calvert, Ben; and Boswell, Steve 2003, *Security+ Guide to Network Security Fundamentals*, Boston, Cisco Learning Institute.
- Canavan, John E. 2000, *Fundamentals of Network Security*, Boston, Artech House.
- Khasnabish, Bhumi 2003, *Implementing Voice over IP*, New Jersey, John Wiley & Sons, Inc.
- Minoli, Daniel and Minoli, Emma 2002, *Delivering Voice over IP Networks*, Indianapolis, Wiley Publishing, Inc.
- Ohrtman, Franklin D 2003. *Softswitch: Architecture for VoIP*, New York, McGraw-Hill.
- Swale, Richard 2001, *Voice over IP: Systems and Solutions*, London, The Institution of Electrical Engineers.
- Wright, David J. 2001. *Voice over Packet Networks*, West Sussex, John Wiley & Sons, Ltd.
- Cisco Systems, Inc., “Cisco 3700 Series MultiService Access Router Data Sheet”, 2004,
<http://www.cisco.com/en/US/products/hw/routers/ps282/products_data_sheet09186a008009203f.html>
- Cisco Systems, Inc., “Configuring Voice over IP for the Cisco 3600 Series”, 2005,
<http://www.cisco.com/en/US/products/hw/switches/ps669/products_configuration_guide_chapter09186a008007f1f1.html>
- Cisco Systems, Inc., “Understanding Packet Voice Protocols”, 2001,
<http://www.cisco.com/en/US/tech/tk652/tk701/technologies_white_paper09186a008009294d.shtml>
- CommuniGate Systems, “CommuniGate Pro Guide”, 2006,
<<http://www.communiGate.com/CommuniGatePro/>>
- Ethereal: A Network Protocol Analyzer, 2006, <<http://www.ethereal.com>>

- Federal Communications Commission U.S.A - Consumer & Governmental Affairs Bureau, “VoIP – FCC Consumer Facts”, 2006, <<http://www.fcc.gov/cgb/consumerfacts/voip.html>>
- GTK+ - The GIMP Toolkit 2006 <<http://www.gtk.org>>
- gtkmm 2006 <<http://www.gtkmm.org>>
- HowStuffWorks, “How VoIP Works”, 2006, <<http://electronics.howstuffworks.com/ip-telephony.htm>>
- International Engineering Consortium On-Line Education. 2005 <<http://www.iec.org/online>>
- Ixia, “H.323 Technology Guide”, 2004, <http://www.ixiacom.com/technology_guides/pdfs/h323.pdf>
- Ixia, “SIP Technology Guide”, 2004, <http://www.ixiacom.com/technology_guides/pdfs/sip.pdf>
- libSRTP 2006 <<http://srtp.sourceforge.net/srtp.html>>
- Minisip 2006 <<http://www.minisip.com>>
- National Institute of Standards and Technology, “Security Considerations for Voice Over IP Systems”, January 2005, <<http://csrc.nist.gov/publications/nistpubs/800-58/SP800-58-final.pdf>>
- Snom Technologies AG 2006 <<http://www.snom.com>>
- The Internet Engineering Task Force (IETF), “RFC 3261 – SIP: Session Initiation Protocol”, June 2002, <<http://www.ietf.org/rfc/rfc3261.txt>>
- The Internet Engineering Task Force (IETF), “RFC 3711 – The Secure Real-time Transport Protocol”, March 2004, <<http://www.ietf.org/rfc/rfc3711.txt>>
- Voice over IP Security Alliance (VoIPSA) 2006 <<http://www.voipsa.org>>
- VOIP-info.org. 2006 <<http://www.voip-info.org/tiki-index.php>>
- Wikipedia - The Free Encyclopedia, “Secure Real-time Transport Protocol”, June 2006, <http://en.wikipedia.org/wiki/Secure_Real-time_Transport_Protocol>
- Wikipedia - The Free Encyclopedia, “Voice over IP”, June 2006, <<http://en.wikipedia.org/wiki/Voip>>

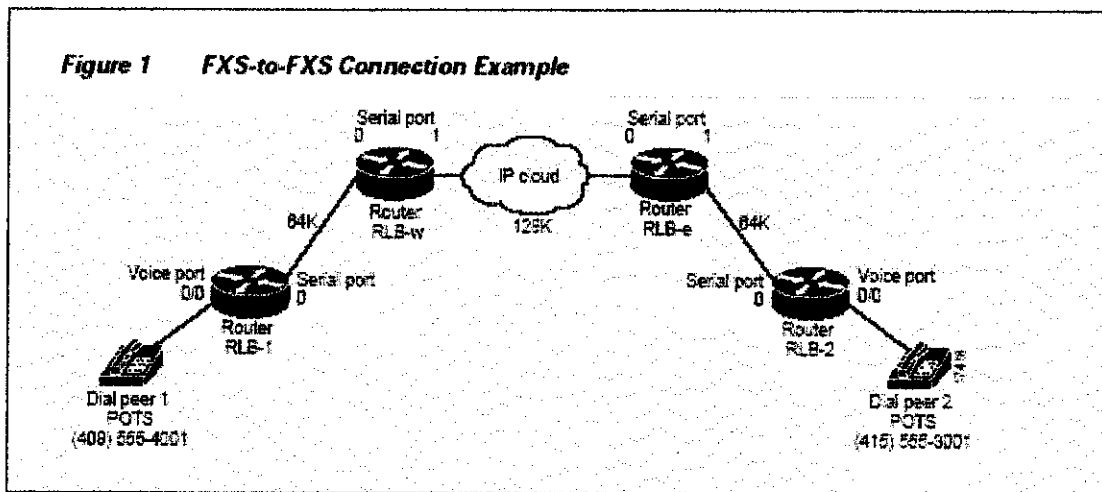
APPENDICES

Appendix I: Commands used to configure and monitor VoIP

Command	Description
acc-qos	Generate an SNMP event if the QoS drops below a specified level.
answer-address	Specify the full E.164 telephone number to identify the dial peer of an incoming call.
codec	Specify the voice coder rate of speech for a dial peer.
comfort-noise	Specify whether or not background noise should be generated.
connection	Specify a connection mode for a specified voice port.
cptone	Configure a voice call progress tone locale.
description	Include a description of what this voice port is connected to.
destination-pattern	Specify either the prefix or the full E.164 telephone number to be used for a dial peer.
dial-control-mib	Specify attributes for the call history table.
dial-peer voice	Enter the dial peer configuration mode.
dial-type	Specify the type of out-dialing for voice-port interfaces.
echo-cancel coverage	Adjust the size of the echo cancel.
echo-cancel enable	Enable the echo cancel feature.
expect-factor	Specify when the router will generate an alarm to the network manager.
fax-rate	Establish the rate at which a fax is sent to the specified dial peer.
icpif	Specify the Calculated Planning Impairment Factor (CPiF) for calls sent by a dial peer.
impedance	Specify the terminating impedance of a voice-port interface.
input gain	Configure a specific input gain value.
ip precedence	Set IP precedence (priority) for packets sent by the dial peer.
ip udp checksum	Calculate the UDP checksum for voice packets transmitted by the dial peer.
music-threshold	Specify the threshold for on-hold music for a specified voice port.
non-linear	Enable nonlinear processing in the echo canceller.
num-exp	Define how to expand an extension number into a particular destination pattern.
operation	Select a specific cabling scheme for E&M ports.
output attenuation	Configure a specific output attenuation value.
port	Associate a dial peer with a specific voice port.
prefix	Specify the prefix of the dialed digits for this dial peer.
req-qos	Specify the desired QoS to be used in reaching a specified dial peer.
ring frequency	Specify the ring frequency for a specified FXS voice port.
ring number	Specify the number of rings for a specified FXO voice port.
session protocol	Establish a session protocol for calls between the local and remote routers.
session target	Specify a network-specific address for a specified dial peer.
show call active voice	Show the active call table.
show call history voice	Display the call-history table.
show controllers voice	Display information about voice related hardware.
show diag	Display hardware information for the router.

show dial-peer voice	Display configuration information for dial peers.
show dialplan incall number	Pair different voice ports and telephone numbers together for troubleshooting.
show dialplan number	Show which dial peer is reached when a particular telephone number is dialed.
show num-exp	Show the number expansions configured.
show voice dsp	Display current status of all DSP voice channels
show voice port	Display configuration information about a specific voice port.
shutdown (dial-peer configuration)	Change the administrative state of the selected dial peer from up to down.
shutdown (voice-port configuration)	Take the voice ports for a specific VIC offline.
signal	Specify the type of signaling for a voice port.
snmp enable peer-trap poor-qov	Generate poor-quality-of-voice notification for applicable calls associated with VoIP dial peers.
snmp-server enable traps	Enable the router to send SNMP traps.
snmp trap link-status	Enable SNMP trap messages to be generated when this voice port is brought up or down.
timeouts initial	Configure the initial digit timeout value for a specified voice port.
timeouts interdigit	Configure the interdigit timeout value for a specified voice port.
timing	Specify timing parameters for a specified voice port.
type	Specify the E&M interface type.
vad	Enable VAD for the calls using this dial peer.
voice-port	Enter the voice port configuration mode.

Appendix II: FXS-to-FXS connection using RSVP



Configuration for Router RLB-1

```

hostname RLB-1
! Create voip dial-peer 2
dial-peer voice 2 voip
! Define its associated telephone number and IP address
destination-pattern 14155553001
sess-target ipv4:40.0.0.1
! Request RSVP
req-qos controlled-load
! Create pots dial-peer 1
dial-peer voice 1 pots
! Define its associated telephone number and voice port
destination-pattern 14085554001
port 0/0
! Configure serial interface 0
interface serial1/0
ip address 10.0.0.1 255.0.0.0
no ip mroute-cache
! Configure RTP header compression
ip rtp header-compression
ip rtp compression-connections 25

```

```
! Enable RSVP on this interface  
ip rsvp bandwidth 48 48  
fair-queue 64 256 36  
clockrate 64000  
router igrp 888  
network 10.0.0.0  
network 20.0.0.0  
network 40.0.0.0
```

Configuration for Router RLB-w

```
hostname RLB-w  
! Configure serial interface 0  
interface serial0/0  
ip address 10.0.0.2 255.0.0.0  
! Configure RTP header compression  
ip rtp header-compression  
ip rtp compression-connections 25  
! Enable RSVP on this interface  
ip rsvp bandwidth 96 96  
fair-queue 64 256 3  
! Configure serial interface 1  
interface serial1/0  
ip address 20.0.0.1 255.0.0.0  
! Configure RTP header compression  
ip rtp header-compression  
ip rtp compression-connections 25  
! Enable RSVP on this interface  
ip rsvp bandwidth 96 96  
fair-queue 64 256 3  
! Configure IGRP  
router igrp 888  
network 10.0.0.0  
network 20.0.0.0  
network 40.0.0.0
```

Configuration for Router RLB-e

```
hostname RLB-e
! Configure serial interface 0
interface serial0/0
ip address 40.0.0.2 255.0.0.0
! Configure RTP header compression
ip rtp header-compression
ip rtp compression-connections 25
! Enable RSVP on this interface
ip rsvp bandwidth 96 96
fair-queue 64 256 3
! Configure serial interface 1
interface serial1/0
ip address 20.0.0.2 255.0.0.0
! Configure RTP header compression
ip rtp header-compression
ip rtp compression-connections 25
! Enable RSVP on this interface
ip rsvp bandwidth 96 96
fair-queue 64 256 3
clockrate 128000
! Configure IGRP
router igrp 888
network 10.0.0.0
network 20.0.0.0
network 40.0.0.0
```

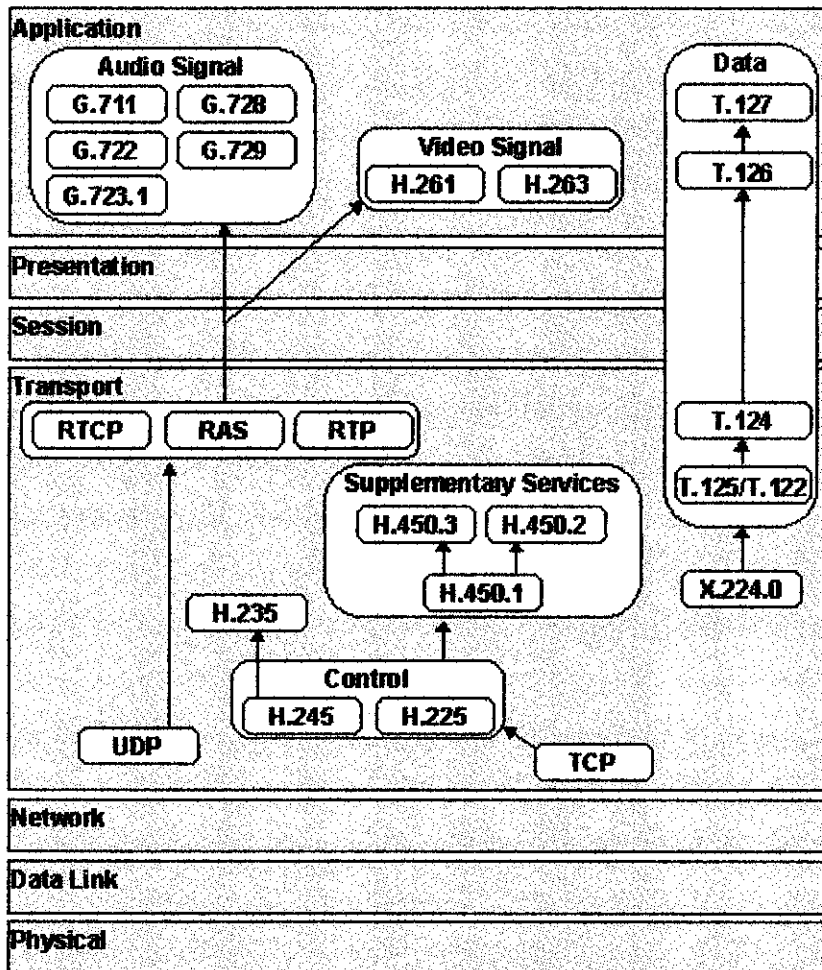
Configuration for Router RLB-2

```
hostname RLB-2
! Create pots dial-peer 2
dial-peer voice 2 pots
! Define its associated telephone number and voice-port
destination-pattern 1415553001
```



```
port 0/0
! Create voip dial-peer 1
dial-peer voice 1 voip
!Define its associated telephone number and IP address
destination-pattern 14085554001
sess-target ipv4:10.0.0.1
! Configure serial interface 0
interface serial1/0
ip address 40.0.0.1 255.0.0.0
no ip mroute-cache
! Configure RTP header compression
ip rtp header-compression
ip rtp compression-connections 25
! Enable RSVP on this interface
ip rsvp bandwidth 96 96
fair-queue 64 256 3
clockrate 64000
! Configure IGRP
router igrp 888
network 10.0.0.0
network 20.0.0.0
network 40.0.0.0
```

Appendix III: H.323 protocol in the OSI layer



Appendix IV: Media security key-exchange protocols comparison table

	Sig. Conf.	Forking	Media before Answer	Shared-key Conf.	PKI	Re-keying	Downgrade attack protection
MIKEY Pre-Shared Key	No	No	Yes	Yes	No	Yes	Yes
MIKEY Public Key	No	No	Yes	Yes	Yes	Yes	Yes
MIKEY-DH	No	No	No	No	Yes	Yes	Yes
MIKEY-DHMAC	No	No	No	No	No	Yes	Yes
MIKEY-RSA-R	No	Yes	No	Yes	Yes	Yes	Yes
SDES	Yes	Yes (Insecure)	No	Yes	No	Yes	No
SDES-EM	Yes	Yes (Insecure)	Yes	Yes	No	Yes	No
EKT	Yes (In SDES mode)	Yes (Insecure)	Yes	Yes	No	Yes	Depends on base handshake
SDP-DH	No	No	No	No	No	No	No
ZRTP	No	Yes	Yes	No	No	Yes	Yes
DTLS	No	Yes	Yes	No	No	Yes	Yes