Bachelor's Thesis (UAS)

Degree Program: Information Technology

Specialization: Internet Technology

2012

Yusuf Ogunjimi

Practical Analysis of Voice Quality Problems of Voice over Internet Protocol (VoIP).

TURUN AMMATTIKORKEAKOULU
TURKU UNIVERSITY OF APPLIED SCIENCES

BACHELOR’S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Degree Programme | Information Technology

2012 | 57

Instructors: Priyadarsan Venugopalan, Patric Granholm

Yusuf Ogunjimi

PRACTICAL ANALYSIS OF VOICE QUALITY PROBLEMS OF VOICE OVER INTERNET PROTOCOL.

The reasons why Voice over Internet Protocol (VoIP) technology is usually adopted by many individuals and enterprises are nothing but cost reduction. However, this reduction in cost of telephone conversation is not without a price which is the quality of the voice in adopting this technology for communication. A good number of factors affect voice quality over the internet of which three are more pronounced than the others. These three factors that are typically creating voice quality problems over the internet are jitter, delay and packet loss.

The purpose of this thesis was to experimentally analyze voice quality problems of VoIP. In analyzing these problems, quantitative research methodology was applied using IP Traffic Generator and NetDisturb which are products of ZTI technologies, a company based in France and Wireshark traffic analyzer, for data gathering. The results obtained from the analysis of data gathered revealed how the main factors that are jitter, delay and packet loss affect voice quality in VoIP system. The findings also showed the relationships among these factors as well as their relationship with bandwidth. This thesis significance was the inter-relationship that was established among the factors that are affecting voice quality when voice traffic is being conveyed over the internet.

KEYWORDS: VoIP, H.323, SIP, PBX

FOREWORD

My gratitude goes to Almighty God for giving me the strength and wisdom to successfully complete this thesis at this trying times. Also, my profound gratitude goes to Priyadarsan Venugopalan of Bymacht Systems Pte, Singapore for believing in me and in this project. Likewise, my appreciation goes to Yves Legendre of ZTI Telecom, France for providing me with free licenses of some software used for the completion of this thesis. In addition, my special thanks go to my classmate in the person of Donald Egbeyon for his endless support.

2012 ,Turku.

Yusuf Ogunjimi.

CONTENTS

FIGURES

TABLES

ACRONYMS, ABBREVIATIONS AND SYMBOLS

VoIP	Voice over IP
PCM	Pulse Code Modulation
IP	Internet Protocol
RTP	Real-Time Transport Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
PBX	Private Branch Exchange
PSTN	Public Switch Telephone Network
SIP	Session Initiation Protocol
IETF	International Engineering Task Force
ISDN	Integrated Service Digital Network
RTCP	Real-Time Transport Control Protocol
UA	User Agent
LAN	Local Area Network
WAN	Wide Area Network
MOS	Mean Opinion Score
PSQM	Perceptual Speech Quality Measurement
PESQ	Perceptual Evaluation of Speech Quality
E-model	Enhanced Model

1 Introduction

The rate at which voice traffic is transferred over packet networks has significantly grown to a larger extent in recent times. In the past, voice traffic utilized frame relay capacity, that is, voice over a frame relay (VoFR) but IP dominance in recent times has shifted most attention from VoFR to Voice over Internet Protocol (VoIP). VoIP involves the transformation and transmission of analogue audio signals to digital signals over the internet. This method of transmitting voice traffic over packet networks can significantly reduce the per-minute cost, which translates into lower rates of long distance call cost.

Most of the dial around calling schemes today rely on VoIP backbones to transmit the voice, passing some of the cost savings to customers. These high-speed connections take advantage of the convergence of internet and voice traffic to create a single managed network. The network convergence also unlocked doors to many novel applications developments. Some of these applications are interactive shopping whereby web pages include applications that enable users to chat or talk with service providers, video or audio streaming, conference calls and some other exciting applications.

Interestingly, as exciting as the VoIP capabilities seem, customers are worried over possible degradation of voice quality when voice traffic is carried over these packets. These genuine concerns may be borne out of previous experience with early applications used for internet telephony. The concerns may also be based on the complete understanding of how the internet data networks work by the customer. What this implies is that acceptability of VoIP services relies heavily on the voice quality.

As a result of this critical parameter in accepting VoIP services which is quality, this thesis will practically analyze the factors affecting voice quality over packet networks which are mainly *Jitter*, *Delay* and *Packet Loss*, and suggest possible techniques for minimizing these factors with the resultant effect of optimizing voice quality in VoIP networks.

1.1 Scope

This thesis introduces and examines how VoIP technologies work, presents the two major signaling protocols used in VoIP technology, factors affecting voice quality problems in VoIP system and the methodologies to analyze these factors. The thesis also presents the empirical analysis of these factors affecting voice quality in VoIP system. In addition, results obtained from the practical analysis are analyzed. Finally, it suggests how to simultaneously minimize the effects of these factors and at the same time optimize voice quality in VoIP.

However, interoperability, security, quality of service (QoS) and transmission media will not be discussed in this thesis. These issues are carefully considered for a complete successful implementation of VoIP solution.

1.2 Chapter Organization

Chapter 1 of this thesis presents the preliminary information about VoIP, the major benefits of adopting the technology as well as the scope of the thesis. Chapter 2 and chapter 3 present VoIP background and the two predominantly used VoIP signaling methods respectively. In Chapter 4, factors affecting voice qualities in VoIP system are presented. Chapter 5 details the methodologies used to practically analyze factors affecting voice quality in VoIP. Chapter 6 presents the design, implementation and testing of the factors presented in Chapter 4. Chapter 7 which is the last chapter analyzes the results of the tests carried out in Chapter 6. It also suggests some techniques to optimize voice quality in VoIP as well as the concluding remarks.

2 VoIP Background

2.1 Voice over IP

According to Ted Wallingford “Telephony is the transmission of spoken data between two or more participants by means of signals carried over electric wires or radio waves” [1]. Telephone systems have been an essential part of every day coordination of activities of both individuals and corporate enterprises. It was seen largely as a staple course of human interaction ever since Alexander Graham Bell invented the telephone circuit and the public telephone system. The advent of internet technologies and high speed data connectivity ushered in different families of telephony technologies and amongst these technologies is the voice over internet protocol (VoIP). VoIP is a technology that allows telephone calls to be made over the internet. It works by transforming analog voice signals into digital data packets and supports real-time two-way communication of exchange using internet protocol (IP).

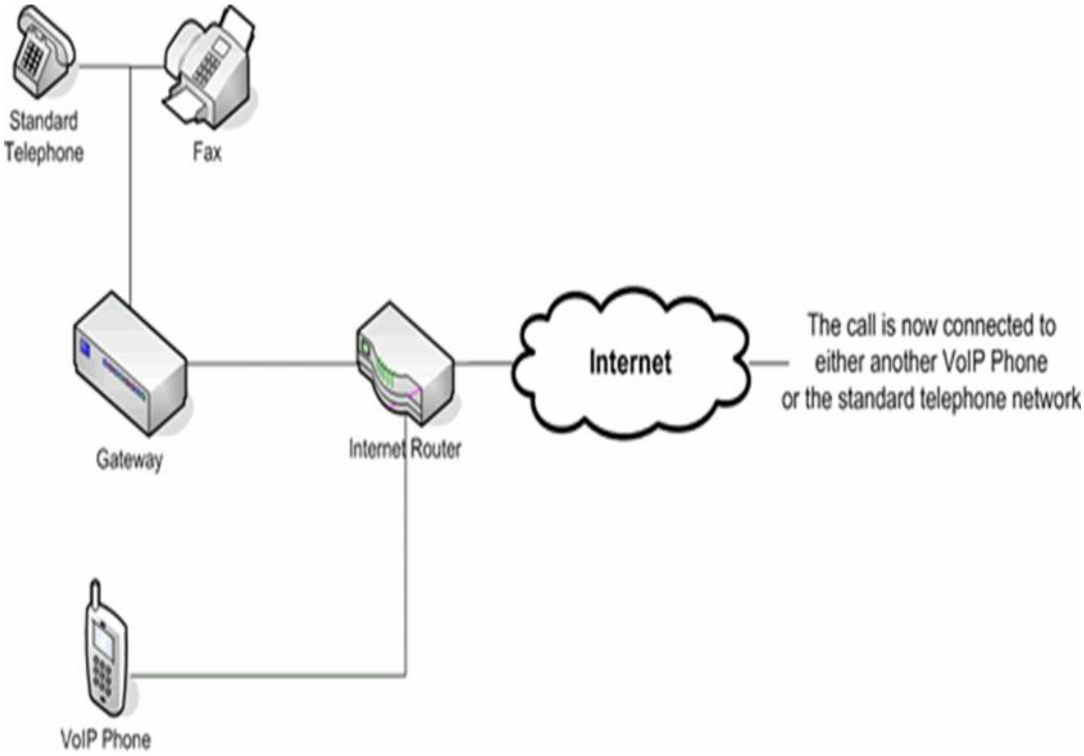

Standard Telephone
Fax
Gateway
Internet Router
Internet
The call is now connected to either another VoIP Phone or the standard telephone network
VoIP Phone

Figure 1. A simple VoIP setup [2].

Figure 1, above is a simple VoIP diagram showing a voice over IP call where both IP phone and traditional handset are deployed.

2.2 Analog and Digital Signaling

2.2.1 Analog Voice Signals

Early telephony systems were built on analog infrastructure and for analog communication until several years ago. Everything audible including physical speech was in analog form. Basically, transmission in analog form is sufficient for human interaction but it is not effective or powerful enough to recover from a line noise. Line Noise can be defined as the introduction of static into a voice network by electrical appliances or radio transmitter [3]. Telephone lines are susceptible to interference in the form of inductance or voltage produced by these handy electric circuits and lines. Amplification of analog signal transmission was a standard procedure during the early development of the telephony system to boost signals.

However, this practice was not efficient enough as it amplified both voice and line noise as well, defeating the amplification effect. Having one input signal going through several amplifiers is known as accumulated noise and this often resulted in most of the time an unusable connection. Line noise amplification is shown by Figure 2, below using amplifiers.

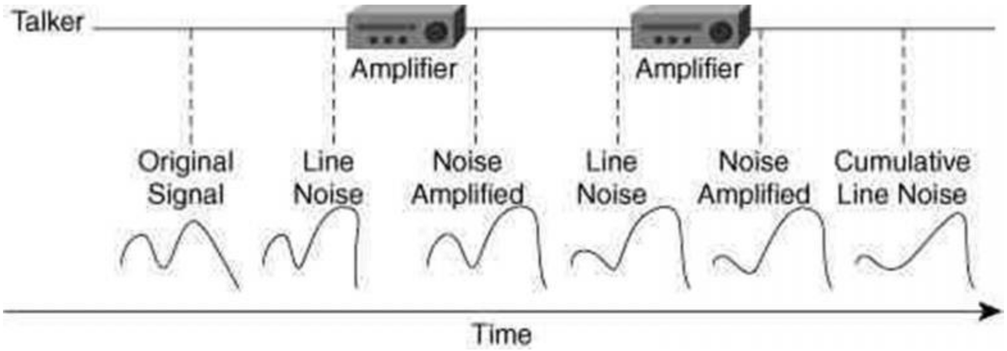


Figure 2. Analog Line Distortion [3].

2.2.2 Digital Voice Signals

Line Noise is not much of an issue in digital networks because repeaters are used instead of amplifiers. The repeaters amplify the voice signal as well as clean it to its original form. This option on digital transmission is achievable due to the fact that digital communication is based on 1s and 0s. Therefore, a clean sound is maintained when signals are repeated. The telephony system was migrated to a pulse code modulation (PCM) when the digital representation benefits were no more in doubt. PCM is the mostly used method in order to convert the analog signal to a digital signal in telephone networks. Figure 3, below shows how voice input is amplified and cleaned to its original form using repeaters.

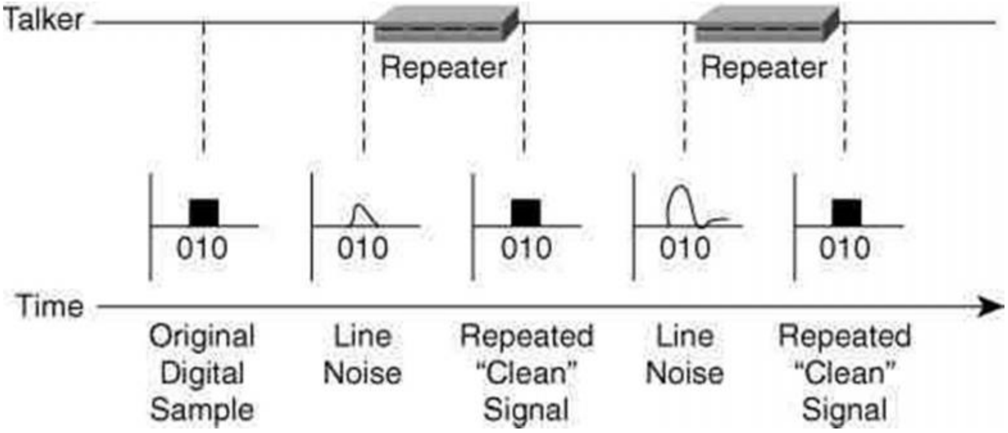


Figure 3. Digital Line Distortion[3].

However, it is noteworthy to mention that digital transmission is not 100 percent free from line noise. This can be caused by different situations one of which is the distance between repeaters. Also, the line noise can be introduced into digital signal transmission as a result of abrupt changes in signal or changes in amplitude are either low or high. A common example of this phenomenon is when digital TV suddenly goes blank when watching a programme. This is because the digital system is confused as to either send a 1 or 0 to the screen thereby resulting in the blank picture.

2.3 IP Transport Mechanisms

There are various features or characteristics that Transport Control Protocol (TCP) and User Datagram Protocol can use for different applications. An application uses TCP/IP to guarantee packet delivery if reliability is more important than delay for the application. In TCP/IP protocol re-transmission is employed to guarantee packet delivery to their destinations reliably. As for applications that make use of UDP/IP protocol, re-transmission is not used which then lowers reliability. However, re-transmission in some cases is not useful, most especially in real-time voice applications or correspondence. Therefore, it is necessary to understand the components of an IP packet.

Understanding the features of an IP packet, gives a better understanding of VoIP signaling protocols, H.323 and *session initiation protocol* (SIP) which are transport layer protocols. Table 1 and Table 2, below present the components of voice packet and IP packets.

Table 1. Voice Packet

Link Header	IP Header	UDP Header	RTP Header	Voice Payload
Variable size depending on link layer protocol	20 Bytes	8 Bytes	12 Bytes	Variable size depending on Codec

Table 2. IP Packet Fields [4]

32 Bits

<table>
<tr><td>Version</td><td>IHL</td><td>Type of Service</td><td colspan="2">Total Length</td></tr>
<tr><td colspan="3">Identification</td><td>Flags</td><td>Fragment Offset</td></tr>
<tr><td colspan="2">Time to Live</td><td>Protocol</td><td colspan="2">Header Checksum</td></tr>
<tr><td colspan="5">Source Address</td></tr>
<tr><td colspan="5">Destination Address</td></tr>
<tr><td colspan="5">Options (+ Padding)</td></tr>
<tr><td colspan="5">Data (Variable)</td></tr>
</table>

2.3.1 TCP

Transmission control protocol (TCP) is a transport layer protocol used for applications that require packet delivery guaranty. It is a connection-oriented, a reliable protocol and provides a guaranteed delivery of data through retransmission. The protocol works by providing a virtual full duplex, acknowledgment and a flow control service to upper layer protocols between two nodes. Data flows in a steady stream byte with a number sequence attached to the first byte for identification, and these are called segments. The continuous flow of the stream is regulated by a flow control scheme known as a sliding window. A TCP port number in a TCP header identifies an upper layer protocol. Some well-known TCP ports are reserved for some upper application layer protocols and examples of these applications are file transfer protocol(FTP) with TCP port 21, World Wide Web (WWW) on TCP port 80, telnet with TCP port 23 and so on.

TCP guarantees data delivery as earlier explained through retransmission and acknowledgement. However, these two factors (acknowledgement and retransmission) put another overhead in the network in form of delay or latency which makes TCP not suitable for real time data transmission such as voice. If TCP was to be used as the transport mechanism for real time voice data on a VoIP call, the voice quality would be rendered unacceptable as a result of latency that would be introduced due to acknowledgment and retransmission. Therefore, latency control takes precedence over reliable packet delivery in VoIP and other real time applications. Nevertheless, within the signaling protocols of VoIP, TCP is typically used to ensure the reliability of call setup. H.323 currently uses TCP and SIP also supports TCP as a transport mechanism.

2.3.2 UDP

User Datagram Protocol (UDP) is also a transport layer protocol used for data delivery where reliability is more or less not needed. This protocol is sometimes called unreliable protocol. Best efforts are used by UDP to deliver packets to their destinations. Unlike the TCP protocol, UDP is a connectionless protocol and does not retransmit packets. For these reasons, UDP is the preferred protocol for multimedia communications. Due to the unreliable nature of UDP, that is, the non-usage of retransmission and acknowledgement, UDP has smaller overhead when compared with TCP. The UDP header has only four fields which are source and destination ports, length and UDP checksum. UDP source and destination port fields act the same way as they do in the TCP header. The UDP header length is specified by the length field and the checksum field guarantees packet integrity although this field is optional in UDP.

UDP is actually the transport mechanism used in VoIP to transmit real-time voice traffic. TCP is not used as a transport mechanism of choice for VoIP applications because retransmission of audio packets as well as flow control are unnecessary. There is a continuous transmission of audio stream regardless of whether 5 percent or 50 percent packet loss is being experienced when UDP is in use for VoIP calls [5].

In summary, TCP/IP is used for the reliability of critical applications while UDP/IP is the preferred protocol when delay is not acceptable and retransmission of a lost packet is unnecessary.

3.0 VoIP Signaling Protocols

Call management servers, public-switched telephone networks (PSTN), legacy private branch exchange (PBX) system and telephones communicate with general languages for setting up and terminating calls in VoIP networks. These common languages are known as VoIP signaling protocols. There are multiple signaling protocols that can be used as the protocol of choice within a VoIP system of which two of these signaling protocols are primarily used. The two signaling protocols are session initiation protocol (SIP) and H.323. The former was developed by an Internet Engineering Task Force (IETF) working group while the latter was developed by International Telecommunication Union (ITU) working group [6]. However, other signaling protocols are Cisco’s skinny client control protocol (SCCP), gateway control protocol (MEGACO/H.248), media gateway control protocol (MGCP) and inter-asterisk exchange (IAX).

3.1 H.323

H.323 is a signaling and control protocol that provides audio and visual communication foundation in any packet network as recommended by the ITU telecommunication standardization sector. The protocol is a peer-to-peer protocol and can be used for multimedia transport and control. It also addresses the issue of bandwidth control for point-to-point and multi-point systems. The protocol is widely deployed by audio-visual equipment manufacturers, real-time application developers, internet service providers (ISPs) and enterprises for both voice and video services over IP networks. The ITU-T has series of H.32x protocols and H.323 is a member of the protocol series. In addition, H.323 is also used as protocol for multimedia communication over ISDN, PSTN or SS7 and 3G- mobile networks.

In November 1996, the ITU-T published the first version of H.323 with video-conferencing capabilities over a local area network (LAN) as its focal point. However, this publication was expressly adopted as a means of voice communication over different IP networks by the industry which includes wide area networks (WANs) and the internet. H.323 versions have been revised several times with the required enhancements to improve voice and video functionality over packet networks and the most interesting feature is that each version is backward compatible with the earlier version. The H.323 name was changed in 1998 to Packet-Based Multimedia Communications Systems and it had since remained unchanged.

The name change was necessitated as a result of H.323 use for communications, not only on LANs, but over WANs and within larger carrier networks. In November 2009, current version of H.323 which is version 7 was approved by ITU-T.

3.1.1 H.323 Elements

The H.323 system has four key elements and the four key elements are terminals, gateway (GW), gatekeeper (GK) and multipoint control units (MCU). The elements are presented in the paragraphs below in the course of this chapter.

a. H.323 Terminal

Terminal is another word for endpoints or nodes and it provides point-to-point or multipoint conferencing for voice, and sometimes video and data packets. A terminal must meet some minimum requirements for it to be regarded as an H.323 terminal and these include *system control unit*, *media transmission*, *audio codec* and *packet-based network interface*. Computers or IP phones with the necessary software and hardware can be used as H.323 terminal.

b. H.323 Gateways (GW)

Gateways are used to provide interoperability between H.323 networks and other networks and interconnect PSTN and ISDN networks for H.323 endpoint internetworking. This internetworking is done by translating between the voice, video and data transmission formats in addition to communication systems and protocols. This includes setting up of calls as well as termination of the calls on both IP networks and switched-circuit networks. Gateways are optional except interconnection between H.323 and other networks is needed. This is because H.323 terminals can communicate directly over packet networks without a gateway.

c. H.323 Gatekeepers (GK)

Address translation services and admission control are provided terminals or gateways by gatekeepers. Gatekeepers and other elements are separated logically in an H.323 environment. The connection between two or more gateways is established in an unspecified manner. Location of remote users can be carried out in two ways, either by a simple query or response sequence (Location Request (LRQ) or Location Confirmation (LCF)).

d. H.323 MCU

A Multipoint controller unit is a component that resides in terminals, gateways or gatekeepers and at least comprises a multicontroller (MC) and one or more multiprocessor (MP). The MC provides information about multipoint conference capabilities to each endpoint present in the multipoint conference and these sets of capabilities can be revised during the multipoint conference. Also, the MP which is one of the components of MCU accepts and transmits the voice, video and/or data streams to terminals participating in the multipoint conferencing.

3.2 H.32 Protocol Suite

The H.323 protocol suite comprises several other protocols and the association of these protocols gives the protocol robust functionalities of call setup, status, call admission, call termination, media stream and a message. Various protocols in H.323 system are supported by both reliable and unreliable transport mechanism for packet transport over packet networks. The protocol suite is divided into three major areas of control and these are

- Registration, Admission and Status (RAS) Signaling which provides pre-call control in H.323 gatekeeper-based network
- Call Control Signaling portion of the H.323 protocol suite that is used to connect, maintain and terminate calls between endpoints.
- Media Control and Transport is responsible for the reliable H.245 channel that carries the media control message using unreliable UDP stream as transport mechanism.

a. RAS Signaling

Where a zone and gatekeepers exist in a H.323 IP network, pre-call control information is usually provided by the establishment of RAS signaling channels between endpoints and gatekeepers across the IP network prior to the establishment of any other channel. This signaling is not dependent on call control signaling and the media transport channel. This pre-call information which is an unreliable UDP connection carries the RAS messages which are procedures for registration, admissions, bandwidth changes, status and disengage. Gatekeeper discovery will only be explained amongst the various messaging procedures for RAS signaling in this thesis.

- Gatekeeper Discovery

Gatekeepers can be discovered by endpoints in two ways, either by manual or automatic method for registration processes. For the manual process, a static IP address of the gatekeeper is configured on an endpoint which enables the endpoint to discover the gatekeeper. Similarly, the automatic process involves endpoint to dynamically discover the gatekeeper using auto discovery mechanism. The auto discovery process works by using a multicast message to enable the endpoints discover the unknown gatekeeper. This reduces the administrative workload or error that accompanies statically configuring endpoints for gatekeeper discovery. UDP port 1718 is used for gatekeeper discovery while UDP port 1719 is used for registration and port status. In addition, a multicast message for gatekeeper discovery uses multicast address 224.0.1.41. Gatekeeper auto discovery in an H.323 system involves three steps for RAS messaging and these are

- Gatekeeper Request (GRQ)- An endpoint sends the multicast message to discover a gatekeeper.
- Gatekeeper Confirm (GCF) - Reply sent from the gatekeeper to an endpoint’s GRQ including the transport address of the gatekeeper’s RAS channel.
- Gatekeeper Reject (GRJ) - This is a notification by the gatekeeper to the endpoint that its registration is rejected.

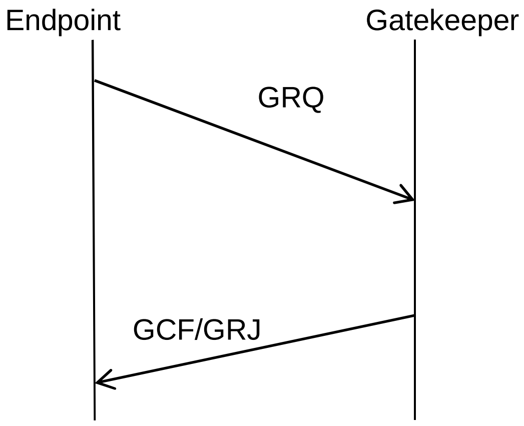


Figure 4. Gatekeeper Auto Discovery Process [3].

a. Call Control Signaling (H.225)

The International Telecommunication Union (ITU) recommends the call control signaling method which is H.225 in H.323 networks. This H.225 recommendation includes the use and support of Q.391 signaling messages. A reliable call control channel across an IP network on TCP port 1720 is created and this port initiates the Q.931 call control messages between two endpoints for the sole purpose of connecting, maintaining and terminating calls.

Although port 1720 is the well-known port for H.323 calls, temporary ports are sometimes assigned by the IP stack of a machine for a specific use after initial call setup for actual call control and keepalive messages. H.225 also specifies the use of Q.932 messages for supplementary services [7]. The commonly used Q.931 and Q.932 signaling messages in H.323 networks are summarized in the table and figure below.

Table 3. Q.931 and Q.932 Messages.

Q.931/Q.932 Messages	Description
CALL SETUP	This is used for call initiation
CALL PROCEEDING	This message shows that call procedure is in progress
ALERTING	This signifies that the called party has been alerted (ringing)
CONNECT	This indicates that the called party accepted the call
RELEASE COMPLETE	This indicates that the call is being released
FACILITY	This shows whether a call is a direct or routed call
STATUS	This indicates the RAS status information message

It is also necessary to point out the fact that the call signaling channel in H.323 networks can be routed in two different ways.

- Direct Endpoint Call Signaling

Call signaling messages are sent directly between two endpoints in the direct endpoint call signaling method. This method is used mainly in setups where endpoints are controlled by a private dial plan. The source information can be provided by endpoints in many ways such as *trunk group* IDs or *trunk groups* (TG). Least call routing is a prominent gatekeeper's application that works based on this method of routing.

- Gatekeeper Routed Call Signaling(GKRCS)

The gatekeeper routed call signaling method differs from the direct endpoint call signaling as a result of the fact that call setup messages are routed through a gatekeeper. A VoIP application where the need for accurate call handling, reporting and centralized provisioning of networks elements is required prefers the GKRCS method of routing.

b. Media Transport (RTP AND RTCP)

Real-time transport protocol (RTP), a media transport mechanism in H.323 networks is responsible for instant end-to-end delivery of a video, data and voice over a unicast or multicast networks. RTP and other lower layer protocols in the OSI model enable the on-time delivery of voice, video or data, response reservation, reliability and QoS. Packetization and transmission services are payload identification, sequencing, timestamping and monitoring.

Real-time transport control protocol (RTCP), a counterpart of RTP, on the other hand, simultaneously monitors data delivery, controls and identifies services. The media channel for RTCP and RTP are created using UDP. RTP streams are transported on even port numbers while the corresponding RTCP streams use the next higher odd port numbers.

It had been mentioned in the beginning of this section that H.323 used to be the most used VoIP protocol and that explained why it is discussed in detail with its underlying protocols in this thesis. However, a less detail and mostly used protocol in today's VoIP networks compared with the H.323 VoIP protocol will be presented in the rest of this chapter and this is session initiation protocol (SIP).

3.3 SIP

The *Session Initiation Protocol* was developed by the Internet Engineering Task Force (IETF) and is described in RFC 2543 and RFC 3261, as a signaling protocol that controls the initiation, modification and termination of interactive multimedia sessions. It is a text-based protocol that is similar to Hyper Text Transfer Protocol (HTTP) and Simple Mail Transfer Protocol (SMTP). SIP is a peer-to-peer protocol, which means that network capabilities such as call routing and session management functions are distributed across all the nodes including endpoints and network servers within the SIP network [6]. SIP does not send many signaling and control messages over the network which is one of the drawbacks of H.323.

Because of its light weight and flexibility, SIP had received more attention from major hardware and software vendors like Microsoft, Cisco, Nortel, SNOM and Lucent in recent past, and research is still on-going for enhancing the protocol. SIP is one of the IETF's multimedia protocols used for interactive communication. Other IETF multimedia protocols are Session Description Protocol (SDP), Session Advising Protocol (SAP), Real-Time Protocol (RTP) Real-Time Control Protocol (RTCP) and Real-Time Streaming Protocol (RTSP). SDP is for maintaining session and flow control for multimedia sessions while SAP advertises multicast conferences. RTP and RTCP, as explained earlier H.323 signaling provides real-time delivery of data. On-demands delivery of real-time data is provided by Real-Time Stream Protocol (RTSP). The detail descriptions of these other IETF multimedia protocols mentioned in this paragraph are not within the scope of this thesis.

The table below presents the messages used in SIP signaling methods.

Table 4. SIP Methods of Signaling.

Method	Description
REGISTER	This message is used to register users with a SIP server
INVITE	This message is used to initiate a call
ACK	This message is used to acknowledge acceptance of a call
BYE	This message is used to end a call
CANCEL	This message is used to reject a call not yet connected
OPTION	This message is used to query the server about its capabilities

The SIP network comprises some key components that interact with themselves by using the signaling messages presented in the Table 4 above. These components are User Agents (UA), Proxy Servers, Registrar Server, Redirect Servers, and Location Servers. User Agents are the same as terminals earlier discussed in H.323 network in this thesis. UAs can be used either as a client or a server in a SIP network. Acting as a client, a UA sends a SIP request to another UA which acts as a server which responds to the SIP demand. UAs can be a PC or an IP phone. The Registrar Server is the SIP element responsible for registering a UA by accepting a register message from it.

The Proxy Server acts like any other proxy server in a network and it functions as an intermediary between the UA and the Registrar Server by accepting a SIP request from a UA and subsequently forwarding this to the appropriate server. The Redirect Server enables a client to call another user directly by providing the calling party with the address information of the called party. The Location Server offers the services that enable both redirect and the proxy server reach the called party by providing the possible address information needed to reach the called User Agent.

SIP as mentioned earlier in this section as being a lightweight protocol when compared with H.323 protocol cannot provide all the capabilities required for interactive multimedia communication and this is the reason why it is implemented with other protocols to deliver the desired capabilities and functionalities expected in a multimedia session. An example of such protocol implemented in a SIP network is the simple traversal of UDP through NAT (STUN) which is used to discover the presence and type of network address translation (NAT) used between a UA client behind a firewall and the public internet. In addition, the protocol is also used to determine the public IP address assigned to the NAT.

3.3.1 SIP Addressing

SIP addresses are referred to as SIP user resource information (URI) and are used to identify a user within a network domain. Typically, SIP URI is written in email address format:

sip:user@domain:port

sip:user@host:port

A user in the URI can be a name such as Yusuf or phone number 3581234567 within the domain or host context. The port field is optional and the default port for SIP URI is port 5060. SIP URIs examples can be written as:

sip:yusuf@bymacht.com

sip:3581234567@proxy1.bymacht.com

Likewise, there is an address-of-record (AOR) which is also known as public SIP URI and it is globally routable pointing to a domain whose location service can map the AOR to another SIP URI, where the user might be located [6].

3.3.2 SIP Responses

Responses to SIP request are sent by a proxy server or UA server to the offer initiated by a UA client to the server indicating the status of the request. SIP responses are grouped in hundreds, e.g.,1xx, 2xx, and the grouping ranges from 100 to 699. SIP responses are classified into two categories; provisional and final responses. A provisional response is usually indicated with status code 1xx and this shows the progress of the server in handling the SIP request, while the final response indicates the status and the termination of a SIP request with status codes from 2xx to 6xx.

Below is a summary of a SIP response table which shows the status code and description of the code.

Table 5. SIP Response Table [3].

Class of Response	Status Code	Description
Informational	100	Trying
	180	Ringing
Success	200	Ok
Redirection	305	Use proxy
Client -Error	400	Bad request
Server-Error	502	Bad gateway
Global failure	603	Decline

SIP servers can be operated in two modes which are Proxy and Redirect. The proxy mode as earlier explained functions as a rallying point between user agents and servers which makes SIP request on behalf of a UA. This mode is similar to the gatekeeper routed call signaling in H.323. Unlike the Proxy mode, a redirect server mode provides the UA with the direct contact information of the called party and this direct contact information is in the form of an IP address. This mode is also similar to the direct endpoint call signaling in H.323 signaling method.

In summary, this chapter explained the underlying technologies and capabilities of VoIP Signaling methods used in VoIP for the interactive media session. The next chapter examines the different factors that affect voice quality in VoIP networks.

4 FACTORS AFFECTING VOICE QUALITY IN VoIP SYSTEM

For an efficient and effective design of a VoIP system with excellent voice quality, it is pertinent to understand the basic principles and networking technologies involved. This chapter provides detail information about various factors that are impacting voice quality in packet networks. Commonly known amongst these factors are *delay/latency, jitter,* and *packet loss*. Others are *echo* and *voice activity detectio*n. These factors account for the reasons why VoIP has not entirely replaced the standard or legacy phone system.

4.1 DELAY/LATENCY

Delay or Latency in a VoIP system is defined as the duration it takes the called party (listener) to hear what the calling party (caller) says during an interactive voice exchange. This factor is, however, the most important factor used in determining the voice quality in VoIP system. The major problem this factor causes is mainly speech overlap. Delay can be introduced into VoIP system as a result of different factors including distance and these produce different delays such as propagation delay, handling delay, queuing delay, packetization and buffering [5]. The ITU-T defines the average network delay for voice applications in recommendation G.114 and this recommendation defines three bands of one-way delay as shown in the table below [3].

Table 6. ITU-T Recommendation G.114 on Delay specification [3].

Ranges in Milliseconds	Description
0 – 150	Acceptable for most applications
150 - 400	Acceptable on the condition that delay effect is known to users
Above 400	This is generally unacceptable

4.1.1 PROPAGATION DELAY

Propagation delay arises as a result of the distance travel through a media by electric signals or light from one end to another. The medium can be fibre networks, copper-based or a wireless network. Propagation delay differs for different media. This delay has a direct relationship with the speed of light which is 3×10^8 m/s when it traverses a vacuum while it is approximately 2×10^8 m/s when an electric signal or light travels through a fibre or copper. Although this delay is not detectable by the human ear, it can affect speech quality significantly when combined with other delays.

4.1.2 HANDLING DELAY

Handling delay is sometimes referred to as processing delay. It describes different causes of delay in VoIP systems when a voice packet is being processed. The different causes of delays include codec processing, packetization and serialization. Serialization impact on delay is infinitesimally small and for this reason it will not be discussed in this thesis.

a Codec Processing Delay

Codec simply means a coder and a decoder. An audio codec converts audio analog to digital at one end and reverses the process at the other end. Codec processing delay arises when voice signal transit through different states which are coding, compression, decompression and decoding. Coding can be said to be the conversion of analog audio signals to digitize signals. Compression is the use of an algorithm to reduce space and bandwidth required for transmitting data. Decoding and Decompression are the opposite of coding and decoding processes respectively.

Low bit rate codecs (G.732.1 and G.729) usually use a more complex algorithm for a voice or audio compression which reduces the bandwidth requirement. However, this introduces codec processing delay as a result of higher computational time which arises as a result of the complex algorithm used, thereby degrading voice quality. In a nutshell, there should be a balance between network bandwidth requirement and voice quality based on the computational power used by codec algorithms.

The table below shows the relationship between different codecs, bit-rate coders and codec processing delay.

Table 7. Relationship between Codecs and Codec Processing Delay [8].

Codec	Bit Rate (kbps)	Codec Processing Delay (ms)
G.711	64	0.75
G.726	32	1
G.728	15	3 to 5
G.729 A	8	10
G.723.1	6.3	30
G723.1	5.3	30

b Packetization Delay

Packetization delay is the time taken for a packet payload to be filled with compressed speech. This delay is a function of the sample block size required by the voice coder and the number of blocks placed in a single frame. Packetization delay is sometimes called Accumulation delay, as the voice samples accumulate in a buffer before they are released [3]. To determine acceptable voice quality in a VoIP network, the packetization delay should not be more than 30 ms. Although, packetization delay increases as packets gets larger, large packets are preferred to smaller ones, because using smaller packets causes more overhead by the header information. As a result of this situation, a compromise between bandwidth utilization and packetization delay is made which determines the voice quality. The size of a voice packet is 40 bytes for every packet as shown in the table below.

Table 8, Voice Packet Header

RTP → 12 Bytes	**UDP** → 8 Bytes	**IP** → 20 Bytes

4.1.3 Queuing Delay

Apart from the different types of delays earlier explained in this chapter, there are other forms of delays in packet networks one of which is queuing delay. Queuing delay occurs when a network device has more packets to process on an outbound interface at a given time than it can actually handle. As a result of processing multiple packets at a time, packets are held in a queue due to congestion on the outbound interface. Prioritization of voice traffic may be deployed in a packet network to minimize this delay. In addition, the queuing delay factor should be kept less than 10 ms whenever it is possible. This is because longer delay is unacceptable in almost all voice networks.

4.1.4 Jitter Buffer Delay

Jitter can be defined as the difference between arrival times of packets in packet based networks. In voice networks, packets are expected to be reliably transmitted at regular time intervals between endpoints, for example, frames sent every 15 ms. However, there is no guarantee that these packets will arrive at their destinations at the regular time specified without delay in the network caused by different factors like propagation delay. To mitigate such inter-arrival time delay between packets, jitter buffer delay with proper sizing is introduced. This buffer conceals the difference between the arrival times between packets. The proper buffer size is very important so as to compensate for delays effectively as setting the jitter buffer too low or too high might cause unnecessary delay and packet loss respectively.

4.2 Jitter

Jitter as explained above is the difference between packet arrival times at their destination. This difference between when a packet is expected and when it is actually received may be caused by routing, as different packets have different routes to their destinations. Because packet networks are unpredictable in nature, an adaptive mechanism like the jitter buffer is needed to neutralize the inter-arrival time difference that might be introduced when packets traverse different routes. RTP timestamps can be used to ascertain the level of jitter in a packet network and appropriately adjust the buffer size.

4.3 Packet Loss

Packet loss is not an uncommon occurrence in data networks especially when the network is saturated or congested. This factor is leveraged on by many protocols in data networks to be able to determine the actual status of network conditions and expressly reduce the amount of packets sent when it is discovered that packets are being lost. Voice traffic is critical and important to many organizations; this is why there is a need to limit the number of packet loss in data networks. Packet loss is a very common issue in voice network or VoIP, because voice traffic is to be real-time and transported with UDP mechanism which does not introduce as much overhead as the transmission control protocol, but unreliable in nature and does not retransmit lost packets.

4.4 Echo

Echo can be said to be a scenario where a caller hears his or her voice back in the telephone receiver after some time during a phone conversation. This is a very common scenario in telephone conversation but can easily become annoying and unbearable. It can also cause interruptions which break the steady flow of conversation; if the voice is heard after a delay of more than about 25 ms. Technically, echo occurs as a result of impedance mismatch from a four-wire network switch conversion to a two-wire local loop network in traditional telephony networks [9]. However, the impedance mismatch in such networks is controlled by echo cancellers.

Echo has two features which render it undesirable in voice networks; these are loudness and how long the echo is. Echo can have one or both of these drawbacks at the same time which renders voice communication unacceptable. The louder and the longer the echo is, the more annoying and incoherent voice conversation becomes. Various networks utilize different methods of canceling out echo during voice conversation. In analog telephony networks, suppressors are used to mitigate echo effects. However, this method causes other forms of problems. In VoIP networks, echo cancellers are built into low bit codecs and are included within a digital signal processor (DSP) to minimize the impact of echo in the system. Echo cancellers work by keeping an opposite pattern of voice signals passing through them for a certain amount of time and amalgamating this inverse pattern with the echo signal bouncing back from the receiving end [10]. The best practice during an initial installation of VoIP equipment is configuring the appropriate amount of echo canceller.

4.5 Voice Activity Detection (VAD)

A mechanism used in VoIP or any other voice system to detect the presence or absence of physical speech is known as *voice activity detection* (VAD). Usually during phone conversation, it is expected to pause intermittently while speaking on the phone. While there is a pause in speech activity, the voice activity detection detects that and suspends voice packet generation and this in turn can transform into bandwidth gain.

VAD implementation leads to bandwidth gain because without it the system cannot detect breaks and pauses in speech activity and continues to generate voice packets during the silent period which results in the wasting of bandwidth. In a VoIP system, this wasted bandwidth can be used for other purpose when VAD is enabled. When VAD detects a quiet period in voice conversation, it waits for about 200 ms before suspending voice packet transmission.

Nevertheless, VAD like every other technology has its own drawbacks: front-end speech clipping and signal-to-noise threshold. Front-end speech clipping means the cutting off or losing the first few sentences made during a voice conversation when VAD detects a speech activity and transits from a silence suppression mode to a packet generating mode. Signal-to noise threshold, on the other hand, is the inability of the VAD to distinguish between speech and a background noise which triggers the VAD to transit from the suppression mode to the packet generating mode and this consumes bandwidth. The table below presents VAD and the bandwidth gain it is enabled for codec G.732.1.

Table 9.~~,~~ Bandwidth gain by Silence Suppression [5].

Codec	Silence Suppression	Background Noise	Number of IP Packets	Number of bytes	IP-level bandwidth (kb/s)	BW gain by silence suppression
G.723.1 5.3 kb/s	ON	Quiet	8047	636,989	5.7	1.88
	OFF	Quiet	15,062	1,203,289	10.7	−−
	OFF	Car Noise	15,053	1,202,545	10.7	1.00
	OFF	Car Noise	15,053	1,202,569	10.7	−−

.

5.0 VOICE QUALITY MEASUREMENT METHODOLOGY

Voice over IP call quality can be impaired or affected by many factors as such as noise, delay, jitter and so on as discussed in the previous chapter. This chapter aims to briefly explain the different methods available for measuring voice quality in VoIP system and present the method to be used in this thesis. Measuring voice quality can be either subjective or objective. Subjective methods of measuring voice quality involves studying human perceptions of calls by allowing users rate the voice samples after listening to such samples. This testing method of voice quality is time and resource-consuming to implement.

However, it is the most commonly used method for measuring voice quality in VoIP system. One of the most commonly used methods which comes under the subjective measurement category is *mean opinion score* (MOS).

5.1 MEAN OPINION SCORE (MOS)

The MOS method expresses voice quality in VoIP systems as numerals. The number indicates the quality or rating of the voice sample after transmission and compression by codecs. These number scale ranges from 1 to 5, 5 being the best and 1 the worst. The table below shows a typical MOS rating scale for VoIP call.

Table 10. MOS Rating Scale [3].

MOS Score	Listening Quality Scale	Listening Effort Scale
5	Excellent	No effort required
4	Good	Attention necessary but not appreciable effort needed
3	Fair	Moderate effort needed
2	Poor	Significant effort needed
1	Bad	Nothing is achieved with significant effort

The objective method of measuring voice quality in VoIP system, on the other hand, does not rely on human interpretation of voice quality according to what they hear and rating of the voice sample. This method is based on voice quality measurement of distortion between the transmitted voice sample and the received signals using computer-based tools or a computational algorithm of combined factors that affect conversational voice quality to facilitate the measurement of real time voice quality in the VoIP system.

5.2 PSQM, PESQ and E-Model

The ITU-T P.861, 862 and G.107 recommendations; Perceptual Speech Quality Measurement (PSQM), Perceptual Evaluation of Speech Quality (PESQ), and Enhanced Model (E-model) are objective methods for voice quality measurement respectively. PSQM is usually used to measure codecs quality used in a VoIP system while the PESQ can be used to measure both voice codec and end-to-end voice quality.

However, these objective methods (PSQM and PESQ) do not take into account network delays and packet losses which are critical factors in measuring voice quality in VoIP system and this is where the E-model becomes useful. The E-model is an objective method that measures voice quality without intrusive conditions in a VoIP system which explains why it is the most suitable objective method used in real time for voice quality measurement. As a result of the complexity of the modern day network, it is recommended that the possible interactions between factors that are affecting voice quality in a VoIP system be measured in all possible combinations.

5.2.1 The R Factor

The R factor in E-model approach simply means "transmission rating factor" and it is a numerical measurement of voice quality on a scale of 0 to 100. The value 100 represents the best voice quality with no impairment and R value of 50 is the minimum needed for VoIP networks. In addition, this R value can be converted to MOS value as in the subjective method which effectively estimates how users perceive voice quality. The R factor comprises the algorithmic computation of all the transmission system parameters and is expressed mathematically as:

$$\mathbf{R = R_O - I_S - I_d - I_{e\text{-}eff} + A}$$

Where RO represents signal-to-noise ratio, IS represents impairments that are simultaneously occurring with useful speech which includes loudness,codec distortion and side tones. Id denotes the combination of different delays while Ie-eff is for low bit-rate codecs and packet loss impairments. The A which stands for Advantage factor represents impairments compensation when the user has other access.

The table below clearly shows ITU-T recommendation G.109 which clearly defines R value ranges and their correlation on users' perception of voice quality.

Table 11. Correlation between R value and User Satisfaction [11].

R- Value Range	User Satisfaction
90≤ R <100	Very Satisfied
80≤ R < 90	Satisfied
70≤ R < 80	Some users dissatisfied
60≤ R < 70	Many users dissatisfied
50≤ R < 60	Nearly all users dissatisfied

This thesis will deploy the combination of PESQ and E-model for measuring the voice quality. The PESQ will be used as the best method of measuring the voice codec and end-to-end voice quality while the background noise, echo, delays and packet loss will be measured by E-model approach.

In summary, this chapter explained the subjective and objective methodologies that can be used for measuring voice quality in VoIP system. The next chapter presents the testing of the various factors affecting voice quality in VoIP systems as explained in Chapter 4 and analysis of the test results obtained using the methodologies presented in this chapter.

6 TESTING AND DATA GATHERING

6.1 VOIP IMPLEMENTATION

This chapter describes a simple VoIP experimental design using internet protocol private branch exchange box (IP-PBX), a broadband internet connection, softphones with SIP protocol to test the factors affecting voice quality in VoIP system which had been explained in Chapter 4 and analyze the results using the methodologies explained in Chapter 5 of this thesis. The main factors which usually compromise voice quality are jitter, delay, packet loss. Other factors are codecs, echo, link- error and voice activity detection for emphasis.

6.1.1 NETWORK DIAGRAM

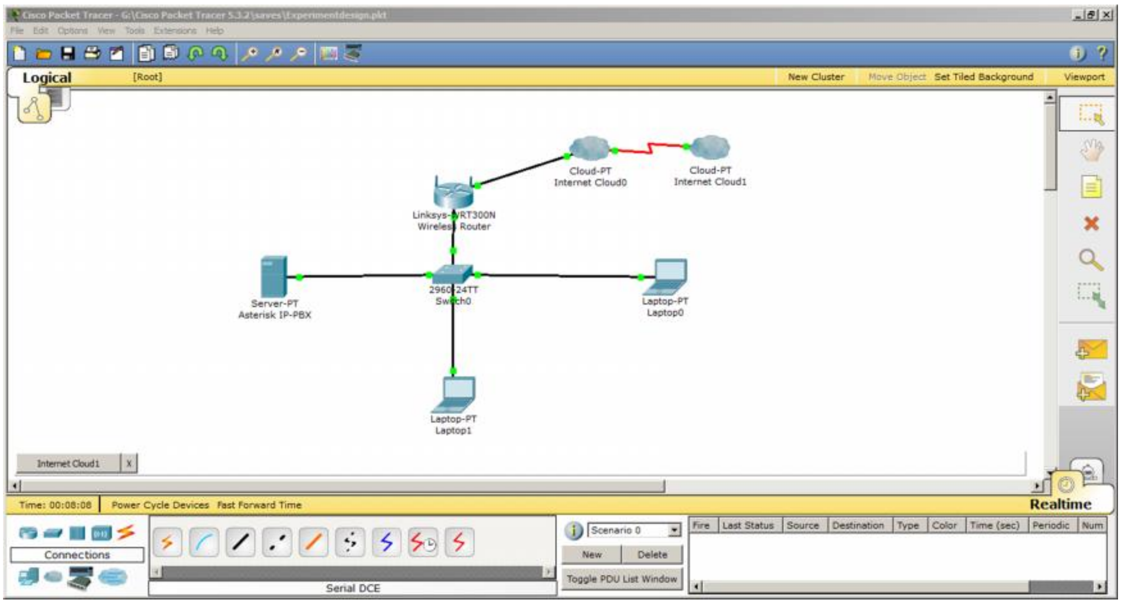


Figure 5. Simple VoIP Topology.

This thesis VoIP implementation is done on a LAN and internet emulation software was used to simulate the behavior of packets in the public internet. As we all know, the internetworking is the interconnection of many networks without any central authority to manage the networks. Also, voice packets traverse different networks and devices in packet networks to reach their different destinations. These factors introduce some forms of delay and other network overhead compared with private LAN which has more manageability and dedicated bandwidth.

6.1.2 Experimental Descriptions

This experiment describes the hardware and software used in the VoIP setup for the testing as well as the analysis of the topical issue in question. This also gives a description of installations and configurations made while the Asterisk IP-PBX was built from scratch using the latest CentOS 6.2 operating system with i386 architecture which is a 32-bit architecture. The compilation of Asterisk 1.8.0.8, an open source telephony server which the IP-PBX derives its telephony capabilities from and the freepbx 2.8 version which is a graphical user interface (GUI) to manage the asterisk telephony server software. Every software mentioned in this paragraph is open source and Linux-based.

The network topology used in this thesis is similar to the screenshot of a packet tracer designed above in Figure 5. This is to give an overview of what the experimental components entail. The components on the left hand side of the first internet cloud are the experimental facilities with full control access rights while the network components behind the Internet Cloud1 are conceptual based on the information extracted from a company. After this design of the VoIP network topology, it is noteworthy to mention both the hardware and software components that made up the topology and the focus are mainly on the testing facility which is on the left hand side on the internet cloud in the topology in the screenshot.

- Hardware Components

Dedicated Server as an IP-PBX with 1Gb memory, 80Gb hard drive and Intel CPU.

Additional Network Interface Card(NIC).

TFT 17 inch monitor

I/O device (Keyboard and mouse)

Gigabyte Router/Switch with 5 ports (1 WAN port and 4 Switch ports)

RJ 45 cables, serial cables and power cables.

- Software Components

Operating system CentOS 6.2 Final release 32-bit architecture

FreePbx v 2.8

Asterisk v1.8.8.0

Softphones (X-lite)

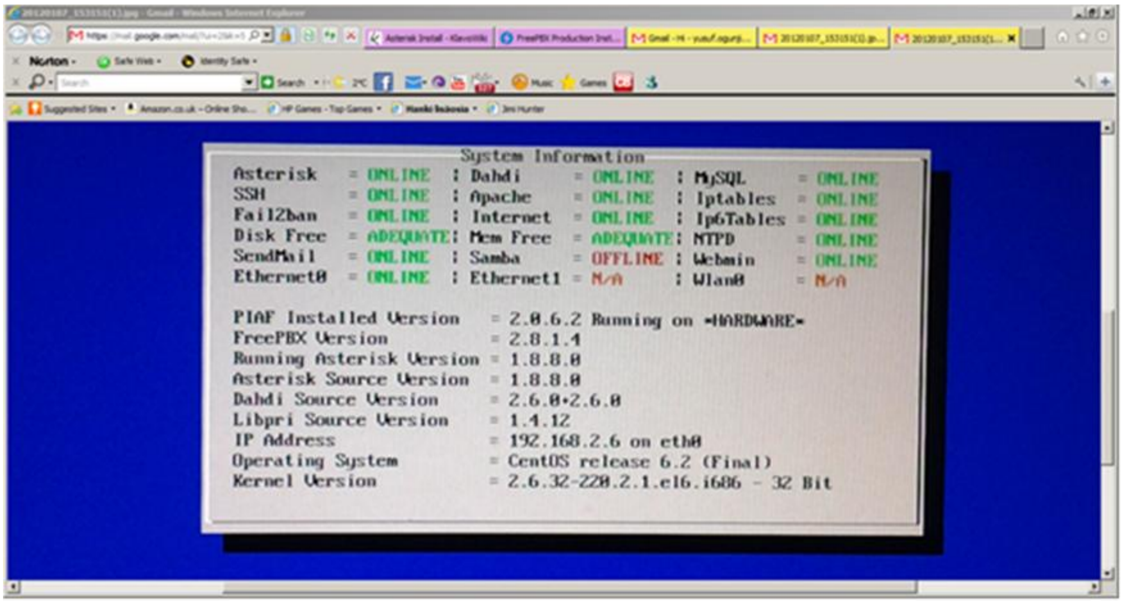
NetDisturb

IP Traffic Generator

NetDisturb and IP Traffic Generator are the products of ZTI Telecom which is a France-based company. NetDisturb is an IP network impairment emulator which is used to introduce different impairments such as jitter, delay, packet loss and so on into an IP network, while the IP Traffic Generator can be used to generate different types of traffic over an IP network. This traffic generated can be either UDP, TCP or internet control message protocol (ICMP).

6.2 ASTERISK INSTALLATION AND CONFIGURATION

After the installation, compiling and building of every necessary package to make the Asterisk PBX server fully functional. A screenshot of the Asterisk server is shown below and some of the services that were up and running were explained before the Asterisk screenshot itself.



```
                          System Information
Asterisk      = ONLINE   : Dahdi      = ONLINE   : MySQL     = ONLINE
SSH           = ONLINE   : Apache     = ONLINE   : Iptables  = ONLINE
Fail2ban      = ONLINE   : Internet   = ONLINE   : Ip6Tables = ONLINE
Disk Free     = ADEQUATE : Mem Free   = ADEQUATE : NTPD      = ONLINE
SendMail      = ONLINE   : Samba      = OFFLINE  : Webmin    = ONLINE
Ethernet0     = ONLINE   : Ethernet1  = N/A      : Wlan0     = N/A

PIAF Installed Version     = 2.0.6.2 Running on *HARDWARE*
FreePBX Version            = 2.8.1.4
Running Asterisk Version   = 1.8.8.0
Asterisk Source Version    = 1.8.8.0
Dahdi Source Version       = 2.6.0+2.6.0
Libpri Source Version      = 1.4.12
IP Address                 = 192.168.2.6 on eth0
Operating System           = CentOS release 6.2 (Final)
Kernel Version             = 2.6.32-220.2.1.el6.i686 - 32 Bit
```

Figure 6. Captured screenshot of an Asterisk IP-PBX.

SSH: This is a secure shell remote login mechanism which is safer when it is compared with telnet.

Fail2ban: This is a daemon used like an access-list to define which IP address is allowed access

to network resources and log system intrusion attempts using python scripts.

Dahdi: This is a Digium Asterisk hardware device interface that replaced the earlier one called Zaptel which is a hardware used for integrating traditional PSTN into VoIP system.

Apache: This is a web server daemon and it is included, should the pbx server be used as a web server in the future.

MySQL: This is mainly used to maintain the database of call detail record (CDR) in VoIP system.

Libpri: Lipbri is used in connecting a VoIP system with a PRI service.

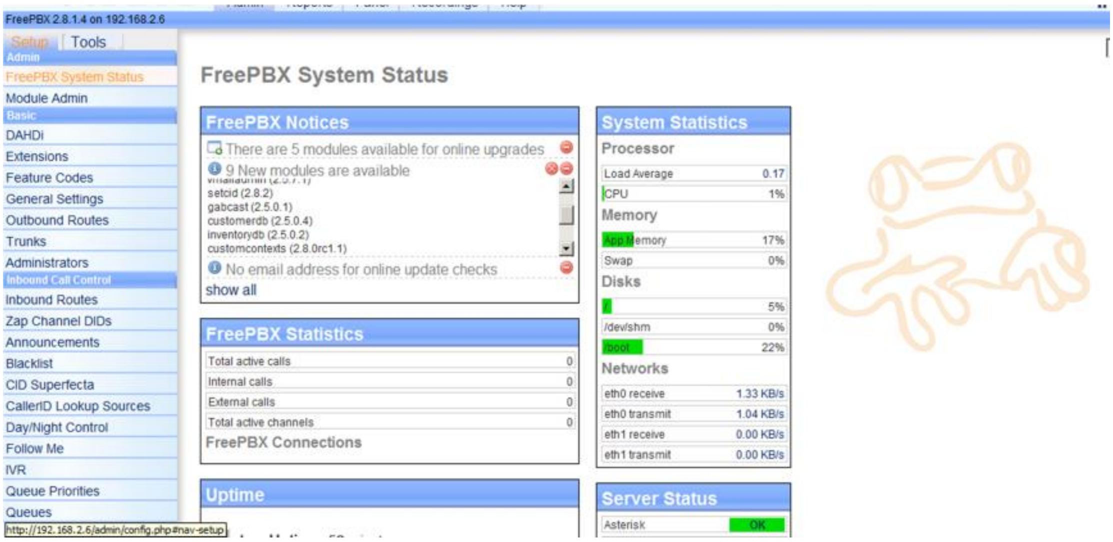
Iptables: This is the firewall in Linux where ports or services can be open, allowed, denied or shut down.

Webmin: This is a web interface for Linux administration on CentOS 6.

FreePBX: This is a graphical user interface that is can be used to manage the Asterisk Server.

Other patches, updates and fixes were carried out to ensure that the Asterisk server is running the latest software needed for fully functional telephony functionality.

After updating the necessary packages and making sure that the Asterisk IP-PBX is fully functional, extensions were created using the FreePBX GUI and were registered on the Asterisk Server. Below are the screenshots of a FreePBX and extension registrations on the Asterisk Server after configuring the X-Lite softphone, so as to be able to dial out or receive incoming calls. Other necessary configurations such as Static IP addressing for the Asterisk server and DNS configuration were also carried out.



FreePBX System Status

FreePBX Notices

There are 5 modules available for online upgrades

9 New modules are available

setcid (2.8.2)
gabcast (2.5.0.1)
customerdb (2.5.0.4)
inventorydb (2.5.0.2)
customcontexts (2.8.0rc1.1)

No email address for online update checks

FreePBX Statistics

Total active calls	0
Internal calls	0
External calls	0
Total active channels	0

FreePBX Connections

System Statistics

Processor

Load Average	0.17
CPU	1%

Memory

App Memory	17%
Swap	0%

Disks

/	5%
/dev/shm	0%
/boot	22%

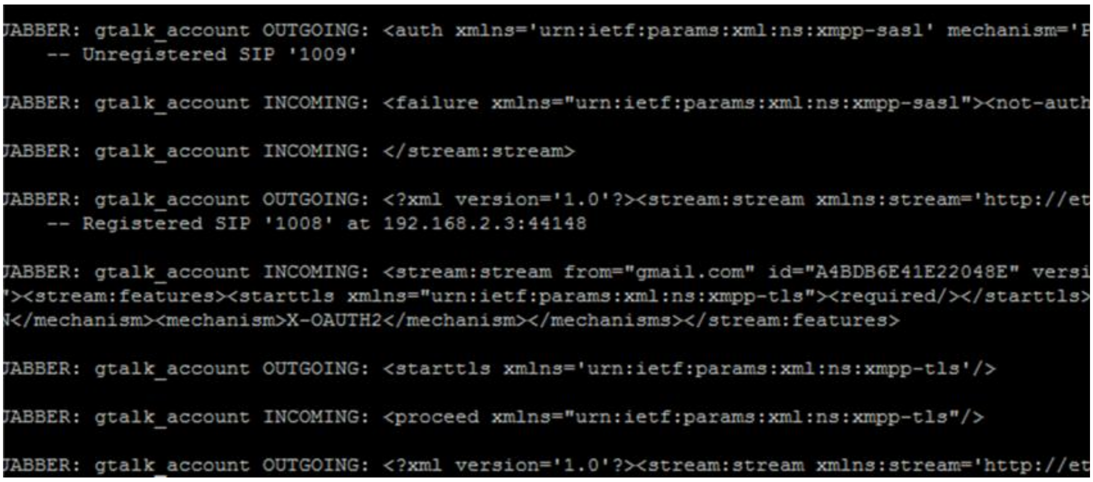
Networks

eth0 receive	1.33 KB/s
eth0 transmit	1.04 KB/s
eth1 receive	0.00 KB/s
eth1 transmit	0.00 KB/s

Server Status

Asterisk	OK

Figure 7. Captured FreePBX screenshot for the Asterisk Server.



```
JABBER: gtalk_account OUTGOING: <auth xmlns='urn:ietf:params:xml:ns:xmpp-sasl' mechanism='P
    -- Unregistered SIP '1009'

JABBER: gtalk_account INCOMING: <failure xmlns="urn:ietf:params:xml:ns:xmpp-sasl"><not-auth

JABBER: gtalk_account INCOMING: </stream:stream>

JABBER: gtalk_account OUTGOING: <?xml version='1.0'?><stream:stream xmlns:stream='http://et
    -- Registered SIP '1008' at 192.168.2.3:44148

JABBER: gtalk_account INCOMING: <stream:stream from="gmail.com" id="A4BDB6E41E22048E" versi
"><stream:features><starttls xmlns="urn:ietf:params:xml:ns:xmpp-tls"><required/></starttls>
N</mechanism><mechanism>X-OAUTH2</mechanism></mechanisms></stream:features>

JABBER: gtalk_account OUTGOING: <starttls xmlns='urn:ietf:params:xml:ns:xmpp-tls'/>

JABBER: gtalk_account INCOMING: <proceed xmlns="urn:ietf:params:xml:ns:xmpp-tls"/>

JABBER: gtalk_account OUTGOING: <?xml version='1.0'?><stream:stream xmlns:stream='http://et
```

Figure 8. Captured Extensions Registration on Asterisk Server.

6.2.1 X-Lite Softphone Configuration

The X-Lite softphone configuration contains the IP address of the Asterisk PBX, the extensions created earlier in the Asterisk PBX using the FreePBX GUI. A sample of the configuration used on the X-Lite software can be seen below:

Display Name: Ossi

Username: 1008

Password: ****

Domain: 192.168.X.6(Asterisk Server IP Address)

All these parameters must match the one earlier set on the Asterisk server using the FreePBX GUI thereby resulting in the output shown by the screenshot in Figure 8 above. A user agent was registered after configuring a softphone with username 1008 and other necessary parameters (Display Name, Domain and Password). It can be seen from line 4 of the command line in Figure 8 that the user extension or username registration was successful with the IP address and port number used by the client to establish the SIP connection shown.

6.2.2 Configuration Difficulties

It is noteworthy to mention that the CentOS 6.2 could not recognize the pre-installed network interface card (NIC) until an extra NIC was installed and this extra NIC was recognized as eth0 while the onboard NIC was later recognized as eth1 after the installation of the new NIC.

In addition, the NetDisturb PC required two NICs and if the wireless network interface were to be used as the second NIC, then a route to the target PC had be added with the Wi-Fi NIC as the gateway in the routing table of the NetDisturb PC. A sample of the route added in the NetDisturb PC can be seen below from the command prompt and this must be run as an administrator.

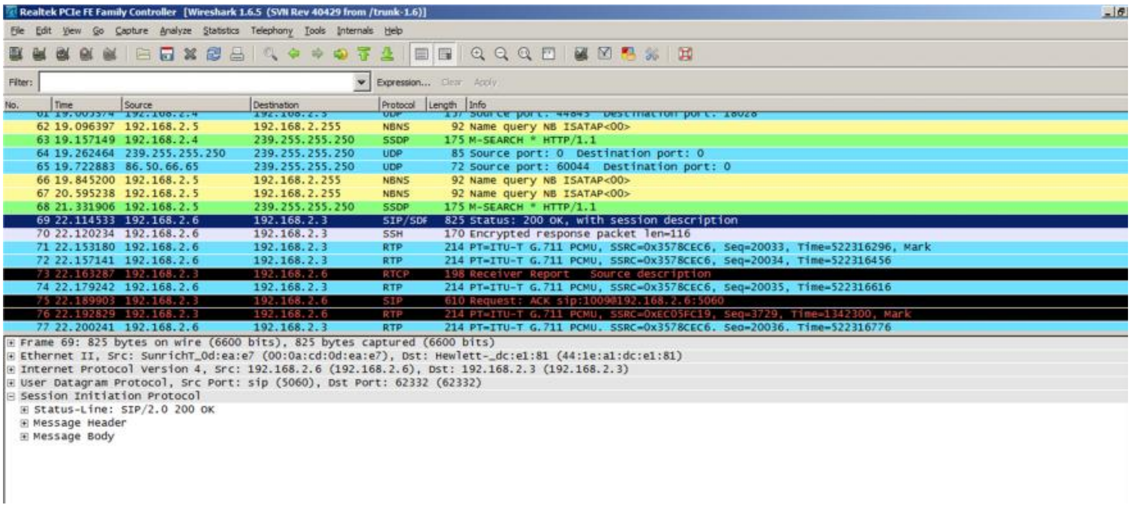
C:\> route add 192.168.2.3 mask 255.255.255.255 192.168.2.5(NetDisturb PC IP Address).

The screenshot of the added route above can be found in the appendix with the caption NetDisturb PC routing table.

6.3 Data Gathering

6.3.1 Testing Without Network Impairment

It was imperative to make a baseline testing to make sure everything is working as expected and do some information gathering before the network impairments were introduced into the system to emulate the internet behavior as previously said. Several phone calls were placed from softphone A on a PC-A(192.168.2.4) to another PC-B(192.168.2.3) with a softphone B and traffic between PCs were captured using wireshark. Below is a screenshot of the captured traffic without impairment.


Realtek PCIe FE Family Controller [Wireshark 1.6.5 (SVN Rev 40429 from /trunk-1.6)]
File Edit View Go Capture Analyze Statistics Telephony Tools Internals Help
Filter:
Expression... Clear Apply

```
No.  Time       Source           Destination      Protocol Length Info
  61 19.003374  192.168.2.4      192.168.2.3      UDP       137 Source port: 44843  Destination port: 18028
  62 19.096397  192.168.2.5      192.168.2.255    NBNS       92 Name query NB ISATAP<00>
  63 19.157149  192.168.2.4      239.255.255.250  SSDP      175 M-SEARCH * HTTP/1.1
  64 19.262464  239.255.255.250  239.255.255.250  UDP        85 Source port: 0  Destination port: 0
  65 19.722883  86.50.66.65      239.255.255.250  UDP        72 Source port: 60044  Destination port: 0
  66 19.845200  192.168.2.5      192.168.2.255    NBNS       92 Name query NB ISATAP<00>
  67 20.595238  192.168.2.5      192.168.2.255    NBNS       92 Name query NB ISATAP<00>
  68 21.331906  192.168.2.5      239.255.255.250  SSDP      175 M-SEARCH * HTTP/1.1
  69 22.114533  192.168.2.6      192.168.2.3      SIP/SDP   825 Status: 200 OK, with session description
  70 22.120234  192.168.2.6      192.168.2.3      SSH       170 Encrypted response packet len=116
  71 22.153180  192.168.2.6      192.168.2.3      RTP       214 PT=ITU-T G.711 PCMU, SSRC=0x3578CEC6, Seq=20033, Time=522316296, Mark
  72 22.157141  192.168.2.6      192.168.2.3      RTP       214 PT=ITU-T G.711 PCMU, SSRC=0x3578CEC6, Seq=20034, Time=522316456
  73 22.163287  192.168.2.3      192.168.2.6      RTCP      198 Receiver Report   Source description
  74 22.179242  192.168.2.6      192.168.2.3      RTP       214 PT=ITU-T G.711 PCMU, SSRC=0x3578CEC6, Seq=20035, Time=522316616
  75 22.189903  192.168.2.3      192.168.2.6      SIP       610 Request: ACK sip:1009@192.168.2.6:5060
  76 22.192829  192.168.2.3      192.168.2.6      RTP       214 PT=ITU-T G.711 PCMU, SSRC=0xEC05FC19, Seq=3729, Time=1342300, Mark
  77 22.200241  192.168.2.6      192.168.2.3      RTP       214 PT=ITU-T G.711 PCMU, SSRC=0x3578CEC6, Seq=20036, Time=522316776

⊞ Frame 69: 825 bytes on wire (6600 bits), 825 bytes captured (6600 bits)
⊞ Ethernet II, Src: SunrichT_0d:ea:e7 (00:0a:cd:0d:ea:e7), Dst: Hewlett-_dc:e1:81 (44:1e:a1:dc:e1:81)
⊞ Internet Protocol Version 4, Src: 192.168.2.6 (192.168.2.6), Dst: 192.168.2.3 (192.168.2.3)
⊞ User Datagram Protocol, Src Port: sip (5060), Dst Port: 62332 (62332)
⊟ Session Initiation Protocol
  ⊞ Status-Line: SIP/2.0 200 OK
  ⊞ Message Header
  ⊞ Message Body
```

Figure 9. Captured Traffic from Between Two Softphones.

Analyzing different frames from the captured traffic in Figure 9 gave an insight into how sessions were setup between the two softphones. Analyzing frame 69-71 for instance indicated that a session initiation protocol had been established between the X-Lite softphone B (192.168.2.3) with the Asterisk server (192.168.2.6) and 200 OK message status was displayed. It can also be seen that both RTP and RTCP protocols which carry encoded streams and communication controls were used respectively. Also, the quality of a voice call made during this capture was very clear and audible in both legs of the calls that are incoming and outgoing at the VoIP terminals.

The data gathered from the wireshark frame analysis is shown in the table below.

Table 12. Gathered data from wireshark before network impairment.

<table>
<tr><td colspan="2">Source IP Addess</td><td colspan="2">Source Port</td><td>Destination Port</td><td>Destination IP Address</td></tr>
<tr><td colspan="2">192.168.2.6</td><td colspan="2">5060</td><td>44148</td><td>192.168.2.3</td></tr>
<tr><td colspan="2">192.168.2.6</td><td colspan="2">5060</td><td>44263</td><td>192.168.2.3</td></tr>
<tr><td colspan="2">192.168.2.6</td><td colspan="2">5060</td><td>62332</td><td>192.168.2.3</td></tr>
<tr><td colspan="6">Port Used</td></tr>
<tr><td colspan="3">5060</td><td colspan="3">44148</td></tr>
<tr><td colspan="3"></td><td colspan="3">44263</td></tr>
<tr><td colspan="3"></td><td colspan="3">62332</td></tr>
</table>

6.3.2 Testing With Network Impairments

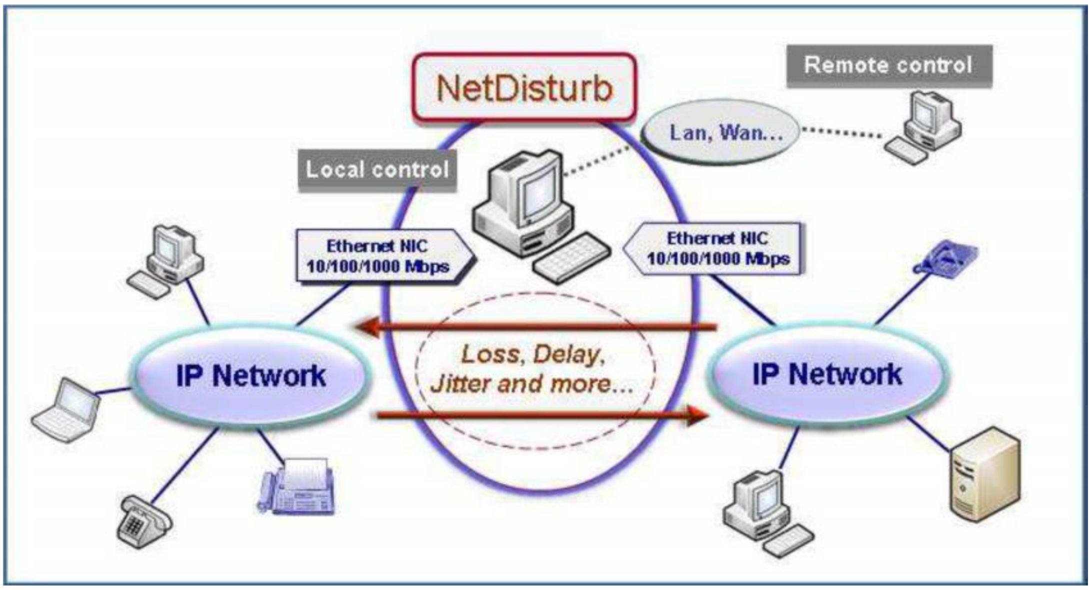
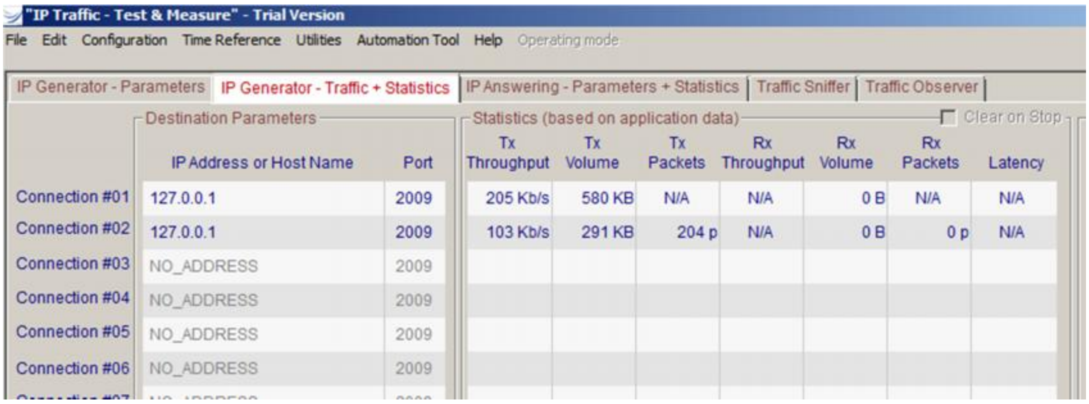

NetDisturb
Remote control
Lan, Wan...
Local control
Ethernet NIC 10/100/1000 Mbps
Ethernet NIC 10/100/1000 Mbps
IP Network
Loss, Delay, Jitter and more...
IP Network

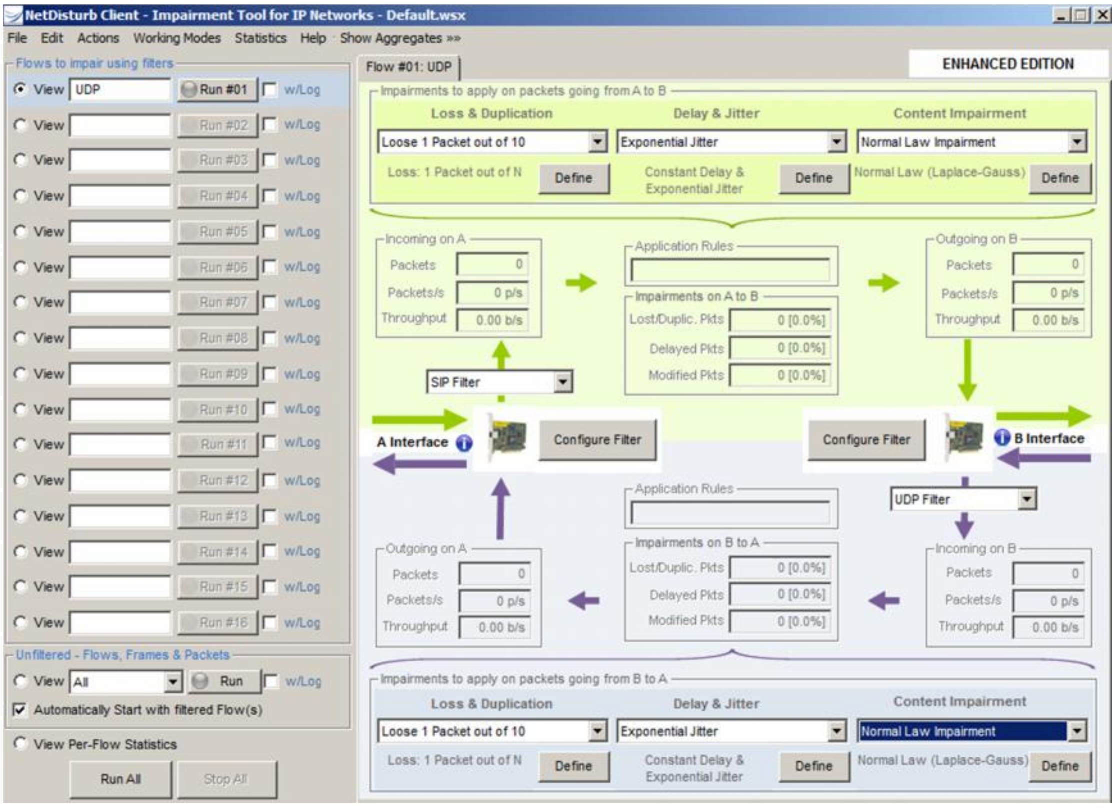
Figure 10. Network Impairment Topology between Two IP Networks red [12].

The testing with impairments was done using the two testing software NetDisturb and IP Traffic Generator. NetDisturb was used, on one hand, to introduce mainly Delay, Jitter and Packet Loss in to the system. The IP Traffic Generator, on the other hand, was used to generate traffic so as to limit the bandwidth on the network and the test was repeated several times and summary of the results obtained can be seen below with some screenshots of the testing parameters.


"IP Traffic - Test & Measure" - Trial Version
File Edit Configuration Time Reference Utilities Automation Tool Help Operating mode
IP Generator - Parameters
IP Generator - Traffic + Statistics
IP Answering - Parameters + Statistics
Traffic Sniffer
Traffic Observer
Destination Parameters
Statistics (based on application data)
□ Clear on Stop

	IP Address or Host Name	Port	Tx Throughput	Tx Volume	Tx Packets	Rx Throughput	Rx Volume	Rx Packets	Latency
Connection #01	127.0.0.1	2009	205 Kb/s	580 KB	N/A	N/A	0 B	N/A	N/A
Connection #02	127.0.0.1	2009	103 Kb/s	291 KB	204 p	N/A	0 B	0 p	N/A
Connection #03	NO_ADDRESS	2009							
Connection #04	NO_ADDRESS	2009							
Connection #05	NO_ADDRESS	2009							
Connection #06	NO_ADDRESS	2009							

Figure 11. Captured throughput from IP Traffic Generator.



● View UDP

☑ Automatically Start with filtered Flow(s)

○ View Per-Flow Statistics

○ View All

Impairments to apply on packets going from A to B

Loss & Duplication	Delay & Jitter	Content Impairment
Loose 1 Packet out of 10	Exponential Jitter	Normal Law Impairment
Loss: 1 Packet out of N	Constant Delay & Exponential Jitter	Normal Law (Laplace-Gauss)

SIP Filter

UDP Filter

Incoming on A	
Packets	0
Packets/s	0 p/s
Throughput	0.00 b/s

Impairments on A to B	
Lost/Duplic. Pkts	0 [0.0%]
Delayed Pkts	0 [0.0%]
Modified Pkts	0 [0.0%]

Outgoing on B	
Packets	0
Packets/s	0 p/s
Throughput	0.00 b/s

Outgoing on A	
Packets	0
Packets/s	0 p/s
Throughput	0.00 b/s

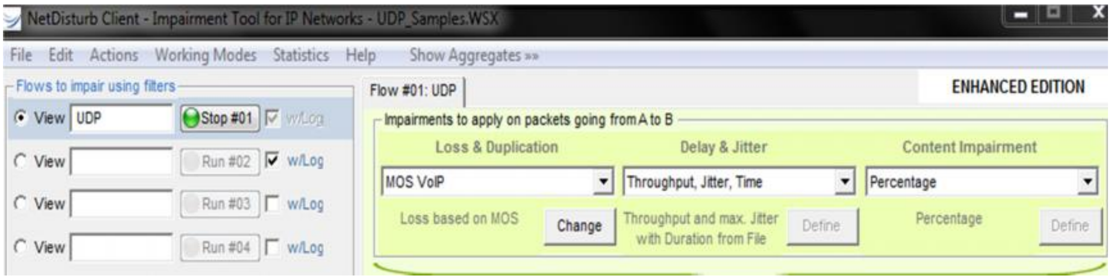
Impairments on B to A	
Lost/Duplic. Pkts	0 [0.0%]
Delayed Pkts	0 [0.0%]
Modified Pkts	0 [0.0%]

Incoming on B	
Packets	0
Packets/s	0 p/s
Throughput	0.00 b/s

Impairments to apply on packets going from B to A

Loss & Duplication	Delay & Jitter	Content Impairment
Loose 1 Packet out of 10	Exponential Jitter	Normal Law Impairment
Loss: 1 Packet out of N	Constant Delay & Exponential Jitter	Normal Law (Laplace-Gauss)

Figure 12. Captured NetDisturb with Exponential Jitter, packet loss and Normal Impairment applied.



NetDisturb Client - Impairment Tool for IP Networks - UDP_Samples.WSX

Flows to impair using filters

● View UDP Stop #01 ☑ w/Log

○ View Run #02 ☑ w/Log

○ View Run #03 □ w/Log

○ View Run #04 □ w/Log

Flow #01: UDP

Impairments to apply on packets going from A to B

Loss & Duplication	Delay & Jitter	Content Impairment
MOS VoIP	Throughput, Jitter, Time	Percentage
Loss based on MOS	Throughput and max. Jitter with Duration from File	Percentage

Figure 13, Captured NetDisturb with Throughput, Jitter ,time and Percentage Impairment content applied.

These test parameters were chosen to observe their effects on voice quality when voice communication is being conveyed over the internet.

Examples of parameters chosen were packet loss, Delay/Jitter Law ’Delay 100ms to 150ms (exponential) and normal impairment law. Packets were delibrately dropped by applying different types of packet loss starting from dropping 1 percent of total transmitted packets. Also, the Delay/Jitter was applied for a specific amount of time.

Table 13. Summary of the data gathered from NetDisturb and IP Traffic Generator capture.

Time (s)	Codec	Jitter (ms)	Delay (ms)	Packet Loss (%)	Throughput (kb/s)
60	G.711	5	103	1	458
120	G.711	10	109	2	400
180	G.711	15	121	5	250
240	G.711	20	150	12	104
300	G.711	40	180	24	52.08
360	G.711	50	200	36	34.72

7 Results and Conclusion

7.1 Data Analysis without Network Impairment

From the data gathered from Table 12 above, it can be seen that the Asterisk Server uses the default port for SIP signaling which is 5060 while the X-Lite softphone dynamically selects port from RTP range of 44148 -62332 every time for incoming or outgoing calls made to and from the softphone. The significance of this result is that most of the VoIP systems in today's network are sitting behind a network address translation (NAT) or firewall and to prevent one way audio problem; the dynamic RTP port must be opened in the firewall or router. This is sometimes called port forwarding and ensures that the SIP signaling protocol is linked with the audio in both ways, that is incoming and outgoing.

7.2 Data Analysis with Network Impairment

Table 13 summarizes the results obtained from NetDisturb and IP Traffic Generator. It can be deduced from the table that Jitter has a linear correlation or relationship with Delay and Packet loss and inverse relationship with the Throughput. For clarity, graphical representations of the data gathered in Table 13 were drawn and it can be seen clearly from the graphs below that Figure 14 (Jitter vs Time), that is, Jitter which is inter-packet arrival or packet that is out of order has a direct or linear correlation with Delay and Packet Loss presented in Figure 15 and 16 respectively and an inverse relationship with Throughput as shown in Figure 17. It can also be seen that as the Jitter value increases, the value of Delay and Packet Loss increased respectively while the Throughput in kb/s decreases.

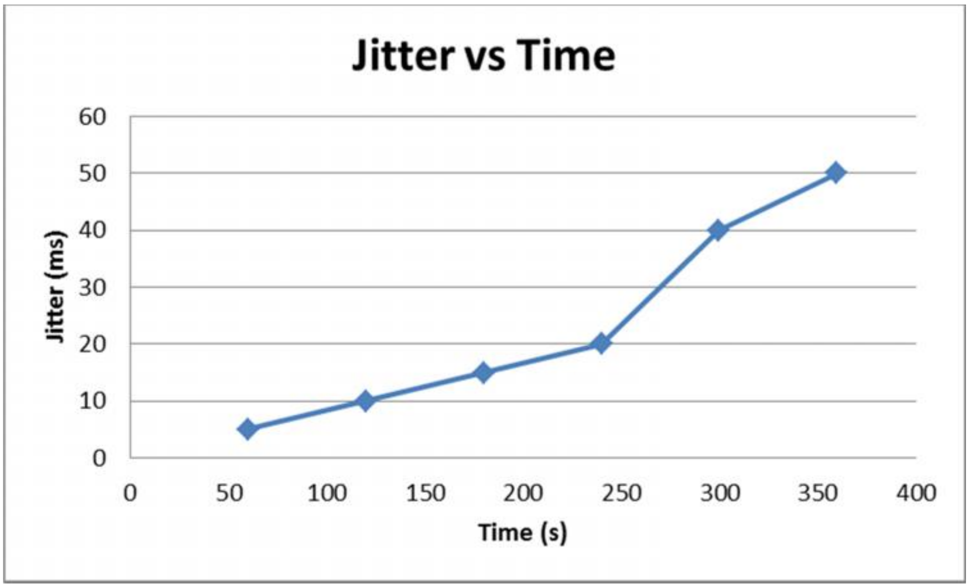


Figure 14. Line Graph showing Jitter over time in a VoIP system.

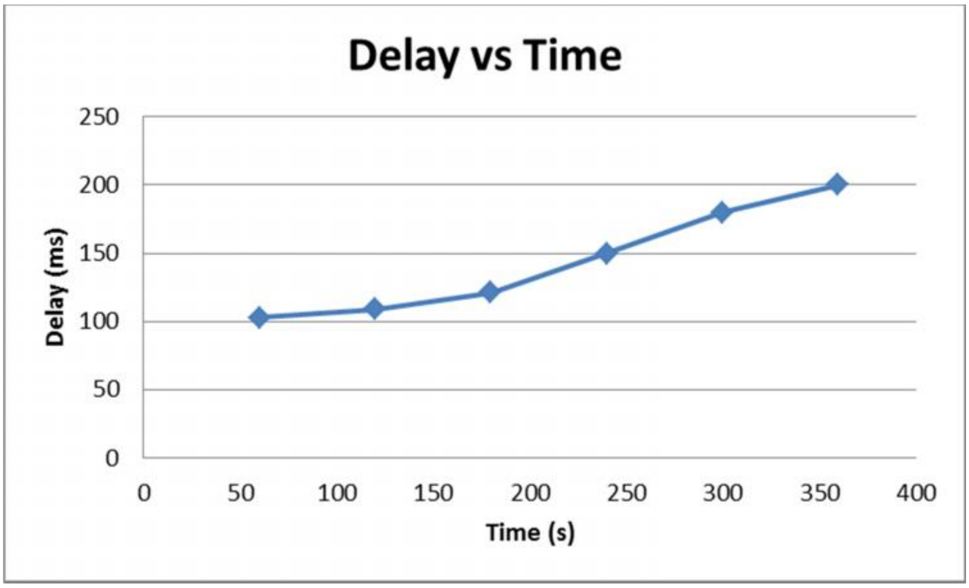


Figure 15. Line Graph showing the Delay in a VoIP system over time.

During the analysis, as Jitter and Delay impairment increases simultaneously in the VoIP System, the quality of the voice call started degrading until the conversation becomes inaudible, at this time the Delay in the system had exceeded about 180 ms. This observation reinforces the fact that Delay and Jitter in a VoIP system must be kept at the barest minimum level in order not to compromise the voice quality.

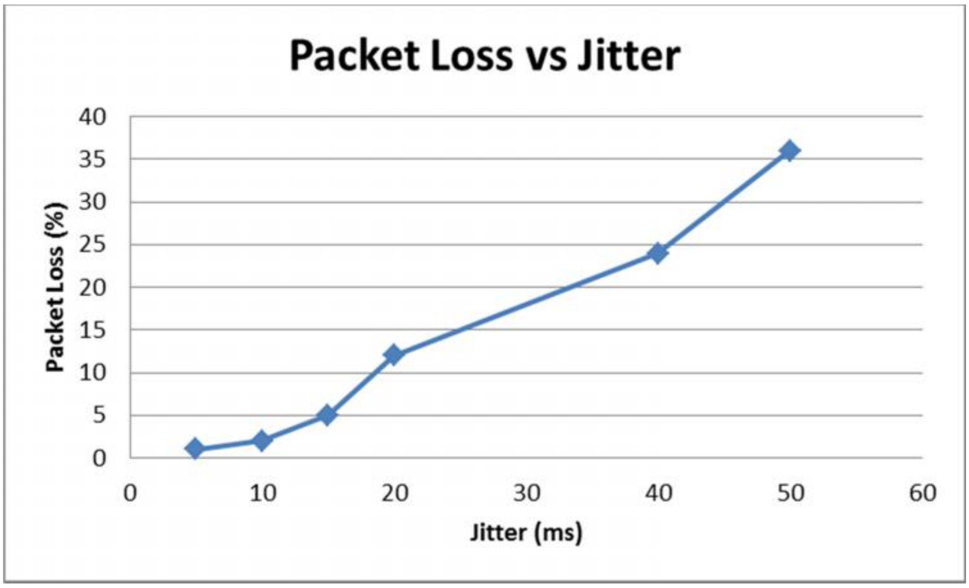


Figure 16. Line Graph showing the relationship between Packet Loss and Jitter.

Looking at the graph in Figure 16 above, it is crystal clear that Jitter has a direct relationship with Packet Loss as both factors increases linearly on the graph. During the test to determine the relationship between Jitter and Packet Loss and the effect of both on voice quality, the amount of Jitter was deliberately increased in the VoIP system which has a linear correlation with Delay and some percentage of voice packets sent from PC-A to PC-B in the VoIP setup were deliberately dropped. Initially, the voice quality was not impacted but as Jitter increases to some level, packets were dropped without deliberately dropping the packets and voice quality was severely degraded. Hence, excessive Jitter causes Delay and too much Delay leads to Packet Loss.

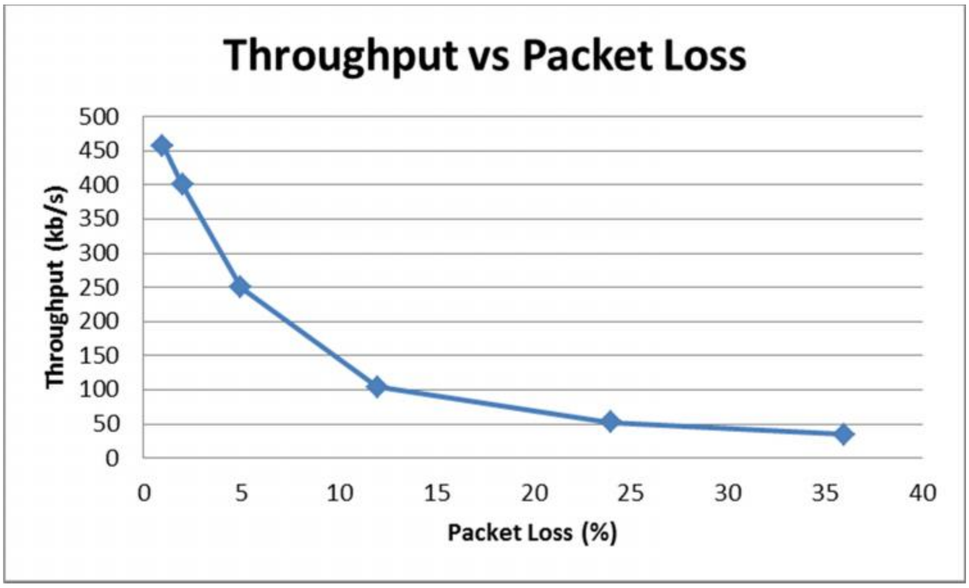


Figure 17. Line Graph showing the relationship between Packet Loss and Throughput.

Examining the graph above in Figure 17, which shows the relationship between Packet Loss and Throughput. This graph unarguably shows the correlation between Throughput, Jitter and Delay as well, since it had been earlier established that Packet Loss has a linear relationship with these factors (Jitter and Delay). It can be seen that Packet Loss has a negative relationship with Throughput which also means that Delay and Jitter have the same relationship as Packet Loss with Throughput since their relationship with Packet Loss is positive. What this means invariably is that, as the Packet Loss and Delay within a VoIP network increases the Throughput or bandwidth decreases and when bandwidth decreases voice call quality over packet networks degrades as it was confirmed during the lab test.

7.3 Limitations

The limitations in this thesis were the lack of hardware to analyze how different *codecs* affect voice quality as well as *echo loss* in packet networks. Voice activity detection (VAD) feature is not available as well in the X-Lite softphone used.

7.4 Suggestions

Relying on the results obtained from the various tests conducted in this thesis, it can be suggested that to obtain an optimal performance for VoIP system or minimizing the effects of the factors that affect voice quality performance in Voice over IP communication which had been extensively described in Chapter 4 of this thesis, the following measures can be taken.

I. A Jitter or a Delay buffer of appropriate size can be used to mask the out-of-order behaviour of the voice packet so that phone conversation can flow as if there were no Jitter or Delay in the system. However, it should be understood that a Jitter/Delay buffer can only mask mild jitter or delay as the buffer can be overwhelmed if the Jitter or Delay becomes severe.

II. Packet loss concealment is another way of minimizing the effect of packet loss. This method can be used to conceal the packet loss during a gap period in phone conversation.

III. Bandwidth should be increased as this will lessen the rate of Packet Loss, Jitter and Delay in VoIP system because bandwidth has an inverse relationship with these other factors.

7.5 Conclusion

For some reasons internet behaviour is unpredictable in comparison with a LAN which has more control as to the route through which packets traverse to get to their destinations, bandwidth and other network parameters. It is expected that due to lack of governing and central authority administering how packets are routed or get to their destination, different packets traverse different networks to arrive at their destinations. Some paths are shorter than others. These paths are paths with very good feasible distances and some are long which makes packets belonging to the same data arrive at different times, unnecessarily delayed or sometimes resulting in partial or complete packet loss.

Therefore, it can be reiterated that the various factors that affect the voice quality in Voice over IP network that had been dwelt so much upon in this thesis cannot be totally eliminated. However the effects of these factors can be lessened by the various methods suggested above.

REFERENCES

[1] Wallingford, T. (2005). Voice and data: Two separate worlds? In K. M. Loukidesl (Ed.), *Switching to VoIP* (First Edition ed., pp. 1). 1005 Gravenstein Highway North, Sebastopol, CA 95472: O'Reilly Media, Inc.

[2] BPSD, L. (2011). *Voice over IP.* Retrieved November 3, 2011, from http://www.businessphonesystemsdirect.co.uk/Voice_over_IP

[3] Davidson, J. (2007). *Voice over IP fundamentals* (2nd ed.). Indianapolis, Ind.,London: Cisco; Pearson Education distributor.

[4] Zavareh, K. (2006). Voice over Internet Protocol. Bachelor's thesis.Turku: Turku University of Applied Sciences.

[5] Jan, J., Vleeschauwer, D. D., Windey, R., Petit, H. G., & Leroy, J. (2002). *Delay bounds for Voice over IP calls transported over satellite access networks. Mobile Networks and Applications, 7*(1), 80.

[6] Wallingford, T. (2005). VoIP signaling protocols. In K. M. Loukidesl (Ed.), *Switching to VoIP* (first edition ed., pp. 131). 1005 Gravenstein Highway North, Sebastopol, CA 95472: O'Reilly Media Inc.

[7] Wallace, K., CCNP. (2005). *Voice over IP first-step*. Indianapolis, Ind.: Cisco.

[8] Held, G. (1998). *Voice over data networks*. New York, London: McGraw-Hill.

[9] Khasnabish, B. (2003). *Implementing Voice over IP*. Hoboken, N.J. ; Chichester: Wiley.

[10] *Deploying Cisco Voice over IP solutions*(2002). In Davidson J., Fox T. (Eds.), . Indianapolis, Ind.; London: Cisco; Pearson Education.

[11] Keagy, S. (Ed.). (August 2003).

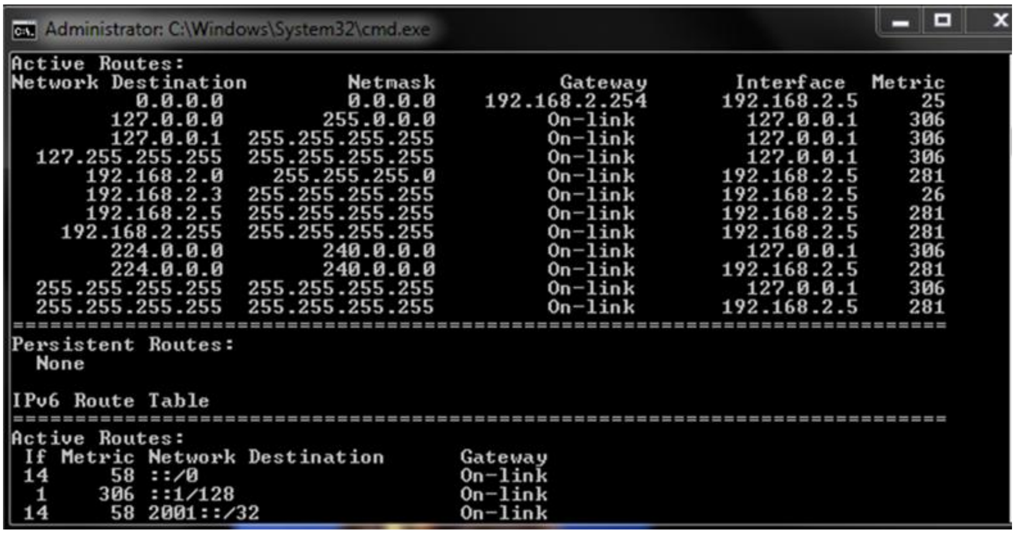
Integrating voice and data networks (Fourth edition ed.). 800 East 96th Street, Indianapolis,IN 46240: CiscoPress.

[12] ZTI, T. (2010). *IP software testing tools* (Enhanced Version 5) [Computer-Application]. 22302 Lannion Cedex, France.

Available from http://www.zti-telecom.com/

[13] Madsen, L., Meggelen, V. J. & Bryant, R. (2011). *Asterisk: The definitive guide.* Retrieved October 30, 2011, from http://www.asteriskdocs.org/en/3rd_Edition/asterisk-book-html/asterisk-book.html#asterisk-UnderstandingTelephony-FIG-1

Appendix



```
Active Routes:
Network Destination        Netmask          Gateway       Interface  Metric
          0.0.0.0          0.0.0.0    192.168.2.254      192.168.2.5     25
        127.0.0.0        255.0.0.0         On-link         127.0.0.1    306
        127.0.0.1  255.255.255.255         On-link         127.0.0.1    306
  127.255.255.255  255.255.255.255         On-link         127.0.0.1    306
      192.168.2.0    255.255.255.0         On-link       192.168.2.5    281
      192.168.2.3  255.255.255.255         On-link       192.168.2.5     26
      192.168.2.5  255.255.255.255         On-link       192.168.2.5    281
    192.168.2.255  255.255.255.255         On-link       192.168.2.5    281
        224.0.0.0        240.0.0.0         On-link         127.0.0.1    306
        224.0.0.0        240.0.0.0         On-link       192.168.2.5    281
  255.255.255.255  255.255.255.255         On-link         127.0.0.1    306
  255.255.255.255  255.255.255.255         On-link       192.168.2.5    281
===========================================================================
Persistent Routes:
  None

IPv6 Route Table
===========================================================================
Active Routes:
 If Metric Network Destination      Gateway
 14     58 ::/0                     On-link
  1    306 ::1/128                  On-link
 14     58 2001::/32                On-link
```

NetDisturb PC routing table.