Chalmers University of Technology

Master Thesis Report

Verification of Voice over IP

(In Network Integration Solution area)

By

Ibrahim Qazzaz

Supervisor

Lars Romin

Examiner

Bengt Wedelin

April 1999

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Content

Foreword

Introduction

Chapter One

Project Assignment

Chapter Two

IP

Chapter Three

What is VoIP

Chapter Four

Approaches of VoIP

Chapter Five

System Description

Chapter Six

Proposed Tests for Integrating VoIP with BC10 PBX

Chapter Seven

Transacted Tests

Abbreviations

References

Annexes (A-E)

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Foreword

This report was written as a prerequisite for the Master degree in Digital Communications and Technology at Chalmers University of Technology – Gothenburg / Sweden. The project was supported and done in Ericsson Business Network AB in Stockholm.
Verification of VoIP meant for me transacting a subjective test and that is what I did.
My conclusion is that VoIP is very promising, but many improvements should have taken place before have it commercially released. Moreover, it will be an attractive choice if a VG is built in a PBX as an integrated part.

Introduction

It is a hot topic. VoIP –according to its hardliners- will alter the used techniques for telephony. VoIP is not very new from theoretical point view, at least in the sense of merging data and voice. VoIP saves both cost and bandwidth, and again for both operators and end users.
This report does not presume specialized background of readers, and hence after first chapter which is the proposal of the thesis, three chapters –could be skipped by experienced reader- were written to clarify this “jargon”
Chapter 5 describes briefly Ericsson’s system of VoIP and relates to ITU-T recommendation H.323. While chapter 6 is a report that was handed to project supervisor as a preliminary idea for tests to be done, chapter 7 is touches the transacted tests and analysis.
Footnotes are used extensively to avoid any expected ambiguity or misinterpretation. The [ ] sign is used to denote a reference number.
A list of abbreviations and another for references are given before few useful annexes that include forms, questions and announcement messages that were used in the subjective test.

Finally, I would like to express my deep appreciation to many people in Ericsson Business Network AB, Ericsson Radio Access AB, Ericsson Telecom AB, and Ericsson Software (Erisoft) AB whom without their supports and advice this work would not see light. For any comment, please write to email: Ibrahim.Qazzaz@etx.ericsson.se.

Chapter One

Project Assignment:[1]

Verification of Voice over IP (In the Network Integration Solution area)

1.1 Background

The most recent approach to the market is going from selling products to selling solutions.

In order to build confidence into a solution, verification of the solution is needed. Integration tests of all products included, needs to be performed. Earlier products have only been tested separately and not together with other products, used in typical customer environments.

The aim with solution verification will be to shorten the time to market and increase sales. Another important aim is to get the Local Companies' and increase customers' confidence in Ericsson products. They should also feel that our products as well as partner products are of high quality, integration tested and are working well with each other as a solution. All work must be documented.

1.2 Problem

The Network Integration Solution area is one critical customer solution area. These solutions generally involve products from many different vendors taking different roles in the network, such as switches, transmission products and gateways.

It is difficult to specify how to verify the solutions. The verification should determine if the solutions have the high reliability and the high capacity expected.

It is also important to determine the results of the verification and what parts should be presented to the customer.

1.3 Objectives

The master thesis project has the following goals:

[1] By Lars Romin (project supervisor).

1. To set up criteria and find methods for measuring the load on the IP Network generated by the voice over IP applications, which are included in the Network Integration solution area using MD110[2] release BC10 and the new IP Gateway product. These criteria and methods should be the basis for future verifications.
2. To set up criteria and find methods for measuring voice transmission quality when transferred over a loaded IP Network.
3. To document the Network Integration solutions to be verified. This documentation must include specifications of all the hardware and software in the test scenarios. Technical descriptions of the servers, release information about all products and the Network Integration related MD110 configurations etc. should also be included. The results from the measurements should also be part of the document.
4. To define scenarios within the Network Integration Solution area that are to be verified.
5. To perform verification and integration tests in the defined scenarios while measuring the IP Network traffic.
6. To document the possible limitations of different methods and scenarios.

1.4 Plan and Follow up

A detailed time schedule and plan regarding pre-study, literature studies needed and course of action shall be documented.

1.5 Equipment

In order to perform measurements, the Network Integration Solution Verification laboratory shall be used.
The laboratory must be equipped with MD110 and IP Gateway products of the latest releases as well as hubs and routers.

1.6 Scope and Limitations

No measurements have to be performed on the MD110, only measurements on the IP Network will be performed. No function tests on products will be performed.

Assistance in building up the laboratory will be provided.

Defining the solution verification process is not an aim of this project.

[2] The commercial name for the Private Branch Exchange (PBX) of Ericsson.

Chapter Two

IP

2.1 IP and Internet

Internet Protocol[3] or IP is the most common data communication protocol. It is the underpinnings of Internet, which witnesses the ever vast booming. Originally, IP is a part of a suit of protocols built by Department of Defense (DoD) Advanced Research Projects Agency (DARPA) Internet Community, better known as TCP/IP[4].
The Internet became widely known and talked about in the early to mid 1990s. It is a world - wide collection of interlinked (hence the word **Inter** is emanated) wide-area networks, with associated local area networks. In the 1970s it was better known as the ARPANET[5], when it was the first network to establish the viability of wide-area computer communication. It later became known as the Darpanet communication. Despite it became known as the Darpanet, the term ***Internet*** is preferred today. Up to the mid-1990s it was largely a research, military and educational network, with a very limited and restricted amount of commercial traffic over it. This situation changed dramatically in the middle of 1990s with the growth of interest in the provision of World Wide Web pages. The use of Transmission Control Protocol (TCP) and Internet Protocol (IP), but particularly the latter characterize communication over the Internet, with a variety of other protocols on top. All the protocols in the suite are generally collectively known as "the TCP/IP protocols" (even if they do not actually use TCP), or more accurately as "the Internet protocols".
Originally, TCP/IP was very much aimed at wide-area networking, but its adoption in the early 1980s by the UNIX developers led to its wide spread use in the late 1980s on local area networks.
As a network layer protocol suite, IP is widely used on Ethernet networks.
IP was defined in RFC[6] 791; the current version of IP is IPv4. A new version, called IPv6 or IPng is under development.
The Internet protocols are vendor independent, and their specification is controlled by open public discussion, that is not dominated by any specific vendor. They are also widely implemented by a variety of vendors and hence

[3] Protocol is a set of formal rules describing how to transmit data, especially across a network. Low level protocols define the electrical and physical standards to be observed, bit- and byte-ordering and the transmission and error detection and correction of the bit stream. High level protocols deal with the data formatting, including the syntax of messages, the terminal to computer dialogue, character sets, sequencing of messages etc.
Many protocols are defined by RFCs or by OSI.

[4] The connection to TCP will be reached soon.

[5] Or network of ARPA.

[6] RFCs are Comments submitted to IETF for the purpose of discussing, setting and ratifying evolution of Internet.

fulfil the definition of "open" as used in OSI. It is remarkable to know that there is no formal document specifying the architecture of the TCP/IP-related (Internet) specifications, so there can be different views by different authors on the de facto architecture, and particularly on its relationship to the OSI architecture.

2.2 Ethernet Frame

Because of close relevance between IP and Ethernet, it might be reasonable to show the composition of an Ethernet frame, as illustrated in figure 2-1 below.

Preamble	Destination Address	Source Address	Length	Data	CRC
8 bytes	6	6	2	0 - 1500	4

Figure (2-1)
Ethernet Frame

The preamble is a predefined sequence of 1s and 0s used for synchronization. CRC is for checking.

2.3 OSI and IP

OSI is a model of network architecture and a suite of protocols (a protocol stack) to implement it, developed by ISO in 1978 as a framework for international standards in heterogeneous computer network architecture. The OSI architecture is split between seven layers (up direction).

1. Physical.
2. Data link.
3. Network.
4. Transmission.
5. Session.
6. Presentation.
7. Application.

This is shown in figure (2- 2) which also reveals the connection to Internet protocols suit. Each layer uses the layer immediately below it and provides a service to the layer above. In some implementations a layer may itself be composed of sub-layers.

Today, when people talk about "open networking", they can mean implementation of either OSI or TCP/IP. Marketing brochures need to be examined carefully to see which is meant.
(Whilst all the early TCP/IP protocols have equivalent or better OSI equivalents, there is no OSI equivalent for the protocol underlying the World-Wide Web, although the markup language used to author pages (HTML - Hyper-Text Markup Language) is based on an ISO Standard (SGML - Standard Generalized Markup Language).

How is the architecture of TCP/IP?

It has strong similarities with parts of OSI, but is broadly much simpler. As with OSI, there is an end-to-end network service provided by the use of IP, with network switches understanding only the IP protocol and associated routing and management protocols. Beneath the IP layer there is whatever real

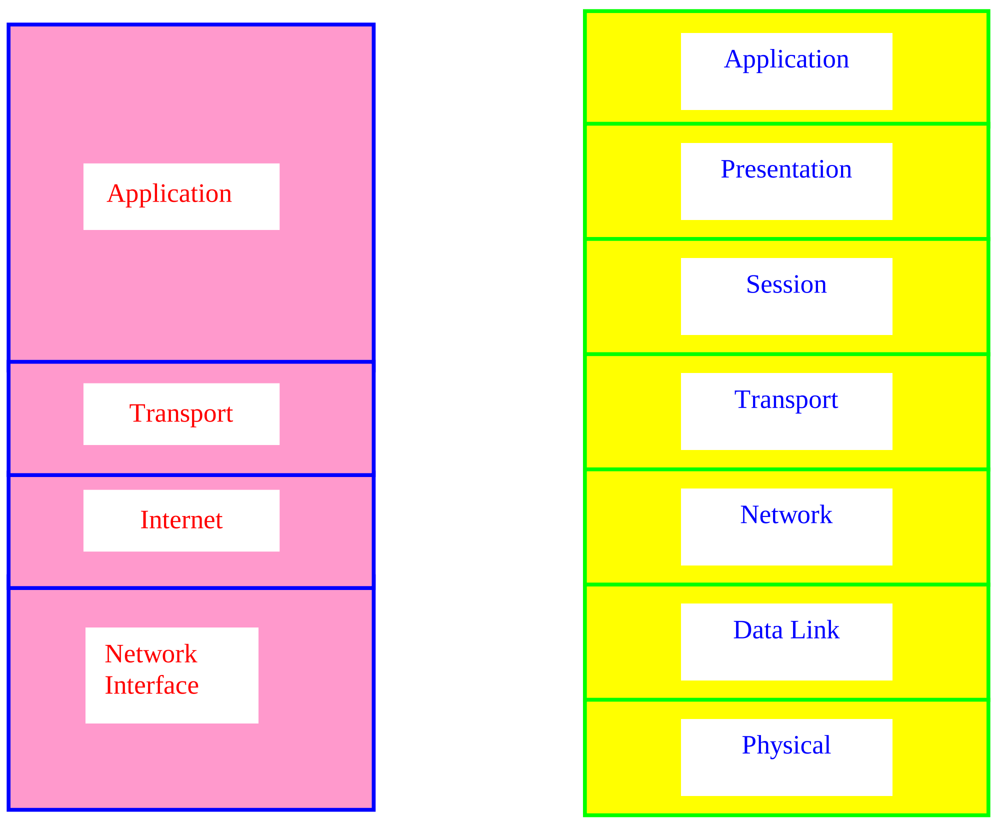

Figure (2-2): IP vs. OSI

networks are around, and there is a series of specifications (Internet specifications are called, somewhat misleadingly, Requests for Comment RFCs[7] that specify how to transmit IP messages over a whole range of real-

[7] Was pointed to previously.

world networks, including some vendor-specific ones). This part of the architecture then, is very similar to the OSI "Internal Organization of the Network Layer" described earlier. Above the IP layer, there is a layer corresponding quite closely to the Transport Layer of OSI, and containing TCP or another (very simple) protocol called User Datagram Protocol (UDP) as we will see later. On top of these there sit monolithic specifications for applications.

The main difference between the OSI and the TCP/IP architectures is that in TCP/IP the Session and Presentation Layer functionality is not factorized into separate specifications, nor are application specifications normally broken down into a set of Application Service Elements or ASEs[8].
It should be pointed out that, Network Interface layer is also referred to as *Link Layer (LL).*

2.4 Mechanism: Packet Switching

Stream of data (in the application layer) is fragmented into sets of bits, widely known as packets. Encapsulation[9] of packets takes place: Header and trailer information is added in this stage to packets in order to form what is called frames.
To state the more accurate notation [24]:

* **Packet,** is the data unit that pass through the interface between Internet layer and the Network Interface layer, which means it includes IP header plus data.
* **IP datagram** (or simply datagram) is the end- to – end unit of transmission in IP. Therefore, an IP datagram comprises IP header followed by transport Layer data. Datagrams are self-contained independent entity of data carrying sufficient information to be routed from the source to the destination computer without reliance on earlier exchanges between this source and destination computers and the transporting network (connectionless[10]). On that basis, it could be said a packet might be an IP datagram, or a fragment of an IP datagram.
* **Frame**, is the unit of transmission in the Network Interface Layer, and is hence composed of the header of this Layer followed by a packet.

---

8 One can discern some elements of the ASE concept in the TCP/IP TELNET protocol, which is used both as an actual application (terminal login) and also to support other application protocols (file transfer and electronic mail). With this exception, however, Internet specifications above TCP or UDP tend to be self-contained.

9 Encapsulation: The technique used by layered protocols in which a layer adds header information to the protocol data unit (PDU) from the layer above. As an example, in Internet terminology, a packet would contain a header from the Network Interface Layer, followed by a header from the Network Layer (IP), followed by a header from the Transport Layer (TCP), followed by the application Protocol data.

10 The data communication method in which communication occurs between hosts with no previous setup.

*** Message**, is the unit of transmission in a Transport Layer protocol. A message consists of Transport protocol header followed by application data. For TCP a message is called *segment.*
In order to be transmitted end - to - end a message should be encapsulated inside a datagram.

The terms Frame, Packet, Datagram and Message are illustrated by figure (2-3) [24].

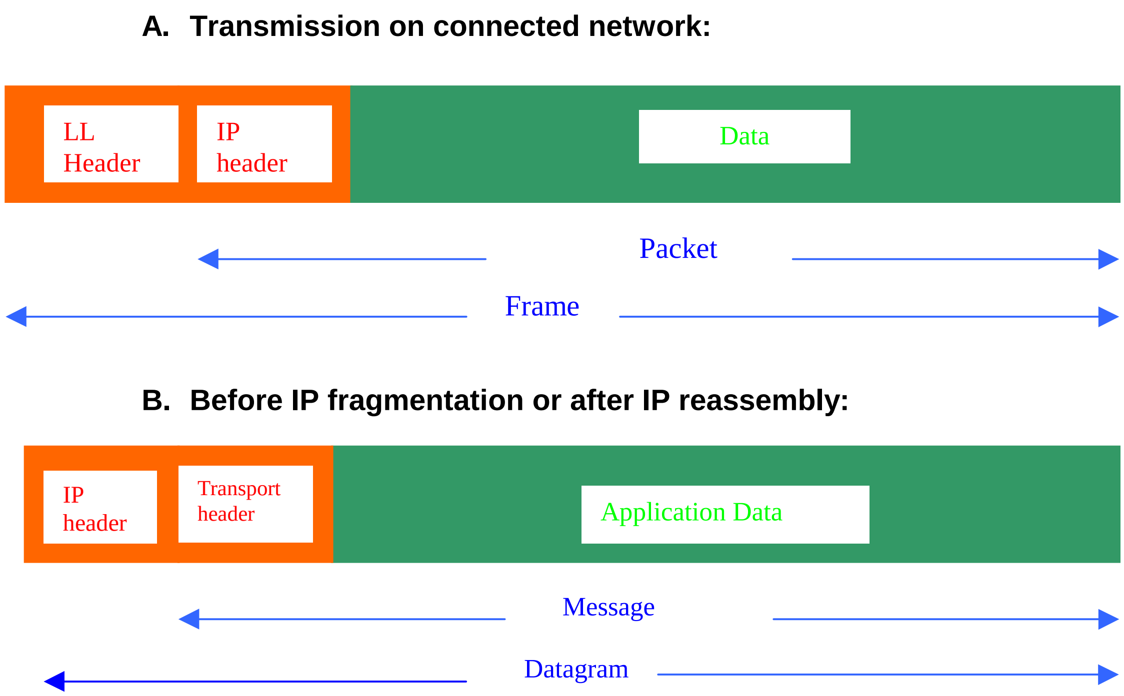

Figure (2-3): Headers and Payloads

Routing is transacted by bridging devices called Routers. Router is a device connecting two different LANs together.
Figure (2-4) depicts this connection.

Transmission or forwarding Datagrams from source to destination is called packet switching. Each packet (or frame) is transmitted individually from one node to another to destination as shown in figure (2-5). Packet can even follow different routes to its destination. Packets within the same burst of data (message or web site) and from same source going to same destination could follow completely different routes.
Once all the packets forming a message arrive at the destination, they are reassembled and recompiled into the original message.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

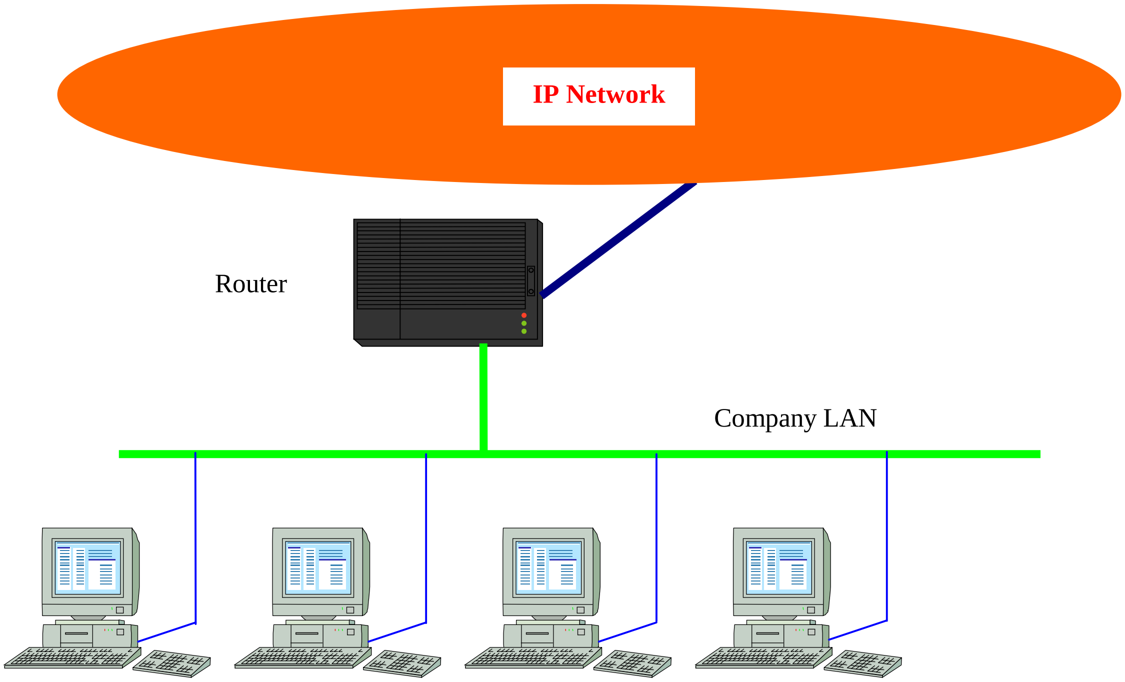

Figure (2 – 4)
LAN connection to Internet

One can apprehend two significant features of packet switching:
1. Since packets are short, the communication links between the nodes are only allocated to transferring a single message for a short period of time while transmitting each packet. Longer messages require a series of packets to be

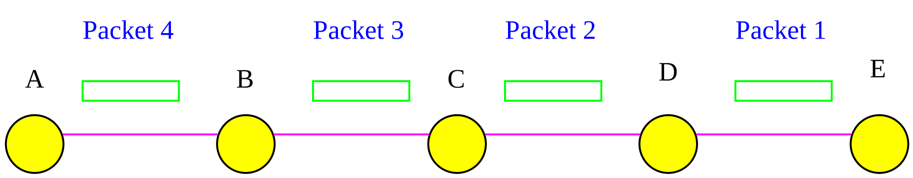

Figure (2-5)
Node to node connection

sent, but do not require a link to be dedicated between the transmission of each packet.
The implication, is that packets belonging to other messages may be sent between the packets of the message being sent from A to E. This provides a much fairer sharing of the resources of each of the links.
2. Assuming a long message is transmitted from A to E in figure (2-5) above; at the time packet 1 is sent from B to C, packet 2 is sent from A to B; packet 1 is

sent from C to D while packet 2 is sent from B to C, and packet 3 is sent from A to B, and so forth. This simultaneous use of communications links represents a gain in efficiency and is denoted as *Pipelining*.

2.5 Higher Level Protocols:

Two protocols are of interest.

2.5.1 TCP: TCP is built on top of Internet Protocol (IP) and is nearly always seen in the combination TCP/IP (TCP over IP). It adds reliable, flow-control, multiplexing and connection-oriented communication. The level of this protocol is at Transmission (and Session); whereas the IP protocol deals only with packets, TCP enables two hosts to establish a connection and exchange streams of data. That is, a connection must be established (using the “connect and accept” functions) before communication begins. This is similar to a telephone call; caller can not just pick up the phone and start talking (Of course he you can, but he would be talking to a dial tone which some would consider a waste of time).
TCP guarantees delivery of data and also guarantees that packets will be delivered in the same order in which they were sent. The data will be received as a stream of bytes--the actual packaging (the number and boundaries of packets of data) of the bytes received may be different from how it was sent.

To clarify last discussion: if a sending process transmits a packet with 100 bytes of data, it may be received as a 100-byte packet or as a 75-byte packet followed by a 25-byte packet.
In any case, the receiving program does not see the packet boundaries. It simply asks (its TCP stack) for a specified number of bytes from the input queue, and up to that number of bytes is given to the receiving program, regardless of whether those bytes happen to span more than one incoming packet or not. The application protocol must therefore have a protocol for identifying the boundaries of the data structures exchanged (since the TCP protocol does not provide this service).
Finally, TCP is the most common transport layer protocol used on Ethernet and the Internet. It was developed by DARPA and defined in RFC 793.
To summarize: It provides full-duplex, process-to-process connections. It is connection-oriented[11] and stream-oriented, as opposed to UDP.

2.5.2 UDP: UDP is an Internet standard (similar to transport layer and session layer protocols in ISO) which provides simple but unreliable Datagram services. That is, packets may be duplicated,

---

[11] It is called, therefore virtual circuit.

lost, or received in a different order than the one in which they were sent.
The receiving program requests a number of bytes (up to the number in the received packet). If less that the full packet is read, then the remainder is discarded, and the next is read from the next packet. In other words, the boundaries of the original packet are preserved (which is different than for TCP).
UDP is defined in RFC 768. It adds a checksum and additional process-to-process addressing information .UDP is a connectionless protocol which, like TCP, is layered on top of IP. UDP neither guarantees delivery, nor does it require a connection. As a result it is lightweight and efficient for application that can tolerate some percentage of packet loss, hence the application program must take care of all error processing and retransmission.

UDP is best suited for small, independent requests, such as requesting a MIB value from an SNMP agent, in which first setting up a connection would take more time than sending the data. Less efficient than TCP, since the full address must be sent with each packet.

2.6 Discussion

As such, IP is something like the postal system. It allows addressing a package and dropping it in the system, but there is no direct link between sender and the recipient. TCP/IP, on the other hand, establishes a connection between two hosts so that they can send messages back and forth for a period of time.

To pave the way to talking about VoIP, it is concluded that we have to compromise between reliability and speed. Bearing in mind that packet switching is slow (may be slowest data communication), and that human ear can tolerate some lost packets more than late speech (late speech could be puzzling, it leads to impulsive interruption and may be misunderstanding), we have no choice but to opt speed. Now it is obvious why UDP was chosen for Internet telephony.

Chapter Three

What is VoIP?

3.1 First: What is the normal (circuit switching) call?

Spectrum of human voice can extend up to a maximum of 8kHz [12]. According to Nyquist theory, such an analog signal should be sampled at double of this frequency in order to be able to reassemble it again. It was found however, that components of human voice are sharply weakening after 3.3 kHz. So it was reasonable to confine sampled signal to that limit. The sampling frequency was chosen to be slightly larger than double of that bound.
Samples are quantized using 8 bits per sample scale.
Consequently, when somebody tries to make a telephone call in the traditional way his voice is sampled at a frequency of 8000 kHz and transmitted at a bit rate of 64000 bits per second (or 64 kbps). This figure is very fundamental in the world of telephony.

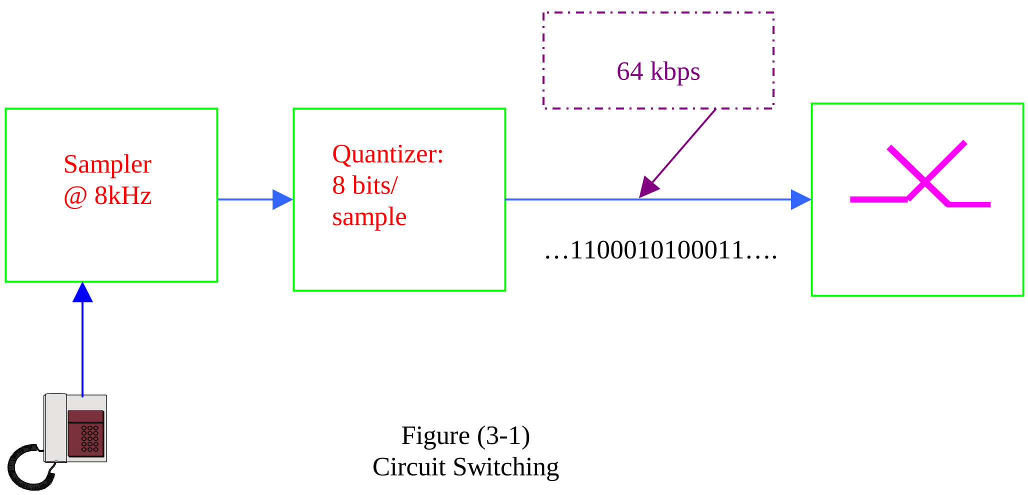

Figure (3-1)
Circuit Switching

The voice is converted to a series of bits that passes from one node (switch) to another until they reach their final destination, wherein bits are transformed again in a counterpart process into audio signal.

3.2 Then what is the VoIP in simplest manner?

VoIP is not completely different!
Plainly, voice is digitized, packetized and then transmitted over IP network. The difference is concealed mainly in transmitting network and transmission format.

[12] Less than this figure for males, while music can go much beyond this frequency.

In the case of circuit switching call, a physical end - to - end channel is occupied as long as the connection (call) is on. Even moments of silence are sampled, quantized and transmitted.

As we have seen in the previous chapter, IP is another story. Performance is not guaranteed: a packet may be received after 100 ms or 10000 ms. Furthermore, since packets could be received of out sequence, real time[13] conversation becomes tricky.
Because of importance of minimizing the time delay, TCP is avoided at the transport layer and UDP is used instead. *The mediator between PSTN and IP network is a device called Voice Gateway or simply Gateway.* Gateways as well as other system ingredients will be discussed later. Figure (3-2) clarifies this question.

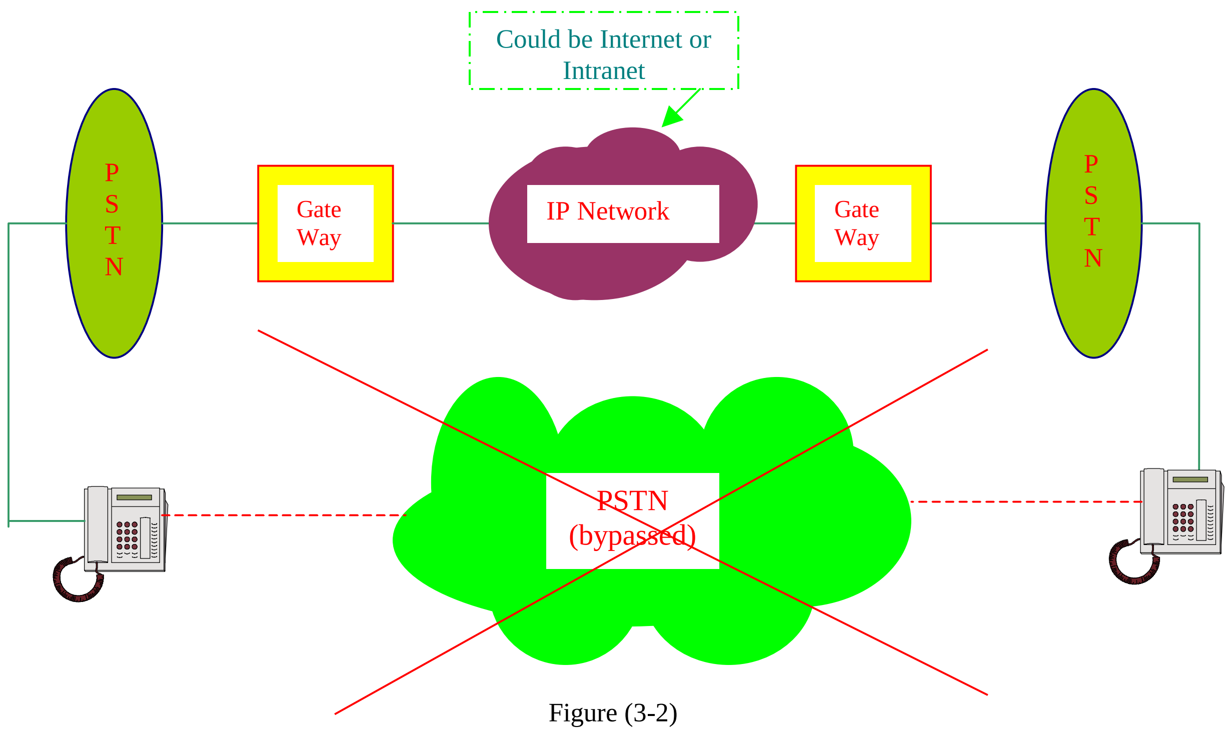

Figure (3-2)
Voice gateway as an interface between IP and PSTN

The following problems are anticipated to be faced when using IP for telephony:

1. Latency (Packet Delay) and bandwidth limitation: Things are getting worse with subscription to small ISPs who oversubscribe their limited access ports. As discusses above usage of UDP could be one part of a solution.

[13] Real-time: Describes an application, which requires a program to respond to stimuli within some small upper limit of response time (typically milli - or microseconds). Used to describe a system that must guarantee a response to an external event within a given time [22].

The complementary part is to compress audio signal from 64 kbps into few kbps before transmission.

How is that done?[14]

We could think of several ways:

a. Silence Cancellation: It is estimated that 60% of any telephone call is nothing but silence, albeit circuit-switching devices keep the process of sampling and transmission running. Detecting moments of silence and pausing transmission during those occasions imply less bit rate and hence less delay.

b. Making use of strong correlation between successive data, next information can be always predicted with some percentage of error. Setting a predictor on both transmitter and receiver, errors (difference between predicted and correct information) can be sent rather than information as such.
Transmitted errors are smaller in magnitude than original data, which means again less needed bits, less bit rates and eventually less delay.

c. Since 8 bits are used to represent each sample, it is sensible to prioritize importance of different bits. MSBs affects intactness of the samples much more than LSBs; loosing first three LSBs will not be easily noticed.
Assigning highest transmission priority to most three bits and allowing LSBs to arrive bit later (even discard them if they arrive later than some predefined threshold) is the second trick to accelerate reception of audio.

Another side of this story is bandwidth saving[15], While a maximum of 30 simultaneous calls per E1 can take place, this number reaches 166 simultaneous calls with VoIP (assuming 12 kbps for each call[16]).

2. Jitter[17] (out of sequence reception of packets): Buffering number of packets before reassembling audio signal back is a common technique. Despite this solution divert back to the problem of delay, it is important to use it in order to have acceptable voice quality.
Here, we should compromise between quality and delay. Experiments were done to find the good threshold for size of the buffer.[18]

3. Loss of packets: Which could stem from real loss of packets or discarded delayed packets. In both cases, result is the same; impairment of voice quality as a result of loosing some information.

---

[14] ITU-T recommendations G.722 – G.729 [10-15] tackle these techniques.
[15] 24 vs.128 for T1.
[16] Could be much less than this number as we will see later.
[17] Jitter: Random variation in the timing of a signal, especially a clock[22]
[18] In release 1.6 of Ericsson's VG, it was chosen to be 2 frames or 80 ms.

Theoretically, 1 – 3% of packet loss can be forgiven by human ear. The more packets lost the more deterioration in voice quality felt.

3.3 Why VoIP?

It is a circle like totem pole: whereas today we use voice network for transmitting data (at least for personal use), it is going to be the other way around.
The question could be: *why IP not the more convenient choice: the ATM*[19]?
First, let us have a closer view of ATM.
ATM is a method for the dynamic allocation of bandwidth using a fixed-size packet (called a cell). Each cell contains 48-byte (plus 5 bytes of overhead). ATM is therefore called Cell-Switching scheme. Because of its high speeds, it is suitable for data as well as digitized voice and video.
It is flexible; called asynchronous, because each cell can be independently addressed to allocate bandwidth between many virtual channels as needed.
It is interesting to know that it eliminates the distinction between LAN and WAN, since ATM can be used for both. ATM transmission media could be any of:

1. 25.6 Mbps (for Token Ring).
2. T3[20]
3. 100 Mbps FDDI
4. SDH: STM 1: 155.52 Mbits/s.

Now, back to the previous question: Why IP?
The fact of the matter, pervasiveness of Internet abetted to try VoIP. On the other hand, commerce and bypassing services of traditional telephone operators could be the raison d'être of VoIP. Internet does not recognize geographic boundaries. When a web site is to be visited or an email is to be sent, it does not make any difference for the user if the server of the other end was within the same town or in another continent. That it to say, unlike the case of normal telephony, wherein callers have to add country code and/ or area code, Internet users do not have to dial any national or international code for specific site except others. In most countries, the cost of a connection to the Internet (per minutes) is only few hundredths of the cost of an international call, thereupon it was intuitive to endeavor a means to load voice over the Internet. The technology to load voice over the Internet was not difficult to build; especially everything is migrating to be software controlled. Sound cards have been used long time ago for multimedia application; therefore starting point was not primeval.
First attempts were for personal use by hobbyists, but the push of cost saving lead to business use.
Still the problems of delay and voice quality coexist. While delay for a typical domestic call is 50 to 70 ms (and 150 to 500 ms for transoceanic calls), Internet delay can range from 400 ms to 2 seconds, well above the 250 ms

---

[19] In fact ATM is one of IP data carriers.
[20] Bit rate is 44.736 Mbits/s

level considered noticeable [10]. Nevertheless many ITSPs claim and guarantee delays of 250 to 400 ms.
Now it is what to opt: high quality at a high price or less quality at a more reasonable price.
That is not everything: Evangelists of IP telephony say:” It is over, everything we should do is finished; standards and technology, but when Internet is ready we are ready”.
The next anticipated driving force is no longer arbitraging of international access; it will be Voice / Data network integration.

3.4 Near Future Solutions

Though some ITSPs have already puzzled out many of VoIP problems by building their own network backbone and are able to ensure a higher quality of services, near future will bear even more affordable solutions.

1. The most viable candidate to replace the current Internet Protocol is (IPv6, IPng: IP next generation) The primary purpose of IPv6 is to solve the problem of the shortage of IP addresses. The following features have been purposed:
   a. 16-byte addresses instead of the current four bytes.
   b. Embedded encryption, a 32 - bit Security Association ID (SAID) plus a variable length initialization vector in packet headers.
   c. User authentication (a 32-bit SAID plus variable length authentication data in headers).
   d. Autoconfiguration (currently partly handled by Dynamic Host Configuration Protocol); support for delay-sensitive traffic - a 24 bit flow ID field in headers to denote voice or video, etc.

   RFC 1550 is a white paper on IPng.
   IPng or version 6 of IP will handle the issue of differential services[21] better (prioritize voice and video to normal data: fourth jargon above), so it will be possible to minimize delay more to reach very good performance.
2. Likewise, installing more ISDN lines and running ADSL[22] technique will hasten reaching comparable efficiency of VoIP by offering higher bandwidths for end users.

---

[21] Differential service mechanism allows provider to allocate different levels of services for different users of the Internet. This issue is still a matter of debate and not easily predictable.

[22] ADSL, or Asymmetric Digital Subscriber Loop is a form of Digital Subscriber (xDSL) Line in which the bandwidth available for downstream connection is significantly larger then for upstream. Although designed to minimize the effect of crosstalk between the upstream and downstream channels, this setup is well suited for web browsing and client-server applications as well as for some emerging applications such as video on demand. The data-rate of ADSL strongly depends on --the length and quality of the line connecting the end-user to the telephone company. Typically the upstream data flow is between 16 and 640 kilobits per second while the downstream data flow is between 1.5 and 9 megabits per second. ADSL also provides a voice channel. ADSL can be configured to carry digital data, analog voice and broadcast MPEG2 video in a variety of implementations to meet customer needs.

3. Similarly, enhancing deployment of corporate networks by ITSPs has positive influence.

3.5 VoIP and markets

Talking in the same context, FCC estimates that Fortune 1000 companies each spends an average of $36 million per year on long distance telephone calls, and that half of this cost is incurred on calls between companies own office. Using organization existing data network or, Intranet, to also carry voice traffic will minimize these charges.
VoIP is pushing to impose its existence in the market.

* Market size (for computer telephony) was $3.5 million in 1995. Overall, 1998 Sales of IP Telephony Companies totaled over US$ 350 Million Dollars [23].
* In March 1998 it was estimated that there at least 15 ITSPs were offering VoIP services [23].
* Internet telephony market is expected to reach $560 million this year (1999), and $1.8 billion by 2001.
* Consultancy Philips Tarifica LTD estimates AT&T will lose $620 millions to $950 million in international calls by 2001 because of IP telephony.
* 4% of U.S. telephone company revenues will be taken by Internet telephony service providers.
* Addressing Minute wise, it is estimated that the current run rate for IP Telephony Minutes is about 50 million minutes a month of commercial international traffic on the Public Internet.
  This figure was estimated to be only 6.3 million minutes a month in December 1997, which indicates almost 800% of growth rate [23] [23].

*** Some VoIP preachers even predict that packet switching will run the entire telecommunications network within 5-10 years!

3.6 VoIP and Regulatory bodies

Seeing the voluminous danger of VoIP, traditional telephony service providers are objecting this new competitor and pushing regulation authorities to forbid using this wise or regulate it. Thus, there is a threat of banning VoIP or even imposing extra tax on using this promising fashion. Unfortunately, while some European countries have already done, laws in other countries, including Italy, Norway and Switzerland let them ban any entity other than state phone company from carrying voice traffic.
Up to now[24], FCC and some other regulators in the world do not consider VoIP as to be a telephone call, and hence it is hoped that this technology will witness substantial improvements until it can afford a satisfactory level of service.

---

23 Total number of VoIP minutes is not that large compared to the total number of minutes in the overall telecom marketplace which expressed in terms of Trillions!

24 December 1998

3.7 Standardization of VoIP

Standardization means interoperability, a keystone to incite prevalence of VoIP. Interoperability introduces some challenges, but the same time it will create new opportunities. It will ensure integration, and marketing benefits. H.323 was issued and ratified[25] by ITU-T in order to set a reference for describing equipments and multimedia communication services over LANs with no guarantee of quality. VoIP is encompassed in this jargon. Annex A is a summary of H.323 recommendation.

In the parallel line European Telecommunications Standard Institute (ETSI), started a project titled: Telecommunications and Internet Protocol Harmonization Over Network (TIPHONE) for details formulating.

Not all vendors managed to comply with H.323 specification, yet it is expected to be by the end of 1999.

Some observers expect that H.323 will not be the sole standard for VoIP.

[25] November 1996

Chapter Four

Approaches of VoIP

4.1 Software Approach

Most vendors have this methodology as one solution, because it is the easiest choice. In fact this approach was the first trial to realize conveying voice over the Internet rather than usual PSTN network[26]. In this system, the processor of the host PC is used to run a software that transact codecs, echo cancellation and call handling with the aid of the existing built–in sound card and TAPI[27] modems.

This approach could be expedient only for the case of a limited-channels system, namely Small Office/ Home Office (SOHO).

4.2 DSP Card Approach

Since one of the keystones of elevating the performance of VoIP is reducing the latency[28], this way was implemented in order to relieve the exertion that burdens processor and limits number of simultaneous call in the previous case. The substantial benefits of this approach are scalability and performance.

4.3 PBX Integrated Module

The ideal solution for business applications could be a completely self-contained card or module that, on one side connects to the PBX and on the other side provides a LAN Ethernet interface. In other words, to construct an add-in module that handles all functions (signaling, compression, etc) of the voice gateway, such that the PBX would view this gateway as a group of trunk lines.

This way facilitates for bridging many distant PBXs into a single system, whereby employees in an enterprise are able to dial extensions without caring

---

[26] By Vocaltec Inc. in February 1995.

[27] TAPI : Telephony Application Programming Interface, an API for connecting a PC running Windows to telephone services. Microsoft and Intel introduced TAPI in 1993 as the result of joint development. The standard supports connections by individual computers as well as LAN connections serving many computers. Within each connection type, TAPI defines standards for simple call control and for manipulating call content.

[28] In networking, the amount of time it takes a packet to travel from source to destination. Together, latency and bandwidth define the speed and capacity of a network.

of physical location. VPN[29] could be integrated for this approach in order to achieve needed security.
The fabulous feature in this approach is that it enables embedding a URL in enterprise's web site that allows visitors to call customer services (or other subjects in the company) via the Internet as they would do on ordinary trunk lines. As a matter of fact this is a hot issue; some companies have already released their products of IPBX[30], while others declared they would soon.
It is believed it will start to gain significant momentum during 1999.
The aforementioned discussion leads to future of the PBX. Many telecommunication consultants foretell that PBXs in their current functionality will perish. It has to be adapted to the vast stream of Internet [21].

---

[29] The use of encryption in the lower protocol layers to provide a secure connection through an otherwise insecure network, typically the Internet. VPNs are generally cheaper than real private networks using private lines but rely on having the same encryption system at both ends. The encryption may be performed by firewall software or possibly by routers [22].

[30] Until November 1998, I tracked the following vendors: Selsius and Lucent, NetPhone, Shoreline Teleworks NBX, Praxon, Calista and Touchwave.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Chapter Five

System Description

5.1 System Components:

In general, we can say the following components could exist:

5.1.1 Information Stream

Information streams for VoIP includes audio, data, communication control and call control.
Audio signal contains speech that has been digitized, compressed and packetized.
Data signal could be a fax message as well as other data streams.
Communication control signals are the control data between end Gatekeepers for capability exchange, mode control etc.
Call control signals are used for call establishment and disconnection in addition to any other control functions.
Figure (5-1) illustrates how control and voice signals merge in the same route in some stages and break up in others.

5.1.2 Terminal

For VoIP, the terminal is the telephone that can handle IP telephony.
The terminal should be able to support full bandwidth (64 kbps: according to ITU-T G.711 including A, µ laws of encoding).
According to H.323, VoIP terminal should also be capable of handling compression (and decompression) [31]according to G.722, G.723, G.728, G.729 and MPEG[32] 1 audio. This is not the case for Ericsson IPT version 1.5, whereby this burden is assigned to the voice gateway.

5.1.3 Voice Gateway (Gateway or GW in H.323)

Voice Gateway or VG is the interface between PSTN network and IP network in a transparent format. VG perform call set up and clearing on both sides. Number of PSTN connections, number of simultaneous calls, inclusion of multipoint functions and number of H.323 terminals (VoIP phones) are not standardized.
Ericsson's release 1.5 of VG has two main functions (depending on transmission direction):

---

[31] In other words, to contain a CODEC.
[32] An ISO committee that generates standards for digital video compression and audio. It also denotes for their algorithm. MPEG-1 is optimized for CD-ROM.

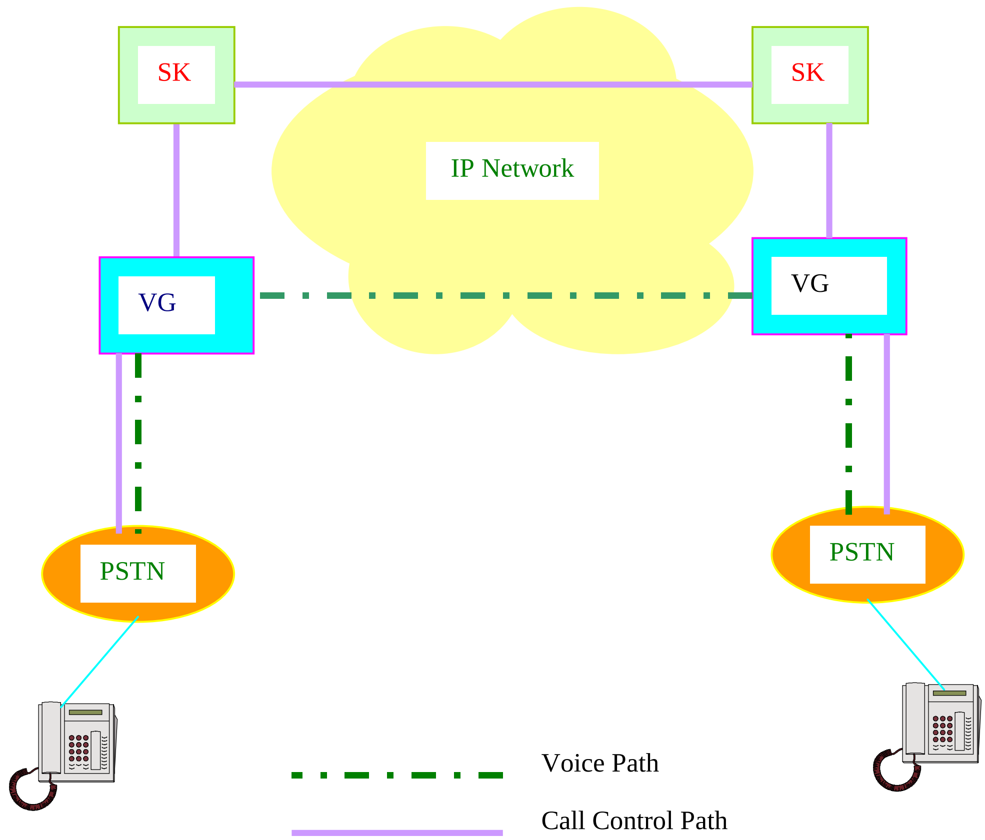

Figure (5-1) Information Stream

1. Packetizing and compressing voice samples coming from PSTN to be in a suitable format for transmission over IP.
2. Decompress and reconstruct received packets from IP network to be transmitted to PSTN network. This task include buffering packets before retransmission to ensure setting packets in the correct sequence, and discarding overdelayed packets (after a predefined time equivalent to buffer size).

Since normal PSTN calls are associated with signaling information (e.g. Q.931 and Q.2931), VG transforms them into control protocol to allow the Site keeper to manage calls.
Ericsson VG tackles the question of echo[33].
In order to allow for higher capacity, more than one VG could be deployed at the same location.

[33] Two techniques could be used:
1.Echo Cancellation: Wherein filters are used to eliminate or minimize reflected wave.
2.Echo Suppression: which implies switching off receiving channel during transmission period.
It is noted that echo suppression requires less capacity.

5.1.4 Site Keeper:

As in the case of PSTN, a switch must be installed in order to manage connexion as well as signaling. For IP telephony this task is met by a virtual switch named Site Keeper or SK. In H.323 SK is named Gatekeeper and is optional.

Restrictively for Ericsson VG, when a caller tries to initiate a VoIP call:

1. As a response to a command from SK, Billing Server- that contains a database for all valid accounts- authenticates user[34].
2. The VG requests for admission to its SK in order to minimize needed bandwidth or rejecting the call attempt in case of overload.
3. The SK authenticates the eligibility of offering service or not, according to preset criteria for this purpose.
4. SK of calling party contacts the SK that supports the called number.
5. SK of the terminating site transacts same admission checks for the call.
6. SK chooses the VG that supports called number if it was within same network, or otherwise any VG that has free channels.
7. Incoming traffic from VG is routed –according to the information in the routing tables for least cost- to the destination VG that completes the process.

NTP[35] is implemented to assure time synchronization.

In the time that same SK can support multiple of VGs that reside in the same location, a hot standby SK can also be utilized as a redundant.

5.1.5 Net Keeper

Net Keeper (NK) is the sole manager for all functions of VoIP platform, but without involvement in call processing.

NK :

1. Allows addition and removal of system components.
2. Stores routing configuration and network topology
3. Handles alarms.
4. Creates master routing tables that controls traffic routing paths by choosing cheapest price at calling instance.

5.1.6 Interactive Voice Response

34 This is only for the case of invalidated traffic as we will see later.

35 A protocol built on top of TCP/IP that assures accurate local timekeeping with reference to some on – line clock located on the Internet. This protocol is capable of synchronizing distributed clocks within milliseconds over long time periods. It is defined in RFC 1119.

Interactive Voice Response (IVR) is responsible for:
1. Receiving and decoding DTMF from clients.
2. Contacting the billing server to authenticate users and get their categories.
3. Sending voice messages for users.

One IVR is shared by all VG in the same site.

5.1.7 Accounting Server

Call Detail Records (CDRs) of different sites are collected and stored by the accounting server. All site keepers could share the accounting server as far as they belong to the same proprietor.

5.1.8 Billing Server

Billing server is an SQL[36] server that stores information that can be used by an external billing system. It is also used as a database for users.

5.1.9 PC- Client

Since many users will utilize their PCs, which are equipped with audio card, PC client takes the data from these cards and put them in the IP formats. PC client can also reconstruct voice packets to PC audio cards. PC client is capable of initiating and receiving VoIP calls.
As expected, PC client is used at the side(s) that uses PC rather than phones.
One type of PC client is Quick call client, which allows subscriber to only dial a predefined number.

5.2 System Architecture:

Figure (5-2) shows the entire configuration for VoIP system developed by Ericsson.

[36] SQL is a standardized query language for requesting information from a database. SQL is the favorite query language for database management systems running on minicomputers and mainframes. Increasingly, however, SQL is being supported by PC database systems because it supports distributed databases (databases that are spread out over several computer systems). This enables several users on a local-area network to access the same database simultaneously.

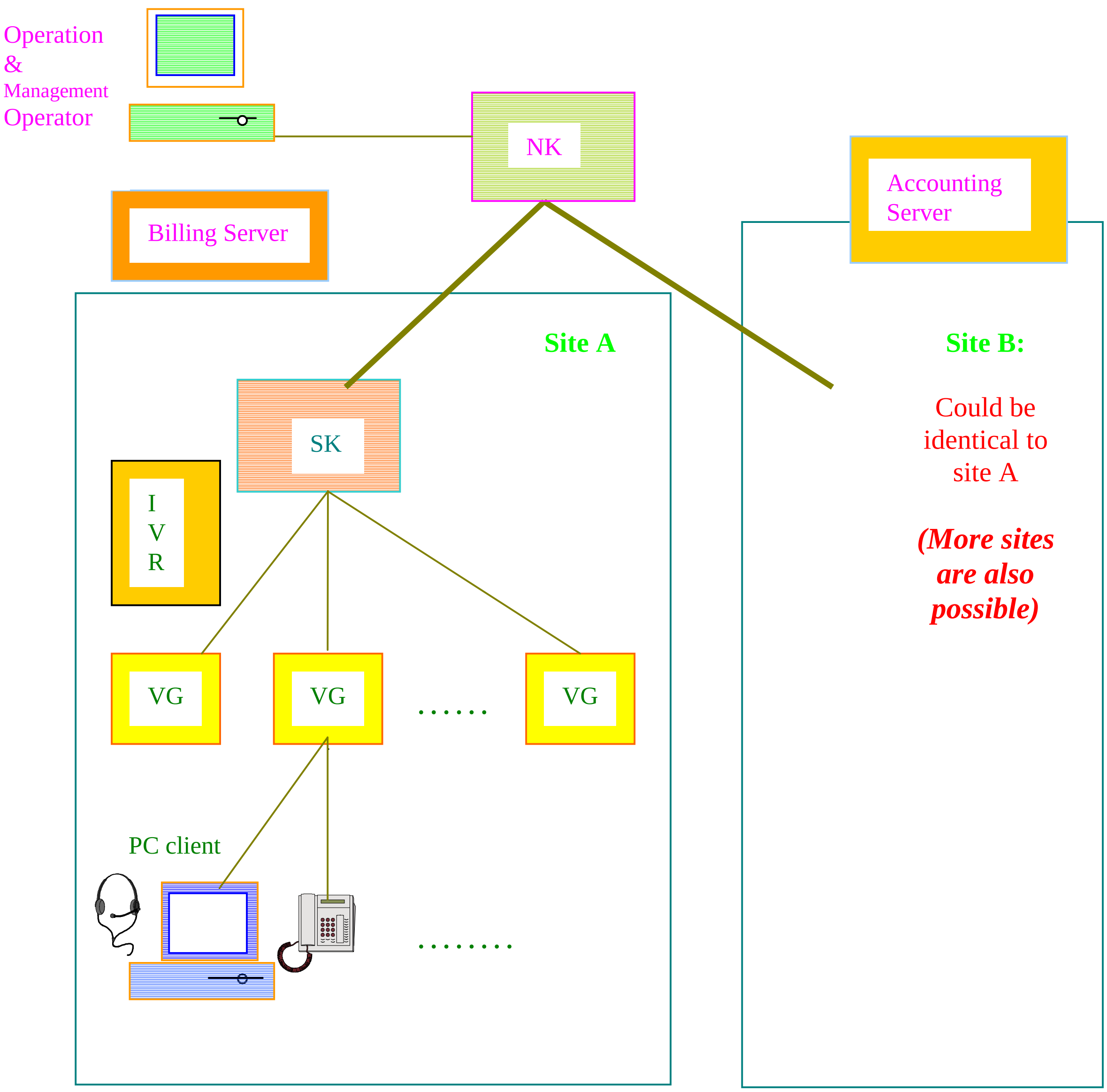

Figure (5-2) Complete system arrangement

5.3 Differences to H.323

1. System components for H.323 include Multipoint Controller (MC), Multipoint Processor (MP) and Multipoint Control Unit (MCU) to support conferences between three or more endpoints in a multipoint conference. Such feature is not implemented –as far as I know- by any VoIP vendor[37], and hence these components are not needed.

[37] **Probably** Multipoint Conferencing will be more vital for video (?). By definition H.323 touches audio and video aspects for visual telephone system for LANs.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

2. H.323 assumes intelligence in the end terminals to conduct compression and decompression (Codecs) as stated before, buffering packets (for jitter control), data multi channel (optional), capability statement, RAS[38] signaling function and call signaling function. In the case of Ericsson's IPT (release 1.5 at least) most of these tasks are performed by the voice gateway, or site keeper.
3. Address translation is expected in H.323 to be conducted by SK, but here it is implemented by the IVR.
4. H.323 mandates the dynamic behavior in assigning an endpoint (VG, terminal, or MCU) to a Gatekeeper (SK in our literature) over time. This behavior is within the context of a process called *Gatekeeper Discovery*. The aforementioned mechanism is not the case for Ericsson IPT, wherein Operation and Management Applets of NK is used manually by an operator to set the configuration.
5. None of IVR, Accounting Server or Billing Server were specified by H.323, nevertheless they can be regarded as complementary parts of the system.

5.4 Scenarios

Different scenarios are implemented for VoIP. System components vary accordingly. Following approaches for VoIP may take place:

1. PC to PC: A call originates from an IPT phone and transmitted over IP network to the called number.
2. PC to phone: A user connected to IPT telephony uses his/ her PC to direct a call over IP network, to a VG and finally to the called number in the PSTN network.
3. Phone to PC: A normal PSTN phone is used to initiate a call forwarded through a VG over IP network to the IPT called number.
4. Phone to phone: An IPT subscriber forwards a call through a VG to the IP network, from then to another VG to the PSTN network again to reach called party.
5. Phone Doubler: An E.164[39] alias is attributed for a PC, such that when a call is directed to the corresponding PSTN alias, IPT platform checks to see if any match between these two numbers. If so the incoming phone call is directed to the PC, rather than the PSTN phone.

---

[38] A feature that was initially built into Windows NT that enables users to log into an NT-based LAN using a modem, X.25 connection or WAN link. RAS works with several major network protocols, including TCP/IP, IPX. To use RAS from a remote node, one needs RAS client program, which is built into most versions of Windows, or any PPP client software. For example, most remote control programs work with RAS.

[39] E.164: The ITU-T standard that specifies the telephone number-type address format used for ISDN. Addresses are a maximum of 15 digits and are a geographically hierarchical structure (which is well suited to worldwide routing). Addresses are assigned by carriers.
In contrast, other addressing schemes (such as for the IP addresses for the Internet) are organizationally oriented. ITU-T standard E.164 (Numbering Plan for the ISDN Era) is the same as ITU-T standard I.331.

This fabulous scenario is promising in the sense that it allows phone calls to be summoned while surfing without interruption, and on the other hand, to be able initiate telephone calls while surfing again without interruption.

5.5 Quality of Service (QoS)

Ericsson's IPT offers different contracted levels of QoS:
1. Default fax.
2. Default voice.
3. Low quality voice.
4. Medium quality voice.
5. High quality voice.

Each VG has a table of 5 IP addresses, one address for each QoS rank.
Offered quality depends on signed contract between ITSP and customer.
Utilizing different combinations of CODECs and IP routes attains different levels of quality.
Supported CODECs are:
1. GSM full rate[40] with/ without silence encoder.
2. G.711 without silence encoder.
3. Fax mode[41] at (4.8, 9.6, 14.4 or 64 kbps).

A unitless factor called *Network Pressure* is of significance. It reflects the level of traffic in the network to a specific site. This factor has one value between 1 and 10; the higher given value for this factor means higher traffic load on the path (route) leading to that site.
For each previously mentioned IP address, a matrix is established to adduce for what QoS and at what network pressure that IP address have calls routed to.
Setting of network pressure, IP addresses for each QoS class is manually done by operator of Operation and Management Applets.

5.6 Types of Traffic

Two Types of traffic are managed by Ericsson's system:

1. **Validated Traffic:**
   Where no authentication or real time billing is needed.
   Billing Server is of no use in this case. In validated traffic:
   1. User dials a prefix followed by B number (called party number).
   2. PSTN network forward the call to IP (namely IPT) network.
   3. Two cases here:
      a. Called party has a number matches PC E.164: in this case, the call will destine in that PC directly without leaving the IPT network to the PSTN network again.

---
[40] 13 kbps without silence encoder and less than this figure with silence encoder.
[41] Fax mode is out of the scope of this thesis.

b. B number refers to a traditional phone: here the call passes the IP network to the PSTN network again according to least cost routing mechanism as well as signed agreements with PSTN proprietors.

2. **Invalidated Traffic:**

Contrary to the former case, no utilization of IPT functions is granted without priori authentication check. Authentication can be assured either by challenging the user to enter account and PIN numbers, or by detecting the calling number. Billing Server contains a database of all valid accounts.
Two major cases are presented in a more detailed fashion:

A. **Phone to phone / Phone to PC**: in this example

1. User dial ITSP number (normal PSTN number).
2. User is validated by either:
   a. Inspecting the dialing number.
   b. Being answered and asked for his account and PIN numbers.
3. Billing Server authenticates the user, presents agreed QoS with that user and informs the remaining value in the account of the user if the service he gets is prepaid.
4. The system could be configured either to:
   a. Allow the user to dial B number (called party) from the very beginning and in this case call is routed directly.
   b. Ask the user to dial B number and the call is routed afterwards.
5. Call is routed as in the case of validated traffic before and according to agreed QoS. Again SK could forward the call in to a PC or a traditional telephone.

B. **PC to Phone / PC to PC**: In this eventuality:

1. User starts with PC - client to ask for a new call.
2. User is asked for user ID and password, which are checked by the bill server.
3. Billing server, again, authenticates the user, presents his/her agreed QoS and remaining amount in his account if his service is prepaid.
4. Call is routed as before.

5.7 Routing

Netkeeper has a tariff table for each site (or even VG) which manifests -for each hour in the day- the cost of placing a call from the concerned VG(s) to certain destination in PSTN network.
Each SK is given a Master Routing Table accordingly.

A comment:
It is intuitive, to emphasize that powerful ITSP is the one who has as many VGs as possible, wisely distributed to support calls directed anywhere. This enables

avoiding threat of having to use many hubs over PSTN, and in this case it would be difficult to compete.
Another choice is to have agreements with other ITSP for call completion within their areas of power.

5.8 Configuration:

As rendered in figure (5-3), VG can be configured to PSTN network using the following interfaces:
1. E1 ISDN-ETSI PRI.
2. T1 ISDN-ANSI PRI.
3. T1.
4. E1.
5. Analog POTS: in multiples of 4 channels.
On the other side VG can be configured to :
1. Ethernet 10 Base – T
2. Ethernet 100 Base – T.

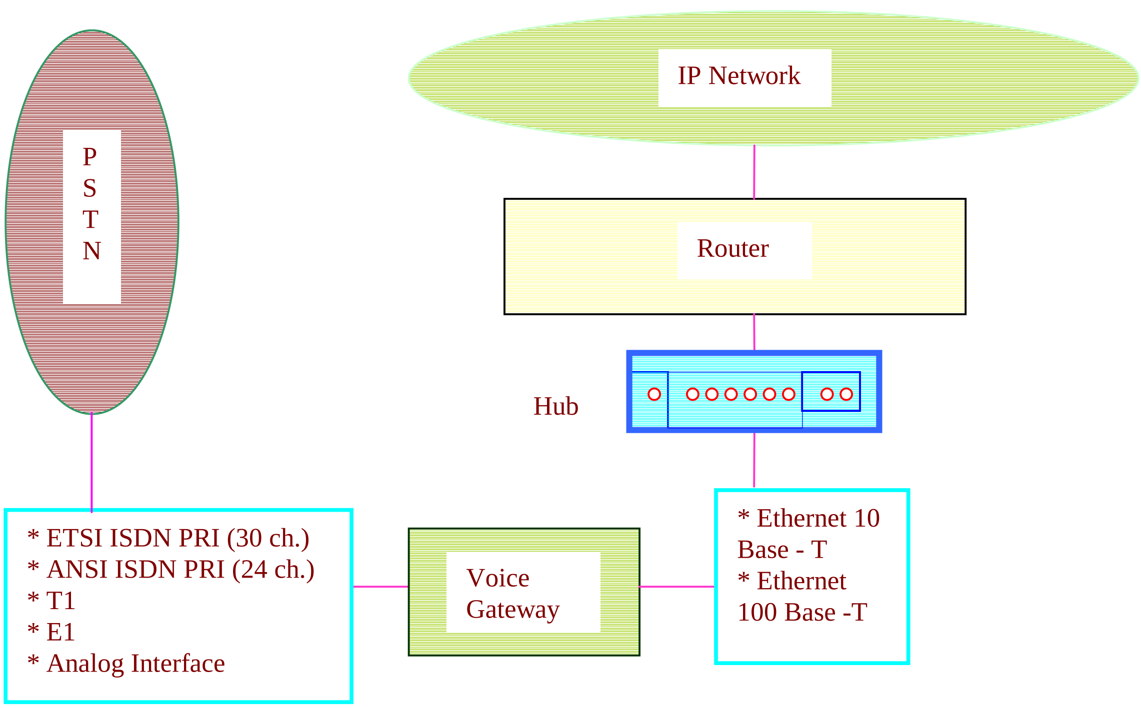

Figure (5-3) Configuration

Chapter Six

Proposed Tests for Integrating VoIP with BC10 PBX[42].

6.1 Delay Test:

Test configuration is as shown in figure (6.1). Here the test will not include any router or Schouter; direct connection of 2 Voice Gateways (VG) back to back.
The significance of this test is to measure the delay the packets suffer by the VG. I failed to find any claimed value for this parameter in any document for release 1.5. I could say that though had it been specified, I should measure it as a field check anyway.

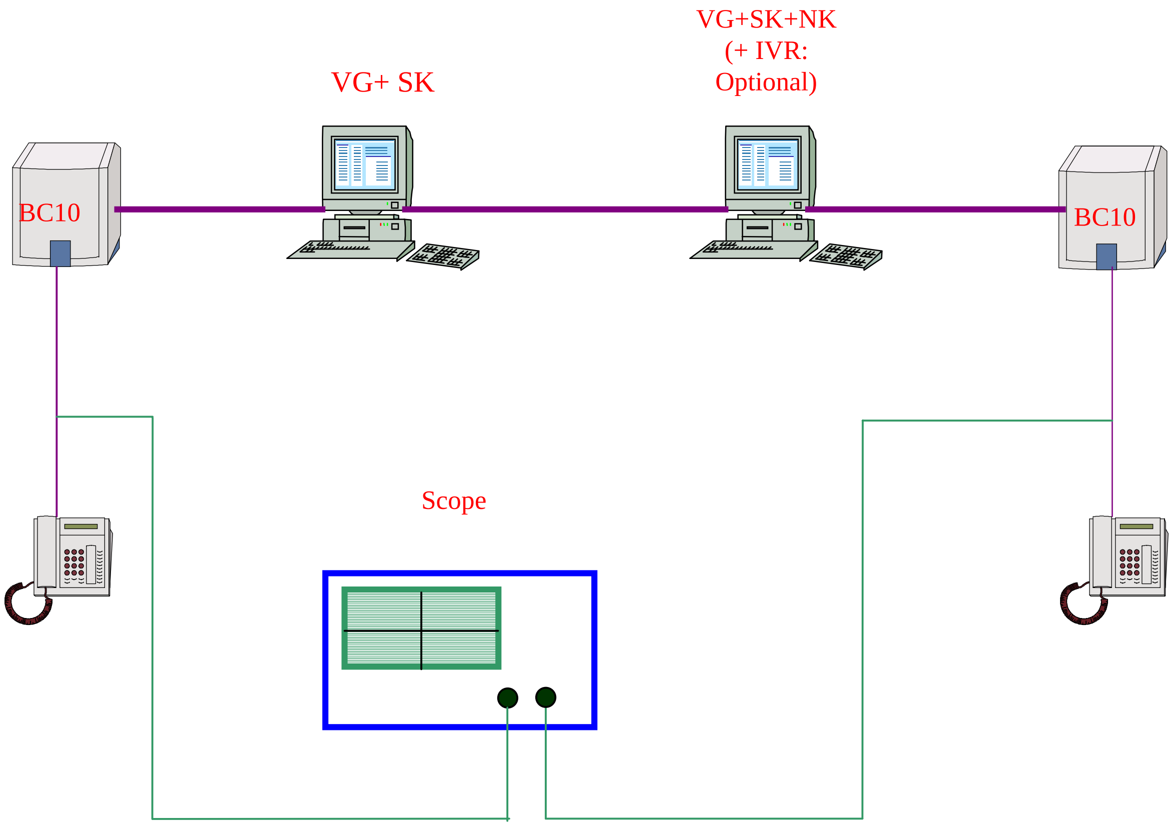

Figure (6-1) Measuring the delay.

[42] A report I submitted to my supervisor at a mid stage of the project. It echoes what tests I intended to transact.

6.2 Traffic Measurements for VoIP

This is the first assignment stated in my Thesis Proposal. As a matter of fact, I am not going to make any direct tests or measurements for the load. This is simply because the Net Keeper (NK) (single point of control for the Operation and Management O&M of the whole IP Telephony IPT platform) plays the role: It has the applet that shows the instantaneous status of all connected circuits as well as maximum and used capacities. Furthermore, the traffic can not exceed the some allowable figure that depends on the chosen configuration (Analog POTS, T1 R1, ISDN – ETSI /ANSI PRI). This fact is the reason for rejecting new call attempts in case of network overload.
It might be wise to compensate for deleting this test by monitoring the CPU Usage of the PC (a facility offered by NT for measurements of performance) that bears the VG (Site Keeper and Net Keeper in one side).

6.3 Objective Tests

In ITU-T recommendation P.861, an algorithm for Objective Test was described on the basis of objective quality measure called Perceptual Speech Quality Measure (PSQM). This approach –according to my best knowledge- has yet to be an accredited test means for Telephonometry.
Therefore, and bearing in mind that transacting such a test require some additional expensive test device[43] which has not been field proven, I opted for making subjective tests.

6.4 Subjective Tests

As expected, subjective test is a time consuming process and mandates assistance of lots of people. Yet this is the traditional way of assessment of Voice Quality. Final arbiter is human ear not measurement devices, knowing this gives the satisfaction feeling and leads to trust this methodology.
As far as I know neither a subjective test nor an objective one had ever been implemented for VoIP[44].
I used the following ITU-T recommendation as a platform for subjective test. P.10, P.84, P.85, P.810, P.830[45].
Test configuration is depicted in figure (6-2).
Subjective tests will comprise:[46]

---

43 After an extensive searching, the only system I could find for this purpose was Hammer System for VoIP Tester. It was claimed to be a P.861 compatible.

44 Until December 1998!

45 It is a point worthy that P.800 characterizes 8 different tests, but I picked up only the ones I felt more expedient for evaluating VoIP.

46 All following tests shall be repeated for 3 different CODECs: GSM full rate with and without silence encoder as well as G.711, and also for different values of packets' delay!

6.4.1 Conversation–opinion Tests[47]

Category

Making real conversations: Illustrative demonstration, Clear instructions and result forms shall be given to people I will ask to make these calls.
• Despite P.800 recommends using soundproof cabinets (with specific requirements of volume, reverberation time and Standing wave pattern ratio) for conversations' doers, I will be satisfied with the normal

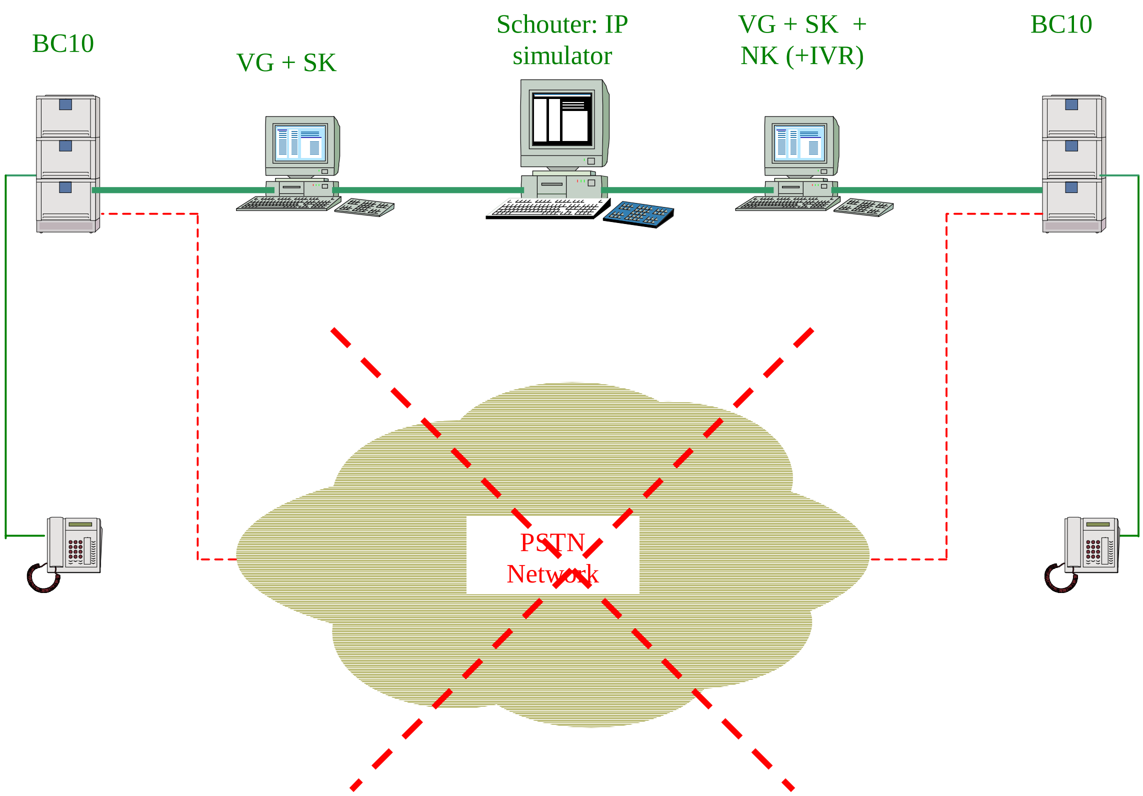

Figure (6-2)
Test configuration for making calls

conference rooms we have at Ericsson facilities.
• Noise could be incurred while conversations are going on; I should set some criteria for this noise: e.g. tape, moving vehicle.
• Since results are based on opinions of those who make conversations, the following implications should be considered:
a. Making as many conversational calls as possible.
b. People of different ages, languages, and sex will make these calls.[48]

---

[47] It includes Dialing and Ringing Tests.
[48] I hope I will able to cope with this by asking some friends to help.

c. Subject should not have participated in any subjective tests for at least 6 months, and have not directly involved in a work connected with assessment of performance of telephone circuits or related work.

Annex B depicts a suggested form to be distributed to conversation performers. It contains two questions to be answered by each subject *individually.*

The arithmetic mean of any collection of opinion scores for the first question is called Mean Conversation – Opinion Score MOS_c or MOSc (or even simply MOS).

6.4.2 Listening Opinion Tests:

6.4.2.1 Degradation test:

6.4.2.1.1Test Description:

The goal of this test is to evaluate the quality of voice when packetized and transmitted over IP network compared to normal voice as a reference.

In this test, two separate recording systems are to be used simultaneously: one for recording the wideband speech in one channel and the other for recording the telephone speech (VoIP) in the corresponding channel. The dual recording system ensures that the same speech is recorded in two forms (codec telephone speech and wideband speech).

Despite the VoIP technique is intended to be used for voice only, nevertheless I should think of recording parts of musical pieces just for the sake of having more accurate comparison (by magnifying the difference). Music occupies higher frequency spectrum than normal speech.[49]

I will make both of recording and playing back of test messages in a digital format on a PC which I think is keep signals more intact than when using normal cassette recorder. This test is also called Degradation Category Rating DCR [ITU: P.800].

* The recording (and testing process) could be repeated with: - Introduced noise. - Different SNR levels.[50]

---

[49] Voice of females could reach a maximum of 8 kHz spectrum (less than this figure for males), while human ear can sense frequencies up to 20 kHz!

[50] These cases will dictate finding some way that fulfils 2 conditions: first doping some noise (white or frequency dependent) and second, variation of the Signal (voice) for a fixed noise pressure.
As a preliminary solution I plan to ask each talker to fix a headphone with some music or speech on his/ her ears in order to force him / her to raise his/ her voice in a sustainable way. For the first case I

6.4.2.1.2 Talkers

Talkers should be at least one male and one female to reduce the possibility of results' dependence on peculiarity of the chosen voices.

6.4.2.1.3 Listeners:

24 persons are needed: This is the trend followed in Ericsson Software (Erisoft). It is important to make sure that all listeners are exposed to exactly same conditions.
According to [ITU-T: P.800: Annex B], the following conditions should be met by chosen listeners:
- Should not have participated in any subjective tests for at least 6 months.
- Have not directly involved in a work related to assessment of performance of telephone circuits or related work.
- Have never heard the sentences before.

6.4.2.1.4 Test and Results

Listeners shall listen to each sentence[51] twice (one is the reference, and the other is the VoIP signal), afterwards they should be asked to answer the assessing question shown in Annex C *for each case*: Also, listeners are encouraged to give comments.

6.4.2.2 Clearness (Transparency) Test:

Category:

The theme of this test is to inspect how perspicuous (clear) the speech is, when transmitted over Internet. In other words it is a measure of faith of IP system when exploited for transmitting voice.

6.4.2.2.1 Test Description:

In this test, several sentences of a fluent talker(s) are to be transmitted over IP and recorded at the RECEIVER.
Suggested Sentences are shown in Annex D [P.85][52].

---

could make RECORDING in a noisy environments (e.g. the fan behind me in Mobility Lab1, I am addict to its rhythm).

[51] Or music.

[52] With some variations!

In order to minimize any individual peculiarity of talkers[53], for each group of sentences recited by a single talker, Listeners will be asked to

1. While listening: answer some sensitive and short questions on "Examination papers" that would reflect the ability to feel the contrast in speech (e.g. distinguish P & T).
   Suggested answer sheet is shown in Annex E [P.85].
2. After finishing listening session: fill in the assessment forms shown in Annex F.

6.4.2.2.2 Talkers:

Same as in 6.4.2.1.2.

6.4.2.2.3 Listeners:

Same as in 6.4.2.1.3

6.4 A Comment: Summary of tests:

It will be a time consuming process indeed:

1. Delay measurement.
2. Conversation – opinion: I would suggest that number of telephone calls to be 10 for each case.
3. Degradation test: (4 talkers[54] * 2 sentences + 2 music pieces → 24 listeners.) * 3 levels of SNRs.
4. Clearness test: 4 talkers * 12 messages

All tests shall be repeated for 3 different modes of Voice quality.

---

[53] Not needed for Music.

[54] 2 males and 2 females.

Chapter Seven

Transacted Tests

7.1 Delay

It was not possible to measure the delay by simply using the intended set up in section 6.1, whereby I assumed that pressing one digit from one phone could be traced on the other one (of course after establishing a call between them) and measure the delay thereupon. In fact the DMFT signals are in terms of kHz and therefore, time span for those signals are in terms of fractions of milliseconds while the delay is assumed to be some tens of milliseconds. On this basis it would be impossible to measure the delay this way.
The methodology I implemented was to inject a very low frequency test signal into one phone, and detect this signal on the other side. Both signals were displayed on a normal oscilloscope. The injected signal was 1.3 Hz saw tooth. The measured delay was around 200 ms, which means the delay for each voice gateway is 100 ms.
I have to admit that precision[55] of this figure could reach 25%, because it is very difficult to have a still signal at this frequency on the oscilloscope.
More accurate delay measurements require special software that measures delay automatically.

7.2 Test Environments

7.2.1 An Important Note

It is crucial to emphasize that if tests were to be transacted according to the configuration depicted in figures (6.1) and (6.2), then the system would collapse. In those scenarios I assumed that Netkeeper as well as Sitekeeper(s) could be installed in the same hardware platform of the Voice Gateways. This means that control information between Netkeeper and Sitekeeper(s) would be exposed to loss and delay like other transmitted packet over IP network (Internet), which implies loosing control and breakdown.
This can be clarified by exploring the case in Figure (7.1), if we imagine that NK and the SK are to be installed in the same PC that bears one of the two VGs. In this hypothetical case (I installed, at the start before noticing the error), Control information between SK and the other VG would pass through the Schouter to the other PC. Since the Schouter is supposed to emulate IP network that includes a small percentage of packet loss, then loosing control can be understood. This could be a killing point in real systems, so unless dedicated leased lines or VPN are used for connecting Netkeeper to Sitekeepers, and Sitekeeper to Voice Gateways the system will be spineless.

[55] By precision I mean, how results of different trials are close to each other.

7.2.2 Set up

It was the first hole I faced: how to emulate the Internet. My understanding of the emulator is the one that allows for:

1. Packet delay.
2. Packet loss.
3. Out of sequence reception of packets (Jitter).

After endeavor, I found software called Schedulable Router (abbreviated as **Schouter**) works under Linux that meets all these requirements. Description of the Schouter is given in a separate subsection. Figure (7.1) shows the set up I built and used for tests; taking into consideration the aforementioned important note.

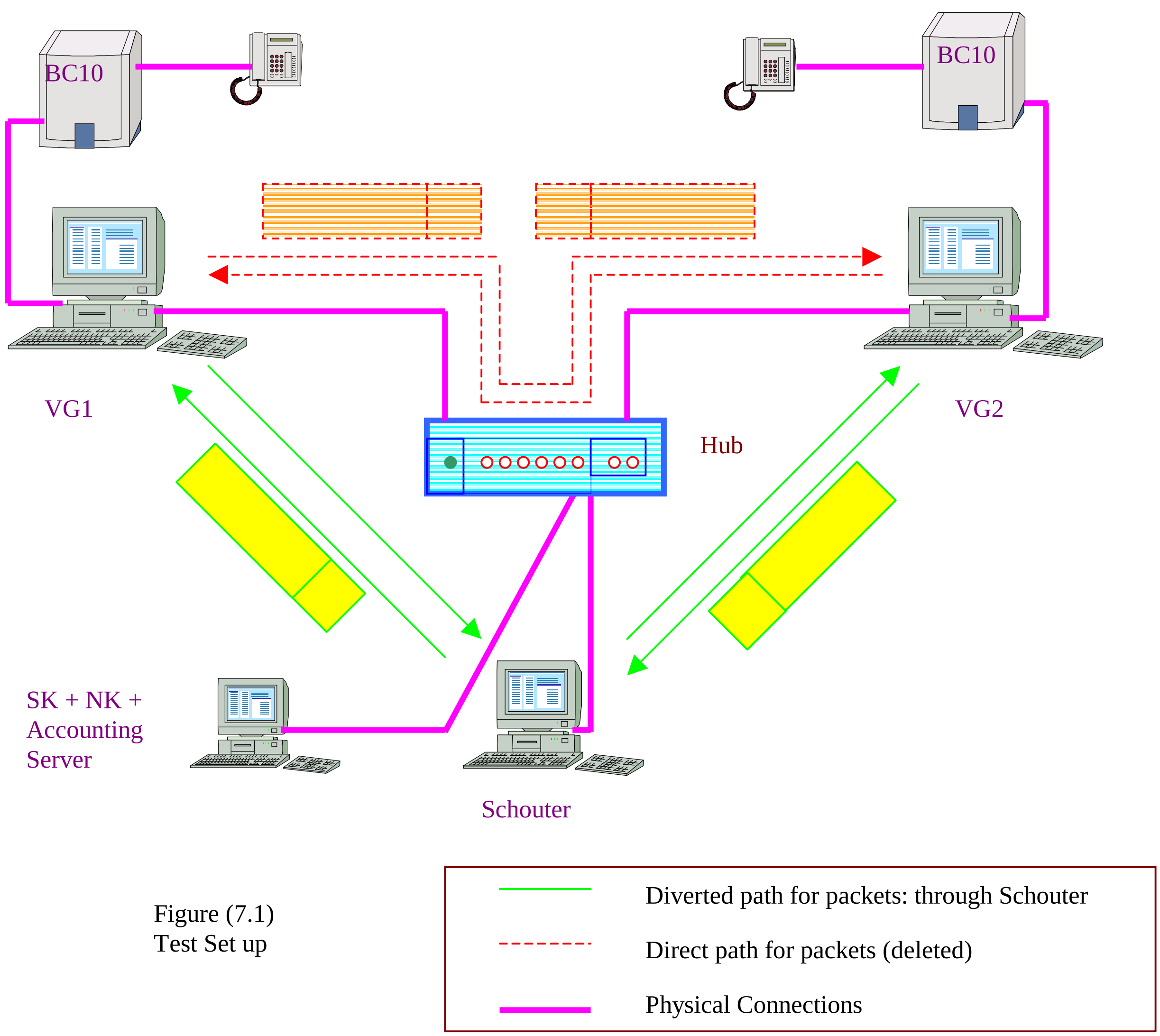

Figure (7.1)
Test Set up

The routing table of this set up forces packets from one VG to the other to pass through the Schouter -as shown in figure (7.1)- wherein they are assumed to adhere to parameters of the activated log file. As described in the context of the Schouter, and according to mode of operation[56], packets are treated individually or collectively in terms of being delayed or thrown away. In my tests, I used two modes of the Schouter:

1. Static mode: Packets are dealt with collectively for fixed value of delay, no packet loss and no jitter.
2. Event mode: Packets are dealt with individually for simulating delays, packet loss as well as jitter.

VG: The used VG's were version 1.5 with Dialogic cards.

Environments:
VG: Windows NT.
SK, NK, and Accounting Server: Windows NT.
Schouter: Linux.

7.2.3 The Schouter[57]

The Schouter is a tool for emulation of packet based networks. It works on the IP level and is therefore application independent. With emulation means that the tool itself does not provide any information of the network behavior, but the tool should be fed by a delay pattern. With this delay pattern the Schouter will, in real-time, delay each packet on the network according to the given delay pattern. The purpose for developing this tool was to get a repeatable IP channel conditions to be used for real-time evaluation of real-time application such as voice. It can work both event driven of time driven. In event driven mode it steps forward in the delay pattern each time a packet is received into the Schouter. In time driven mode it monitors the offset in time between two consecutive packets and by comparing the time with the resolution in the delay pattern it decides whether or not to step forward in the delay pattern.
The word Schouter comes from a play with the words schedulable and router.

Delay patterns

The delay pattern files which are used to feed the Schouter is in a proprietary format. The proprietary format is a 512 bytes header and the delay values are given by the number of tenth of ms. Each value is stored as a 16 bit integer which gives the maximum delay for one packet to be just about 6.5 seconds. If a larger delay value is wanted a

[56] Specified in the configuration file.
[57] Quoted from a descriptive literature for the Schouter written by Ericsson Software (Erisoft).

scale factor is available in the header. By scale the delay values the resolution will be reduced with the scale factor.
From ASCII files with one value in number of tenths of millisecond per row that particular packet should be delayed, an application taking care of the conversion is also distributed.
The delay patterns could be generated in several different ways:
1. Derived on a statistical basis with Excel, Matlab or any other tool providing statistic mathematics.
2. Through measurements in networks. If the measurement points are separated, the clocks should be synchronized to get the real one-way delay values. If measure the round-trip delay by echo the packets at a reflector, the delays should be halved.
3. By simulations of IP networks and logging of a certain IP activity. A usable tool for simulation of IP networks is Plasma IP.

Hardware and software requirements

The Schouter has been developed as a software to run in user space. It utilizes the library "libpcap" which gives an interface to the network hardware e.g. the Ethernet network board. It has been written in C and has been compiled and run successfully on several different PCs running Linux / Red Hat 5.X and on Alpha workstations running Digital Unix. To be able to achieve an acceptable performance the kernel must also be modified.

Performance

During development of the Schouter the target for the performance was that the accuracy between actual and wanted delay should be less that 1ms.
When running on a Pentium 200 and loading the Schouter with traffic (UDP) 50 packets per second and a application data load of 35 bytes (a typical GSM over IP call) during approximately 28 minutes, 99 per cent of the packets had a maximum difference less than 0.951 ms between wanted and actual delay. The mean of all packets where 0.343 ms with a standard deviation of 0.245 ms. The maximum difference during the call was 3.9 ms (probably due to a command to see that everything was up and running).

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

7.3 Subjective Tests for the first Codec[58]

The first Codec was full GSM with silence detection, and hence the occupied bandwidth is always less than 13kbps. Intensified care was taken for this instance, because it is the most bandwidth saver[59] and thus of high interest for service providers accordingly.

7.3.1 Conversation - Opinion Test

7.3.1.1 One line delay:

Although it might be a fanciful test criterion, I was curious to know the consequences of having one way delay on this connection.
In this test each person in the group of people who made conversations[60] made two calls, namely one from delayed route and another from the delay free source.
Before discussing quantitatively, it was interesting to know that in the contrary to my early expectations, the party that received delayed packets was more satisfied than the side that had his/ her speech delayed before transmission.
This test comprised the following values of delay:

1060 ms (860 ms from the Schouter + 200 ms from 2 VGs):
Average evaluation, or *Mean Opinion Score (MOS) as* called in ITU recommendations (in percentage format[61]):

Receiver of delayed packets	55%
Delay causing side	43%

Difficulty :

Receiver of delayed packet	28%
Delay causing side	39%

620 ms (420 ms from the Schouter + 200 ms from 2 VGs):
Average evaluation (MOS):

Receiver of delayed packet	63%
Delay causing side	52%

No *difficulties* were specified on either side.

7.3.1.2 Tests with identical conditions in both lines.

---

58 During all test for all Codecs, Dialing and ringing between phones over the built set up worked successfully. Similarly, no loading on any of the PC's was detected.
59 Of Ericsson's solution.
60 For each single conversation test case, at least 8 different people were asked to go through the trial.
61 These figures are the arithmetic mean for group replies for each case. It is the average for answers to the first question in annex B, expressed in percent.

The tests were transacted by making combinations of:
1. Packet loss.
2. Packet delay[62]
3. Out of sequence reception of packets (Jitter).

Results are shown in Table 7.1

Case	*Delay*	*Loss %*	*Jitter %*	*MOS %*	*Difficulty %*
1	*200*	*0*	*0*	***73***	***7***
2	*200*	*5*	*0*	***60***	***18***
3	*200*	*20*	*0*	***18***	***67***
4	*200*	*30*	*0*	***11***	***73***
5	*400*	*5*	*0*	***56***	***25***
6	*500*	*3*	*0*	***60***	***15***
7	*700*	*10*	*0*	***24***	***48***
8	*500*	*1*	*1*	***45***	***21***
9	*500*	*0*	*3*	***31***	***42***

Table 7.1 Evaluation (MOS) and difficulty results for the first Codec

Exploring this table closely, I would like to point out the following observations:

1. When we compare the results of evaluation (MOS) for different cases, we find a misleading inconsistency. The explanation I can give for this paradox is differences in personal evaluation “or peculiarity” of evaluators:

For each single case, a minimum of 8[63] different volunteers did the task and it was insensible to ask them to make all tests[64]. I was undoubtedly concerned about getting awkward outcome, and hence tried to give a clear explanation of the first question in Annex B. That is, “Perfect” in that question means general evaluation (MOS) is comparable to quality of normal telephony. Inferior opinion ought to be assessed in accordance to the degree of inferiority. Yet personal attribute remains.

[62] As noted in the table the minimum examined delay was 200 ms, which is the value of delay imposed by the two Voice Gateways. In other words despite the value I specified in some trials for Schouter log files was 0, still the delay caused by the Voice Gateways is inevitable.

[63] For some cases the number reached 14.

[64] Number of people who did conversation tests exceeded 50 persons at different ages, languages and sex.

Nevertheless, the moment of these tests has been fulfilled –I guess- by achieving moderate degree of sensing quality improvement.

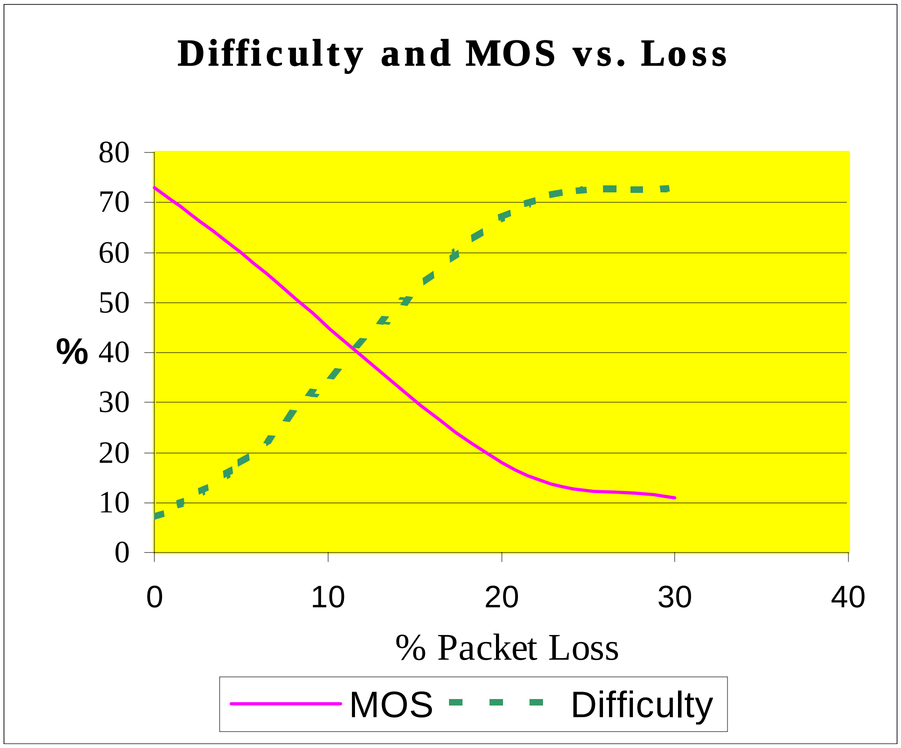

Figure (7.2)

In figure (7.2), the relation between MOS (as well as difficulty) and packet loss is drawn for a fixed value of delay (200 ms). It is somehow similar to relation between MOS and SNR (of course when side-reversed), as shown in figure 9/P.830 in ITU-T recommendation P.830.

2. When it comes to understandability of speech, jitter is more serious than packet loss and packet delay. VGs in general have buffers to ensure correct sequence of packets before decoding, however the size of the buffer can not exceed some value to avoid extra delay. After all, it is a trade off between quality and delay.

3. A 700ms-delay can be acceptable by normal people, likewise a packet loss of up to 10% can be endured. However, delay and packet loss at those levels are easily detectable.

4.Strangely enough, people can still talk at 20% packet loss, though quality is bad and substantial parts of words are missing.

5.The case of 400 ms delay and 5% packet loss is interesting. I -for few months- followed up a web site[65] [25] that monitors the Global status of Internet traffic and found this benchmark the most frequent. It could be even argued that VoIP uses UDP/IP not TCP/IP (the more used protocol for Internet) which would suggest slightly less delay and slightly more packet loss. Figure (7.3) images the global packet loss for the period (23rd January – 23rd February 1999), while figure (7.4) materializes global packet delay for the same period.[66]

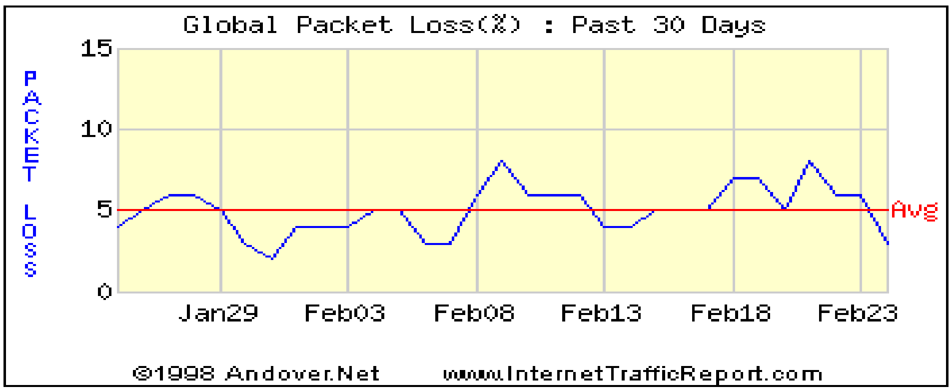

Figure (7.3): Global Packet loss for one month

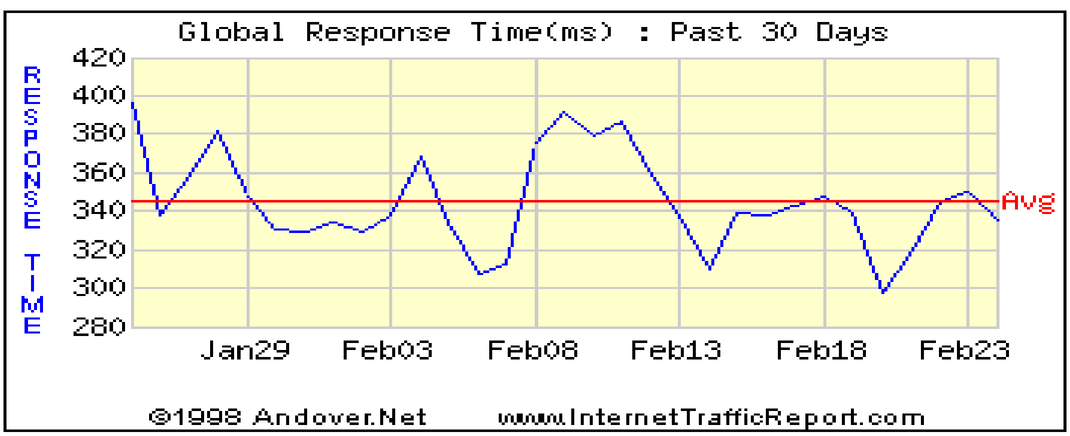

Figure (7.4): Global Packet delay for one month.

6. These tests were done in the following languages: American English, Arabic, Bosnian, Chinese, Croatian, Finnish, Persian,

65 http://www.InternetTrafficReport.com/index.html
66 In fact this time is better than few previous months.

Portuguese, Spanish and Swedish[67]. No discrepancy -that can be referred to because of language difference- was found.
7. Both of general evaluation (MOS) and difficulty are affected by both of packet loss and delay. No single factor acts on one a single parameter.
8. An interesting comment I received, described the voice to be altered such that it was similar to the voice of a drunken person!
Another comment claimed, consonants were more deteriorated than vowels.

7.3.2 Degradation Test:

In spite of what I planned before[68], to do in this test I found it more realistic to hold the comparison of the recorded material at the receiver side of VoIP telephone to same material recorded at the receiver side of normal telephone. More precisely:
1. Materials were recorded in digital format on a PC. Playing and recording of messages were done using software tools called "Cool Edit". Cool Edit is a flexible means; it caters for choosing sampling frequency between 6 – 48 kHz and 8/16-bit quantizer[69]. For all messages, 8 kHz sampling rate and 16 bits quantizer were chosen; that is 128 kbps or double single channel telephony bandwidth.
2. A normal telephone line connection was established
3. The recorded materials were played at one side of the aforementioned connection.
4. Another recording system (also in digital format) was initiated at the other side of the connection.
5. Procedures from 2 to 5 were repeated using a loss free and delay free[70] VoIP connection.
6. Outcomes of steps 4 and 5 are to be compared.
Luckily, I managed to have 30 listeners at the same time.

Two "raw" materials used were:
1. A non-technical and non-specialized Swedish text, that was recited by a fluent speakers for few minutes.
2. An Arabic song with light music. Again I should emphasize that VoIP (and even normal telephony) was not designed to bear this sort of high frequency signals, albeit my main interest was just to feel the difference.
Results of this poll were as follows:

---

67 All by native speakers, to minimize the possibility of misunderstanding due to different accents.
68 Section 6.4.2.1 ; in the report I submitted to my supervisor before starting tests
69 Analysis of recorded messages in frequency domain was also possible.
70 Except 200 ms caused by the two VGs, though delay is not expected to have consequences on this one-way voice transmission.

1. Swedish Text

Results are summarized in Table (7.2).

% difference felt between 2 materials	Replies	%
Less than 5%	3	10
5% - 10%	4	13.3
11% - 18%	5	16.7
19% - 25%	9	30
26% - 33%	7	23.3
34% - 42%	1	3.3
43% - 50%	1	3.3
> 50%	0	0

Table (7.2)

Considering the weight of each category, the difference between two materials was about 20%.

2. Arabic Song

Table 7.3 depicts the Replies.

In this case the average felt difference exceeds 33%.

% difference felt between 2 materials	Replies	%
Less than 5%	0	0
5% - 10%	0	0
11% - 18%	2	6.6
19% - 25%	3	10
26% - 33%	14	46.7
34% - 42%	4	13.3
43% - 50%	5	16.7
> 50%	2	6.6

Table (7.3)

Probing result distribution graph in figure (7.5) shows how the difference was felt better in the second case, and even how it was closer to natural distribution graph. Of course less difference is expected for other Codecs, as they offer higher bandwidths.

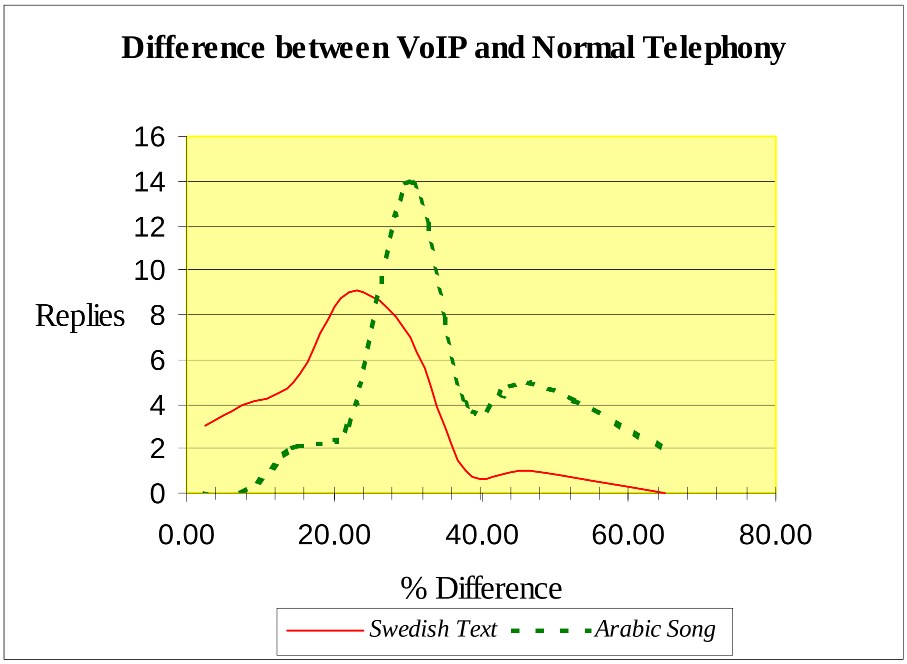

Figure (7.5)

7.3.3 Clearness Test

7.3.3.1 Comparison between VoIP and Normal Telephony

Once again, I credited to hold a comparison between normal telephony and VoIP clearness wise. The used Codec for the case of VoIP was full GSM with silence detection (as before, bandwidth is less than 13kbps), without introduction of any external error or delay.
I have selected the messages in annex D.5 (read by Marie Gabrielsson) to be transmitted over telephony network, while other 4 groups (D.1 - D.4 inclusive) were transmitted over IP network.

Results for Telephony case:

Results can be summarized in table (7.4):

Item	% of people answered correctly
Name of product Namnet på varan	91,1
Reference Number Referens nummer	95
Price Pris	100
Delivery Period Leverastid	100
Train Number Tåg numret	95
From or To Till eller Från	95.2
Time Tid	88
Platform Plattform	96
Train Spår	96

Table (7.4) results of clearness when using normal telephony line.

The table may tell that the overall clarity was 95.1%.

Results for VoIP case:

Table (7.5) depicts the outcome of clearness tests for four different readers.

Item	% of people who answered correctly : For sentences read by :				
	Ingela %	**Boston %**	**Åsa %**	**Niklas %**	**Average %**
Name of the product Namnet på varan	75	89.7	87	78.2	**82.5**
Reference Number Referens nummer	76.2	63.2	93	86.2	**79.7**
Price Pris	92.9	93.1	97.7	95.4	**94.8**
Delivery Period Leverastid	95.2	96.6	100	98.9	**97.7**
Train Number Tåg numret	61.9	90.8	93.1	83.9	**82.4**
From or To Till eller Från	76.2	82.8	95.4	81.6	**84.0**
Time Tid	80.9	85	89.6	87.4	**85.7**
Platform Plattform	85.7	72.4	93.1	89.7	**85.2**
Train Spår	70.2	64.4	92	59.8	**71.6**
Average %	**79.4**	**82.0**	**93.4**	**84.6**	**84.8**

Table 7.5 results of clearness test when using VoIP connection line.

7.3.3.2 Acceptability of VoIP

This subsection brings out results of answers to the questions in annex F[71].

Overall Impression:

Results are tabulated in Table (7. 6)
Examining the overall average in this case, we find it 10 points outperformed by normal telephony line.
Probably, it is reasonable to comment on these results:

[71] 29 persons out of 30 filled this form.

1. The mean of clarity for all items in VoIP case is less than that for normal telephony in nearly all individual cases (for different reciters).
2. Some letters like G in announcement M4 (annex D.1: read by Ingela), H in announcement M6 (annex D.4 : read by Boston) and J in announcement M5 (annex D.5 : read by Niklas), were puzzling in deed.[72]

- 21% could recognize the G. received answers included P, C ,D,E,T.
- Only 10% managed to distinguish the H. Other answers were O, and 2.
- Less than 7% identified the J. D, E, Y and I were given in wrong answers.

Impression (MOS)	Replies	%	Cumulative
10.Perfect	0	0	0
9.Excellent	0	0	0
8.Very Good	0	0	0
7.Good	2	6,9	6,9
6.Moderately Good	7	24,1	31,0
5.Fair	4	13,8	44,8
4.Somewhat Poor	11	37,9	82,8
3.Poor	2	6,9	89,7
2.Very Poor	1	3,4	93,1
1.Bad	1	3,4	96,6
0.Useless	1	3,4	100,0
Total	29	100	

Table (7.6)

Multiplying the weight of each rank by accompanied repetition, and dividing by total number of replies (29) gives an average of 4.45 out of 10: or 44.5%).

[72] They were of course read in the Swedish way.

Plotting the cumulative graph as suggested in figure 1/P.85 in ITU-T recommendation P.85, we find 50% of cumulated assessments give around 4 or 40%[73] which is not very good!

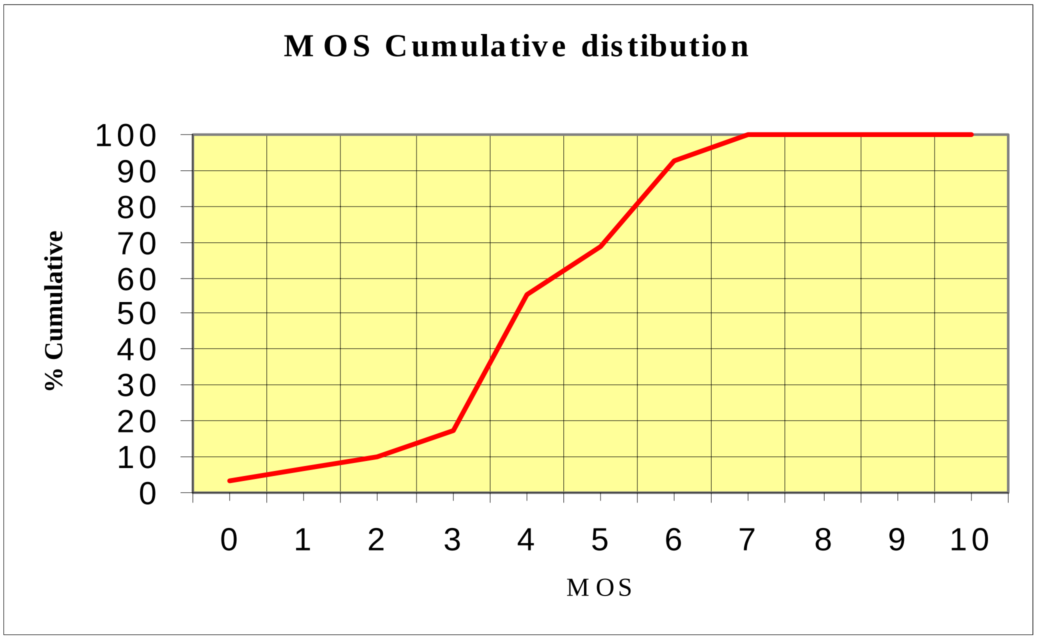

Figure (7.5)

7.3.3.2 Listening Effort

Results are as shown in table (7.7)

Item No.	Description	Replies	%
5	Complete Relaxation	0	0
4	Attention Necessary: No appreciable effort required	4	13.8
3	Moderate Effort Required	15	51.7
2	Effort required	10	34.5
1	No meaning understood with any feasible effort	0	0

Table (7.7)

[73] Not far from calculated average.

If we are to infer a value for concentration measure and assuming -as shown in the table- point 1 means necessity of high concentration and 5 denotes for complete relaxation, then the value we can derive from the table is 2.79 which is not good either!

7.3.3.3 Comprehension Problem

As described in item number 3 in annex F, this issue discusses the difficulty of understanding some words. Results are presented in Table (7.8)

Item	Replies	%
Never	0	0
Rarely	6	20.7
Occasionally	19	65.5
Often	4	13.8
All of the time.	0	0

Table (7.8)

From this table, it can be deduced that the measure of this difficulty is 3.07 out of 5 , which is quite high.
Two points should be recalled when discussing Comprehension, Articulation and Acceptance[74].

1. This case was packet loss free, results could be worse with even 5 % packet loss.
2. Despite I made my best to make it good, listening conditions may be were not perfect. In fact, I used a PC with a sound card and two big external speakers[75]. There were no complains that could be referred to the system but to speed of talkers, as will be discussed in a later subsection.

7.3.3.4 Articulation

The articulation was interpreted as the distinguishability of sounds. Outcome is given in Table (7.9)

In other words Articulation was 2.665 out of 5, which is also not very promising.

[74] In the following subsections
[75] 120 Watts was written on the speakers, I find it difficult to believe!

Item No.	Description	Replies	%
5.	Very clear	0	0
4	Clear enough	1	3.5
3	Fairly Clear	17	58.6
2	Not very clear	11	37.9
1	Not at all	0	0

Table (7.9)

7.3.3.5 Acceptance (Usability) of VoIP

Result is the answer for recommending VoIP as an information service.
Table (7.10) clarifies.

Item	Replies	%
Yes	11	37.9
No	18	62.1

Table (7.10)

7.3.3.6 Speaking Rate:

It is clear that this question is good only for analyzing listening conditions of questions of relevance to quality. As such, speaking rate is not an assessment parameter for VoIP, but it gives a clue on possibility of making mistakes while listening because of the fast reader.
Result is shown in Table (7.11)

Item No.	Description	Replies	%
5	Much faster than preferred	7	24.2
4	Faster than preferred	17	58.2
3	Preferred	5	17.2
2	Slower than preferred	0	0
1	Much slower than preferred	0	0

Table (7.11)

This means that speed was 4.07 out of 5, which indicates unacceptable listening conditions. This impression -probably- had affected the overall results.

7.4 Subjective Tests for the second Codec

The second Codec is full GSM without silence detection, or 13 kbps bandwidth. The tests done were only conversations. 4 cases that are thought to be more frequent were tested.

Results of tests are tabled in table (7.12)

.

Case	***Delay***	***Loss %***	***Jitter %***	***Quality %***	***Difficulty %***
1	*200*	*0*	*0*	***81***	***<5***
2	*400*	*5*	*0*	***59***	***8***
3	*500*	*3*	*0*	***68***	***5***
4	*500*	*1*	*1*	***51***	***18***

Table (7.12)

7.5 Subjective Tests for the third Codec

The third Codec is full 64kbps (without silence detection, of course). Three tests were transacted for this Codec as presented in table (7.13).

Ser.	***Delay***	***Loss %***	***Jitter %***	***Quality %***	***Difficulty %***
1	*200*	*0*	*0*	***93.6***	***0***
2	*400*	*5*	*0*	***63***	***7.5***
3	*500*	*1*	*1*	***57***	***13.5***

Table (7.13)

7.6 Comments

Comparing Tables (7.1), (7.12) and (7.13), the following remarks can be stated:

1. Despite I can not claim that results were accurate, one can find voice compression feasible. No substantial degradation to voice quality happens when using Codecs 1 or 2 (<13 and 13 kbps respectively) compared to

Codec 3 (64 kbps) while bandwidth saving is very large. Codec number 1 is very efficient even compared to Codec number 2, that is because -as discussed before- at least 50% of the time is silence for both parties consequently, occupied bandwidth is minimized to a maximum of 6.5 kbps.

2. The downgrading in Full GSM with silence detector compared to normal Full GSM stems from the function of silence detector, which senses the existence of speech in order to enable other processes (speech quantizing, compression, packetization..etc). It could be understood that a slight amount of speech loss would necessarily come off during those critical events. This drawback of this Codec can be thought of to be comparable to slight additional packet loss
3. Drawing a graph for three Codecs using the three most frequent cases[76]-as shown in figure (7.6)- MOS improvement can be imaged. In fact it is not very similar to the figure 3/P.830 in ITU-T recommendation P.830. Main difference is that the slope in figure (7.6) is less, a thing eventually means efficient Codecs!

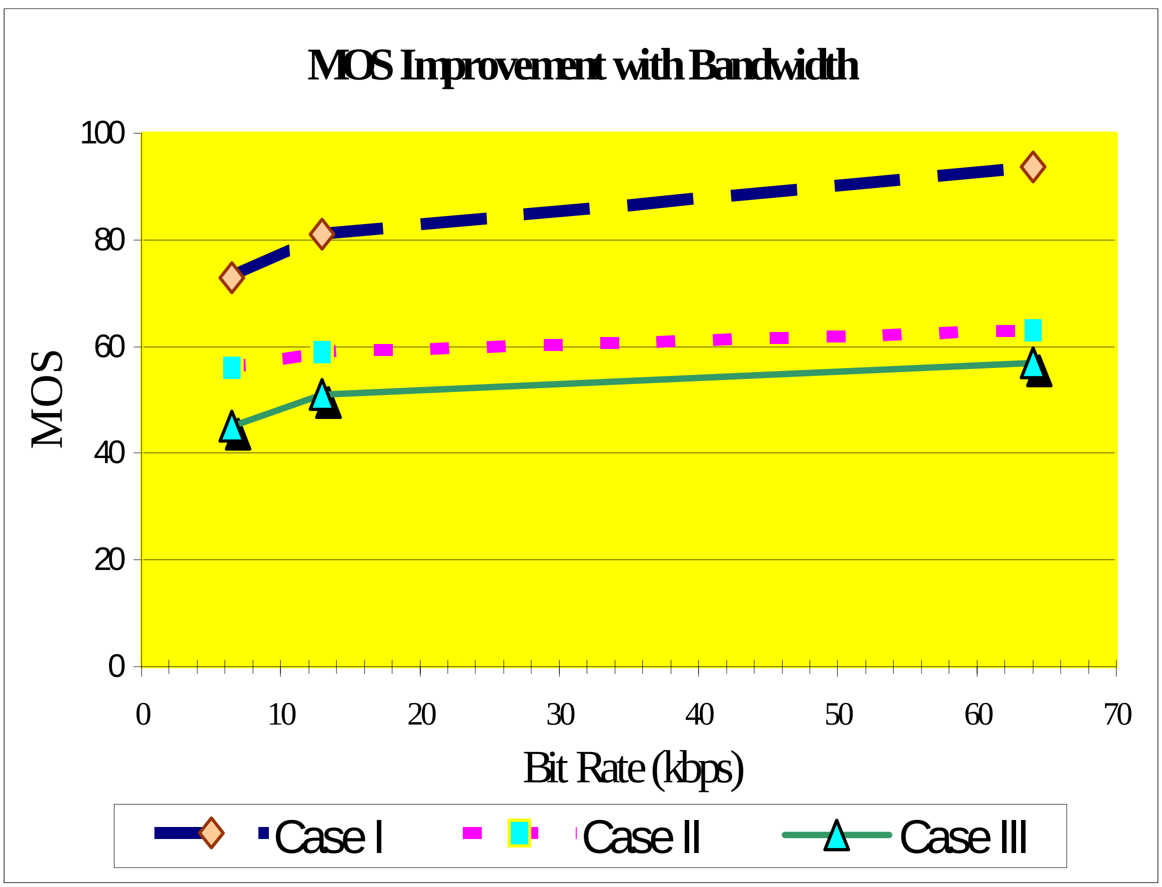

Figure (7.6)

[76] The ones I tested in the third Codec in table (7.13).

Case I means, tests transacted at a delay value of 200 with zero packet loss and zero jitter. Case II touches the conditions of 400 ms, 5% packet loss and zero jitter. Finally, case III is the one underwhich delay was 500 ms, 1% was the value for both of jitter and packet loss.

4. In all conversation tests I mentioned only the average value for MOS's but not Standard deviations. Despite they are important to state from statistical point of view, I found those value –in nearly all cases- very small and hence discarded them.

7.7 Complementary Test

Making a subjective test for VoIP as well as verifying integration of VG to the PBX means making more test. Had I more time, I would make the following additional tests.

1. Adding Noise to both of conversation and listening tests. That is repeating test for different levels of SNR.
2. Loading of the VG, by making simultaneous telephone calls at the same time.

In the light of the given time for this project, I would say that most important test have been done.

Abbreviations:

ADSL: Asymmetric Digital Subscriber Loop

ANSI: American National Standards Institute

ARPA: Advanced Research Projects Agency

ARPANET: Advanced Research Projects Agency Network

ASE: Application Service Element.

CDR: Call Detail Records

CRC: Cyclic Redundancy Check

CODEC: COmpressor / DECompressor
or COder / DECoder

DARPA: Defense Advanced Research Project Agency.

DSP: Digital Signal Processing.

DTMF: Dual Tone Multi Frequency

ETSI: European Telecommunications Standards Institute

IETF: Internet Engineering Task Force.

IPBX: Internet Private Branch eXchange.

IPng: next generation of IP.

ISDN: Integrated Services Digital Network
.

ISP: Internet Service Provider.

ITSP: Internet Telephony Service Provider.

IVR: Interactive Voice Response.

LAN: Local Area Network.

LL: Link (or Network Interface) Layer.

LSB: Least Significant Bit.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

MCU: Multipoint Control Unit.

MPEG: Moving Pictures Experts Group

MSB: Most Significant Bit

NK: Net Keeper

NTP: Network Time Protocol

OSI Open Systems Interconnect

POTS: Plain Old Telephone Service

PRI: Primary Rate Interface

RAS: Remote Access Services

RFC: Request for Comment.

SDH: Synchronous Digital Hierarchy

SK: Site Keeper

SNR : Signal to Noise Ratio

SQL: Structured Query Language

TAPI: Telephony Application Programming Interface.

VG: Voice Gateway

VoIP: Voice over IP

VPN: Virtual Private Network

References

1. International Telecommunication Union: ITU-T P.10, Vocabulary of terms on telephone transmission quality and telephone sets, March 1993.
2. ITU-T P.84, Subjective (listening) test methods for evaluating digital circuits multiplication and packetized voice systems, March 1993.
3. ITU-T P.85, A method for subjective performance assessment of the quality of speech voice output devices, June 1996.
4. ITU-T P.800, Methods for subjective determination of transmission quality, August 1998.
5. ITU-T P.810, Modulated Noise Reference Unit, February 1996.
6. ITU-T P.830, Subjective Performance assessment of telephone-band and wideband digital CODECs, February 1996.
7. ITU-T P.861, Subjective quality measurements of telephone band (300-3400Hz) speech codecs, August 1996.
8. ITU-T H.323, Visual telephone systems and equipment for local area network which provides a non-guaranteed quality of service, November 1996.
9. ITU-T G.711, Pulse code modulation (PCM) of voice frequencies,1988.
10. ITU-T G.722, 7 kHz audio-coding within 64 kbit/s, 1988.
11. ITU-T G.723, Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s, 1996.
12. ITU-T G.726, 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM), 1990.
13. ITU-T G.727, 5-, 4-, 3- and 2-bits sample embedded adaptive differential pulse code modulation (ADPCM), 1990.
14. ITU-T G.728, Coding of speech at 16 kbit/s using low-delay code excited linear prediction, 1992.
15. ITU-T G.729, Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP), 1996.
16. ITU: CCITT, Handbook on Telephonometry, Geneva 1987.
17. European Telecommunications Standard Institute, Telecommunications and Internet Protocol Harmonization Over Network (TIPHONE): Documents (V0.3.4, V1.1.1, V.1.1.3, V.1.3.2). Available on the web:
    http://tnweb.tn.etx.ericsson.se/x_s/x_si/tiphoner/
    ETSI web site is on:
    www.etsi.fr or www.etsi.org.
18. Data Communications Magazine: available on the web
    http://www.data.com/
19. Webmedia pc: available on the web.
    http://webopedia.internet.com/
20. Phone Zone on line magazine: available on the web:
    http://www.phonezone.com/tutorial/ip-phone.htm
21. Computer Telephony magazine: available on the web:
    www.computertelephony.com
22. Computer Dictionary of Swedish University NETwork (SUNET): available on the web:
    http://ftp.sunet.se/foldoc/
23. Pulver Web Site: http://pulver.com

24. IEFT, RFC 1122, Requirements for Internet Hosts: Communication Layers, October 1989.
25. Internet Traffic Report web site: http://www.InternetTrafficReport.com/index.html

Annex A

Summary of ITU-T Recommendation: H.323[77]

Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service

Recommendation H.323 describes terminals, equipment and services for multimedia communication over Local Area Networks (LAN) which do not provide a guaranteed quality of service. H.323 terminals and equipment may carry real-time voice, data and video, or any combination, including Videotelephony. The LAN over which H.323 terminals communicate, may be a single segment or ring, or it may be multiple segments with complex topologies.
It should be noted that operation of H.323 terminals over the multiple LAN segments (including the Internet) may result in poor performance. The possible means by which quality of service might be assured on such types of LANs/ internetworks is beyond the scope of this Recommendation.H.323 terminals may be integrated into personal computers or implemented in stand-alone devices such as videotelephones.
Support for voice is mandatory, while data and video are optional, but if supported, the ability to use a specified common mode of operation is required, so that all terminals supporting that media type can interwork. This Recommendation allows more than one channel of each type to be in use.
Other Recommendations in the H.323-Series include H.225.0 packet and synchronization, H.245 control, H.261 and H.263 video codecs, G.711, G.722, G.728, G.729, and G.723 audio codecs, and the T.120-Series of multimedia communications protocols.
This Recommendation makes use of the logical channel signaling procedures of Recommendation H.245, in which the content of each logical channel is described when the channel is opened. Procedures are provided for expression of receiver and transmitter capabilities, so transmissions are limited to what receivers can decode, and so that receivers may request a particular desired mode from transmitters.
Since the procedures of Recommendation H.245 are also used by Recommendation H.310 for ATM networks, Recommendation H.324 for GSTN, and V.70, interworking with these systems should not require H.242 to H.245 translation as would be the case for H.320 systems.H.323 terminals may be used in multipoint configurations, and may interwork with H.310 terminals on B-ISDN, H.320 terminals on N-ISDN, H.321 terminals on B-ISDN, H.322 terminals on Guaranteed Quality of Service LANs, H.324 terminals on GSTN and wireless networks, and V.70 terminals on GSTN.

[77] Quoted from ITU-T literatures.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Annex B

Questionnaire for Conversation – opinion Test:

Name: -----------------------------------

Please answer the following questions carefully and **individually**.
Your opinion is very important!

1. **Opinion of the connection you have just been using.**
   Please choose the most accurate description of what you think of the quality of voice and goodness of connection from following 11 points scale

   10 Perfect
   9 Excellent
   8 Very Good
   7 Good
   6 Moderately Good
   5 Fair
   4 Somewhat poor
   3 Poor
   2 Very Poor
   1 Bad
   0 Useless

2. **Did you –or your partner- have any difficulty in talking or hearing over the connection?**
   Yes
   No

If the answer is yes please, state a percentage approximation for the difficulty (next to the word yes) with comments if you have.

Comments: ---

Thanks for your time!

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Annex C

Questionnaire for Degradation Test:

Name: -----------------------------

Please answer the following question carefully and **individually**.
Your opinion is very important!

* **What difference you feel between the first and second sentences: in terms of voice quality?**

Possible Answers:

¤ Less than 5%: Identical: They are confounded

¤ 5% -10% :Undetectable: It is difficult to find difference

¤ 11% -18%: Just detectable: difference can be detected by listening carefully.

¤ 19% - 25%: Slight: Difference is detectable but not very clear.

¤ 26%-33%: Moderate: You feel the difference easily.

¤ 34% -42%: Somewhat big: Difference is substantial.

¤ 43% 50%: Big: Difference is very clear.

¤ Greater than 50%: Very big: The difference is shouting.

Comments

Thanks for your time!

Annex D

Suggested Messages for Degradation Test

Annex D.1: Read by Ingela

(M): Order beställning
(R): Trafik information tåg

M1: Fru Larsson, löparskors färg : vit , storlek : 11, referensnummer : 501-97-52 pris : 500 kronor, kommer att levereras inom en vecka.

M2: Herr Johnson, multistandard TV-set med fjärrkontroll, 36 cm skärm, referensnummer: 811-61-32, pris: 3000 kronor, kommer att levereras inom 3 veckor.

M3: Herr Moore, den elektriska borren D162, energi: 550 watt 2 hastigheten, referensnummer: 481-20- 30, pris : 500 kronor, kommer att levereras inom 2 veckor.

R1: Tåg nummer 9783 från Göteborg anländer kl: 9.24, till plattform nummer 3, spår G.

R2: Tåg nummer 7826 till Helsingborg avgår kl: 12.20, från plattform nummer 9, spår A.

R3: Tåg nummer 4320 från Stockholm anländer kl:5.44, till plattform 2, spår C.

Translation of Annex D.1[78]

(M): Mail Order Shopping
(R): Railway Traffic Information
Three messages are given for each application.

M1: Mrs. Larsson, the running shoes color: white, size: 11, reference: 501-97-52, price: 500 kronors, will be delivered to you in 1 week.

M2: Mr. Johnsson, the multistandard TV set with remote control, 36 cm screen, reference: 811-61-32, price: 3000 kronors, will be delivered to you in 3 weeks.

M3: Mr. Moore, the electric drill D162, power: 550 watts, 2 speeds, reference: 481-20-30, price: 500 kronors, will be delivered to you in 2 weeks.

R1: The train number 9783 from Gothenburg will arrive at 9:24, platform number 3, track G.

R2: The train number 7826 to Helsinborg will leave at 12:20, platform number 9, track A.

R3: The train number 4320 from Stockholm will arrive at 5:44, platform 2, track C.

[78] Refer to Annex A in ITU-T, P.85.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*	Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz			
Dokansv/Godk - *Doc respons/Approved* Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin			

Annex D.2

Read by : Åsa.

(M): Order beställning
(R): Trafik information tåg

M1: Fru Andersson löparskors färg : vit , storlek : 7, referensnummer: 396-85-47, pris : 400 kronor, kommer att levereras inom fyra veckor.

M2: Herr Persson, multistandard TV-set med fjärrkontroll, 54 cm skärm, referensnummer: 622-59-32, pris: 5000 kronor, kommer att levereras inom 6 veckor.

M3: Herr Svensson, den elektriska borren B162, energi: 320 watt 3 hastigheten, referensnummer: 598-32-40, pris : 700 kronor, kommer att levereras inom 8 veckor.

R1: Tåg nummer 1056 från Borlänge anländer kl:13.50, till plattform nummer 7, spår D.

R2: Tåg nummer 2001 till Kalmar avgår kl:18.30, från plattform nummer 12, spår I.

R3: Tåg nummer 5610 från Fagersta anländer kl:4.20, till plattform 7, spår E.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Translation of Annex D.2

(M): Mail Order Shopping
(R): Railway Traffic Information
Three messages are given for each application.

M1: Mrs. Andersson, the running shoes color: white, size: 7, reference: 396-85-47, price: 400 kronors, will be delivered to you in four weeks.

M2: Mr. Persson, the multistandard TV set with remote control, 54 cm screen, reference: 622-59-32, price: 5000 kronors, will be delivered to you in 6 weeks.

M3: Mr. Svensson, the electric drill B162, power: 320 watts, 3 speeds, reference: 598-32-40, price: 700 kronors, will be delivered to you in 8 weeks.

R1: The train number 1056 from Borlänge will arrive at 13:50, platform number 7, track D.

R2: The train number 2001 to Kalmar will leave at 18:30, platform number 12, track I.

R3: The train number 5610 from Fagersta will arrive at 4:20, platform 7, track E.

Annex D.3

Read by : Niklas

(M): Order beställning
(R): Trafik information tåg

M1: Tant Olga, löparskors färg : rosa, storlek : 8, referensnummer :861-99-40, pris : 700 kronor, kommer att levereras inom sex veckor.

M2: Herr Torsson, multistandard TV-set med fjärrkontroll, 60 tums skärm, referensnummer: 912-20-11, pris: 4000 kronor, kommer att levereras inom 4 veckor.

M3: Herr Pettsson, den elektriska borren C200, effekt: 400 watt 5 hastigheten, referensnummer: 691-76-50, pris : 900 kronor, kommer att levereras inom 5 veckor.

R1: Tåg nummer 2347 till Ystad anländer kl:18.15, till plattform nummer 9, spår E.

R2: Tåg nummer 3856 från Karlstad avgår kl:20.20, från plattform nummer 15, spår J.

R3: Tåg nummer 8320 till Skellefteå anländer kl:03.18, till plattform 10, spår F.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Translation of Annex D.3

(M): Mail Order Shopping
(R): Railway Traffic Information
Three messages are given for each application.

M1: Aunt Olga, the running shoes color: rose, size: 8, reference: 861-99-40, price: 700 kronors, will be delivered to you in six weeks.

M2: Mr. Torsson, the multistandard TV set with remote control, 60 cm screen, reference: 912-20-11, price: 4000 kronors, will be delivered to you in 6 weeks.

M3: Mr. Pettsson, the electric drill C200, power: 400 watts, 5 speeds, reference: 691-76-50, price: 900 kronors, will be delivered to you in 5 weeks.

R1: The train number 2347 to Ystad will leave at 18:15, platform number 9, track E.

R2: The train number 3856 from Karlstad will arrive at 20:20, platform number 15, track J.

R3: The train number 8320 to Fagersta will leave at 03:18, platform 10, track F.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Annex D.4

Read by : Boston

(M): Order beställning
(R): Trafik information tåg

M1: Fröken Nora, löparskors färg : blå, storlek : 90, referensnummer :105 - 83 - 50, pris : 60 kronor, kommer att levereras inom 12 dagar.

M2: Herr Olsson, multistandard TV-set med fjärrkontroll, 27tums bild skärm, referensnummer: 911 - 85 - 61, pris: 4000 kronor, kommer att levereras inom 50 timmar.

M3: Herr Nilsson, den elektriska borren M581, effekt: 1000 watt 5 hastigheten, referensnummer: 392 - 91 - 78, pris : 900 kronor, kommer att levereras inom 1 månad.

R1: Tåg nummer 3028 till Malmö anländer kl: 18.10, till plattform nummer 23, spår A.

R2: Tåg nummer 365 från Uppsala avgår kl: 11.45, från plattform nummer 18, spår S.

R3: Tåg nummer 4362 till Luleå anländer kl: 02.49, till plattform 4, spår H.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Translation of Annex D.4

(M): Mail Order Shopping
(R): Railway Traffic Information
Three messages are given for each application.

M1: Miss Nora, the running shoes color: Blue, size: 90, reference: 105-83-50, price: 60 kronors, will be delivered to you in 12 days.

M2: Mr. Olsson, the multistandard TV set with remote control, 27 inch screen, reference: 911-85-61, price: 4000 kronors, will be delivered to you in 50 hours.

M3: Mr. Nilsson, the electric drill M581, power: 1000 watts, 5 speeds, reference: 392-91-78, price: 900 kronors, will be delivered to you in 1 month.

R1: The train number 3028 to Malmö will leave at 18:10, platform number 23, track A.

R2: The train number 365 from Uppsala will arrive at 11:45, platform number 18, track S.

R3: The train number 4362 to Luleå will leave at 02:49, platform 4, track H.

Annex D.5

Read by Marie Gabrielsson

M1: Fröken Nora, löparskors färg : blå, storlek : 18, referensnummer: 597 - 20 - 10, pris : 30 kronor, kommer att levereras inom 15 dagar.

M2: Herr Olsson, multistandard TV-setet med fjärrkontroll, 6 dm skärm, referensnummer: 830 - 49 - 40, pris: 5550 kronor, kommer att levereras inom 40 timmar.

M3: Herr Nilsson, den elektriska borren N690, effekt: 1500 watt 6 hastigheten, referensnummer: 453 - 82 - 52, pris : 1000 kronor, kommer att levereras inom 6 veckor.

R1: Tåg nummer 5166 till Växjö anländer kl: 19.45, till plattform nummer 18, spår U.

R2: Tåg nummer 635 från Umeå avgår kl: 21.30, från plattform nummer 105, spår P.

R3: Tåg nummer 3256 till Borås anländer kl: 1.48, till plattform 6, spår O.

Uppgjord (även faktaansvarig om annan) - *Prepared (also subject responsible if other)*		Nr - *No.*		
NA/EBC/EN/DAS Ibrahim Qazzaz				
Dokansv/Godk - *Doc respons/Approved*	Kontr -	Datum - *Date*	Rev	File
NA/EBC/EN/DAS Lars Romin				

Translation of Annex D.5

(M): Mail Order Shopping
(R): Railway Traffic Information
Three messages are given for each application.

M1: Miss Nora, the running shoes color: Blue, size: 18, reference: 597-20-10, price: 30 kronors, will be delivered to you in 15 days.

M2: Mr. Olsson, the multistandard TV set with remote control, 6 decimeter screen, reference: 830-49-40, price: 5550 kronors, will be delivered to you in 40 hours.

M3: Mr. Nilsson, the electric drill N690, power: 1500 watts, 6 speeds, reference: 453-82-52, price: 1000 kronors, will be delivered to you in 6 weeks.

R1: The train number 5166 to Växjö will leave at 19:45, platform number 18, track U.

R2: The train number 635 from Umeå will arrive at 21:30, platform number 105, track P.

R3: The train number 3256 to Borås will leave at 1:48, platform 4, track O.

Annex E

Frågor (Enkät) till **förtydligande** testet

Namn: --

Var vänlig svara på nedanstående frågor noggrant och enskilt.
Ditt svar är mycket betydelsefullt!

De tre uppgifterna relaterade till orderbeställnings formuläret.

Namnet på varan
(1-3 ord): --

Referens nummer --

Pris ---

Leveranstid ---

De tre uppgifterna relaterade till trafik informations formuläret.

Tåg numret ---

Till eller från --

Tid --

Plattform --

Spår --

Translation of Annex E

Questionnaire for Clearness Test

Name: ----------------------

Please answer the questions below carefully and Individually.
Your answer is very important.

These three questions are related to the Mail Order Shopping.

Name of Item (1-3) words --

Reference Number ---

Price --

Availability ---

These three questions are related to the Railway Traffic Information.

Train Number --

To or From --

Time --

Platform --

Track --

Annex F

Suggested form for Clearness Test

Questionnaire for Clearness Test:

(To be completed after finishing all listening sessions)

Name: ---------------------------------

Please answer the following question carefully and **individually**.
Your opinion is very important!

1. Overall Impression[79]:

10 Perfect
9 Excellent
8 Very Good
7 Good
6 Moderately Good
5 Fair
4 Somewhat poor
3 Poor
2 Very Poor
1 Bad
0 Useless

2. Listening Effort:

How would describe the effort you were required to make in order to understand the message?

Options:

Complete Relaxation possible; no effort required.
Attention Necessary: No appreciable effort required.
Moderate effort required.
Effort required.
No meaning understood with any feasible effort.

1 of 2

[79] Here I have widened the scale of P.85 and quoted the one in “Handbook on Telephonometry” again for a wider range of marking.

Cont. - Annex F

3. Comprehension Problem:

Did you find certain words hard to understand?

Options:

Never.
Rarely.
Occasionally.
Often.
All of the time.

4. Articulation:

Were the sounds distinguishable?

Options:

Yes, very clear.
Yes, clear enough.
Fairly clear.
No, not very clear
No, not at all.

5. Acceptance Test:

Do you think this voice could be used for such an information service by telephone?

Options:

Yes.
No.

6. Speaking Rate:

The average speed of talker was:

Much faster than preferred.
Faster than preferred.
Preferred.
Slower than preferred.
Much slower than preferred

Comments:

Thanks for your time!

2 of 2