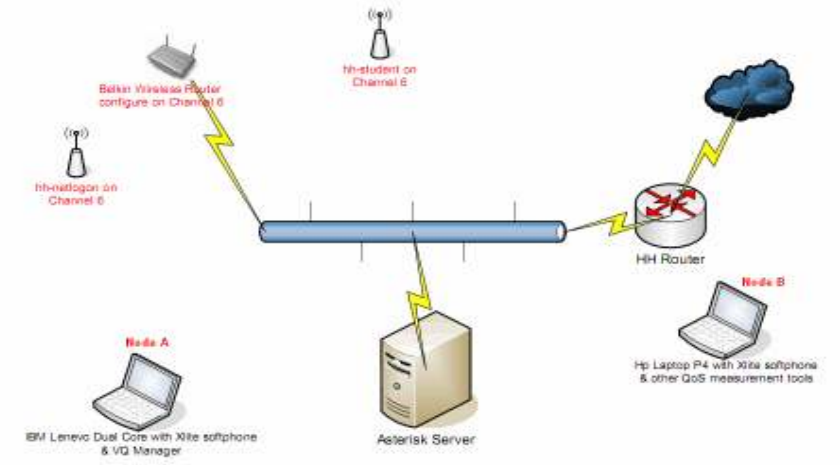
Technical report, IDE0958, November 2009

Evaluation of VoIP Codecs over 802.11 Wireless Networks

(A Measurement Study)

Master's Thesis in Computer Network Engineering

Arbab Nazar



School of Information Science, Computer and Electrical Engineering
Halmstad University

Preface

This report is submitted in partial fulfillment of the requirements for the degree of Master in Computer Network Engineering in the School of Information Science, Computer and Electrical Engineering at Halmstad University Sweden.

I express my profound and cordial gratitude to offer thanks to my learned, kind and experienced supervisor **Per-Arne Wiberg**, for his kind behaviour, encouraging attitude, constructive suggestions and ever-helping supervision.

My sincerest thanks are to my kind friend **Sohail Akhtar** who willingly helped me during my project work.

Arbab Nazar
Halmstad University, November 2009

Abstract

Voice over Internet Protocol (VoIP) has become very popular in recent days and become the first choice of small to medium companies for voice and data integration in order to cut down the cost and use the IT resources in much more efficient way. Another popular technology that is ruling the world after the year 2000 is 802.11 wireless networks. The Organization wants to implement the VoIP on the wireless network. The wireless medium has different nature and requirement than the 802.3 (Ethernet) and special consideration take into account while implementing the VoIP over wireless network.

One of the major differences between 802.11 and 802.3 is the bandwidth availability. When we implement the VoIP over 802.11, we must use the available bandwidth is an efficient way that the VoIP application use as less bandwidth as possible while retaining the good voice quality. In our project, we evaluated the different compression and decompression (CODEC) schemes over the wireless network for VoIP.

To conduct this test we used two computers for comparing and evaluating performance between different CODEC. One dedicated system is used as Asterisk server, which is open source PBX software that is ready to use for main stream VoIP implementation. Our main focus was on the end-to-end delay, jitter and packet loss for VoIP transmission for different CODECs under the different circumstances in the wireless network. The study also analyzed the VoIP codec selection based on the Mean Opinion Score (MOS) delivered by the softphone. In the end, we made a comparison between all the proposed CODECs based on all the results and suggested the one Codec that performs well in wireless network.

Table of Contents

1. INTRODUCTION:

VoIP (Voice over Internet Protocol) is a technology that gives you the feature of making calls using the internet connection instead of using your analog phone line. Some VoIP services allow you to make call to other people who are using the same VoIP services while other give you chance to make call to anyone who has telephone number (including local and international). Some VoIP services only work on computer or special type of VoIP phones while other services work on traditional phone if you connect it with VoIP adapter. [1]

VoIP technology used IP based networks to carry the voice. With VoIP, the service providers can offer the telephony services as well as the traditional data service using the same existing IP infrastructure. By doing this, it will not only increase their revenue but customers also save a lot of money while comparing to buy these services individual.

In recent years, growth of VoIP has been increased dramatically which gains a lot of attention from the network engineering research communities' consequence the result of rapid commercial solution as well as the network improvements. Other factors that make this popular are ongoing decrease in quality differences between the existing PSTN (Public Switched Telephone Networks) telephony and VoIP, as well as the increase in the bandwidth now a day that are used by commercial and home users for transportation of VoIP services.[2]

Today most IP networks are not designed for carrying the real time and delay sensitive data (voice and video). Current IP networks only provide the best-effort service and also there is no guarantee that the VoIP speech quality will be the same like that PSTN speech quality. [2]

The implementation of VoIP in wireless networks are rapidly increasing and the reason behind it is mobility. The wireless IP network gives the benefit of mobility to its users while the wireless networks have their own special characteristics. It is more challenging task when wireless networks are used to carrying the real-time traffic with the data traffic. [2] Because VoIP is real-time application and it is particularly sensitive to packet loss that can be caused in the wireless networks by the weak signals, limitation in coverage area, and interference from other device that are using the same frequency range as the wireless networks are using. [3]

In order to deploy the VoIP over wireless networks, one should meet a lot of specific requirements for efficient transmission of voice over wireless networks. Because of the delay sensitivity of VoIP applications, competition for transmission with the data on the same wireless medium causes the degradation in voice quality. So it is important to properly implement QoS in wireless networks, because it gives the priority to voice packets over data packets. Security is also a major problem in VoIP; wireless adds

another layer of security concerns, as the packets are transmitting over the open medium (air) instead of cables. So it is easy to capture these packets and convert them back to wav form. Common VoIP protocols like SIP (Session Initiation Protocol) have their own security vulnerabilities. [3]

End-to-end delay and jitter have significant impact on the quality of voice in VoIP. In our project, end-to-end delay as well as jitter will be investigated and analyzed for different codec and suggested the best codec upon these parameters.

There are a number of CODEC used for VoIP these days and each of them have its different and unique characteristics. When different codecs are used and the information that we want to transmit is arranged into different frame sizes, then the transmission result also change. This result is more prominent when the conversion of transport stream changes into the Real-time Transport Protocol (RTP) packets.

Normally, one or more voice frames are put into one RTP packet, RTP packets are then put into the UDP packets and at the end into IP packets before transmitting across the network. IP packets are then encapsulated into MAC frames switch from one node to another. Delays are added during the processing that occurs at each node while across the path from source to destination. The measurement of this delay is of our main interest in this project and to check its impact on the quality of voice.

2. Wireless Network:

A wireless network is any kind of computer network that is connected wirelessly, meaning that the nodes are connected to each other or to the telecommunications network (which connected them to the internet or backbone wired network) without the need of wires. Wireless networks use the electromagnetic waves (commonly radio waves) for carrying the signals and data between the nodes and it is implemented at the physical layer meant to replace the wires. [4]

2.1 Types of Wireless Network:

Some common types of wireless networks are as follows:

1. Personal Area Network (PAN)
2. Wireless Local Area Network (WLAN)
3. Wireless Metropolitan Area Network (WMAN)

2.1.1. Personal Area Network (PAN):

Wireless PAN connects the devices in a relatively small area, which is generally a few meters, using radio waves. It can be used to communicate among the devices (e.g. Bluetooth uses PAN to connect the wireless headset to a laptop, or wireless mice to laptops) or connects the device to the backbone network and Internet. We can also create Wireless PAN with other network technologies like Z-Wave, Bluetooth and IrDA. [5]

2.1.2. Wireless Local Area Network (WLAN):

Wireless Local Area Network (WLAN) is an alternative way to connect computers and devices in Local Area Network (LAN) by using radio waves, while LAN technology uses Ethernet cable (e.g. Cat 5) to connect devices and computers for the purpose of communication within a small area such as home, office or a device within one building. The series of 802.11 are referred to as Wireless LAN.

2.1.3. Wireless Metropolitan Area Network (WMAN)

Wireless MAN is used to connect the two or more networks that are a distance away from each other, like in different cities. It is also used to provide the WLAN services, like Internet, to the entire city. **WiMAX** is usually used as the reference of the Wireless MAN.

2.2 Wireless LAN

Wireless LAN connects two or more devices using **Orthogonal Frequency-Division Multiplexing (OFDM)** or **Direct Sequence Spread Spectrum (DSSS)** modulation techniques to establish communication between devices within a limited range. This gives the advantage that users can move freely within the prescribed area without the fear of disconnection to the network, or the burden of changing the position of the wire from one jack to another. [6]

At home, the Wireless LAN is used because of its easy to installation, to get rid of wires, to move freely and to avoid the drill of every time having to add a new jack when adding a new computer. It is also less expensive then the Wire LAN. At coffee shops and malls, it is used to attract the costumers so that they can surf the Internet while having a coffee or shopping.

Fig 2.1:Typical use WLAN at home

2.2.1 Facts about Wireless LAN:

A Wireless LAN Access Point works like a HUB, which means that only one user, can send/receive data at any given time, Shared Signal and Half Duplex. It uses the unlicensed bands of Radio Frequency (RF). Wireless LAN uses the **Carrier Sense Multiple Access / Collision Avoidance (CSMA/CA)** instead of **Carrier Sense Multiple Access / Collision Detection (CSMA/CD).** [7]

2.2.2 Benefits of Wireless LAN:

Wireless LAN has a lot of benefits, here are some of them:

2.2.2.1 Mobility:

With the development of public wireless network, a user can access the Internet even if he is not within his normal working area. For example, almost all the coffee shops and big malls are offering free of charge Internet to their customers.

2.2.2.2 Cost Stability:

As compared to the Wire LAN, every time you add a new device, you have to run cable, while with Wireless LAN, include an AP/Wireless Router and you can connect a sufficient number of devices to it without any further cost.

2.2.2.3Easy to Install:

It is really easy to install the Wireless AP/Router, even for a home user. To install the wireless equipment at home or SOHO (Small Office Home Office), there is no need for a technician. On the other hand, you need a technician every time for installing an RJ45 jack, or running a ceiling cable.

2.2.3 Deficiency in Wireless LAN:

There are also some weaknesses in Wireless LAN, but we can overcome these weaknesses by using the correct design model and strategy:

2.2.3.1Security:

In Wire LAN, the hacker or malicious person must have to be inside the building, and have the access to the RJ-45 jack, to attack the network or sniff the packets.

However, for Wireless LAN, the situation is totally different because of the nature of radio transmission the intruder does not have to be inside the building. Radio signals leak outside the building and anyone within the range can use them to access the internal network of the company, due to which we must have to implement proper security, so that no unauthorized person can access our Wireless LAN. [9]

2.2.4 Types of Security:

2.2.4.1 WEP (Wired Equivalent Privacy):

WEP is the original wireless security standard, release in 1997, to secure the wireless networks, but it is the weakest form of wireless security. It consist of 40/104 bits WEP static key + 24 bits Initialization Vector (IV).

40/104 bits WEP static key	24 bits (IV)

Static WEP key is always the same, while the IV changes for every packet. It sends the keys in a beacon header and if you are able to catch enough packets, then you can break the security. In 2005, an FBI team made an experiment with WEP and broke it in 3 minutes. [7]

2.2.4.2 WPA (Wi-Fi Protected Access):

It was released by the Wi-Fi Alliance to overcome the weakness of WEP in 2003. It uses the same hardware as WEP but it gives the far better security than WEP. It gave the three solutions: Temporal Key Integrity Protocol (TKIP), Message Integrity Code (MIC), and 802.1x. [7]

2.2.4.3 WPA2 (IEEE 802.11i):

This was released by the Wi-Fi Alliance in 2004. It needs hardware upgrade, and uses the AES (Advance Encryption Standard) as encryption. It is backward compatible with TKIP-hardware also, which means that if client connect to AP that only support TKIP encryption, it provides that, and if the client supports the AES, then it will provide this. [7]

2.2.5 Interference:

Because the Wireless LAN uses the unlicensed band, so the other devices in the wireless network that are using the same channel, such as microwave ovens and cordless phones, can interfere the wireless signal, and this can create significant impacts on the performance of Wireless LAN. Properly designed networks can mitigate the impact of interference. [10]

2.2.6 Architecture of Wireless LAN

2.2.6.1 Station:

All devices that have the ability to connect to the wireless medium are called “stations”. All stations have wireless NICs. Wireless station can be Access Point (AP) or clients. Access Points are devices that are capable of sending and receiving RF to the

backbone network for wireless clients that associate with it. Wireless clients are any devices that have Wireless NIC, e.g. laptops, IP phones, PDA etc.

2.2.6.2 Basic Service Set:

Basic Service Set (BSS) is a collection of all the Access Point and clients that can communicate with each other. There are two kinds of BSS, Independent BSS and Infrastructure BSS. Each BSS has its ID, which is known as SSID. [8]

2.2.6.3 Extended Service Set & Distribution System:

Extended Service Set is a collection of more than one BSS and Distribution System is used to connect the Access Points of BSS in Extended Service Set. [8]

2.2.7 Roaming in Wireless LAN:

There are two types of roaming in Wireless LAN:

2.2.7.1 Layer 2 Roaming:

In layer 2 roaming, mobile nodes migrate from one access point (AP) to other APs, but within the same network. The migration may be due to the fact that the node missed too many beacons, data reaches maximum retry count or data rate is shift down. Mobile node retains its previous open sessions using some software mechanism. [11]

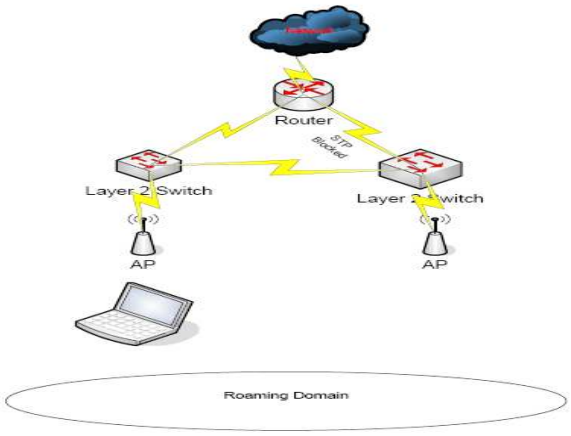


Fig 2.2: Layer 2 Roaming

2.2.7.2 Layer 3 Roaming (a.k.a Mobile IP):

In layer 3 Roaming, mobile node left its home network and goes to the foreign network (network other than the one the node belongs to). There is some special method used to authenticate the mobile node in the foreign network. The mobile node lost all of its previous open sessions. [12]

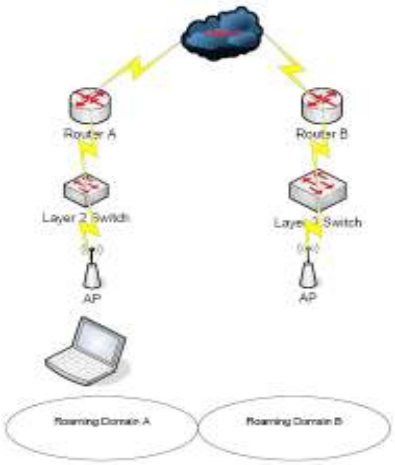


Fig 2.3: Layer 3 Roaming

2.3 802.11 Standards:

The IEEE 802.11 is a collection of standards that are used to carry the data in Wireless LAN using the RF waves in the range of 2.4 GHz and 5.0 GHz. A lot of amendments have been made in original 802.11 standards to make it more efficient and feasible for home users and industry. Some of them are here:

2.3.1 802.11a:

Officially, this was released in September 1999. It uses the RF range of 5.0 GHz. It uses the **Orthogonal Frequency-Division Multiplexing (OFDM)** with 52 subcarriers, to carry the data and gives the data rate of 54 Mbps theoretically with 8 Data Rates of 6, 9, 12, 18, 24, 36, 48 and 54 Mbps, but it actually gives the throughput of 27 Mbps. It is not cross-compatible with 802.11b/g but it has more "Clean" channels of 12 to 23. [13]

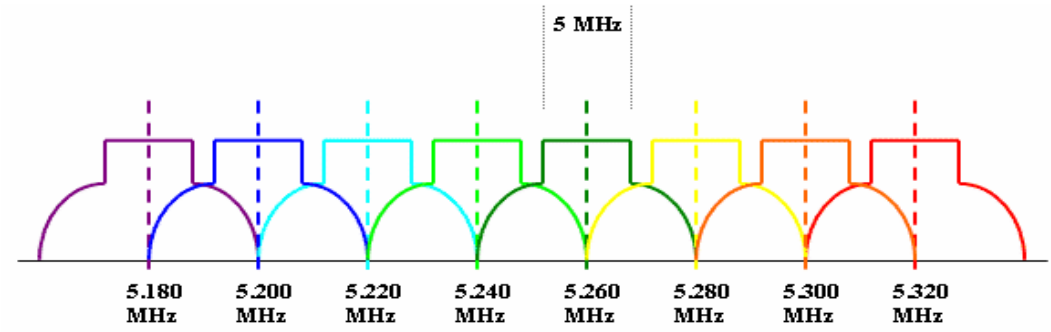


Fig 2.4:802.11a channel [13]

2.3.2 802.11b:

This was also officially released as of September 1999. It uses the RF range of 2.4 GHz and **Direct Sequence Spread Spectrum (DSSS)** to carry the data, and gives the data rate of 11 Mbps theoretically, with 4 Data Rates of 1,2,5.5 and 11 Mbps but it actually gives the throughput of 5 Mbps. It is the most popular standard in 802.11 line-up but it only gives 3 "Clean" channels. The major problem is that 802.11b station cannot decode 802.11b radio signals. [13]

2.3.3 802.11g:

This was released in June 2003 by IEEE. It uses the same RF range of 2.4 GHz as used by 802.11b with **Direct Sequence Spread Spectrum (DSSS) & Orthogonal Frequency-Division Multiplexing (OFDM)** to carry the data, and gives the data rate of upto 54 Mbps, with 12 Data Rates, but throughput of 22 Mbps. It is totally backward compatible, with 802.11b hardware, and gives the same three "Clean" channels. It also suffers the same kind of interference as 802.11b, because this RF range is used by all kinds of devices like microwave ovens, cordless phones etc. [13]

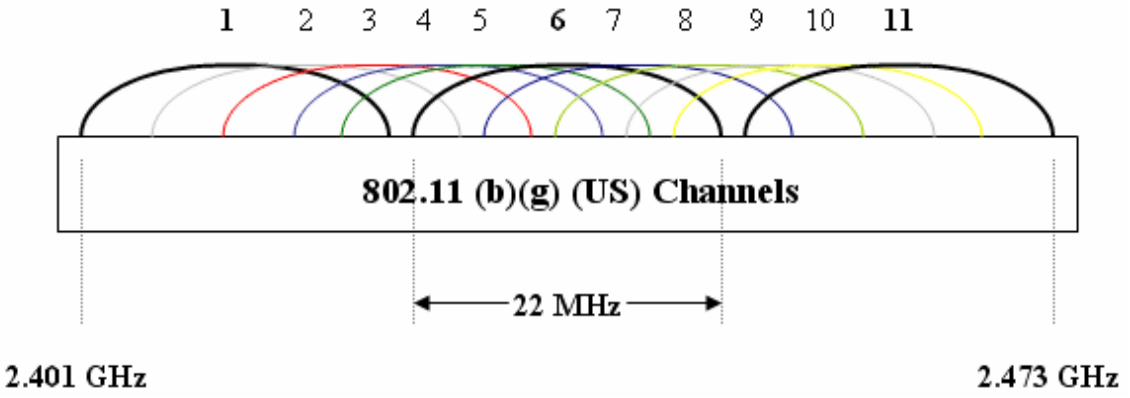


Fig 2.5: 802.11b/g channel [13]

3. Voice over Internet Protocol:

VoIP is a technology that is meant to replace the analog PSTN phone technology. It is a technology that makes it possible to make a phone call over the Internet in the same way as we can do it over standard analog PSTN phone; this is the simplest definition of VoIP. It is a revolutionary technology that converts the standard internet connection in such a way that we can make free phone calls on it. [14] For a technologist, the VoIP is a combination of software and hardware that enables us to use the Internet as the medium to place a call, which is analog signal, into the form of digital data packets (discuss later), instead of PSTN network.

3.1 Ways to make a VoIP Call:

There is not a single way to make a VoIP call. There are a lot of ways to make a VoIP call. Some of the well-known types are as follows:

3.1.1 Analog Telephony Adaptor (ATA):

This is the simplest way that uses the existing analog phone. With the help of ATA, we connect the analog phone with the VoIP network or computer on the LAN. ATA is a two way solution; it converts the analog signal into a digital signal that came from the phone and sends it to the VoIP network, and also converts it back from the digital signal into analog when there is a response/answer from the VoIP network [15]

3.1.2 IP Phones:

These are special types of phones, which are designed to work directly with the VoIP network or Internet gateway, and without the need of ATA or an additional device. It uses the RJ-45 (Ethernet) jack instead of RJ-11 jack that analog phones normally use, but it has the entire feature that the old phones have, and some additional features. [14]

3.1.3 Computer-to-Computer:

In this case, we simply have to install a freeware (mostly) software with the existing speaker, microphone and there is no need to pay any additional cost other than the monthly fee to the ISP. In this case the long distance is also not a problem because the packets have to be carried totally in the Packet Switching Network between computers to computers. [14]

3.2 Methods of Transmitting Voice:

Before discussing that how voice is transmitted over the IP network, we should look into the Circuit Switching and Packet Switching networks for better understanding.

3.2.1 Circuit Switching:

Circuit Switching technology is used in legacy telephone networks. We can say that it is the based of PSTN. When we make a call between two phones, we reserve the line/path between the two phones that is only dedicated to these two, and it is referred to as "**Circuit"**, and the technology used to establish this circuit is called "circuit switching".

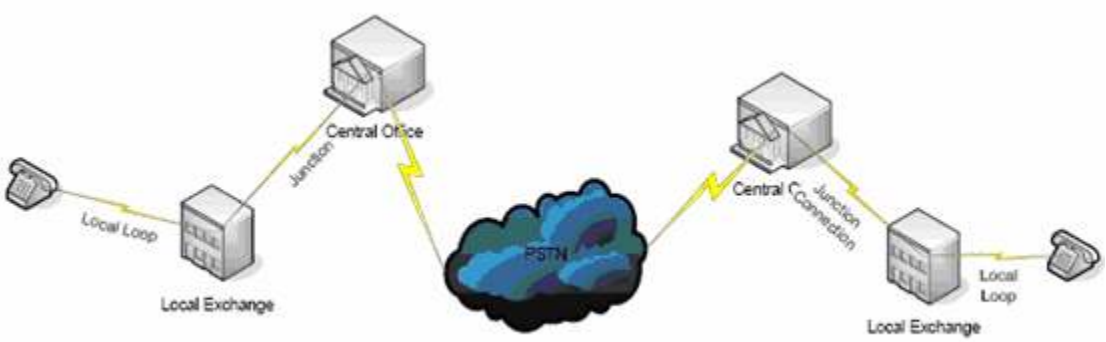


Fig 3.1 PSTN Network

These are steps involved in a typical old telephone call: [14]

1. When we pick the phone, it gives us the dial tone, which is the indication that we are connected to local exchange.
2. Then we dial the phone number of the other end, where we wish to reach.
3. Our call is taking route through the local exchange switch to the central office.
4. The central office then connects our local exchange to the other end party local exchange.
5. A connection has been established between us and other party that is taking a path through the several interconnected switches.
6. The phone on the other end rings and other party answer the phone call.
7. A dedicated circuit is opened between two parties.
8. We talk for specific time period and then end the call.
9. When we close the call, the dedicated circuit is terminated between us and other party, and releases our line as well as all the lines in between the path that it was using.

For the time that we talked on the phone, a dedicated circuit was opened between us and other party, and the call was forwarded by using a fixed rate of 64 Kbps in each direction; a total of 128 Kbps in both directions.

3.2.2 Packet Switching:

Packet Switching Network works in this manner that when one party is talking, the other party is only listening to him, so it only uses half of the bandwidth of the connection at any time and this increases the efficiency. There are also some moments when no party was speaking, and this also reduces these silent periods from conversation and uses even less bandwidth. [14]

It does not establish a dedicated circuit between the two parties (the Internet connection will be really slow if it establishes a circuit between our computer and Web server at any time we were viewing the web page); instead it sends the information in the form of packets over the network (internet) with thousands of redundant paths. That is why it is called **Packet Switching**. Circuit switching opens a connection constantly, while packet switching opens a connection shortly while it has to send a block of data, which is called packet, from one side to other, and it works in this manner: [16]

1. Sending node (computer) breaks the data into small blocks, called “packets”, and assigns an address of destination node to each of them.
2. Inside the packet, it has a payload that contains the actual information that it wants to send to the destination node.
3. The computer sends these packets to its default gateway (a nearby router).
4. The default gateway sends the packets to another router along the way and every router is doing the same until the packets reach the default gateway of the destination computer.
5. The default gateway of the destination computer takes out the address label and sends the packet to the desired computer that uses the information contained inside the packets to resemble these packets, so that it gets the data in its original format.

This provides very efficient, redundant and cheap lines for data transmission. Its also frees the two computers that are communicating with each other, so that meanwhile they are also communicating with other computers and sharing information also with them.

3.3 Transmission of Voice in VoIP Network:

The voice in the VoIP network (Internet) is transmitting in the form of data packets. The figure 3.2 represents the communication between the IP phone and analog phone, but any combination of devices can be involved in the real world example: transmission between analog to analog devices as well as digital to analog etc.

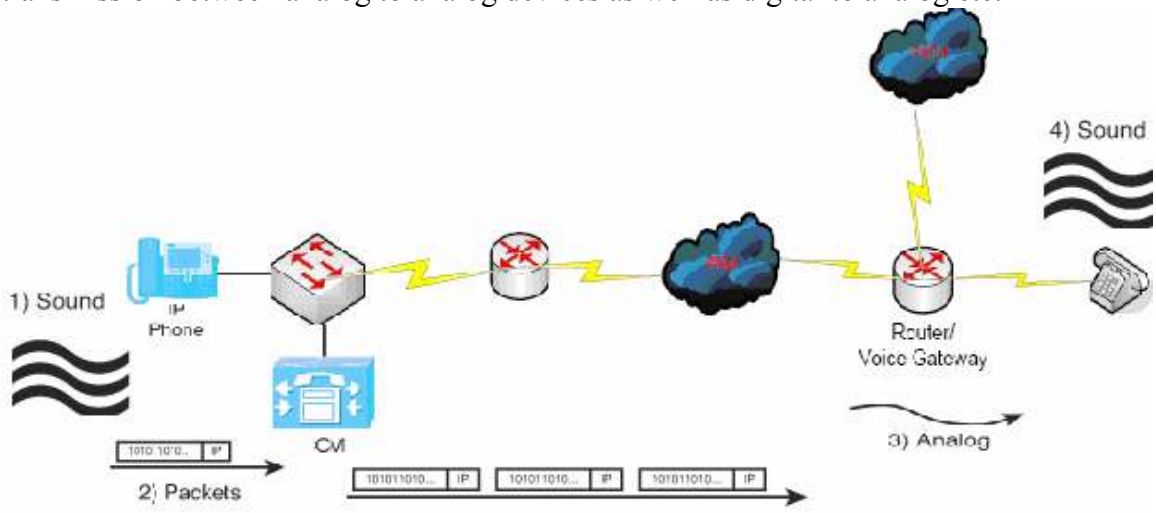


Fig 3.2: Transmission of Voice packets in VoIP network

These are the general steps that are involved in the transmission of voice over VoIP network.

1. Sounds that came from IP phones are grouped into small packets of sound (usually 20ms), then sampled and converted into digital format.
2. Each of these packets of sound is then assigned header that consists of data link, IP, UDP, and Reliable Transport Protocol (RTP), and shipped on the VoIP network.

3. The analog phone on the other end is unable to understand the packets, so the gateway (in this case router) will do the job and convert the packets back into analog signal.
4. Analog signal is forwarded to the phone, and then the phone converts the signal into audio, and plays it on its speaker.

A typical voice packet contains all this information, regardless what codec we are using:

Layer 2 Header	Layer 3 Header	Layer 4 Header	RTP Header	Voice Sample	CRC

Fig 3.3: A normal voice packet

3.4 Packetization:

As we discussed above, the voice is an analog signal and the Internet is only capable of transmitting the digital signals in the form of bits. Here we shall discuss how this is possible?

3.4.1 How the Voice Signal becomes Packets (bits):

There are a lot of factors involved in the conversion of voice signal into packets, but we can summarize these factors into four steps:

1. Take too many samples of the analog signal
2. Calculate the number representing each sample by using the **Pulse Amplitude Modulation (PAM),** which is called the "quantization."
3. Convert these numbers into binary (0 & 1).
4. Compress the signal by using the suitable codec scheme, but this step is totally optional.

In the conversion of Signal to Packet, one theory plays an important role and that is **Nyquist Theorem**.

3.4.2 Nyquist Theorem:

"If you sample a signal in regular intervals of at least twice the highest channel frequency, the samples will contain enough information to accurately reconstruct the signal." [17]

Nyquist Theorem deals the frequency range of 300 – 4000 Hz or 300 – 4 KHz

3.5 Process of Quantization:

Here is the Quantization process that is used to calculate the number that represents each sample by using PAM.

A phone system was designed to produce and capture the signals that have frequency less than **4 KHz** and, according to the Nyquist Theorem, at least twice the highest channel frequency. So

Nyquist: 4000 is higher frequency
Sample Size: (4000 x 2) = 8000
Thus we sample 8000 lines / second
Each 1 = 1/8000 of a second

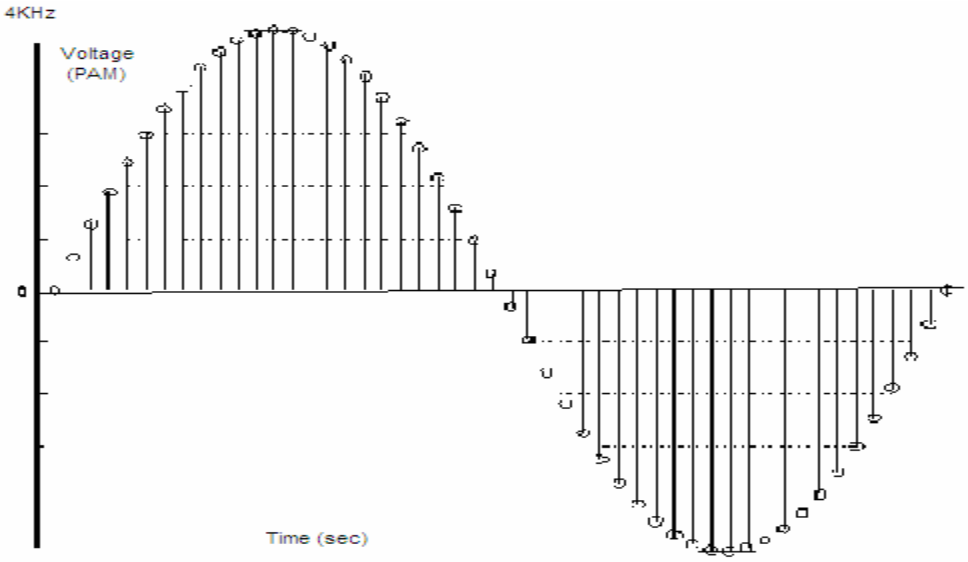


Fig 3.4: Quantization Process

So each 1 contain 1 byte so 8000 x 8 = 64000 bps, or 64 Kbps, which is equal to one voice channel (DS0).

3.6 Codec:

A normal call that uses the VoIP infrastructure needs to go through two different stages. The first stage is that takes the standard analog signal and converts it into digital to transmit on the internet. The second stage is that, when it reaches the other end, it needs to be converted back to analog, otherwise it is not recognizable. [18]

There are two types of algorithms used for the coding/decoding. These are:

I. Waveform Algorithm (It encode everything)
- PCM (Pulse Code Modulation)
- ADPCM (Adaptive Differential Pulse Code Modulation)

II. Source Algorithm (It only encode the changes)
- CS-ACELP (Conjugate Structure Algebraic Code-Excited Linear Prediction)
- LDCELP (Low Delay-Code Excited Linear Prediction)

Some of the famous codecs are there:

3.6.1 G.711 Codec:

This uses the Pulse Code Modulation (PCM) to code, encode, compress and decompress, the analog telephony signal by using 8 KHz sampling and 64Kbps bit rate. It literally codes each and every sample into its own binary calculation for the transmission on the Internet. [18]

3.6.2 G.729 Codec:

This operates at 8 Kbps and 8 KHz sampling frequency by using the **Conjugate Structure Algebraic Code-Excited Linear Prediction** algorithm. It also uses a human voice "Codebook" as a dictionary to work and Look-Ahead of 5 ms. Special mathematical algorithms are used for voice synthesis. The complexity (requirement of processing power) lies at 15 (Appendix B), because of its low bandwidth requirement. It is the most favorable codec that is used in VoIP applications nowadays. [18]

3.6.3 Mean Opinion Score:

Mean Opinion Score (MOS) is a scale that is used to describe the quality of the perceived media after compression and transmission. It lies between 1 and 5, where 1 indicates the lowest quality and 5 indicates the best quality. Here is the MOS used for the codec for VoIP applications. [19]

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

Table 3.1: Mean Opinion Score Table

3.6.4 Speex:

Speex is a free open source speech codec that can be used for VoIP applications and it is absolutely free, licensed under the Berkeley Software Distribution (BSD) license. It can be transmitted directly over UDP/RTP. It is a "lossy" codec, which means that, during the compression and then decompression processes, the original data reduces its quality. [20]

The designers of the Speex codec have been focused to make it best for VoIP to provide the high quality sound at a low bit rate. It uses multiple sampling rates of 32 KHz, 16 KHz and 8 KHz. Speex uses the Code Excited Linear Predication (CELP) for coding and encoding. The choice of CELP for Speex is, because of its performance at both high and low bit rates. [20]

It is flexible in such a way that it uses multiple sample rates and variable bit rates (from 2 Kbps to 44 Kbps), which make it so powerful that it continuously maintains the quality of voice. It also provides the facility of VAD (Voice Activity Detection) and it is not so demanding on the processor. Furthermore, some effort has been made to reduce the noise during the coding/encoding process.

3.7 Transmission:

For the transmission of VoIP, there are three steps:

- Signaling is used to setup and end the call. For this purpose SIP, H.323 and MGCP protocol are used.
- Packetization is used for sending the voice in the form of packets. (As discussed above)
- QoS is used to give the priority to VoIP traffic. (discuss in next chapter)

There are four reasons that the VoIP calls are not as clear as PSTN calls: these are:

- Packet loss
- Latency
- Jitter

- Echo

3.7.1 Packet Loss: packet loss occurs due to the network congestion and it can be solved with proper QoS policy.

3.7.2 Latency: This is the time delay between end-to-end VoIP conversations. It should not be more than 150 ms for one-way.

3.7.3 Jitter: The variable delay that can cause the voice packet to arrive late, or out of order, and it can be solved by using jitter buffer.

3.7.4 Echo: This is the phenomenon in which the callers hear their own voice back, and it can be solved by using the echo-cancellation (G.168).

Here we only discuss the SIP because of its simplicity and features.

3.7.5 Session Initiation Protocol (SIP):

SIP is the most widely used signaling protocol to control voice and video call over the Internet. It is an Application Layer protocol which is based on TCP/IP, but it is not dependent on the underlying transport layer, it can run on anything like UDP or TCP. It based on many of the previous protocols that are already in use, like HTTP, SMTP, and DNS etc. It is more of an "all-in-one" protocol rather than the protocol suite of H.323. It uses the same model that HTTP uses for request/response, as Fig 3.4 shows. For each request of the client, a particular function is invoked on the server. [21]

SIP can work together with many other protocols but it only does the signaling for communication session. SIP client uses UDP/TCP port 5060/5061 to connect to the other SIP device or SIP server. Port 5060 is used for un-encrypted signaling traffic, while 5061 is used for encrypted signaling.

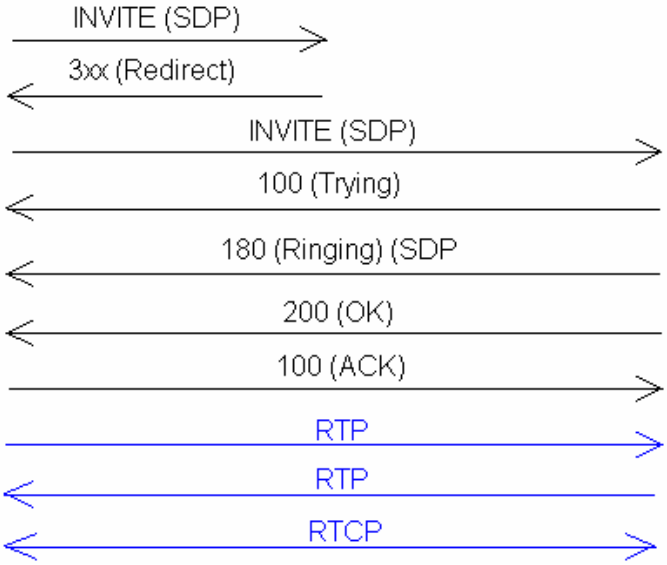


Fig 3.5: SIP Call Setup

Its main purpose is to set up and tear down the voice/video call, but it is also found in many other applications. Voice/video stream in SIP application is carrying by RTP (Real-time Transport Protocol) and parameters such as port number, codec and

protocol for this stream are defined by SDP (Session Description Protocol). SIP is a peer-to-peer protocol but it requires simple network infrastructure with some intelligence in the endpoints. [21]

3.8 Voice Activity Detection:

This is a feature in VoIP that stops sending traffic during silence. This means it saves our bandwidth. On average, phone calls are about 35% silence, though it really does depend on the call types, and also languages matter in VAD. Background noise decreases the efficiency of VAD, and it does not work with MOH (Music on Hold).

4. Quality of Service (QoS):

This is the mechanism used in data communication and networking field to prefer or prioritize some traffic over other, so that the preferred traffic moves through the network as quickly as possible. It gives different level of preference to different applications. It is specially used for VoIP, IPTV and video streaming. If the network is not congested then there is no need of QoS. [22]

For VoIP, the definition of QoS is the ability to transmit and receive the voice continuously and clearly, and without any disturbance.

4.1 Why We Need QoS?

There are four major problems that force us to implement the QoS, these are:

- Lack of Bandwidth
- Delay
- Jitter
- Packet Loss

i) Bandwidth:

When a lot of applications are running then these multiple applications flows are using the same links, which means that the available bandwidth for each application is even smaller, and it is equal to the bandwidth of the smallest link divided by the number of application data flows. If there is insufficient bandwidth, then it will severely damage the time and delay sensitive application like VoIP and video streaming. One way is to increase the bandwidth but this is expensive and an alternative way is to implement the QoS, which ensures the bandwidth for sensitive applications. [23]

ii) Delay:

Delay is the time that a packet takes to reach its destination. There are four types of delay that network traffic faces:

- **Processing Delay**
- **Queuing Delay**
- **Serialization Delay**
- **Propagation Delay**

iii) Jitter:

When the packets reaches from source to the destination with variance delay, then it is known as jitter and it can create a serious problem for audio and video traffic. The major cause for the jitter is varying delay in the router's queues as well as the varying path between source and destination.

iv) Packet Loss Issues:

Packets drop happens due to fact that the router fails to deliver some packets because its buffer memory is full. A router can drop some packets, or all packets, and this depends upon the status of the network, i.e. whether it is fully congestion or not. This packets loss can degrade the performance of voice and video traffic. The router in its

default behavior drops all packets once its queue become full. This leads towards some serious problems: [23]

- TCP global synchronization
- TCP buffer starvation

4.2 Models used for QoS:

Basically, there are three types of models that are used to implement QoS:

- Best Effort
- Integrated Services (IntServ)
- Differentiated Services (DiffServ)

4.2.1 Best Effort:

This is a default model in which the traffic is being sent in the order in which it arrives. It gives the equal treatment to all traffic types and does not give any guarantee of delivery. The advantage of its usage is that its scalability. Internet used the best effort model for the delivery of all kinds of traffic.

4.2.2 IntServ:

IntServ guarantees some specific level of service to each flow of traffic throughout the network for specific period of time. It uses Resource Reservation Protocol (RSVP) to reserve the path throughout the network. An RSVP enabled router requests specific level of service, to its next hop router and each router along the way, reserving the specific bandwidth for that flow for some length of time. If the network is not able to provide the specific level of service, then the session is not allowed. RSVP can work with any type of traffic, but is mainly used for delay and time-sensitive traffic, like VoIP. [23]

4.2.3 DiffServ:

DiffServ is the most efficient and widely used QoS model. It classifies the network traffic into classes, and each class consists of the traffic that needs the same type of QoS e.g. VoIP traffic needs different QoS than email, but email traffic can have the same QoS requirement as web traffic. So we can put the email and web traffic into the same class. The distinction between the classes is based on the certain bits value in the Layer 2 and Layer 3 header. We can treat the traffic as we want each hop long the way to the destination and it is referred to as per-hop behavior (PHB). [23]

4.3 Applications that need QoS:

There are two types of applications; one is called “inelastic”, that has a specific bandwidth requirement and maximum latency to perform its function, and other is elastic applications that can work well with little available bandwidth.

These are some important inelastic applications that require a certain amount of bandwidth:

- VoIP
- Online games
- Multimedia streaming
- IPTV

4.4 Currently, Problem with QoS:

Currently, the Internet is not governed by any central authority. It is administered by many independent authorities and currently the Internet only uses best effort for transmission and Internet2 (a non-profit, advanced networking research community) QoS working group gave a suggestion that increasing bandwidth is a more suitable solution than to implement the QoS. [24]

4.5 Proposed Solution for QoS:

Multi Service Access Everywhere (MUSE) proposed the QoS concept first and that we must agree on discrete jitter value for each class that is implementing in the network. This solution has some benefits, and these are: [25]

- End users are able to notice the difference in service quality
- Easy to implement
- We are able to predict the end-to-end delay

4.6 How is QoS Implemented?

There are a lot of factors that are taken into account while implement the QoS. Some of them are:

4.6.1 Classification and Marking:

The first step in the implementation of the QoS is to classify the traffic. Until the traffic is classified, it is not possible to give the specific level of service, and traffic is often classified by IP address (source or destination) or application. [23]

After the classification, the next step is to mark the classified traffic to the appropriate marking and the location where the traffic is marked is known as the "trust boundary". If the device that marked the traffic is trust, then that marked traffic passes through the network, and each device in the network gives it defined service level and, if the device is not marked, then some trusted device must re-mark this traffic again. [23]

Classification and marking should be done as close as possible to the source that generates the desired traffic so that it saves the resources and it need to happen once only in the network. All other devices just look at the marking and set the policy according to it.

4.6.1.1 Layer 2 Marking:

CoS uses three bits in the layer header, and provides eight levels of markings, from 0 to 7. [26]

COS value	Binary	Application
7	111	Network Control
6	110	Internetwork Control
5	101	CRITIC/ECP
4	100	Flash Override
3	011	Flash
2	010	Immediate
1	001	Priority
0	000	Routine

Layer 2 Marking [26]

4.6.1.2 Layer 3 Marking:

The original TCP/IP standard defined a ToS byte. The first implementation of marking using the ToS byte was IP precedence and it has only used the leftmost three bits and provides the same eight levels of markings as CoS. When the classes of applications increased, and we required more level of markings, then they introduced the DSCP (Differentiated Service Code Point), maintaining backward compatibility with IP precedence. DSCP uses the left most 6 bits currently, three bits for per-hop behavior (PHB), three bits for drop probability, in which last is always zero, and last two bits for flow control, but it is not used now.

Expedited Forwarding (101 110)			
PHB	**High Drop**	**Medium Drop**	**Low Drop**
AF4	11	10	01
AF3	11	10	01
AF2	11	10	01
AF1	11	10	01

Best Effort (000 000)

Layer 3 Marking [26]

4.6.2 Queuing Management:

Queuing mechanism gives us the possibility to control the congested and send it to interface by putting the congested traffic in its own, assigned queue, according to the configured policy.

There are two types of queues: **Hardware Queue**, in which it keeps the packets that are ready to put from transmit ring to media (wire) and that is always FIFO (First In First Out). **Software Queue** is a memory assigned to each interface, where the traffic waits in case the transmit ring is full. When the traffic is put into the queues, this means that the network is congested and it is caused by the speed mismatch and link aggregation. [23]

4.6.2.1 Queuing Strategies:

There are some queuing strategies that manage the traffic during network congestion. Some famous queuing strategies are:

FIFO Queuing
Priority Queuing (PQ)
Round Robin Queuing (RRQ)
Weight Fair Queuing (WFQ)
Class-Based Weighted Fair Queuing (CBWFQ)
Low Latency Queuing (LLQ)

4.6.3 Congestion Avoidance:

Congestion Avoidance is the strategy to implement some technique to avoid the network congestion. To fulfill this purpose, we used two techniques:

- Random Early Detection

- Weighted Random Early Detection

4.6.3.1 Random Early Detection (RED):

This tries to avoid the congestion by random drops packets from TCP flows to minimize the synchronization. Once the queue is filled above the threshold level (the value at which the maximum packets can occupy by the queue), it starts to drop packets randomly from the queue. Dropping becomes more aggressive as the queues fill. [23]

4.6.3.2 Weighted Random Early Detection (WRED):

Random Early Detection is not able to make the difference between different flows of traffic, so it is not suitable for real time traffic. On the other hand, Weighted RED drops the traffic on the value of its IP precedence or DSCP. It combined with CBWFQ to implement the DiffServ's PHB in which each PHB has a unique WRED profile to identify a minimum threshold, maximum threshold and MPD (Mark Probability Denominator) for that profile. The decision of packets drop is taken on the value of IP precedence if the DSCP-based value is not configured. [23]

5. QoS for Wireless Networks:

Sure and guaranteed service is required for audio and video transmission. Many service providers have been offering IP telephony and video services in addition to internet service to their customer and the motivation behind offering these services is the flexibility of packet-switched networks. But the challenges involved in the efficient audio or video transmission are those of compression methods that encode/decode the streams at VBR (variable bit rate) and also assigning the highest bit rates to these streams.

5.1 Challenges Involved in Wired QoS:

There are a lot of challenges involved in providing guaranteed services in a network and the main challenge in providing QoS is congestion in the network; it increases the end-to-end delay because the packet has to stay for long in the queue at each hop in the network. Packets loss also increases when the queue becomes full, and it limits the throughput because we have to retransmit the lost packets.

Multi-Path routing is also a critical problem. When two packets are sent, there is no guarantee that they follow the same path to reach the destination. Maybe one packet takes a path that contains fewer hops, or is less congested, so in these cases the packets will not reach the destination at the same time and this causes an unacceptable delay/jitter. [27]

5.2 Additional QoS Challenges involved in Wireless Networks:

Wireless networks have the same challenges that wired networks have, but they also involve additional challenges. Due to these additional challenges involved in wireless networks, QoS methods used in the wired networks are not feasible for wireless networks. There are a lot of additional challenges involved, some of them are:

5.2.1 Interference:

In wired networks, the cause of packets loss is due to severe network congestion. Only a negligible amount of data is lost due to corrupt transmission wire. On the other hand, the wireless link typically suffers more packets loss because of interference during transmission. This interference is because of using the free range of frequency bands that other applications use, and these create interference with wireless signals. Another cause of interference is electrical noise which causes the severe damage to the wireless signal. [28]

5.2.2 Hidden Node Problem:

In wireless, hidden node is also a serious problem that occurs when a node is visible to the central point (Access Point) but it is not visible for other nodes that are communicating with the AP and it leads towards the media access control difficulties. This problem can, however, be solved by using RTS/CTS messages, but it is still taken into account when implement QoS.

5.2.3 Multipath Propagation:

The signal may take different paths between the sender and the receiver due to diffraction, scattering and reflection, which causes various problems like time dispersion (it can cause problems if high data rate digital modulation is employed) and the signal,

can reach at the receiver directly or phase shifted. The distorted signal depends on the phase of the different parts. [28]

5.2.4 Handoff:

In wireless networks, handoff is a mechanism whereby one AP gives the control of its associated node to another AP. During handoff, it has to re-establish the route reservation that starts at the mobile node.

5.2.5 Propagation Delay:

Propagation delay is another obstacle for implementing QoS in wireless networks because some wireless networks are spread over area of square Kilometers, and propagation delay is an extra burden for the real-time communication that requires guarantee on delay. The problem exists mostly in MAN (Metropolitan Area Networks).

Finally, it is really difficult to maintain the guaranteed service in wireless networks because the nodes are mobile. The scheme that involves the resource reservation did not work well with it, because the new route had to establish to the destination if the sender moved and there was also a possibility that the new route did not provide the level of service that the previous route guaranteed. The problem was also similar if the destination moved. [27]

5.3 802.11 Medium Access Method:

802.11 originally used the DCF (Distributed Coordination Function) as the basic medium access control mechanism but it could also implement PCF (Point Coordination Function) on the top of DCF optionally.

5.3.1 DCF (Distributed Coordination Function):

DCF was originally used by the 802.11 MAC layer to share the access to the medium between the multiple nodes. DCF uses CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) to share the access of medium between nodes, but it also optionally uses RTS/CTS (Request to Send / Clear to Send). It has limitations, but those limitations mean you cannot use it for QoS. [29]

If many nodes start sending data at the same time, many collisions occur which causes the lower available bandwidth, and there is no concept of high and low priority traffic in DCF. Only once can a station get access to the medium; it can keep the medium as long as it wants and, if it has a low bit rate, then it will create problem for all other stations by taking a long time to send its traffic. DCF does not provide any QoS guarantee. [29]

5.3.2 PCF (Point Coordination Function):

802.11 standards optionally defined the PCF (Point Coordination Function), which is used for the transmission of time-sensitive traffic. It only works with infrastructure mode, in which AP acts as the coordinator; controlling which station can transmit during any given period of time by sending “beacon” at regular intervals (~ 0.1 second). [29]

It also defines two periods: Contention Free Period (CFP) and Contention Period (CP). In CPF, the AP sends the Contention Free-Pool (CF-Poll) packets to all the nodes that are operating in PCF mode to give them the chance to send the traffic. In CP, it simply used DCF. Thus, PCF is a contention-free protocol and it gives chance to the stations to transmit data frames with regular time delays between data frame transmissions and, due to this, it provides better management for QoS. Unfortunately, PCF has very limited support for QoS and it does not provide the support for classes of traffic. [29]

To overcome these weaknesses, the IEEE defines the new wireless standard 802.11e to provide the better QoS for time-sensitive applications.

5.4 IEEE 802.11e Wireless Standard:

802.11e is an amendment to the original IEEE 802.11 standard that introduced the QoS support for wireless networks by modifying the Layer 2 (Media Access Control) of OSI model. It has critical importance for time and delay-sensitive applications like VoIP, video streaming, IPTV etc.

5.4.1 802.11e MAC Layer Operation:

802.11e uses HCF (Hybrid Coordination Function) to support QoS. HCF further defines the two medium access mechanisms:

- EDCA (Enhanced Distributed Channel Access), also referred to as Contention-based medium access.
- HCCA (HCF Controlled Channel Access), also known as Controlled medium access.

802.11e also operates in two modes (CP and CFP). EDCA is only used in Contention Period, where as can be used in both modes. The reason it is called "hybrid" is it combined both DCF and PCF methods.

5.4.1.1 EDCA (Enhanced Distributed Channel Access):

EDCA (Enhanced Distributed Channel Access) gives chance to high priority traffic to be sent first, then low priority traffic afterwards, which means that the node with the low priority traffic has to wait more than high priority traffic. In addition, when defining priority, it also assigns Transmit Opportunity (TXOP) to each priority level. TXOP is a specific time interval during which a node can traffic as much traffic as possible. If the frame is too big that it can be not be transmitted in one TXOP, then it should be fragmented into smaller frames. The uses of TXOP also solve the problem for low rate stations that gained the access to the medium for a long time to transmit its frames in the original legacy 802.11 DCF MAC. [30]

The basic purpose of EDCA is to provide the QoS mechanism upon the classes types (to send high priority data first than low priority data) but there can also be situations in which data that belong to the same priority class have to protected from each other, e.g. where the network can only accommodate 10 VoIP call and if 11^{th} call is made. This kind of problem is solved by using Admission Control in EDCA in which AP advertise the available bandwidth in its beacons and client check the available bandwidth before transmitting more traffic on the network. [31]

5.4.1.2 HCCA (HCF Controlled Channel Access):

HCCA works quite similarly to PCF but, instead of waiting for idle time and back-off mechanism, it uses the access point (in this case acting as hybrid coordinator) as central control authority that can guarantee the time, as well as the time during which each connected node can send data(transmission time). Every station that wants to join the network must make a request to the AP; the request includes the QoS parameter (indeed the traffic type), the AP (Hybrid Coordinator) analyzes the QoS parameters to decide whether it can or can not provide the required QoS to the requested node and then admit/deny. The AP also maintains the centralized schedule of the entire register device to it and it grants the permission to each node to access the wireless medium, according to the scheduler, as all the parameters were determined at the time of registration. [30]

The other differences between PCF and HCCA are that the HC is not limited to per-station queuing, and it can also provide the per-session service. [31] These are some problems with HCCA. However, the biggest problem is that the HCCA does not have an ability to work with legacy networks in its neighborhood because HCCA AP takes the access over the wireless medium when working with the legacy network in both CFP and CP situations, so it will interfere with the legacy network. [30] Fig 5.3 shows the enhancement at the data link layer to provide QoS for wireless network.

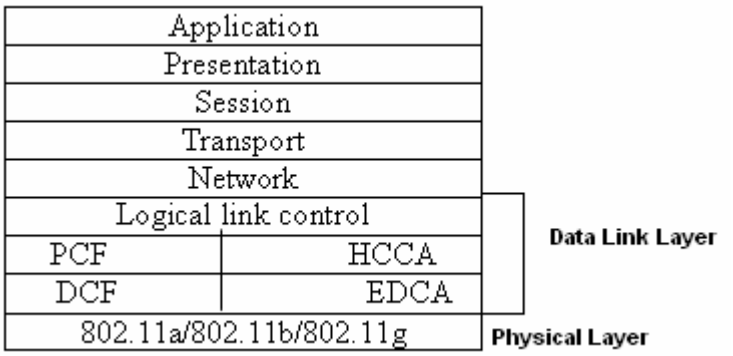


Fig 5.1: OSI model with modification for 802.11e standard

5.4.2 Additional 802.11e Specifications:

There are some additional specifications that are defined by 802.11e in addition to EDCA, HCCA and TXOP to enhance the MAC layer QoS:

5.4.2.1 Automatic Power Save Delivery (APSD):

The Automatic Power Save Delivery is a mechanism to save power efficiently as compared to the legacy power saving method used in original 802.11 standards. APSD works very well with VoIP phones because the data rates in both directions are approximately the same. When voice data are sent to Access Point, the AP start sending the buffered voice data in the other direction so after the voice data transmission is completed, the IP phone goes into a stand by mode until the next voice data has to be sent to the AP. [31]

5.4.2.2 Block Acknowledgement:

Block acknowledgement allows that the block of total TXOP will be acknowledge instead of single frame, by doing this it provide less protocol overhead when longer TXOPs are used. [31]

5.4.2.3 NoAck:

In QoS, there are two values for the service class frames that are QoSAck and QoSNoAck. The frames with the QoSNoAck will not be acknowledged, and this avoids the retransmission of time-sensitive data. [31]

5.4.2.4 Direct Link Setup:

Direct Link Setup allows the direct transmission of frames from node to node within the Basic Service Set (all the devices are associated local wireless LAN) and this design is really useful where the node to node transfer is frequently used. [31]

6. Implementation of QoS in Wireless Networks:

In this chapter, we look at the implementation of QoS in wireless networks and it is definitely different to all the mechanisms we have talked about so far. We shall look at the wireless standard that is built by the wireless association the Wi-Fi Alliance and its partners (Cisco & Microsoft) for the implementation of wireless QoS, which is different, instead of scheduling the packet it really scheduling time.

QoS becomes more important for the wireless access world as we get new devices that use wireless technology. The idea of QoS is needed because the clients, as we can see in Fig 6.1, such as Wi-Fi phones, laptops, PDAs, handheld computer run applications, like the critical data service or VoIP, that needs different levels of services than simply web surfing. Wireless technology works on CSMA/CA that is same as token ring technology, in which a token moves around the ring and, whatever node gets it, can send packets. The same situation is in wireless world: whatever node gets access to the medium can send packets as long as it wants. This means that the more nodes which are connected to the AP (access point), the more the bandwidth is saturated and the fewer time slots there are for the transmission of traffic. [35]

Fig 6.1: A typical Wi-Fi network today

6.1 Why is QoS Necessary?

Legacy wireless networks give equal access to all the devices that are connected to them and, when the demand of traffic exceed from the available bandwidth, throughput is reduced for all data streams, regardless of type of traffic type. This behavior is strongly affected by the application type. A one or two second delay in sending a printing job from computer to a printer is not noticed by the user, but even a one second delay can disturb the VoIP call. It can even drop the call. In the residential and industrial markets, multimedia applications potentially create a new need for QoS. [36]

6.2 QoS for Wi-Fi Networks offered by WMM:

The Wi-Fi Alliance played an active role in the development of QoS for multimedia applications by developing WMM (wi-fi multimedia). The main advantages of WMM are:

6.2.1 Relationship with IEEE 802.11e: The 802.11e standard is approved but WMM is used still everywhere because it provide the base to 802.11e standard. Although 802.11e has additional features, it uses the WMM as for its core functions. E.g. The Wi-Fi Alliance has already developed the test plan for the scheduled access capability. [36]

6.2.2 Industry Support: WMM was developed by Wi-Fi Alliance with their partners (that are all the world leading industries), so all the major industry player adopted the WMM. [36]

6.2.3 DiffServ (Differentiated services): WMM defines QoS classes' structure that is based on the IETF DiffServ architecture, which means that it is cross compatible with wired networks. Individual packet is labeled with DSCP or 802.1d tags. [36]

6.2.4 Universal Plug and Play (UPnP) Compatibility: DiffServ enables UPnP QoS to maintain WMM, and allows administrators to develop network-wide policies that can be applied to wired and wireless infrastructure. [36]

6.3 Access Categories:

The Wi-Fi Alliance came up with a completely new standard, called the WMM. It was meant to replace the QoS mechanism used by the wired network. Layer 2 switches use the 802.1p tag as a QoS mechanism, which defines 8 classes of service by using 3 bits to transfer data across the trunk links.

The APs, when they communicating wirelessly with the clients they do not have "tags" to use between the client and APs and we end up with degraded service when the AP is saturated. WMM boils down the eight levels of services into four levels (voice, video, best effort and background) and the first vendor that came into action was Cisco, who called these classes "platinum", "gold", "silver" and "bronze".

WMM gives flexibility to the administrator to choose the appropriate policy for network-wide use, and also gives permissions to give preferences to one class over another. WMM also defines the protocol that will be used between the AP and the QoS enabled client. [36] Tables 6.2 shows the AC (Access Categories) defined by the WMM.

Access Categories	Description	802.1d Tags
Voice Priority	Highest Priority	7,6
Video Priority	Prioritize traffic over other best effort traffic	5,4
Best Effort Priority	Traffic from legacy or not QoS enabled devices	0,3
Background Priority	Low Priority	2,1

Table 6.2: WMM Access Categories [35]

Fig 6.3 and Fig 6.4 show the effect on throughput if we implement the WMM. Fig 6.3 shows that the WMM gives the higher priority to the voice traffic over the other. Both voice and other (low priority) have enough resources during the initial 10 seconds and the third data stream creates exceeded network capacity due to it transmission

demands. But WMM gives the same priority to the voice traffic for its smooth transmission so that it will not affect it. In the fig 6.4, all the streams have been given the same priority to access the wireless medium, and the introduction of the third stream affects the transmission for all previous streams because, without WMM, all data streams have equal priority.

Fig 6.3: With WMM[35]

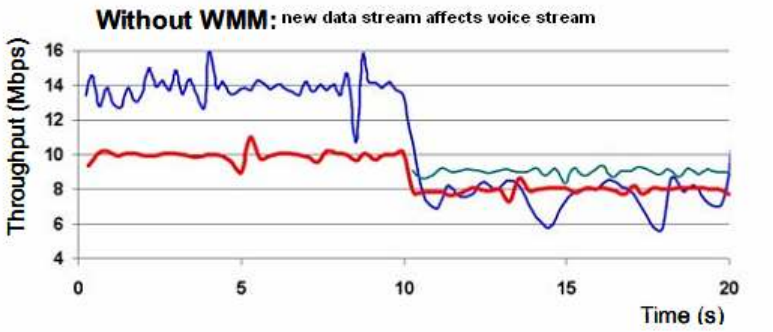


Fig 6.4: Without WMM[35]

6.4 WMM Operation:

WMM gives priority to the traffic according to the four classes that are defined in table 6.2, which means the higher the access category you are, the higher the chance to transmit. By doing that, it overcomes the weakness of DCF (Distributed Coordination Function), which was unable to give the desired service to the multimedia applications. The Access Categories (AC) are designed in such a way that they map to 802.1d (3 bit prioritization tag) to the same QoS policy across wired and wireless network segments. [37] If the packet is not assigned a specific AC, then it is categorized by default the best effort priority. [36]

Applications are assigned an AC to each data packet, and then this packet is added to one queue out of four (voice, video, best-effort and background) independently in the client (as shown in fig 6.5). The client also has an internal collision resolution mechanism, to resolve the collision between different queues, as it selects the frame from the highest priority queue first, and then lowers it for transmitting. The same kind of mechanism is used for external collision resolution, to find which client should be granted the access to the medium. [36]

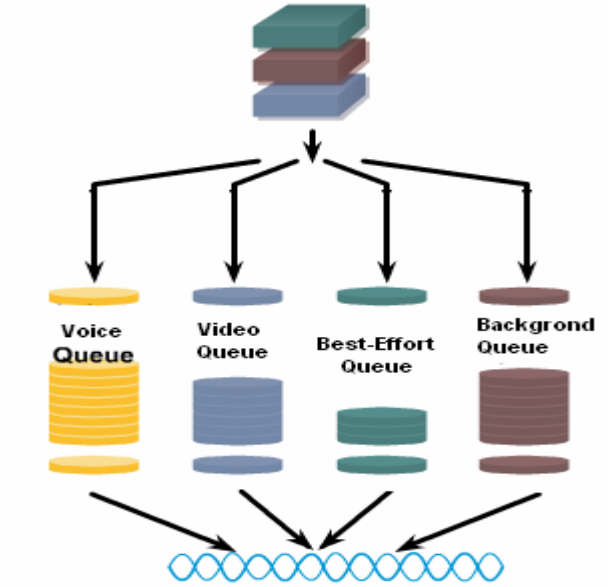


Fig 6.5: Transmit Queues inside the Client [37]

Fig 4.6 shows an algorithm which is responsible granted the access to the medium based upon priority, and it is totally dependent on two timing parameters that are varying for each AC.

- Minimum wait time
- Contention window or random backoff time

Both values are larger for low priority traffic and smaller for high priority traffic. For each AC, the random backoff timer is calculated as the sum of minimum wait time and random value from 0 to the contention window. [36]

After each collision, the value of random backoff time becomes doubled until it reaches to the maximum defined value for each AC. After each successful transmission of packets, it comes to its initial position at 0. The higher AC packets get the access to the medium quickly because they have smaller random backoff time.

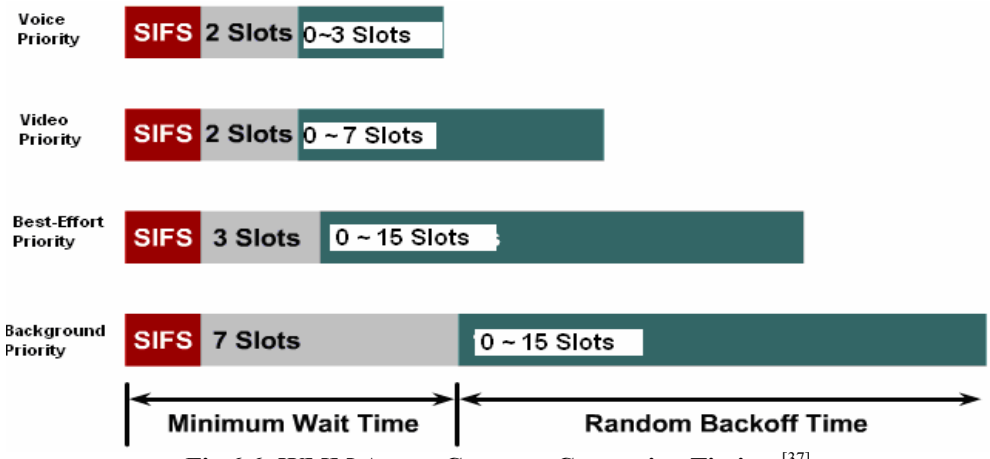


Fig 6.6: WMM Access Category Contention Timing [37]

Once the client gets access to the medium, it is allowed to send the packets for a specific time, which depends on the AC and physical rate. Typically, the medium access

time limits are from 0.2 to 3 ms for background priority and video priority respectively in 802.11 a/g network. For 802.11b network, it is ranged from 1.2 to 6 ms.

6.5.1 Scheduled Access:

Scheduled access is a mechanism used by applications in which a client sends a request to the AP to reserve the network resources based on the traffic types. By using the central scheduling control mechanism, average network latency can be reduced. It is designed to meet the latency and throughput requirement of the applications by assigning different times to the applications as to when they can transmit. The client sends a reservation request to the AP, and the AP assigns the access to the medium by checking different Transmission Specification (TSPEC) parameters of the request like packet sizes, data rate, service interval etc. [36]

In scheduled access, the client knows what kind of resources it needs in advance and the AP makes the assumptions of all the parameters (as defined above) to effectively schedule concurrent traffic. [36]

6.5.2 EDCA plus Admission Control:

EDCA assures that high priority traffic gain access to the medium quickly, as compared to the low priority traffic, and it does not degrade the performance of high priority traffic. Another unique addition to the admission control also prevents the traffic from the same priority class from disturbing the already admitted traffic if the network is not capable of handling both traffic streams. E.g. EDCA assures that the voice traffic stream gets priority over the data stream, while the addition of admission control prevents another voice stream from entering the network, if it is not capable of providing the same level of service to both, and both streams crashing. [38]

Admission control continuously evaluates the network's resources, and only allows the additional streams if the resources are available. It is mandatory for the AP while optional for the end station. [38]

7. Shortcomings of Wireless VoIP Today:

The VoIP is deploying successfully on wireless these days but still there are a lot of challenges to be faced when it come to the wireless medium. In this chapter we briefly describe some of them:

7.1 Service Area:

Coverage area is a huge problem for VoIP over wireless. Even latest versions of access points have ranges of only some yards or feet. For the real wireless coverage for VoIP however, the coverage needs to be extensive. To overcome this problem, we need to deploy a huge number of access points within a coverage area, which also leads us to another problem, namely that this solution is too expensive.

7.2 Reliability:

Most of the users can tolerate delay with non real-time applications. For example, we do not want to receive our emails instantly; even if we get after some minutes, it should not be too much of a problem. When we talk about the voice, the PSTN provides us with the reliable and high quality transmissions, also the system is always on. The reliability is the major reason for companies not converting their existing PSTN phones to VoIP quickly- real time applications are so sensitive for packet loss as compared to other data applications as they do not bear the packet loss and delay. [40]

Wireless adds another layer of problems, like RF interference and signals strength. The latter can cause the poor quality VoIP calls, or even calls can be dropped or packets delayed. [40] Signal strength is another problem that may be due to user roaming. As the user moves away from the access point, the signals become weaker and this degrades the quality of the VoIP call.

7.3 Emergency 911 calls:

Emergency 911 services are also a hurdle for VoIP in the sense that all service providers are not offering the e911 service as standard. Even the service providers that offer the e911 service are often unable to route the calls to the intended location of the user after hours of waiting. Many users see it as a major problem when they compare it with PSTN phones. However, FCC (Federal Communication Commission) imposed the rule that all the service providers must provide the facility of e911 call service but this has not yet been implemented completely. [41]

7.4 Load Balancing:

Because of the asynchronous and untimed nature of wireless communication, clients (pc or handset) need to be constantly monitoring the signal strength due to the fact that the transmitter and receiver must leave a dedicated channel open. That leads towards the problem that open channel have a huge load on the access points. Wi-Fi is capable of handling of a specific amount of clients simultaneously on each access point although theoretically, the value is much larger. During peak communication demand, the use of VoIP takes the number of real connections to hundred(s) and the consequential result is that the access point becomes easily overloaded. There is no automatic mechanism that will help to balance the load, to switch the signals to idle or less busy access points when possible. [42]

7.5 Limited No of Calls Support:

The capacity of taking the calls by the wireless standard is also the shortcoming in the way of wireless VoIP. A popular standard of wireless is 802.11b, which provides the

basis of 802.11e, has the capacity of 11 Mbps. A typical VoIP call consumes about 10Kbps. So, theoretically, the number of VoIP calls that it can take simultaneously is 11Mbps/10Kbps =1100, which correspond to 550 calls each with two streams. If we use the GSM 6.10 codec then it only supports about 12 calls because of a lot of overheads in the packet-header. [43]

7.6 A Lot of Chunks in Each Packet:

A typical VoIP packet contains 40 Bytes of IP/UDP (User Datagram Protocol)/RTP (Real-time Transport Protocol), plus 6 to 22 Bytes of Data Link layer overhead for carrying the 20 Bytes of actual voice payload. This means that we waste 70% of bandwidth in the form of overheads, although the voice is real-time traffic and it does not need all these overheads for all the packets. Fig 7.1 shows the typical voice packet.

Layer 2 Header	Layer 3 Header	Layer 4 Header	RTP Header	Voice Sample	CRC

Fig 7.1: Typical Voice Packet

7.7 Collisions:

VoIP signals are unpredictable in nature due to factor of asynchronous data flow. In busy wireless networks, the access points could be overloaded during the peak flow of data. The key issue arising with this is collisions: that is when too many signals arrive same time at the access point and some get delayed. The delays should not be more than 30 ms for true voice quality. [42]

7.8 Battery Life:

Battery life is an important factor in VoIP devices. IP-based communication is asynchronous, which means that the VoIP device has to be active to properly handle the signal, both in a call and out of a call, as compared to a cellular device. Cellular devices wake up at regular intervals of every 30 ms. This 30 ms enables the device to save power briefly between receiving and transmitting information. As a result, the battery life increases. VoIP handsets cannot preserve the power cyclic. [42]

7.9 Security:

VoIP that is transmitted on the wireless network requires three levels of security: one for voice transmission, one for its associated control signaling and configuration and one for WLAN channel on which voice traffic is transmitted. [44] Wireless security creates hurdles for many companies in implementing 802.11 networks and the same issues exist for VoIP to transmit over wireless. This is because the VoIP packets are carried by airwaves and it is easier for a hacker to hack into them while moving across the network. [40]

7.10 Three-way War:

The wireless market was already established even before the VoIP came. Cellular companies are much more active as compared to the VoIP service providers. VoIP service providers see the opportunity to get back into the market by acquiring cellular companies. There is tough competition between VoIP service providers and cellular companies, and this competition is in favor of consumers in the long run. However, in the short run, to spend a lot of cash for wireless technology could prove to be the wireless phone equivalent to someone investing in a Betmax video recorder a few decades ago. [42]

8. Network Scenario and Measurement Tools:

In this chapter, we will discuss the scenario that we built up for measuring the values of different speech codecs and also the tools that we used to accomplished our task.

8.1 Network Topology:

To conduct this experiment we setup a real network environment that shown in Fig 8.1 to measure the value of VoIP codec, we use three computers all have Windows Xp, out of them one is dedicated Asterisk server and other two laptops are acting as clients with Xlite softphone is installed on it to make a calls between each other. We also use VQ Manager (Voice Quality Manager) to measure the necessary parameters for voice on which we can show that which codec can perform better in wireless networks.

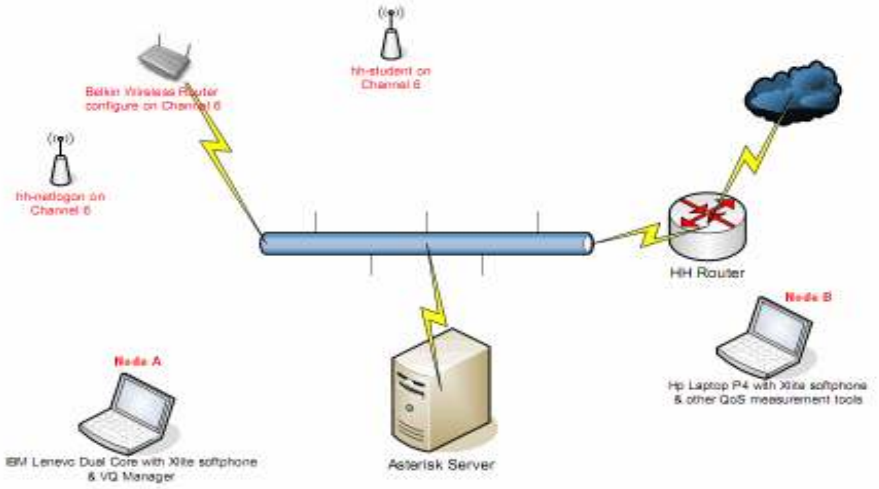


Fig 8.1 Measurement Scenario

8.2 Interference Creation:

Figure 8.2 gives the overview of our working place.

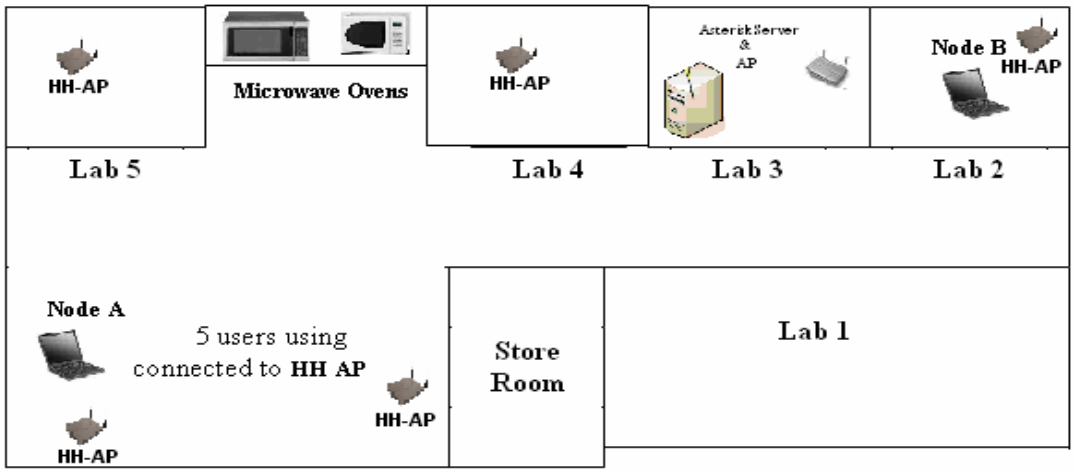


Fig 8.2 Working Environment

We create a lot of interference for taking the results. For this, we setup our experiment wireless router (Belkin) on the same wireless channel (Channel 6) as of the other two wireless APs (**hh-netlogon** and **hh-student**), are already configured on channel 6 as we can see in Figure 8.3.

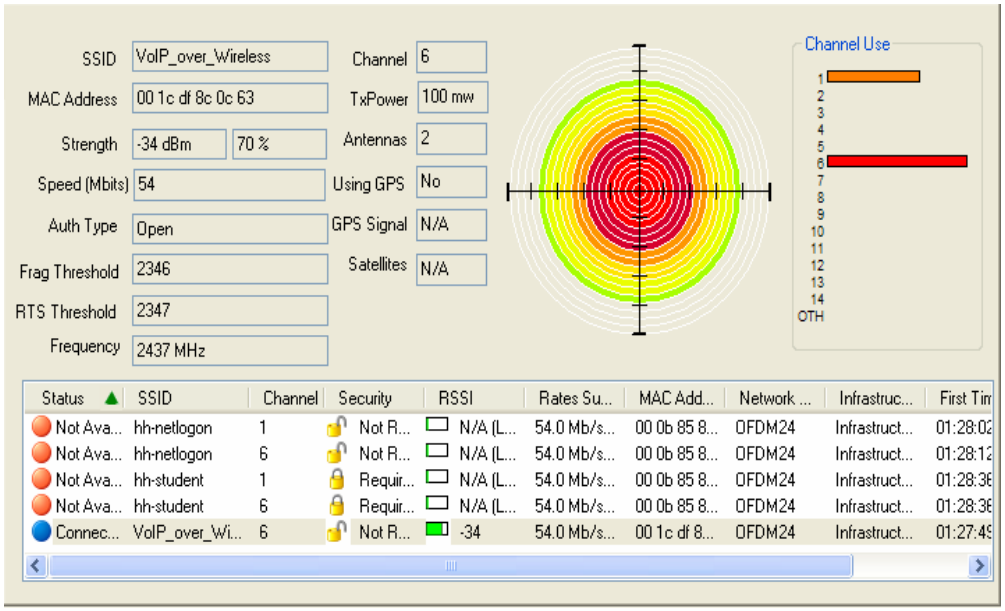


Figure 8.3: Access Point Channel Description

Five users are connected these access points and doing the video streaming continuously while we were making the call. We did this because we do not want to create an ideal situation as we want to create the terrible situation and then want to check that which codec is perform better in such a situation although we know that in working (or industry) environment these things are taken under consideration that two AP are not operate on the same channel.

Figure 8.4 gives the summary of the average utilization of the link between access point and one of the node (from 5 nodes that connect to the HH access points), which shows that how much packets it send and receive in one second.

In our project, we used the sample size of 20 ms for each codec and according to this the total packet that each codec send in one second:

One packet size (sample size) = 20 ms

Total packet in 1 second = $1/(20 \times 10^{-3})$ = 50 packets/second

When we take that value of average packet per second from fig: 8.4 (140 packet/second) and multiplied it with 5, we get 700 packets per second, which means that we have a lot of interference. Table 8.1 gives the overview of the average packets/second for each codec and the interference packets/second on the network.

These calculations are based upon the parameters that we taken defined during our project rather than the measurement, so if we change these parameters than the calculated values also change.

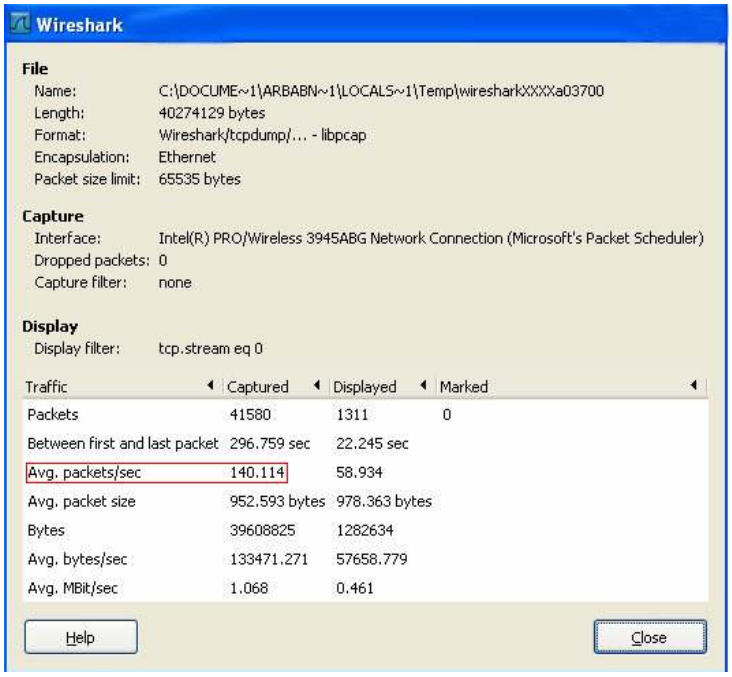


Figure 8.4: Average Packets per Second

Codec	Sample Size	Packet/Second	Average Interference Packet/Second
G.711	20 ms	50	600-700
Speex	20 ms	50	600-700
G.729	20 ms	50	600-700

Table 8.1: Summarized table of codec and interference packets/second

8.3 QoS Model and Queuing Mechanism:

We implement the best effort QoS model because we just treat the voice as regular data. Because in other QoS models implementation, voice perform reasonable well no matter what codec we are using, so we also neglect this factor and do in normal situation (implementing Best Effort model).

Further more we use the FIFO as a queuing mechanism because we want to forward the traffic in the order as it came and do not want to give the priority to the voice traffic. In case if want to prioritize the voice traffic than we use the LLQ (Low Latency Queue).

8.4 Asterisk Server:

Asterisk is free open source software based telephone private branch exchange (PBX) implementation, which was originally created by **Mark Spencer** in 1999. It allows the connected telephones (or softphones) to make calls to each other and also allows connecting to other telephone services like VoIP services.

It was originally designed for Linux system but now it is also available for Microsoft Windows, which is known as **AsteriskWin32.**

In our network environment, we use AsteriskWin32, because of its user friendly graphical interface and configuration. It did not give as much as option that it originally

give on Linux system but still it gives all the options that we need in our study. Figure 8.4 shows the Startup window of AsteriskWin32 server.

Fig 8.4 AsteriskWin32 Startup window

8.5 Softphone:

Fig 8.5: Xlite Softphone

Fig 8.6: Xlite Softphone Codec selection option

On the other hand the software (Xlite) that we used for making a call is free softphone and it did not have any feature except to that we can adjust the speech codec on it which we want to use as shown in Figure 8.5 and 8.6. It cannot perform any error correction or voice quality enhancement. Everything is operating at normal situation.

8.6 Voice Quality Measurement Software:

To measure the voice quality parameters, we used the VQ Manager. VQ Manger is the most reliable and trusted software for monitoring the VoIP network for voice quality, call traffic, bandwidth utilization and keep track of active calls and failed calls. VQ Manager can monitor any device or user-agent that supports SIP, Skinny and RTP/RTCP. [45] It gives the summary of all the parameters for specified time period as well as it show in the form of graphs. It also gives the Information on 'what is going on' in VoIP network and 'how it performs' are presented in the form of comprehensive, intuitive and informative reports. [45]

Fig 8.7: VQ Manger

8.7 CPU Utilization Measurement:

To measure the CPU utilization, we used the NTGM freeware software. Although, it is not a very popular software but still it gives us the idea that how much processor resources used by each codec, we used these result as comparison not to precisely defined that how much processor dependent is each codec.

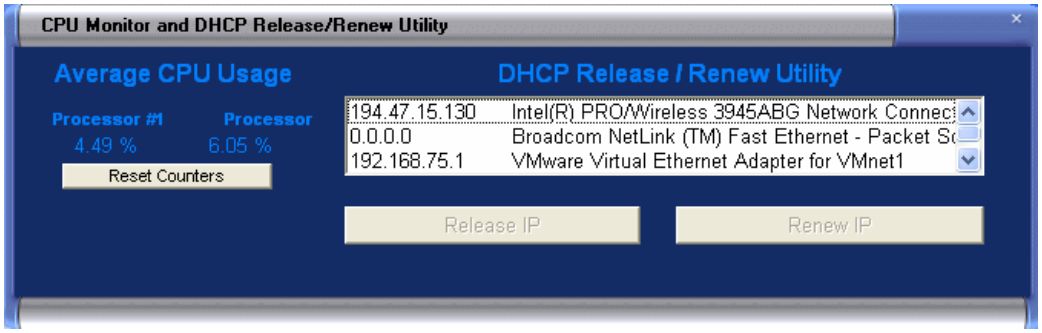


Fig 8.8: NTGM CPU Utilization window

8.8 Wireshark Protocol Analyzer:

Wireshark is free protocol analyzer that is mainly used for troubleshooting in computer networks. It also provides support to troubleshoot the VoIP call setup.
We used Wireshark to check the initial call setup in our scenario. It's really easy to use as it gives us GUI (graphical user interface). We just need to select the interface on which we have to analyze the traffic.

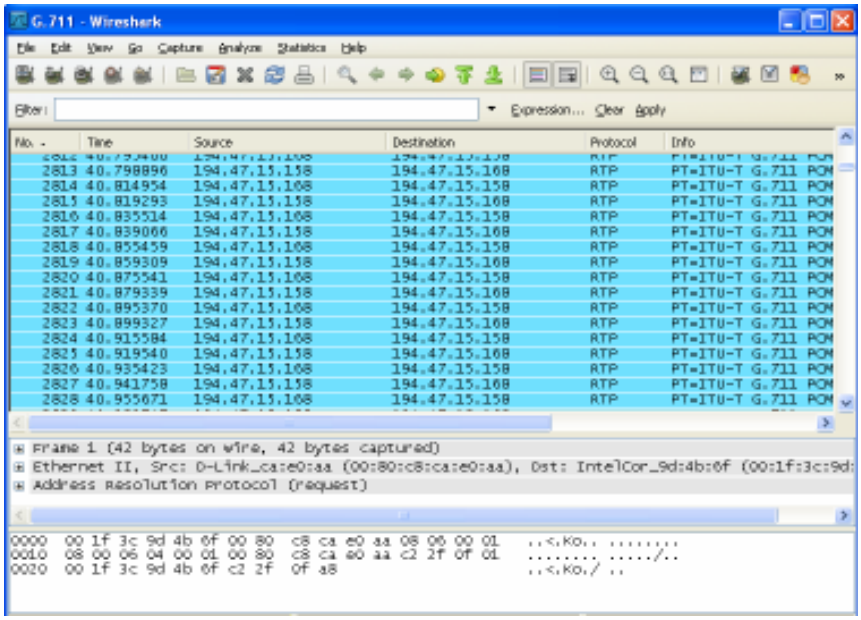


Fig 8.9: Wireshark Protocol Analyzer

9. Results Discussion and Conclusion:

In this chapter we discuss the results that we take from the scenario that we setup in the last chapter. We established a VoIP call from **Node A** to **Node B** and then **Node B** to **Node A** so that we can find the more precise results. The duration of each call was about 10 minutes approximately. We took the time period of each call so long, because we wanted to notice the more and more changes during the call. The results that we got did not reflect any RFC or standard. It is our own described scenario and we want to choose that which codec can do well under bad condition (means interference).

We created a heavy interference as described early; the interference was so strong that sometime we drop the wireless connection. So there is also chance that there is some error in our results but we conducted these test three to four times before finalizing these results but still there is chance that we get some wrong results. But all the results match to the defined characteristics of the codecs.

We have a lot of VoIP codecs that are used these days, but we have chosen these famous codecs for our test G.711, G.729 and Speex. The average CPU utilization before starting any test was 10%.

Now we described the result of each codec one by one that we got in our experiments.

9.0 Available Bandwidth:

Before running any experiment, we checked the available bandwidth with the Cisco Speed Meter software, which is the most authenticated and widely used software in the industry to check the wireless available bandwidth and the result is shown in fig 9.0:

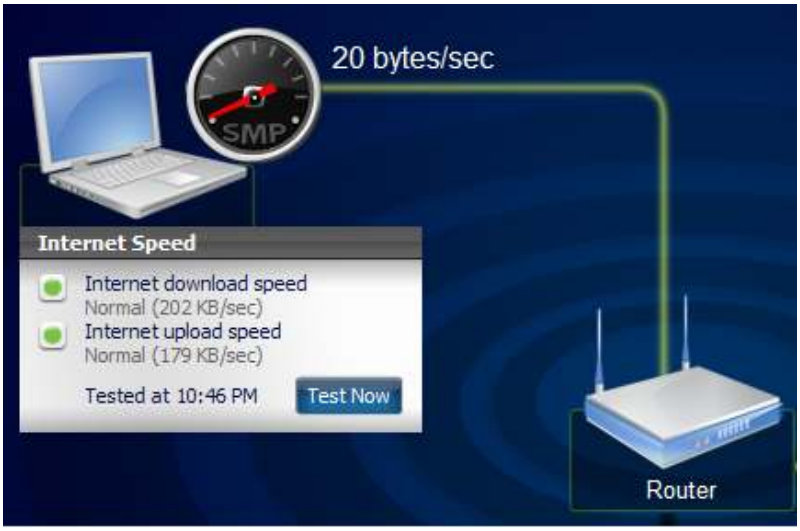


Fig 9.0: Available Bandwidth

As from the above figure, we can see that the available bandwidth is not too high, it is really low but we did this by purpose because we want to check the performance in the terrible situation and the bandwidth was even less when we create more interference as we described before.

9.1 G.711 Pulse code modulation (PCM):

First we enabled the G.711 codec in Xlite softphone on both nodes and established calls between them and find some results which are as follows:

9.1.1 Codec Details:

For the brief introduction, G.711 digitizes the analog voice into 64000 bps or 64Kbps and it did not use any voice compression. This is used as the default standard for the PBX vendors.

CODEC Details			
Name	PCMU	Payload Type	0
Bits per sample	8	Sampling Rate	*Variable*
Frame Size	*Not Applicable*	Packet Size	20 ms
RTP Clock Rate	8000 Hz	No. of Channels	1

Fig 9.1.1: PCM Codec Details

The formal name that use for G.711 codec is Pulse code modulation (PCM) of voice frequencies. G.711 uses the logarithmic pulse-code modulation (PCM) samples for the signals of voice frequencies and it samples at the rate of 8000 samples/second. [46]

The reason for variable sampling rate can be this because we used the latest version of freeware software for testing purpose: "A recent extension to G.711, G.711.1, allows the addition of narrowband and/or wideband (16000 samples/s) enhancements, each at 25 % of the bit rate of the (included) base G.711 bit stream, leading to data rates of 64, 80 or 96 Kbit/s. G.711.1 is compatible with G.711 at 64 Kbit/s, hence an efficient deployment in existing G.711-based voice over IP (VoIP) infrastructures is foreseen. The G.711.1 coder can encode signals at 16 kHz with a bandwidth of 50-7000 Hz at 80 and 96 Kbit/s, and for 8-kHz sampling the output may produce signals with a bandwidth ranging from 50 up to 4000 Hz, operating at 64 and 80 Kbit/s" [46]

In codec detail, this is also mentioned that the frame size is not applicable for this codec and the reason behind this is that G.711 is a sample based codec and this field is not implemented on this codec and the codec that works with the frame size is G.723.1 because it is a frame based codec.[48] There is also a field for the payload type, which indicate that which kind of data the packet is containing and the payload type 0 means that it is PCM audio data according to the IANA (Internet Assigned Numbers Authority).[49]

The formula and table for the values that we used for calculating the bandwidth can be found in Appendix A with reference. The bandwidth calculated using this formula is not absolute because it is theoretical and can be added or subtracted.

$$\text{Bandwidth usage} = \frac{(6+40+160)*64}{160}$$

$$\text{Bandwidth usage} = 82.4\text{Kbp}$$

9.1.2 Initial Call Setup:

Fig 9.1.1 explains the SIP session that it establishes between the two softphones. In the first step, the calling softphone sends the invitation in which it gives the information that it is using G.711 codec and it is representing with **INVITE** keyword. Then the called softphone sends back the response, which represent with **100-Trying**. When the called softphone begins ringing, it sends back response which is represented as **180-Ringing**. When the caller picks up the phone, the called softphone send back its response as **200-OK**. The calling softphone sends the response as **ACK**. Then the actual conversation is transmitted via RTP (G.711). When one end hangs up, the softphone sends the response as **BYE** and the other end is responding with **200-OK.**

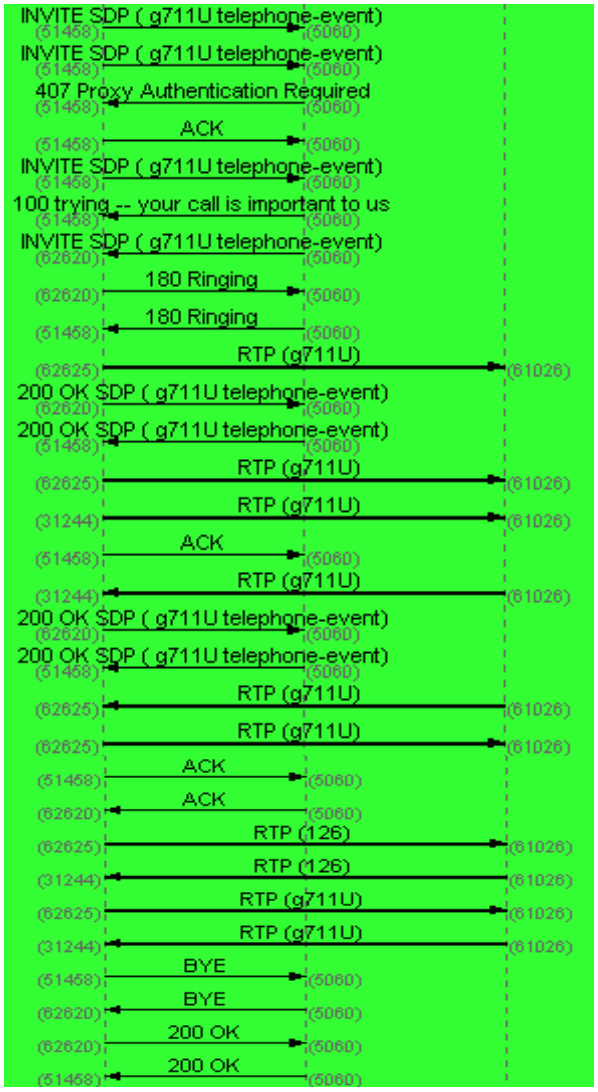
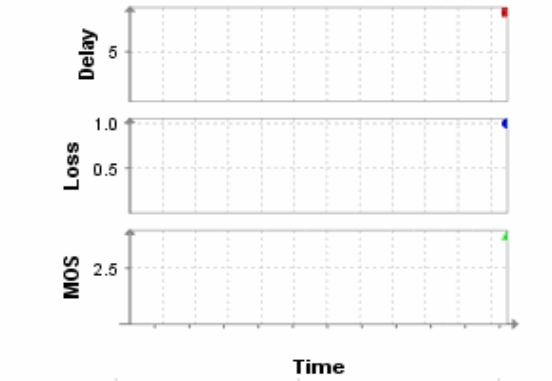


Fig 9.1.2: G.711 Call set between two Nodes

9.1.3 Voice Quality:



	Min	Max	Avg
Delay (ms)	9	9	9
Loss (%)	1	1	1
MOS	4	4	4

Hop	Err	PL%	IP	DNSName	Avg	Min	Max	Cur	Jttr	Graph
1	1	10	192.168.2.1	wl.Belkin	1	1	3	2	0.50	12
2	1	10	192.168.10.1	----------	3	2	12	12	1.38	
3	1	10	194.47.15.130	netlogon130.hh.se	2	2	3	3	0.38	
				Round Trip:	2	2	3	3	0.38	

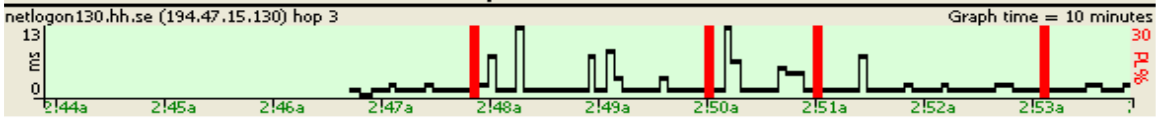


Fig 9.1.3: Voice Quality Parameters

The figures above show the average delay, mean opinion score (MOS) and packet loss. The delay was 9 ms which is acceptable because the maximum delay permitted for VoIP call is 150 ms for one end as well as the packet loss is also acceptable and fulfill the described requirement for VoIP. On the other hand the MOS is really good which is 4 on the scale of 5. The jitter graph shows that the average jitter value is not too much; it is only 0.38 ms, where the red bars represent the delay variation. According to definition of jitter, it is variation in end to end delay and red bars represent where these variations take place during the call. X-axis represents the call duration time and Y-axis represents the random scale for jitter. In the graphs of delay, loss and mos, we see that there are only dots which represent that there was not too much variation in these parameters during the call and only for these and situation or time where the changing took place.

9.1.4 Incoming and Outgoing Call Quality:

These sub graphs represent the delay, loss and mos separately for both incoming and outgoing instead of average. They just give the overview of incoming voice quality and outgoing while on the other hand the graphs in 9.1.3 give the detail information about these parameters in combined form.

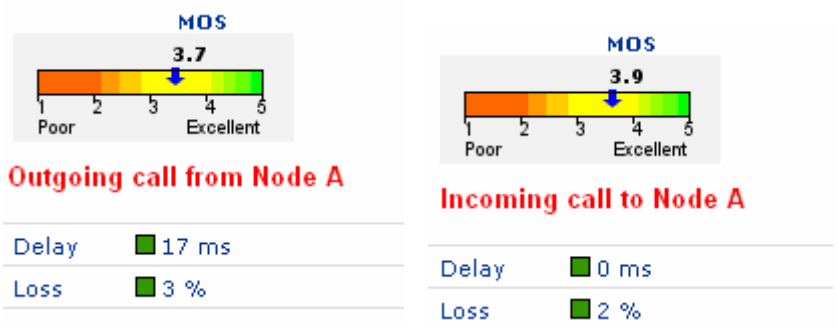


Fig 9.1.4: Incoming and Outgoing Call parameters

9.1.5 CPU Utilization:

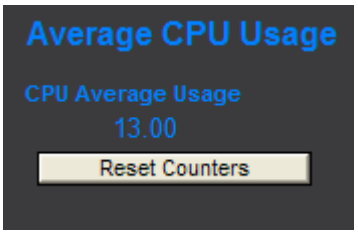


Fig 9.1.5: CPU Utilization

As in the beginning of this chapter, we explained that the average CPU utilization was at 10% before establishing the call. After establishing the call, we noted that the G.711 only creates a load of 3% on the CPU.

9.2 Speex:

Then we enabled the Speex codec in Xlite softphone on both nodes and established between them and find some results which are as follows:

9.2.1 Codec Details:

Speex is free open source codec for audio compression for speech that originally designed for VoIP applications. It is built on the speech coding algorithm named Code excited linear prediction. It is a lossy format, which means the quality is permanently degraded while reducing the file size. Speex compress the voice at the bit rates ranging from 2 to 44Kbps. It uses the sampling rate of 8 KHz (Narrowband), 16 KHz (wideband) and 32 KHz (ultra-wideband).

CODEC Details			
Name	SPEEX	Payload Type	97
Bits per sample	*Not Applicable*	Sampling Rate	8000 Hz
Frame Size	*Not Applicable*	Packet Size	20 ms
RTP Clock Rate	8000 Hz	No. of Channels	1

Fig 9.2.1: Speex Codec Details

The field of bits per sample is not applicable for this codec because it did not use exact bits per sample, it uses variable bit rate. [51] Speex uses the sampling rate of 8 KHz, which means it uses the narrowband here. There is also mention that the frame size is not applicable for this codec and the reason behind this is that speex is a sample based codec

and this field is not implement on this codec and the codec that work with the frame size is G.723.1 because it is a frame based codec.[48] There is also field for the payload type, which indicate that which kind of data the packet is containing and the payload type 97 is from the dynamic range that defined by the IANA (Internet Assigned Numbers Authority) for the voice and video application.[49]

The formula used for calculating the bandwidth can be found in Appendix A with reference and parameters used in this formula are taken from reference 5. The bandwidth calculated using this formula is not absolute because it is theoretical and can be plus or minus.

$$\text{Bandwidth Usage} = \frac{(6+40+30)*11}{30}$$

$$\text{Bandwidth Usage} = 27.8 \text{ Kbps}$$

9.2.2 Initial Call Setup:

Fig 9.2.1 shows the SIP session that it establishes between the two softphones using the speex codec.

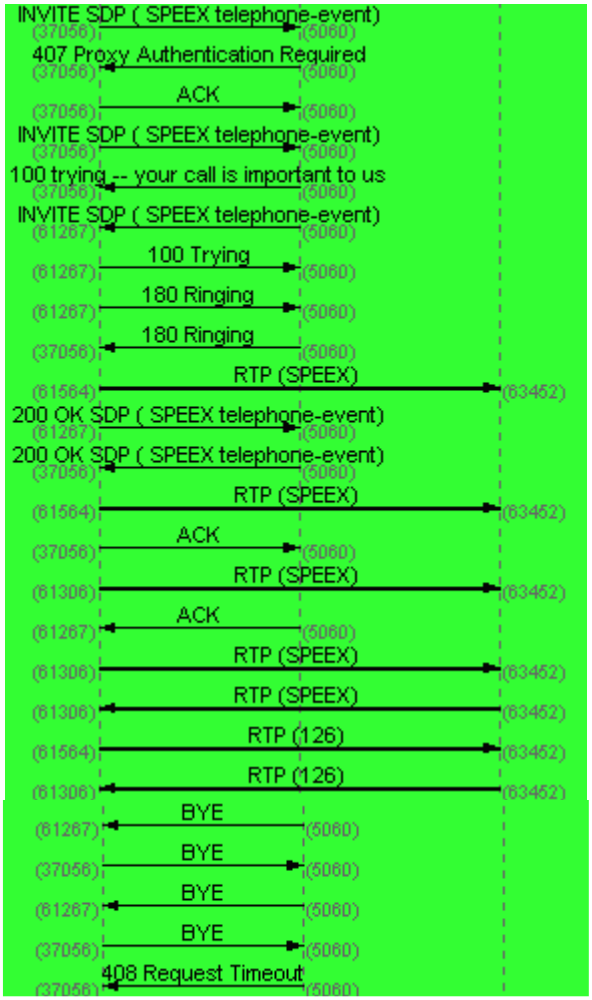
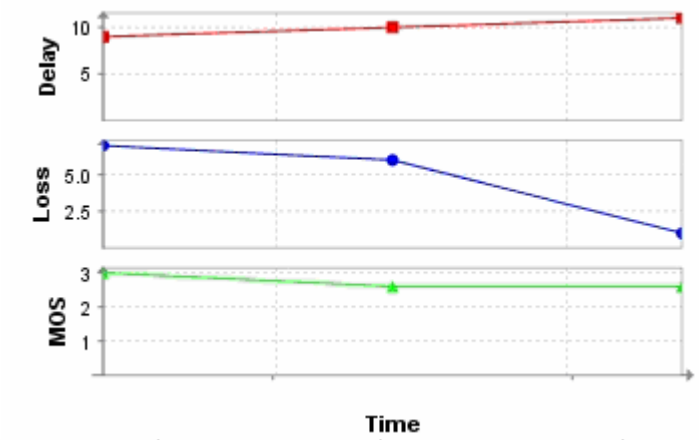


Fig 9.2.2: Speex Call set between two Nodes

9.2.3 Voice Quality:



	Min	Max	Avg
Delay (ms)	9	11	10
Loss (%)	1	7	6
MOS	2.6	3	2.8

Hop	Err	PL%	IP	DNSName	Avg	Min	Max	Cur	Jttr	Graph
1			192.168.2.1	wl.Belkin	5	1	23	1	6.57	
2			192.168.10.1	----------	2	2	5	3	0.71	
3			194.47.15.130	netlogon130.hh.se	6	2	22	3	6.00	
				Round Trip:	6	2	22	3	6.00	

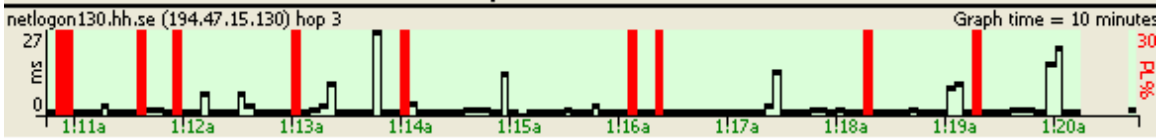


Fig 9.2.3: Voice Quality Parameters

The figures above show the average delay, mean opinion score (MOS) and packet loss. The delay was 10 ms which is acceptable because the maximum delay permitted for VoIP call is 150 ms for one end but the packet loss is not acceptable for good voice quality, it should not be more than 1% otherwise it create jerk in the voice call. On the other hand the MOS is 2.8 on the scale of 5 which is not too good but just acceptable. The jitter graph shows that the average jitter value that is acceptable and is 6 ms, where the red bars represent the delay variation. According to definition of jitter, it is variation in end to end delay and red bars represent where these variations take place during the call. X-axis represents the call duration time and Y-axis represents the random scale for jitter. All the units are in millisecond. In the delay graph, it shows that the delay for voice is almost same because it did not show any change in the curve while in packets loss graph, the curve goes down in the middle which shows that the interference gradually decrease after that point and mos graph, the curve also goes straight which shows that the even the decrease in the packet loss does not affect on it.

9.2.4 Incoming and Outgoing Call Quality:

These sub graphs represent the delay, loss and mos separately for both incoming and outgoing instead of average. They just give the overview of incoming voice quality

and outgoing while on the other hand the graphs in 9.2.3 give the detail information about these parameters in combined form.

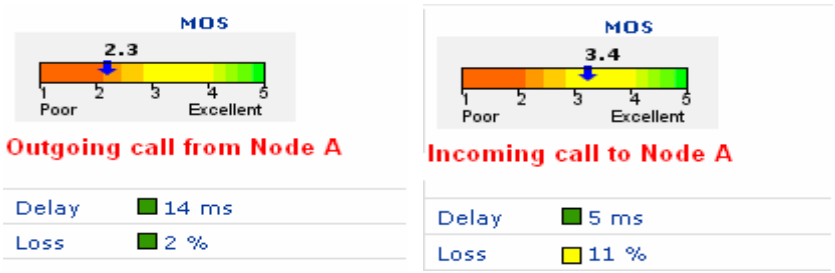


Fig 9.2.4: Incoming and Outgoing Call parameters

9.2.5 CPU Utilization:

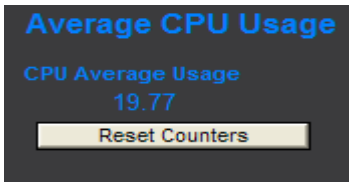


Fig 9.2.5: CPU Utilization

The average CPU utilization was at 10% before establishing the call. After establishing the call, we noted that the Speex creates a load of about 10% on the CPU utilization.

9.3 SpeexFEC:

Then we enabled the SpeexFEC codec in Xlite softphone on both nodes and established between them and find some results which are as follows:

9.3.1 Codec Details:

SpeexFEC (Forward Error Correction) is the enhancement to the speex codec in a way that SpeexFEC sends some error correction code in each packet to avoid the retransmission and detect the receiving system to sends some error correction code in each packet to improve the reliability of data by putting some known code into the data before transmission. This feature enables the receiving system to detect and try to correct the errors due to corruption from the receiver. This technique enables the decoder to correct the errors without retransmission of the original information.

CODEC Details			
Name	SPEEX-FEC	Payload Type	105
Bits per sample	*Not Applicable*	Sampling Rate	8000 Hz
Frame Size	*Not Applicable*	Packet Size	20 ms
RTP Clock Rate	8000 Hz	No. of Channels	1

Fig 9.3.1: SpeexFEC Codec Details

The field of bits per sample is also not applicable for this codec because it did not use exact bits per sample, it uses variable bit rate. [51] It also uses the sampling rate of 8 KHz. There is also mention that the frame size is not applicable for this codec and the reason behind this is that SpeexFEC is also a sample based codec and this field is not implement on this codec and the codec that work with the frame size is G.723.1 because

it is a frame based codec.[48] There is also field about the payload type, which indicate that which kind of data the packet is containing and the payload type 105 is from the dynamic range that defined by the IANA (Internet Assigned Numbers Authority) for the voice and video application.[49]

9.3.2 Initial Call Setup:

Fig 9.3.1 shows the SIP session that it establishes between the two softphones using the SpeexFEC codec.

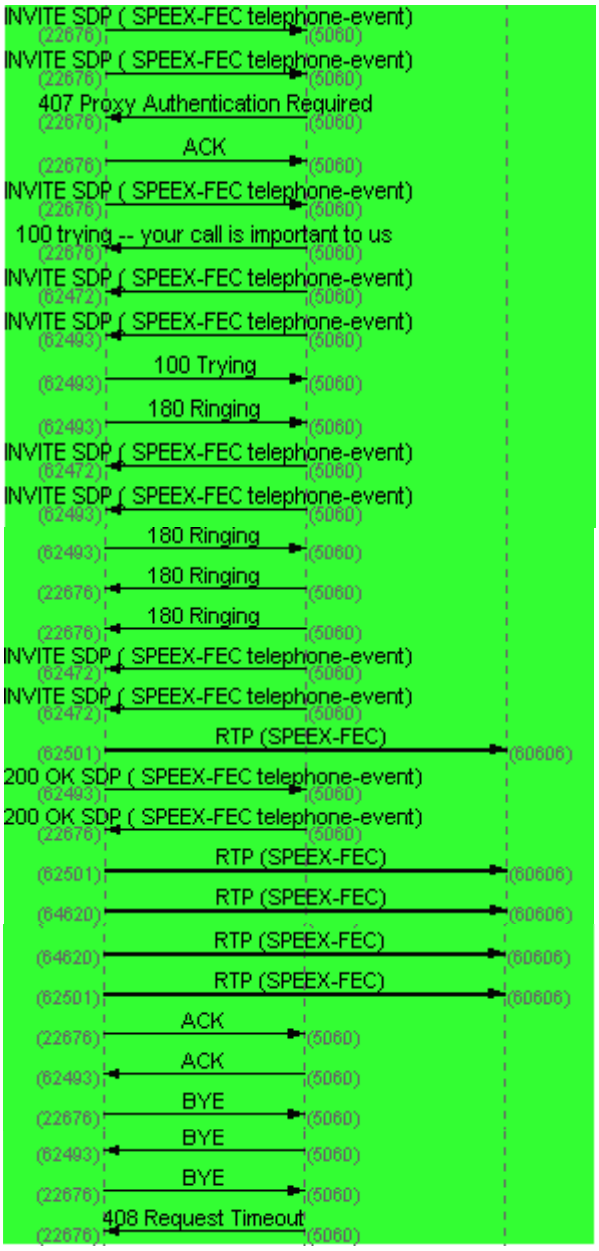
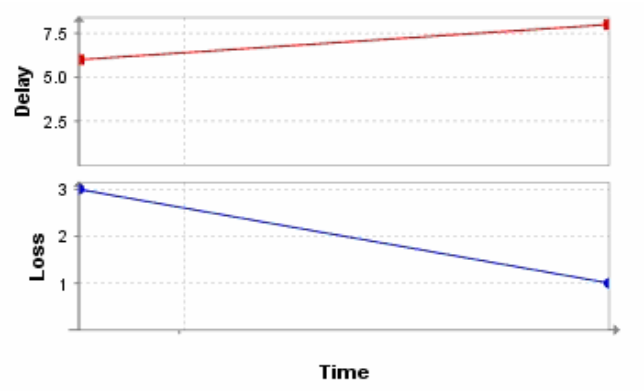


Fig 9.3.2: SpeexFEC Call set between two Nodes

9.3.3 Voice Quality:



	Min	Max	Avg
Delay (ms)	6	8	7
Loss (%)	1	3	3
MOS	3.4	3.5	3.5

Hop	Err	PL%	IP	DNSName	Avg	Min	Max	Cur	Jttr	Graph
1	9	90	192.168.2.1	wl.Belkin	1	1	1	ERR	0.00	
2	9	90	192.168.10.1	----------	2	2	2	ERR	0.00	
3	9	90	194.47.15.130	----------	2	2	2	ERR	0.00	
				Round Trip:	2	2	2	ERR	0.00	

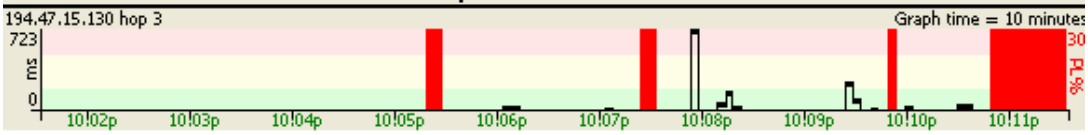


Fig 9.3.3: Voice Quality Parameters

The figures above show the average delay, mean opinion score (MOS) and packet loss. The delay was 7 ms which is acceptable because the maximum delay permitted for VoIP call is 150 ms for one end. The loss rate is about 3% which is not good but acceptable in the environment where there is lot of interference. The MOS is 3.5 which is fairly good when we compare to the Speex codec. The value of jitter is 0 which is incredible, which shows that the SpeexFEC is able to manage the end to end delay efficiently. X-axis represents the call duration time and Y-axis represents the random scale for jitter. All the units are in millisecond. In the delay graph, it shows that it has control in variation of delay because the curve slightly changes during the call. In the packet loss graph, it shows that the packet loss is less because it tries to recover the corrupted packets during transmission.

9.3.4 Incoming and Outgoing Call Quality:

These sub graphs represent the delay, loss and mos separately for both incoming and outgoing instead of average. They just give the overview of incoming voice quality and outgoing while on the other hand the graphs in 9.3.3 give the detail information about these parameters in combined form.

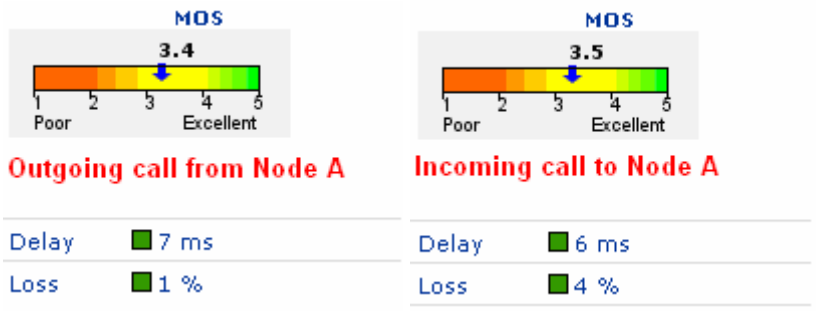


Fig 9.3.4: Incoming and Outgoing Call parameters

9.3.5 CPU Utilization:

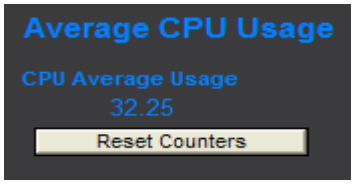


Fig 9.3.5: CPU Utilization

The CPU utilization was high with the SpeexFEC codec and it was about 22% (32%-10%), before the call establishment the CPU utilization was about 10 to 11%. The reason behind the high CPU utilization is, because the CPU has to work on the error correction also during the encoding and decoding process.

9.4 SpeexWideband:

Then we enabled the SpeexWideband codec in Xlite softphone on both nodes and established between them and find some results which are as follows:

9.4.1 Codec Details:

CODEC Details			
Name	SPEEX	Payload Type	100
Bits per sample	*Not Applicable*	Sampling Rate	16000 Hz
Frame Size	*Not Applicable*	Packet Size	20 ms
RTP Clock Rate	8000 Hz	No. of Channels	1

Fig 9.4.1: SpeexWideband Codec Details

Speex Wideband uses the 16 KHz sampling rate and it divide into two bands of 8 KHz signal in which one is representing the low band (0-4KHz) and other high band (4-8 KHz). [51] There is also stated that the frame size is not applicable for this codec and the reason behind this is that Speex Wideband is also a sample based codec and this field is not implement on this codec.[48] There is also field about the payload type, which indicate that which kind of data the packet is containing and the payload type 100 is from the dynamic range that defined by the IANA (Internet Assigned Numbers Authority) for the voice and video application.[49]

9.4.2 Initial Call Setup:

Fig 9.4.1 shows the SIP session that it establishes between the two softphones using the Speex Wideband codec.

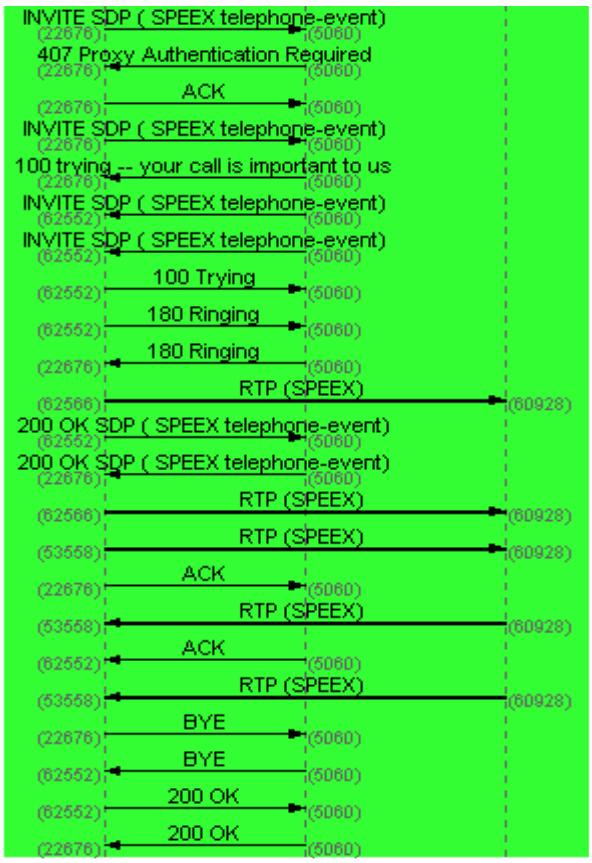
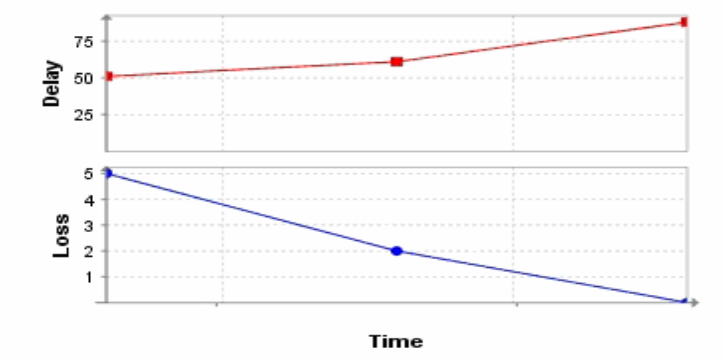


Fig 9.4.2: SpeexWideband Call set between two Nodes

9.4.3 Voice Quality:



	Min	Max	Avg
Delay (ms)	51	88	56
Loss (%)	0	5	3
MOS	2.4	3.8	2.8

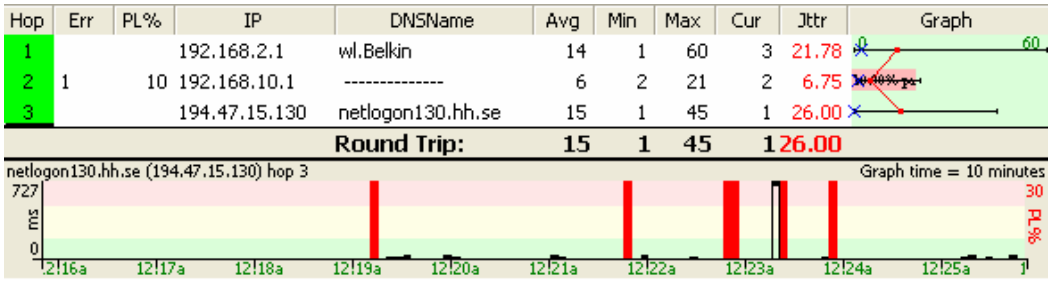


Fig 9.4.3: Voice Quality Parameters

The graphs above show the average delay, mean opinion score (MOS) and packet loss. The delay was 56 ms that is high when we compare to the other version of the Speex codec but still within the range of 150 ms. The loss rate is about 3% which is not good but acceptable in the environment where there is lot of interference. The MOS is 2.8 which are not good. The value of jitter is 26 ms which reflect the delay. X-axis represents the call duration time and Y-axis represents the random scale for jitter. All the units are in millisecond. In the delay graph, it shows that the delay was gradually increase during the call while on the other hand the packet loss decrease which may conclude that the codec take more time for encoding and send few packets that is reason the loss graph gradually decrease.

9.4.4 Incoming and Outgoing Call Quality:

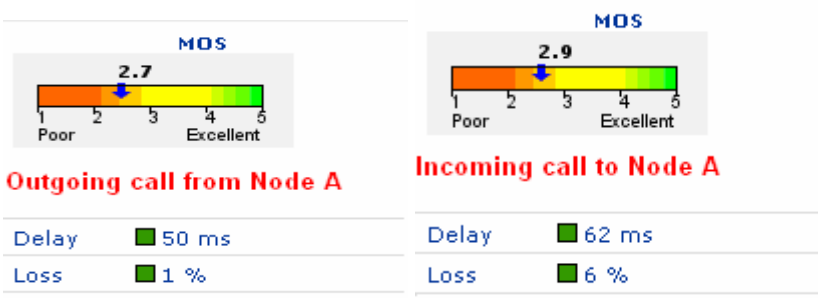


Fig 9.4.4: Incoming and Outgoing Call parameters

These sub graphs represent the delay, loss and mos separately for both incoming and outgoing instead of average. They just give the overview of incoming voice quality and outgoing while on the other hand the graphs in 9.4.3 give the detail information about these parameters in combined form.

9.4.5 CPU Utilization:

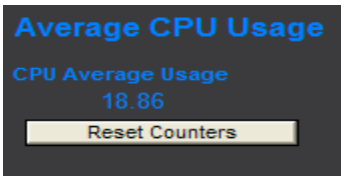


Fig 9.4.5: CPU Utilization

The CPU utilization shows that the wideband did not create too much load on the CPU; it created almost the same load as the Speex narrowband created on the CPU.

9.5 SpeexWidebandFEC:

In this section, we show the result that we take after enabled the SpeexWidebandFEC codec.

9.5.1 Codec Details:

Speex WidebandFEC also uses the same 16 KHz sampling rate and it is also a sample based codec as indicated by the frame size field.[48] There is also field about the payload type, which indicate that which kind of data the packet is containing and the payload type 106 is from the dynamic range that defined by the IANA (Internet Assigned Numbers Authority) for the voice and video application.[49]

CODEC Details			
Name	SPEEX-FEC	Payload Type	106
Bits per sample	*Not Applicable*	Sampling Rate	16000 Hz
Frame Size	*Not Applicable*	Packet Size	20 ms
RTP Clock Rate	8000 Hz	No. of Channels	1

Fig 9.1.1: SpeexWidebandFEC Codec Details

9.5.2 Initial Call Setup:

Fig 9.4.1 shows the SIP session that it establishes between the two softphones using the Speex WidebandFEC codec.

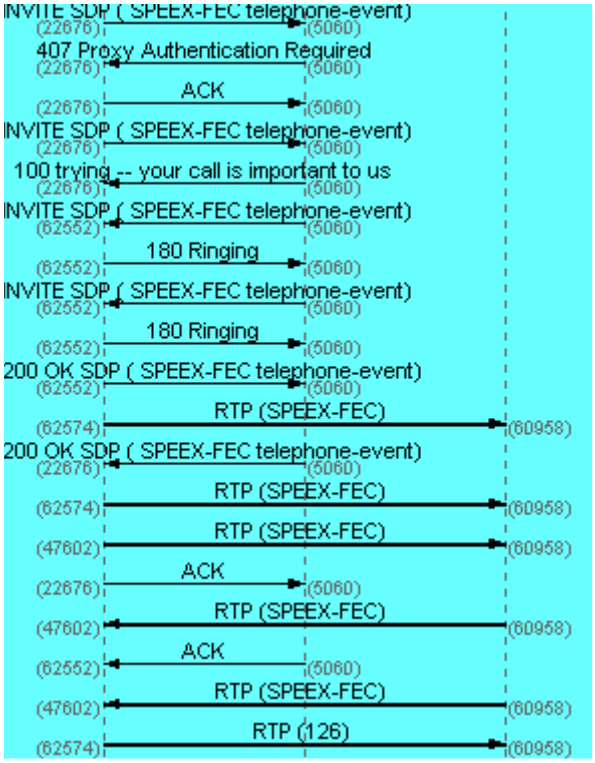
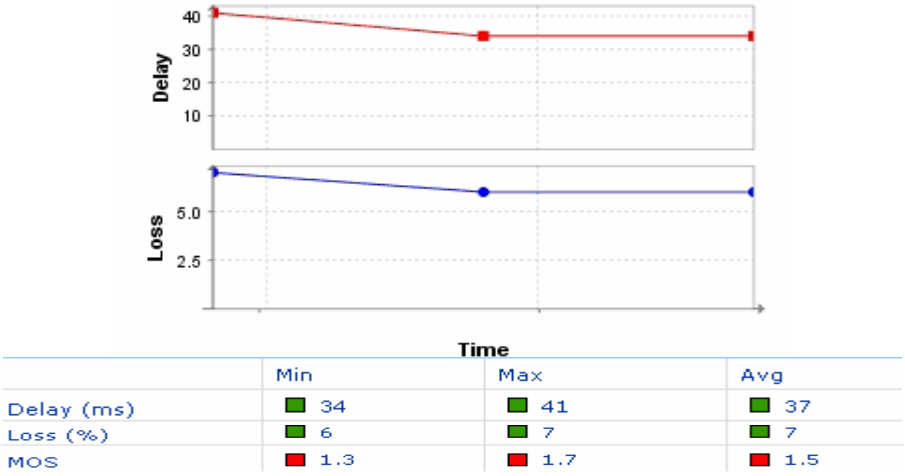


Fig 9.5.2: SpeexWidebandFEC Call set between two Nodes

9.5.3 Voice Quality:



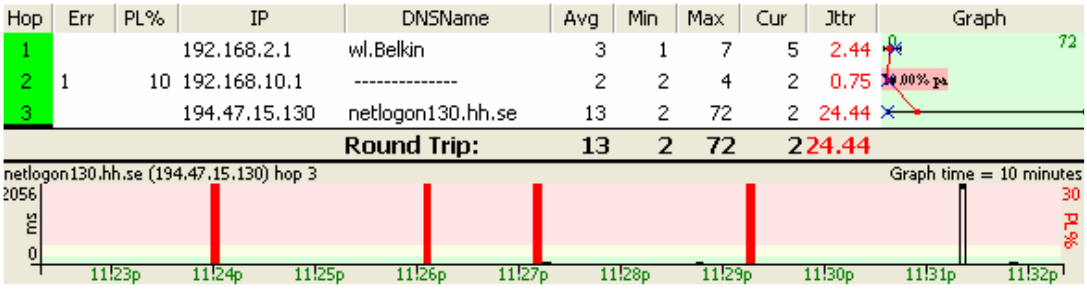


Fig 9.5.3: Voice Quality Parameters

The graphs above show the average delay, mean opinion score (MOS) and packet loss. The delay was 37 ms that is not too high when we compare to the other version of the Speex codec. The loss rate is about 7% which is too high for voice traffic and is not acceptable and the reason behind this loss may be interference that affects the wireless link quality. The MOS is 1.5 out of 5 which shows that it perform really bad and not acceptable. The value of jitter is 24 ms and it reflects the variation in end to end delay. X-axis represents the call duration time and Y-axis represents the random scale for jitter. All the units are in millisecond. In the delay graph, it shows that the delay was almost constant during the call session and packet loss also remains the same during this time when the call session was established. The reason may be the same that the link quality may be too bad that the delay time and packet loss rate did not improve during the call.

9.5.4 Incoming and Outgoing Call Quality:

These sub graphs represent the delay, loss and mos separately for both incoming and outgoing instead of average. They just give the overview of incoming voice quality and outgoing while on the other hand the graphs in 9.5.3 give the detail information about these parameters in compact form.

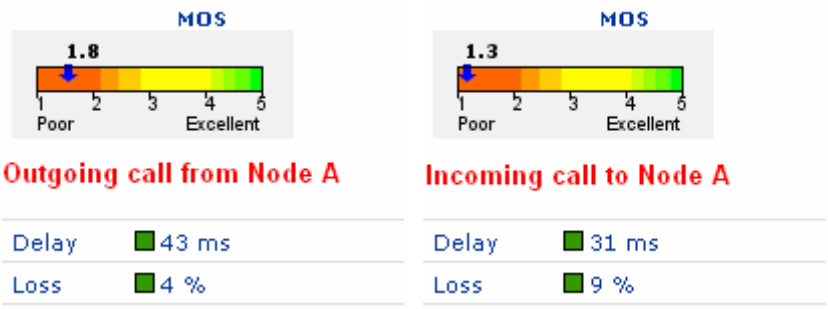


Fig 9.5.4: Incoming and Outgoing Call parameters

9.5.5 CPU Utilization:

The CPU utilization of Speex WidebandFEC is high when compare to other codecs and is about 15% when minus from the initial CPU utilization before establishing the call. When we compare its CPU utilization with its described features, it is not too much.

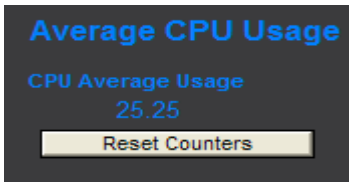


Fig 9.5.5: CPU Utilization

9.6 G.729 codec:

G.729 is mostly used these days for VoIP application. It is proprietary codec; we enabled it on our softphone on both sides and take some results that are as follows:

9.6.1 Codec Details:

This operates at 8 Kbps and 8 KHz sampling frequency by using the **Conjugate Structure Algebraic Code-Excited Linear Prediction** algorithm. It also uses a human voice "Codebook" as a dictionary to work and Look-Ahead of 5ms. Special mathematical algorithms are used for voice synthesis. The complexity lies at 15, because of its low bandwidth requirement. It is the most favorable codec that is used in VoIP applications nowadays. [53]

CODEC Details			
Name	G729	Payload Type	18
Bits per sample	*Not Applicable*	Sampling Rate	8000 Hz
Frame Size	10 ms	Packet Size	20 ms
RTP Clock Rate	8000 Hz	No. of Channels	1

Fig 9.6.1: PCM Codec Details

It requires 10 ms input frames as indicated by the frame size field and generates frames of 80 bits long.[54] Bits per sample are not implementing for this codec as it works

with frames. It uses the payload type of 18 which is reserved for G.729 codec by the IANA (Internet Assigned Numbers Authority). [49]

The formula and table for the values that we used for calculating the bandwidth can be found in Appendix A with reference. The bandwidth calculated using this formula is not absolute because it is theoretical and can be plus or minus.

$$\text{Bandwidth usage} = \frac{(6+40+20)*8}{20}$$

Bandwidth usage =26.4 Kbps

9.6.2 Initial Call Setup:

Fig 9.5.1 shows the SIP session that it establishes between the two softphones using the G.729 codec.

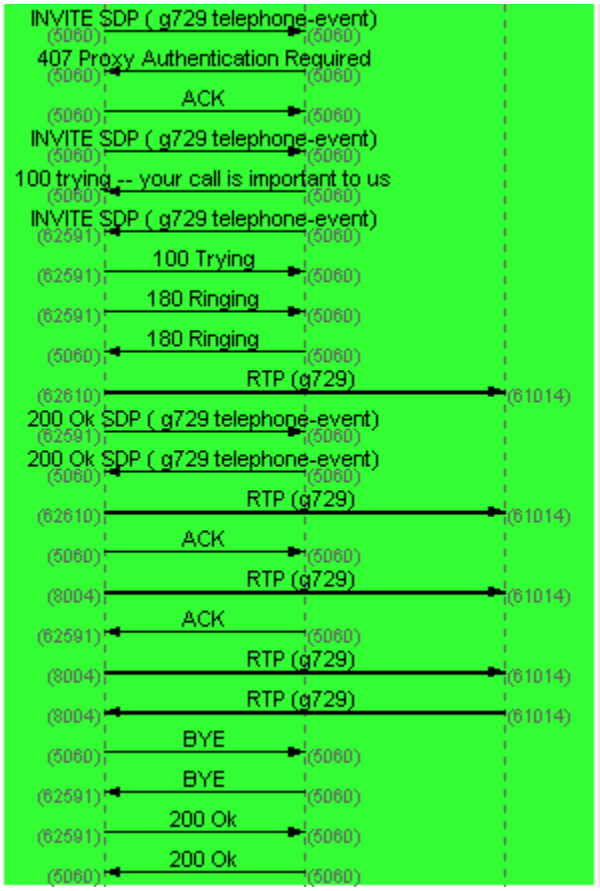


Fig 9.6.2: G.729 Call set between two Nodes

9.6.3 Voice Quality:

The graphs below represent the jitter, loss and average delay for G.729 codec. The delay of 11 ms is acceptable and within the range of 150 ms for one end of VoIP call. The MOS for this codec is really good that is 4 out of 5. Jitter graph represent the average jitter value that is about 2% which shows that the there is not too much delay variation. . X-axis represents the call duration time and Y-axis represents the random scale for jitter. In the graphs of delay, loss and jitter, we see that there are only dots which represent that there was not too much variation in these parameters during the call and only these and situation or time where the changing took place.

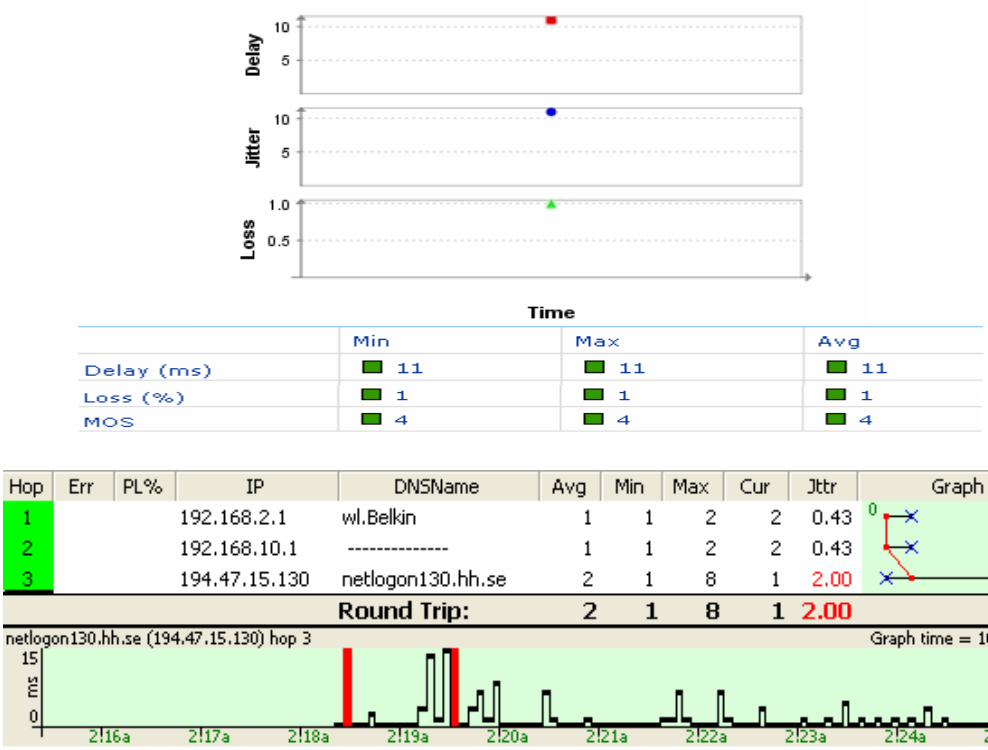


Fig 9.6.3: Voice Quality Parameters

9.6.4 CPU Utilization:

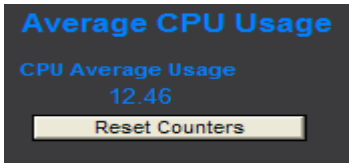


Fig 9.6.4: CPU Utilization

The CPU utilization of this codec is not too much when compare to pervious codec. It seems to be CPU friendly codec, it only create a load of about 2 to 2.5% load on the CPU.

9.7 General Comment on Results:

We established twelve calls to get these six codecs results (in reality we established more than 60 calls to get the average of these 60 calls). It was a long process and we noticed a lot of changes during the calls. The most notable change was, that when we established a call using G.729 codec the call dropped if the interference was too heavy while on the other hand Speex codec never dropped the calls no matter how much the interference was (this is true in our case, in reality it is not promised) but the quality of the call was degraded a lot due to the interference. While we were making the call using Speex codec, we create an extra interference using Microwave oven (because it also operates on 2.4 GHz), still the call was not dropped but the quality was so poor. We record some of these calls are uploaded at **http://www.ziddu.com/download/6904326/CodecsResults.rar.html** in the form of wav file.

In general all the codec perform according to standard description but our purpose was not to check that the standard says right or wrong. Our main purpose was to check that which codec perform reasonable even in horrible environment. When we established a call using G.729 Codec, the quality was good but there was some echo. At some stage during the call the echo problem even dominated and it also performed badly. So it quality was also varying during call.

SpeexFEC codec was the one which performed really well although it create the load on CPU was about 21% but it has jitter value of zero many time and sometime 0.01 but on the average in has jitter value of zero which seems to be unbelievable. When we establish the call using SpeexWideband with FEC (forward error correction) codec, the thing was even worse. We heard the voice of the caller after long time, it was really annoying. This codec was the only that perform really badly. But still in this condition the call was established which proved one thing that Speex is a robust codec.

The highest CPU utilization was for SpeexFEC codec and lowest was for G.729 codec. These values are not absolute, when we change the environment the values will also change. So all these values are dependent on system and environment in which the wireless network is operating.

9.8 Conclusion:

From the results, discussion and most importantly after hearing all the voice codec draw the conclusion that if we have enough bandwidth then G.711 is the best choice because of its less delay, low packet loss and voice quality. But is not robust codec and it cannot be able to adapt the change in the bandwidth and drop the call. On the other hand Speex perform reasonable well, by mean of reasonable that it did not drop the call and adopt the change in the bandwidth. It is also a free open source codec but it did not give the voice quality like G.711 because the reason behind it compresses the voice using Code excited linear prediction algorithm that is a lossy format, which means the quality is permanently degraded while reducing the file size while on the other hand it gives the flexibility by compressing the voice at the bit rates ranging from 2 to 44Kbps. It uses the sampling rate of 8 KHz (Narrowband), 16 KHz (wideband) and 32 KHz (ultra-wideband). If we are not too much quality conscious but we want a robust codec that adapt the change in our wireless environment then SpeexFEC is the reasonable codec. It creates some load on the processor and has little bit more delay then other proprietary

codecs but this load did not create a significant effect on the modern PC because we were using old computer in our environment, may be that is the reason the CPU utilization value is too much.

If we want constant and low bandwidth codec then G.729 is the best choice but we have to figure out the solution for echo cancellation as well as it is the proprietary codec, which means we have to buy a licensed for it. For the echo cancellation G.168 is the best choice. G.168 address the performance issue of echo canceller in PSTN by strictly limits the convergence time allow residual echo, tolerance for varying signal levels and allowed divergence in the presence of destabilizing narrow-band energy. For VoIP the echo canceller can be put after an audio codec to minimize the echo from local hardware. Which means more money we have to spend but if we look it in long run, we can save a lot of bandwidth which mean save money.

We can increase the efficiency and save the bandwidth by using the RTP header compression. By using this technique, we compress the RTP, UDP and IP headers from 40 bytes to in between 4 - 6 bytes and the bandwidth consumption reduced by 60% for compressed voice packets.

To sum up, we can say this that each codec had it own unique properties and there is not even a single codec which can be 100% perfect for wireless network. Although Speex was develop for VoIP but not for Wireless VoIP but still it perform really well.

10. Reference:

[1] www.fcc.gov/voip/
[2] "Performance Evaluation of VoIP Services using Different CODECs over a UMTS Network" by Jianguo Cao and Mark Gregory.
[3] blogs.techrepublic.com.com/
[4] www.cambridge.org/us/catalogue/wireless
[5]www.wirelessnets.com/resources/downloads/wireless_industry_report_2007
[6] oreilly.com/
[7] **CCNP Quick Reference** by Denise Donohue, Brent Stewart and Jerold Swan "ISBN-13: 978-1-58720-236-0"
[8] en.wikipedia.org/wiki/Wireless_LAN
[9] "**WIRELESS LAN SECURITY AND IEEE 802.11I"** by **"**JYH-CHENG CHEN, MING-CHIA JIANG, AND YI-WEN LIU". IEEE Wireless Communications • February 2005
[10] www.wi-fiplanet.com/tutorials/
[11] www.ciscopress.com/articles/ wireless
[12] "**Computer and Communication Networks"** by ***Nader F. Mir***. *ISBN*-10: 0-13-174799-1, Pub Date: November 02, 2006.
[13] www.technology.ku.edu/network/services/data/wireless/fcc_802.shtml
[14] communication.howstuffworks.com/ip-telephony.htm
[15] searchunifiedcommunications.techtarget.com/
[16] www.livinginternet.com/i/iw_packet_inv.htm
[17] **"Signal Processing First"** by **James H. McClellan**, **Ronald W. Schafer** and **Mark A. Yoder**. **ISBN**-10: 0130909998 and Publisher: **Prentice Hall**
[18] www.ip-voip-service.com/
[19] www.cisco.com/
[20] speex.org/
[21] RFC: 2543 "SIP: Session Initiation Protocol"
[22] http://www.com.dtu.dk/teletraffic/handbook/telenook.pdf
[23] "**CCNP ONT Reference Sheets"** by **Denise Donohue, ISBN: 1-58705-315-2**
[24] www.oreillynet.com/pub/a/network/
[25] en.wikipedia.org/wiki/Quality_of_service
[26] RFC: 791 "DARPA INTERNET PROGRAM PROTOCOL SPECIFICATION"
[27] www.cs.wustl.edu/
[28] **"ACOE422-Wireless Networks Notes"** by **Dr Haris Haralambous**
[29] www.wi-fiplanet.com/
[30] "**QoS over WLAN for the CE World**" by www.metalinkBB.com
[31] en.wikipedia.org/wiki/IEEE_802.11e-2005
[32] [LIU05]Qingwen Liu, Shengli Zhou, Georgios B. Giannakis, "**Cross-Layer Scheduling With Prescribed QoS Guarantees in Adaptive Wireless Networks**," IEEE Journal on Selected Areas in Communications v. 23 no. 5 (May 2005) p. 1056-66
[33] [SWAN]"**SWAN: Service Differentiation in Stateless Wireless Ad Hoc Networks**," http://www.iks.inf.ethz.ch/education/ss04/seminar/41.pdf,[SWAN: Service Differentiation in Stateless Wireless Ad Hoc Networks]
[34] [ZHA05]Baoxian Zhang, Hussein T. Mouftah, "**QoS Routing for Wireless Ad Hoc Networks: Problems, Algorithms, and Protocols**," IEEE Communications Magazine v. 43 no. 10 (October 2005) p. 110-17

[35] www.wi-fi.org/
[36] “Wi-Fi CERTIFIED™ for WMM™ - Support for Multimedia Applications with Quality of Service in Wi-Fi ® Networks”
[37] CCNP ONT Official Exam Certification Guide by **Amir Ranjbar**
[38] “QoS over WLAN for the CE World” by www.metalinkBB.com
[39] msdn.microsoft.com/
[40] http://www.gss.co.uk/
[41] http://www.voipresource.net/
[42] http://www.voip-news.com/
[43] **“Solutions to Performance Problems in VoIP Over a 802.11 Wireless LAN” by** Wei Wang, Soung Chang Liew, and Victor O. K. Li. **From** “IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 54, NO. 1, JANUARY 2005”
[44] ntrg.cs.tcd.ie/
[45] “VQ Manager User manual” http://www.manageengine.com/products/vqmanager/
[46] en.wikipedia.org/wiki/G.711
[47] www.bandcalc.com
[48] www.iana.org/assignments/rtp-parameters
[59] people.xiph.org/~jm/papers/**speex**.pdf
[50] Speex manual from speex.org
[51] www.aero.org/
[52] www.ip-voip-service.com/
[53] www.vocal.com/

11.APPENDIX A:

Packet voice transmission requirements (Bits per second per voice channel)							
Codec	Voice bit rate	Sample time	Voice payload	Packets per second	Ethernet	PPP or Frame Relay	
						RTP	cRTP
G.711	64 Kbps	20 msec	160 bytes	50	87.2 Kbps	82.4 Kbps	68.0 Kbps
G.711	64 Kbps	30 msec	240 bytes	33.3	79.4 Kbps	76.2 Kbps	66.6 Kbps
G.711	64 Kbps	40 msec	320 bytes	25	75.6 Kbps	73.2 Kbps	66.0 Kbps
G.729A	8 Kbps	20 msec	20 bytes	50	31.2 Kbps	26.4 Kbps	12.0 Kbps
G.729A	8 Kbps	30 msec	30 bytes	33.3	23.4 Kbps	20.2 Kbps	10.7 Kbps
G.729A	8 Kbps	40 msec	40 bytes	25	19.6 Kbps	17.2 Kbps	10.0 Kbps
Note: RTP assumes 40-octets RTP/UDP/IP overhead per packet Compressed RTP (cRTP) assumes 4-octets RTP/UDP/IP overhead per packet Ethernet overhead adds 18-octets per packet PPP/Frame Relay overhead adds 6-octets per packet							

[52]Bandwidth requirement $= \frac{(L2 + L3 + payload) * codec\_bandwidth}{payload}$

12. Appendix B:

Medium Complexity	High Complexity
G.711 (a-law and m -law)	G.728
G.726 (all versions)	G.723 (all versions)
G.729a, G.729ab (G.729a AnnexB)	G.729, G.729b (G.729-AnnexB)
Fax-relay	Fax-relay

Codec Complexity Table [53]