BSCS FINAL P

Final Project-2 Documentation

Quality of Service for VOIP

Submitted Date:

June 4, 2007

Project Advisor:

Mr. Imran Ahmad

Group Members:

M Zahid Aslam Mughal	032424-014 (Group Leader)
Muhammad Saleem Butt	032424-016
Sharif Arif	032424-005
Yasir Shabbir	032424-020

School of Science and Technology
University of Management and Technology

Certificate of Approval

Certified that the work contained in this thesis entitled.

Study of Qos of VoIP

Is carried out by Zahid Aslam Mughal (032424014), Muhammed Saleem Butt (032424016), Sharif Arif (032424005), and Yasir Shabbir (032424020) under my supervision and that in my opinion, it is fully adequate, in scope and proper, for the degree of BSCS.

Approved by

Project Advisor: Mr. Imran Ahmad

Signature: ----------

Dated: ----------

We dedicate this project
to our beloved parents and sincere
teachers who always guide us at every
stage of our life.

Acknowledgment

In completing our Final Year Project we can not afford to forget the names of the following personals. They gave us precious time and advice instead of their busy schedule. They cordially provided us required information and coaching at various stages of our Project.

We are very thankful from the core of our Heart to

Mr Imran Ahmad (Project Advisor)
Assistant Professor of UMT

Mr Haroon Ahmad Malik
Executive Corvit Systems

Group Members:

Muhammad Zahid Aslam Mughal
032424-014

Muhammad Saleem Butt
032424-016

Sharif Arif
032424-005

Yasir Shabbir
032424-020

ABSTRACT

Since Voice over Internet Protocol (VoIP) has emerged, there has been a need for a better quality for a voice call. If anyone among you has ever used "Net to Phone" service, will better understand the poor quality of a voice call.

In this Project we have studied different Quality of Service (QoS) techniques in order to improve the VoIP on CISCO platform and made a critical analysis of major QoS techniques. We have also examined the quality of VoIP both with QoS and without QoS. On the basis of study and hand on experience we have made certain conclusions and recommendations for a better QoS.

This Thesis has not only covered the theoretical concepts of different QoS techniques but also shows VoIP QoS configurations on Cisco network. We will also implement all this work in FP-2.

List of Figures

List of Tables

Chapter # 01

Introduction to PSTN

1.1 The Beginning of the PSTN

The Public Switched Telephone Network (PSTN) has been evolving ever since Alexander Graham Bell made the first voice transmission over wire in 1876. [1] The first voice transmission, sent by Alexander Graham Bell, was accomplished in 1876 through what is called a *ring-down* circuit. A ring-down circuit means that there was no dialing of numbers; instead, a physical wire connected two devices. Basically, one person picked up the phone and another person was on the other end (no ringing was involved).

Over time, this simple design evolved from a one-way voice transmission, by which only one user could speak, to a bi-directional voice transmission, whereby both users could speak. Moving the voices across the wire required a carbon microphone, a battery, an electromagnet, and an iron diaphragm. It also required a physical cable between each location that the user wanted to call. The concept of dialing a number to reach a destination, however, did not exist at this time. To further illustrate the beginnings of the PSTN, see the basic four-telephone network shown in Figure 1-1. As we can see, a physical cable exists between each location.

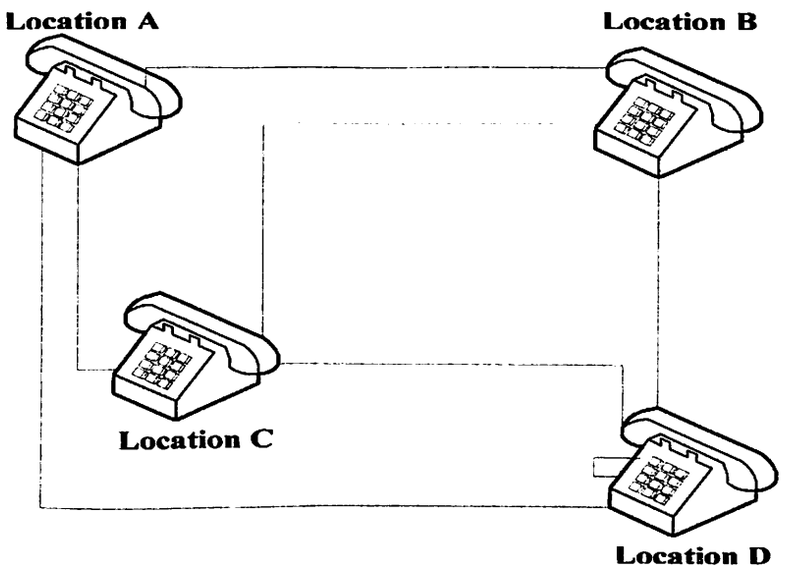


Figure 1-1 Basic Four-Phone Network

Place a physical cable between every household requiring access to a telephone, however, and we'll see that such a setup is neither cost-effective nor feasible. To determine how

many lines we need to our house, think about everyone we call as a value of N and use the following equation:
$N \times (N-1)/2$. As such, if we want to call 10 people, we need 45 pairs of lines running into our house.

Due to the cost concerns and the impossibility of running a physical cable between everyone on Earth who wanted access to a telephone, another mechanism was developed that could map any phone to another phone. With this device, called a *switch*, the telephone users needed only one cable to the centralized switch office, instead of seven. At first, a telephone operator acted as the switch. This operator asked callers where they wanted to dial and then manually connected the two voice paths.

Figure 1-2 shows how the four-phone network example would look today with a centralized operator to switch the calls.

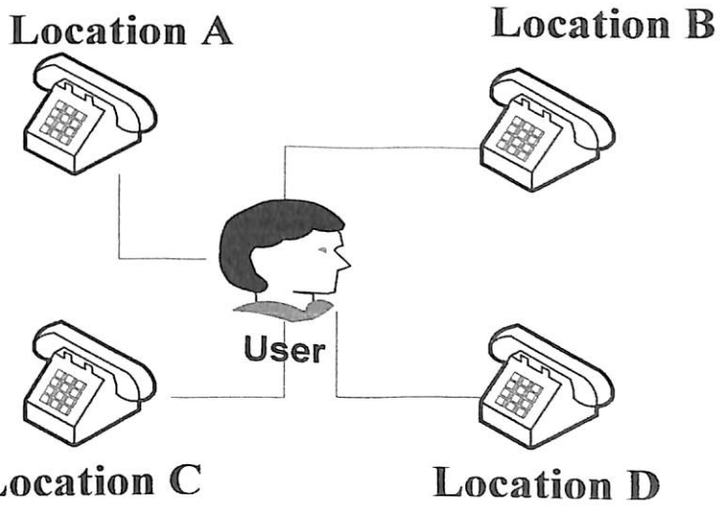


Figure1-2 Centralized Operator: The Human Switch

Now, skip ahead 100 years or so—the human switch is replaced by electronic switches.

1.2 Understanding PSTN Basics

It's not possible to totally understand PSTN basics here. We'll discuss how voice is transmitted in a PSTN, basic switching circuit and why phone number has 10 digits. [2]

1.2.1 Analog and Digital Signaling

In early days voice was transmitted through the channel in analog form. But there were problems associated with it. Main of which was noise and transmitting the signal over longer cables. The problem of distance was solved by passing the signal through amplifiers to boost the signal, but it also boosted the signal.

Analog communication is a mix of time and amplitude. Figure 1-3, which takes a high-level view of an analog waveform, shows what voice looks like through an oscilloscope.

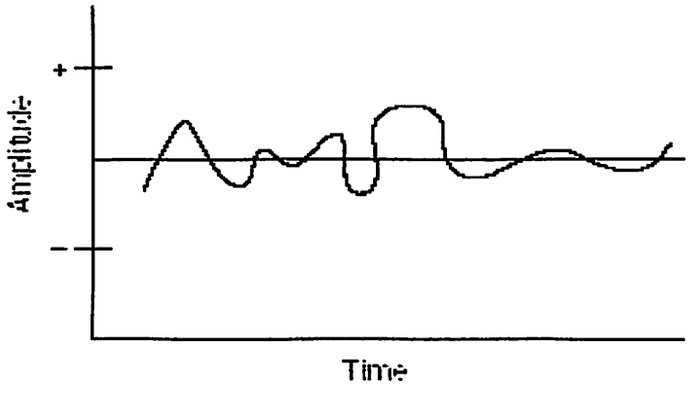


Figure 1-3 Analog Waveforms

As voice is transmitted to our house through an *end office* switch (which provides the physical cable to our home) an amplifier might be required to boost the signal. Analog signals that receive line noise can distort the analog waveform and cause garbled reception. Figure 1-4 shows that an amplifier does not clean the signal as it amplifies, but simply amplifies the distorted signal. This process of going through several amplifiers with one voice signal is called *accumulated noise*.

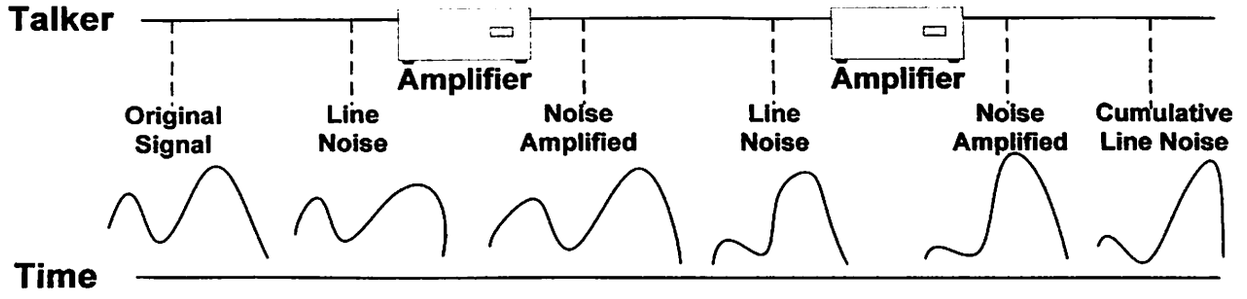


Figure 1-4 Analog Line Distortion

In digital networks, line noise is less of an issue because repeaters not only amplify the signal, but clean it to its original condition. This is possible with digital communication because such communication is based on 1s and 0s. So, as shown in Figure 1-5, the *repeater* (a digital amplifier) only has to decide whether to regenerate a 1 or a 0.

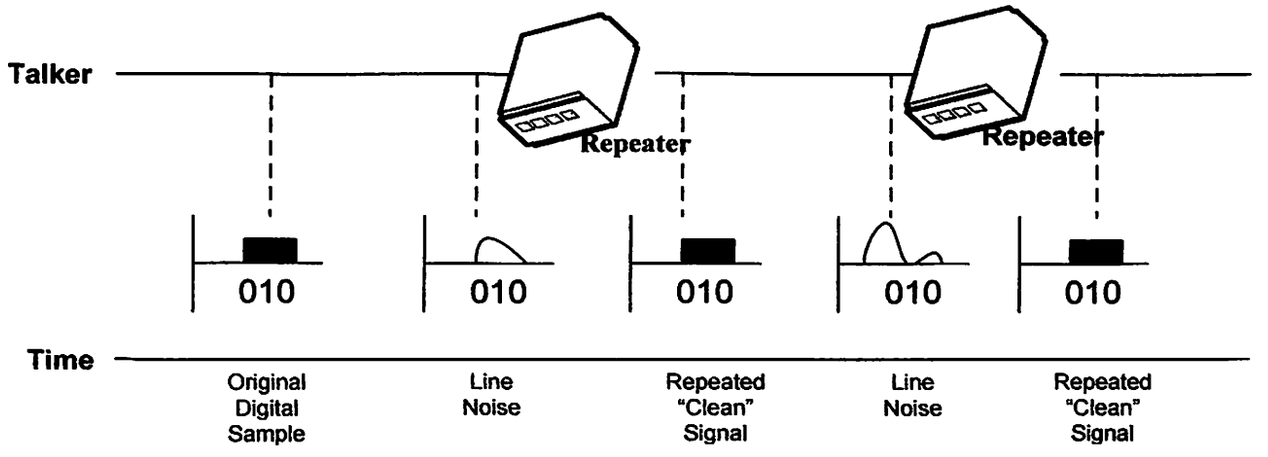


Figure 1-5 Digital Line Distortion

Therefore, when signals are repeated, a clean sound is maintained. When the benefits of this digital representation became evident, the telephony network migrated to *pulse code modulation* (PCM).

1.2.1.1 Digital Voice Signals

PCM is the most common method of encoding an analog voice signal into a digital stream of 1s and 0s. All sampling techniques use the *Nyquist theorem*, which basically states that if we sample at twice the highest frequency on a voice line, we achieve good-quality voice transmission.

The PCM process is as follows:

- Analog waveforms are put through a voice frequency filter to filter out anything greater than 4000 Hz. These frequencies are filtered to 4000 Hz to limit the amount of crosstalk in the voice network. Using the Nyquist theorem, we need to sample at 8000 samples per second to achieve good-quality voice transmission.

- The filtered analog signal is then sampled at a rate of 8000 times per second.

- After the waveform is sampled, it is converted into a discrete digital form. This sample is represented by a code that indicates the amplitude of the waveform at the instant the sample was taken. The telephony form of PCM uses eight bits for the code and a logarithm compression method that assigns more bits to lower-amplitude signals. If we multiply the eight-bit words by 8000 times per second, we get 64,000 bits per second (bps). The basis for the telephone infrastructure is 64,000 bps (or 64 kbps).

Two basic variations of 64 kbps PCM are commonly used: μ-law, the standard used in North America; and a-law, the standard used in Europe. The methods are similar in that both use logarithmic compression to achieve from 12 to 13 bits of linear PCM quality in only eight-bit words, but they differ in relatively minor details. The μ-law method has a slight advantage over the a-law method in terms of low-level signal-to noise ratio performance.

1.2.2 Local Loops, Trunks, and Interswitch Communication

The *local loop* is the physical cabling between our homes to a central office switch. The communication path between the central office switch and our home is known as the p*hone line,* and it normally runs over the local loop. The communication path between several central office switches is known as a *trunk.* See figure 1-6.

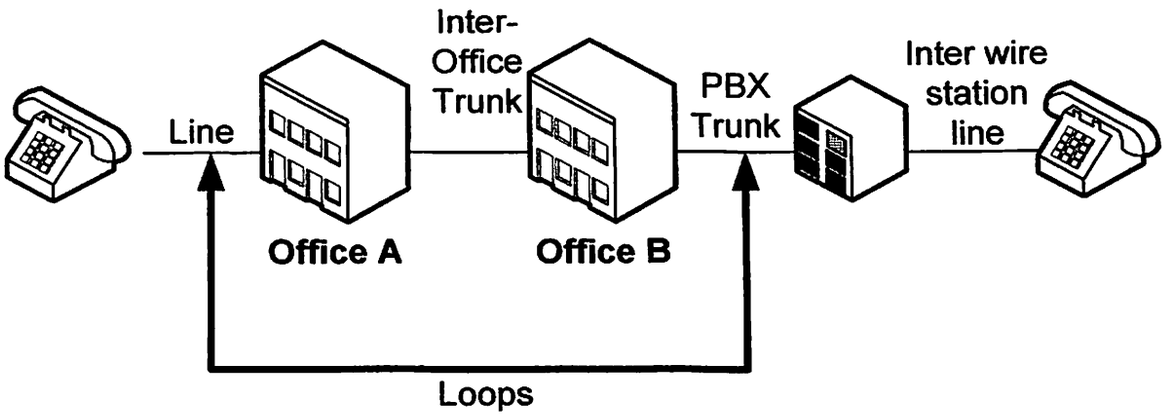


Figure 1-6 Loop, Line and Trunks

Switches establish communication in a hierarchy. End office switches interconnect through trunks tandem switches as shown in Figure 1-7

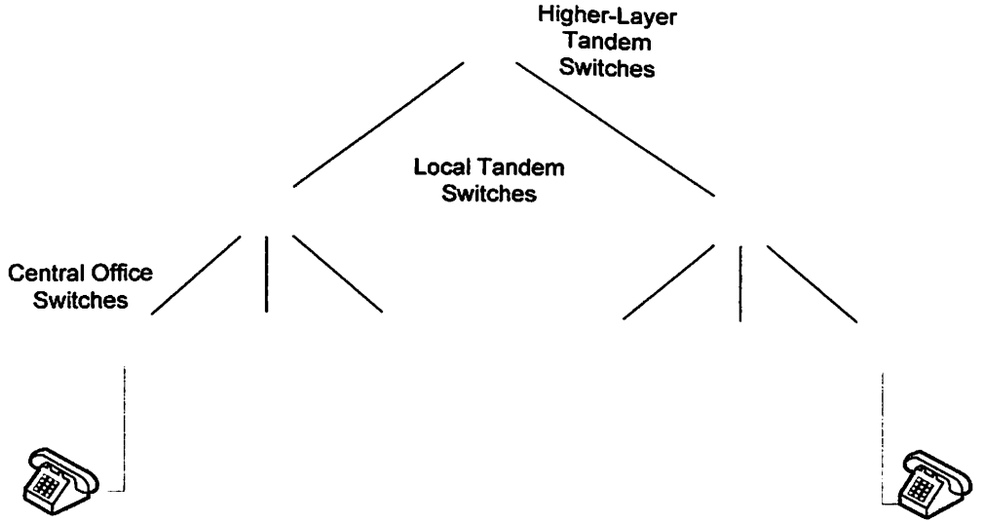


Figure 1-7 Circuit Switching Hierarchy

1.3 PSTN Signaling

There are two types of signaling that run over in a PSTN, these are: [3]

- *User-to-network signaling* —This is how an end user communicates with the PSTN.
- *Network-to-network signaling* —This is generally how the switches in the PSTN intercommunicate.

1.3.1 User-to-Network Signaling

Usually a user is connected to PSTN, using twisted pair cable through analog, Integrated Services Digital Network (ISDN), or through a T1 carrier.

The most common signaling method for user-to-network analog communication is *Dual Tone Multi-Frequency (DTMF)*. DTMF is known as in-band signaling because the tones are carried through the voice path.

ISDN uses another method of signaling known as *out-of-band*. With this method, the signaling is transported on a channel separate from the voice. The channel on which the voice is carried is called a *bearer* (or B channel) and is 64 kbps. The channel on which the signal is carried is called a data channel (D channel) and is 16 kbps.

Figure 1-8 shows a Basic Rate Interface (BRI) that consists of two B channels and one D channel.

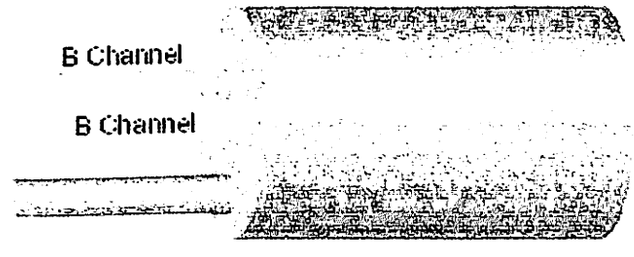


Figure 1-8 Basic Rate Interface

Out-of-band signaling offers many benefits, including the following:

- Signaling is multiplexed into a common channel.
- Glare is reduced (glare occurs when two people on the same circuit seize opposite ends of that circuit at the same time).
- A lower post dialing delay.
- Additional features, such as higher bandwidth, are realized.
- Because setup messages are not subject to the same line noise as DTMF tones, call completion is greatly increased.

1.3.2 Network-to-Network Signaling

Network-to-network communication is normally carried across the following transmission
Media:

- T1/E1 carrier over twisted pair

  T1 is a 1.544-Mbps digital transmission link normally used in North America and Japan.

  E1 is a 2.048-Mbps digital transmission link normally used in Europe.

- T3/E3, T4 carrier over coaxial cable

  T3 carries 28 T1s or 672 64-kbps connections and is 44.736 Mbps.

  E3 carries 16 E1s or 512 64-kbps connections and is 34.368 Mbps.

  T4 handles 168 T1 circuits or 4032 4-kbps connections and is 274.176 Mbps.

- T3, T4 carrier over a microwave link

- Synchronous Optical Network (SONET) across fiber media

  SONET is normally deployed in OC-3, OC-12, and OC-48 rates, which are 155.52 Mbps, 622.08 Mbps, and 2.488 Gbps, respectively.

Network-to-network signaling types include in-band signaling methods such as Multi-Frequency (MF) and Robbed Bit Signaling (RBS). These signaling types can also be used to network signaling methods.

MF is similar to DTMF, but it utilizes a different set of frequencies. As with DTMF, MF tones are sent in-band. But, instead of signaling from a home to an end office switch, MF signals from switch to switch.

Network-to-network signaling also uses an out-of-band signaling method known as *Signaling System 7* (SS7).

The benefits of SS7 are: [4]

- Reduced post-dialing delay

  There is no need to transmit DTMF tones on each hop of the PSTN. The SS7 network transmits all the digits in an initial setup message that includes the entire calling and called number.

- Increased call completion

  SS7 is a packet-based, out-of-band signaling protocol, compared to the DTMF or MF in-band signaling types. Single packets containing all the necessary information (phone numbers, services, and so on) are transmitted faster than tones generated one at a time across an in-band network.

- Connection to the IN

This connection provides new applications and services transparently across multiple vendors' switching equipment as well as the capability to create new services and applications more quickly.

1.4 Working of PSTN

Figure 1-9 further illustrates the working of a PSTN connection as follows. It shows how phone call is established when A is calling B.

1. A pick up the phone and send an off-hook indication to the end office switch.

2. The switch sends back a dial tone.

3. A dial the digits to call a house (they are sent in-band through DTMF).

4. The switch interprets the digits and sends an Initial Address Message (IAM) to the SS7 network.

5. The SS7 network reads the incoming IAM and sends a new IAM to B's switch.

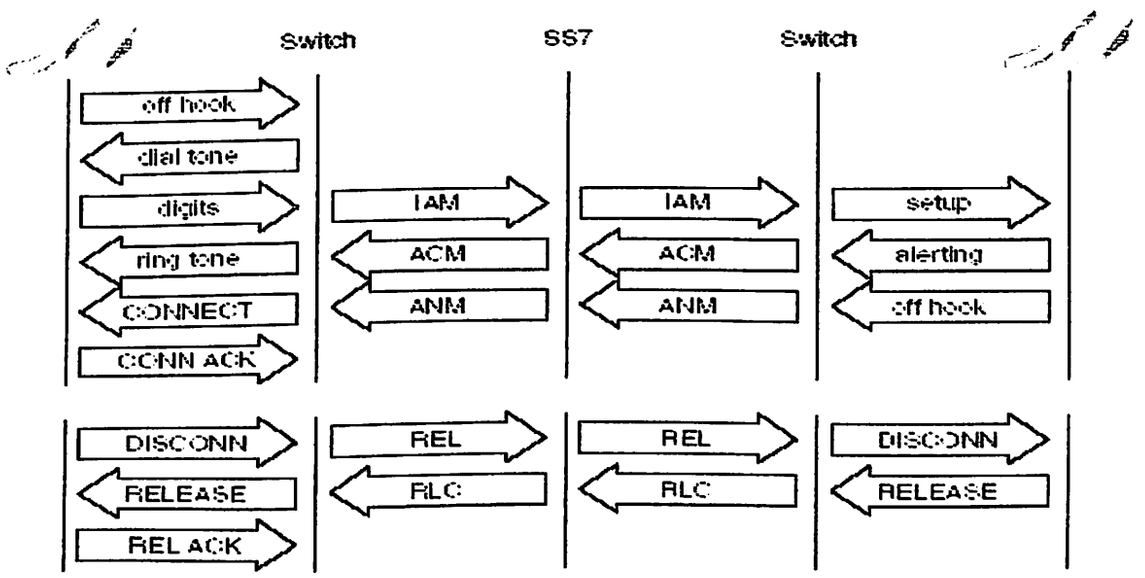


Figure 1-9 Working of PSTN

6. B's switch sends a setup message to B's phone (it rings his phone). An alerting message (alerting is the same as the phone ringing) is sent from B's switch (not from his phone) back to the SS7 network through an Address Complete Message (ACM).

7. The SS7 network reads the incoming ACM and generates an ACM to my switch.

8. A can hear a ringing sound and know that B's phone is ringing. (The ringing is not synchronized; our local switch normally generates the ringing when the ACM is received from the SS7 network.)

9. B picks up her phone, sending an off-hook indication to his switch.

10. B's switch sends an answer Message (ANM) that is read by the SS7, and a new ANM is generated to A's switch.

11. A connect message is sent to A's phone (only if it's an ISDN phone) and a connect acknowledgment is sent back (again, only if it's an ISDN phone). (If it is not an ISDN phone, then on-hook or off-hook representations signal the end office switch.)

12. A can now talk to B until A hang up the phone (on-hook indication).

If B's phone was busy, A could use an IN feature by which A could park on B's line And have the PSTN call A back after B got off the phone.

1.5 PSTN Numbering Plans

There are two numbering plans usually used; North America Numbering Plan (NANP) or Inter Telecommunication Union Telecommunication Standardization Sector (ITU-T). Both are discussed below: [5]

1.5.1 NANP

NANP is an 11-digit dialing plan that contains three parts: the Numbering Plan Area (NPA, also referring to as area code), Central Office Code (NXX), and Station Number. This plan is often referred to as ***NPA-NXX-XXXX.***
NPA codes use the following format: NXX, where N is a value between 2–9 and X is a value between 0–9.
NANP is also referred to as 1+10. This means that when a 1 is the first number dialed, it will be proceeded by a 10-digit NPA-NXX-XXXX number. This enables the end office switch to determine whether it should expect a 7- or 10-digit telephone number.

1.5.2 ITU-T International Numbering Plan

ITU-T Recommendation E.164 specifies that a Country Code (CC), National Destination Code (NDC), and Subscriber Number (SN) be used to route a call to a specific subscriber.

The CC consists of one, two, or three digits. The first digit (1–9) defines world numbering zones.

NDC and SN vary in length based on the needs of the country. Neither one has more than 15 digits.

Although dial plans might not seem extremely important at the moment, they are crucial to the successful deployment and implementation of Voice over IP (VoIP) or traditional circuit-switched networks.

1.6 Introduction to VoIP

1.6.1 Overview

In modern world the end user wants an efficient and foolproof way to get access to multimedia and data services through one access link. The IP network is one such link which users around the world have access to. Further more in existing telephone networks the efficiency of phone call is not more than 50%. To achieve bandwidth efficiency and

cost reduction, the traditional way of circuit switching is obsolete. The packet switched provides the solution. With this approach we can mix data traffic along voice traffic. Thus VoIP provides us with the best solution.

The main reason thus to why use VoIP is the penetration of IP in today's market. In North America and Europe most homes can be reached by a simple dial-up modem. The second main reason is the cost. The cost of construction and maintenance is much less, because the setup will acquire less space. Also, the system can be efficient by improving the network and reducing the maximum hopes. This can be done sitting in an office and configuring and administrating the network. While in traditional PSTN, mostly the efficiency reducing factors are related to physical attributes. Like cables, and physical connections etc.

According to Insight Research, the Voice over IP (VoIP) market is expected to grow from $13 billion in 2002 to $197 billion by 2007 [1]. Fueling this demand is the need for businesses to reduce telephony costs by leveraging existing data networks. As more businesses examine best practice methodologies of adopting and implementing VoIP solutions, one critical factor to evaluate is voice Quality of Service (QoS). Therefore, as network and systems managers evaluate VoIP deployment strategies, they should understand what factors would have the greatest impact to QoS.

1.7 What is VoIP?

Voice over IP (VoIP) can be set up on any existing data network that uses Internet Protocol (IP), such as internet or Local Area Networks (LAN). Voice is first converted from analog to digital form using DSP techniques then converted to packets, then transmitted over an IP network. One of the main motivations for VoIP is the very low cost involved.

UDP, which is datagram-based (or connectionless), best suits the specific requirements of voice traffic. UDP is preferred as the transport for VoIP, even though TCP, which is connection-oriented, is generally considered as a more ideal transport mechanism because of its built-in reliability.

So the voice packet to be sent over IP is split into two portions: (1) the control signaling and (2) the actual packetized voice traffic. The control signaling then runs on top of TCP, and the real-time voice is sent across UDP.

The figure 1-10 shows the VoIP packet.

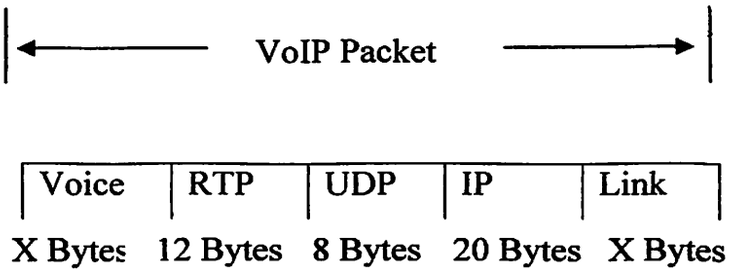


Figure 1.10 VoIP Packet

TCP offers reliability in that it guarantees retransmission of lost frames, so is used to carry the control signaling, but this reliable delivery is useless in the inter-network transportation of packetized voice because a frame that arrives late as a result of retransmission is as useful as no frame at al. In other words, retransmission of packets is not meaningful. By the time the resent packet arrives at the end user endpoint, the required delivery time has long been transgressed.

1.8 Voice Transmission over IP

In comparing the telephone network to a data network, we could say that the telephone network is a very large circuit-switched network. Comparing to any other Network PSTN exists everywhere, it is dependable and easy to use, and everyday life would be inconceivable without it.

Data networks, on the other hand, are privately owned and administered, so architecture, size, and complexity may vary. Data networks were likely originally designed for a different purpose other than to transport voice traffic, which is more sensitive to delay and can demand more bandwidth than data.

VoIP enables Cisco routers and switches to carry telephony-style voice traffic over IP-based data networks (intranetworks or internetworks) rather than the PSTN. To configure VoIP first we shall check the IP connectivity on the existing network.

Cisco routers and switches are equipped to handle the origination, transport, and termination of VoIP traffic. When transmitting voice signals across a data network, the router on the network performs some of the same basic functions as switches, which are as follows:

- The analog signal of the originating call is converted into a digital signal in the digital signal processor in the router.
- The voice signal is then segmented into frames, which are then coupled in groups of two and stored in voice packets that conform to the standard IP format.
- The IP packets are sent over the Internet in compliance with ITU-T specification H.323, explained in a later chapter.

At the receiving end, this process is reversed, with the packets being converted to digital signaling, then back to analog to be received and heard on the other end of the call.

1.9 Problem Definition

Since Voice over Internet Protocol (VoIP) has emerged, there has been a need for a better quality for a voice call. If anyone among you has ever used "Net to Phone" service, will better understand the poor quality of a voice call.

In this Project we have studied different Quality of Service (QoS) techniques in order to improve the quality of VoIP.

1.10 Conclusion

The purpose of this chapter is to understand the basic concept of PSTN, PSTN Signaling, Working of PSTN, PSTN Numbering Plans, VoIP and Voice Transmission over VoIP. We have given overview of main topics PSTN and VoIP in this chapter.

The first voice transmission, sent by Alexander Graham Bell, was accomplished in 1876 through what is called a *ring-down* circuit. A ring-down circuit means that there was no dialing of numbers; instead, a physical wire connected two devices. In early days voice was transmitted through the channel in analog form. But there were problems associated with it. Main of which was noise and transmitting the signal over longer cables. The problem of distance was solved by passing the signal through amplifiers to boost the signal, but it also boosted the signal. In digital networks, line noise is less of an issue because repeaters not only amplify the signal, but clean it to its original condition. PCM is the most common method of encoding an analog voice signal into a digital stream of 1s and 0s. There are two types of signaling that run over in a PSTN, these are:

- **User-to-network signaling** — This is how an end user communicates with the PSTN.

- **Network-to-network signaling** — This is generally how the switches in the PSTN intercommunicate.

In modern world the end user wants an efficient and foolproof way to get access to multimedia and data services through one access link. The IP network is one such link which users around the world have access to. Further more in existing telephone networks the efficiency of phone call is not more than 50%. To achieve bandwidth efficiency and cost reduction, the traditional way of circuit switching is obsolete. The packet switched provides the solution. With this approach we can mix data traffic along voice traffic. Thus VoIP provides us with the best solution.

Voice over IP (VoIP) can be set up on any existing data network that uses Internet Protocol (IP), such as internet or Local Area Networks (LAN). Voice is first converted

from analog to digital form using DSP techniques then converted to packets, then transmitted over an IP network. One of the main motivations for VoIP is the very low cost involved.

Chapter # 02

IP Telephony Protocols

2.1 Overview

This Chapter briefly describes following protocols and their interaction with Cisco call manager.

- IP Protocols
- Analog Telephony Protocols
- Digital Telephony Protocols

2.2 IP Protocols

Cisco call manager performs signaling and call control tasks such as digit analysis, routing, and circuit selection within the PSTN gateway infrastructure. To perform these functions, Cisco call manager uses industry standard IP protocols including H.323, MGCP, SCCP, and SIP. Use of Cisco call manager and these protocols gives service providers the capability to seamlessly route voice and data calls between the PSTN and packet networks.
Four IP protocols are discussing in this section:

- H.323 Protocol
- Media Gateway Control Protocol (MGCP)
- Skinny Client Control Protocol (SCCP)
- Session Initiation Protocol (SIP)

2.2.1 H.323 Protocol

H.323 is a protocol standard for multimedia communications. H.323 was designed to support real-time transfer of audio and video data over packet networks like IP. The standard involves several different protocols covering specific aspects of Internet telephony. The International Telecommunication Union (ITU-T) maintains H.323 and these related standards [6].

Most voice over IP (VoIP) applications utilize H.323. H.323 supports call setup, teardown and forwarding/transfer. Architectural elements of a H.323 based system are Terminals,

Multipoint Control Units (mcus), Gateways, an optional Gatekeeper and Border Elements. Different functions of H.323 run over either TCP or UDP. Overall, H.323 competes with the newer Session Initialization Protocol (SIP), another proven standard often found in VoIP systems.

A key feature of H.323 is Quality of Service (QoS). QoS technology allows real-time prioritization and traffic management constraints to be placed on "best-effort" packet delivery systems like TCP/IP over Ethernet. QoS improves the quality of voice or video feed.

2.2.2 Media Gateway Control Protocol (MGCP)

Media Gateway Control Protocol (MGCP) is a protocol used for controlling Voice over IP (VoIP) Gateways from external call control elements. MGCP is the emerging protocol that is receiving wide interest from both the voice and data industries. MGCP is a protocol for controlling media gateways from call agents. In a VoIP system, MGCP can be used with SIP or H.323. SIP or H.323 will provide the call control functionality and MGCP can be used to manage media establishment in media gateways.

Characteristics of MGCP:

- A master/slave protocol.
  - Assumes limited intelligence at the edge (endpoints) and intelligence at the core (call agent).
  - Used between call agents and media gateways
  - Differs from SIP and H.323 which are peer-to-peer protocols.
- Interoperates with SIP and H.323.

  For example
  - A call agent accepts SIP or H.323 call setup requests.
  - The call agent uses MGCP to control the media gateway.
  - The media gateway establishes media sessions with other H.323 or SIP endpoints.
- MGCP divides call setup/control and media establishment functions.

- MGCP does not replace SIP or H.323. SIP and H.323 provide symmetrical or peer-to-peer call setup/control [7].

2.2.3 Skinny Client Control Protocol (SCCP)

Skinny Client Control Protocol (SCCP) is a Cisco proprietary standard for terminal control for use with voice over IP (VoIP). The SCCP was originally developed by Celsius Corporation. The term "skinny" reflects that SCCP is a simple and uncomplicated ("lightweight") protocol requiring relatively little computer processing.

With Cisco's version of Skinny Client Control Protocol, the end stations in a network, which can be VoIP phone sets or personal computers with VoIP capability, run a program called the Skinny Client. The lightweight Skinny Client helps to minimize the cost and complexity of VoIP end stations. The H.323 call setup processing is done by a proxy known as the Call Manager. The audio communication between end stations makes use of the User Datagram Protocol (UDP) and the Internet Protocol (IP). Variants of SCCP are used by several companies other than Cisco [8].

2.2.4 Session Initiation Protocol (SIP)

The Session Initiation Protocol (SIP) is an Internet Engineering Task Force (IETF) standard protocol for initiating an interactive user session that involves multimedia elements such as video, voice, chat, gaming, and virtual reality.

Like HTTP or SMTP, SIP works in the Application layer of the Open Systems Interconnection (OSI) communications model. The Application layer is the level responsible for ensuring that communication is possible. SIP can establish multimedia sessions or Internet telephony calls, and modifies, or terminates them. The protocol can also invite participants to unicast or multicast sessions that do not necessarily involve the initiator. Because the SIP supports name mapping and redirection services, it makes it possible for users to initiate and receive communications and services from any location, and for networks to identify the users wherever they are.

SIP is a request-response protocol, dealing with requests from clients and responses from servers. Participants are identified by *SIP urls*. Requests can be sent through any transport

protocol, such as UDP, SCTP, or TCP. SIP determines the end system to be used for the session, the communication media and media parameters, and the called party's desire to engage in the communication. Once these are assured, SIP establishes call parameters at either end of the communication, and handles call transfer and termination [9].

2.3 Analog Telephony Protocols

Analog telephony signaling, the original signaling protocol, provides the method for connecting or disconnecting calls on analog trunks. By using direct current (DC) over two-wire or four-wire circuits to signal on-hook and off-hook conditions, each analog trunk connects analog endpoints or devices such as a PBX or analog phone.
To provide connections to legacy analog central offices and PBXs, Cisco call manager uses analog signaling protocols over analog trunks that connect voice gateways to analog endpoints and devices. Cisco call manager supports these types of analog trunk interfaces:

• Foreign Exchange Office (FXO)—Analog trunks that connect a gateway to a central office (CO) or private branch exchange (PBX).

• Foreign Exchange Station (FXS)—Analog trunks that connect a gateway to plain old telephone service (POTS) device such as analog phones, fax machines, and legacy voice-mail systems.

We can configure loop-start, ground-start, or E&M signaling protocols for FXO and FXS trunk interfaces depending on the gateway model that is selected. We must use the same type of signaling on both ends of the trunk interface to ensure that the calls properly connect. The following sections describe the types of analog signaling protocols that Cisco call manager supports:

2.3.1 Loop-Start Signaling

Loop-start signaling sends an off-hook signal that starts a call and an on-hook signal that closes the loop and ends the call. Loop-start trunks lack positive disconnect supervision, and users can experience glare when two calls seize the trunk at the same time.

2.3.2 Ground-Start Signaling

Ground-start signaling provides current detection mechanisms at both ends of the trunk to detect off-hook signals. This mechanism permits endpoints to agree on which end is seizing the trunk before it is seized, and minimizes the chance of glare. Ground start provides positive recognition of connects and disconnects and is the preferred signaling method for PBX connections. Some PBXs do not support ground-start signaling, so then we have to use loop-start signaling for the trunk interface.

2.3.3 E&M Signaling

E&M signaling uses direct current (DC) over two-wire or four-wire circuits to signal to the endpoint or CO switch when a call is in receive or transmit (E&M) state. E&M signaling uses signals that indicate off-hook and on-hook conditions. When the connection is established, the audio transmission occurs. The E&M signaling type must be the same for both ends of the trunk interface for successful connections. Cisco call manager supports following types of E&M signaling:

Wink-start signaling—The originating side sends an off-hook signal and waits to receive a wink pulse signal that indicates the receiving end is ready to receive the dialed digits for the call. Wink start is the preferred signaling method because it provides answer supervision. Not all PBXs support wink-start signaling.

Delay-dial signaling—The originating side sends an off-hook signal, waits for a configurable time period, and then checks if the receiving end is on hook. The originating side sends dialed digits when the receiving side is on hook. The delay allows the receiving end to signal when it is ready to receive the call.

Immediate-start signaling—The originating side goes off hook, waits for a finite time period, and then sends the dial digits without a ready signal from the receiving end.

2.3.4 Channel Associated Signaling (CAS)

Channel associated signaling (CAS) sends the on hook and off hook signals as bits within the frames on the same channel as the audio transmission. CAS is often referred to as robbed bit signaling because CAS takes bits from the voice channel for signaling. These signals can include supervision, addressing, and tones such as busy tone or dial tone.
We can use T1 CAS and E1 CAS digital trunk interfaces to connect a Cisco call manager call to a CO, a PBX, or other analog device.

T1 CAS

The T1 CAS trunk interface uses in-band E&M signaling to carry up to 24 connections on a link. Both ends of the T1 link must specify T1 CAS signaling. Cisco call manager provides the T1 CAS signaling option when we configure ports on some MGCP and H.323 voice gateways and network modules.

E1 CAS

Some Cisco gateways in H.323 mode can support the E1 CAS trunk interface that provides up to 32 connections on the link. We must configure the E1 CAS signaling interface on the gateway, not in Cisco call manager Administration. Both ends of the E1 link must specify E1 CAS signaling [10].

2.4 Digital Telephony Protocols

Digital telephony protocols use common channel signaling (CCS) , a dedicated channel that carries only signals. In a T1 link, one channel carries the signals while the other channels carry voice or data. The latest generation of CCS is known as Signaling System

045458

7 (SS7) and provides supervision, addressing, tones, and a variety of services such as automatic number identification (ANI).
Integrated Services Digital Network (ISDN) is a set of international standards for user access to private or public network services. ISDN provides both circuit-based and packet-based communications to users.

Cisco call manager can support these ISDN protocols:

- Basic Rate Interface (BRI)
- T1 Primary Rate Interface (T1 PRI)
- E1 Primary Rate Interface (E1 PRI)
- Q.Signaling (QSIG)

2.4.1 Basic Rate Interface (BRI)

Basic rate interface (BRI), which is used for small office and home communications links, provides two B-channels for voice and data and one D-channel for signaling.

2.4.2 T1 Primary Rate Interface (T1 PRI)

T1 Primary rate interface (PRI) is used for corporate communications links in North America and Japan. T1 PRI provides 23 B-channels for voice and data and one D-channel for common channel signaling. T1 PRI uses a communication rate of 1.544Mbps.

2.4.3 E1 Primary Rate Interface (E1 PRI)

E1 PRI Primary rate interface (PRI) is used for corporate communications in Europe. E1 PRI provides 30 B-channels for voice and data, one D-channel for common signaling, and one framing channel. E1 PRI uses a rate of 2.048 Mbps.

2.4.4 Q.Signaling (QSIG)

Because enterprises maintain existing telecommunication equipment from a variety of vendors, the protocol system, Q signaling (QSIG), provides interoperability and feature transparency between a variety of telecommunications equipment.

The QSIG protocol, a series of international standards, defines services and signaling protocols for Private Integrated Services Networks (PISNS). These standards use Integrated Services Digital Networks (ISDN) concepts and conform to the framework of International Standards for Open Systems Interconnection as defined by ISO/IEC. The QSIG protocol acts as a variant of ISDN D-channel voice signaling. The ISDN Q.921 and Q.931 standards provide the basis for QSIG protocol, which sets worldwide standard for PBX interconnection.

The QSIG protocol enables Cisco voice switching services to connect to PBXs and key systems that communicate by using QSIG protocol. For QSIG basic call setup, Cisco devices can route incoming voice calls from a private integrated services network exchange (PINX) device across a WAN to a peer Cisco device that can transport the signaling and voice packets to another PINX device, which are PBXs, key systems, or Cisco call manager servers that support QSIG protocol.

In a basic QSIG call, a user in a PINX can place a call to a user that is in a remote PINX. The called party receives the caller name or number as the call rings. The calling party receives the called name and number when the user phone rings in the remote PINX. All the features that are available as a PBX user operate transparently across the network. QSIG protocol provides supplementary and additional network features, as defined for PISNS, if the corresponding sets of QSIG features are supported by both ends of the call. To make supplementary features available to network users, ensure that all PBXs in the network support the same feature set.

Cisco call manager supports the following QSIG features:

- Annex M.1 (Message Tunneling for QSIG)
- Basic Call for QSIG
- Call Completion
- Call Diversion
- Call Transfer
- Facility Selection and Reservation
- Identification Services
- Path Replacement

2.4.4.1 Annex M.1 (Message Tunneling for QSIG)

The Annex M.1 feature uses intercluster trunks to transport (tunnel) non-H.323 protocol information in H.323 signaling messages between Cisco call managers. Annex M.1 supports QSIG calls and QSIG call independent signaling connections. After completion of intercluster trunk configuration in Cisco call manager Administration, QSIG tunneling supports the following features: Call Completion, Call Diversion, Call Transfer, Identification Services, Message Waiting Indication, and Path Replacement.
If any gateway in the network does not support QSIG tunneling, calls drop at the intercluster trunk that is configured for QSIG tunneling

2.4.4.2 Basic Call for QSIG

QSIG basic call setup provides the dynamic establishment of voice connections from an originating PINX (PBX or Cisco call manager) across a private network or virtual private network (VPN) to another PINX. Note: We must use digital T1 or E1 primary rate interface (PRI) trunks to support QSIG protocol.

2.4.4.3 Call Completion

The following Call Completion services provide Cisco Call Back functionality over QSIG enabled trunks:

• Completion of Calls to Busy Subscribers (CCBS)—When a calling party receives a busy tone, the caller can request that the call complete when the busy destination hangs up the phone and becomes available.

• Completion of Calls on No Reply (CCNR)—When a calling party receives no answer at the destination, the calling party can request that the call complete after the activity occurs on the phone of the called party. Cisco Call Manager and the Call Completion services use the Call Back soft key on supported Cisco IP Phone models 7940, 7960, and 7970 in a Cisco Call Manager cluster or over QSIG trunks. Likewise, the following devices support QSIG Call Completion services:

• Cisco IP Phone Models 7905, 7910, 7912, 7940, 7960, 7970

• Cisco VGC Phone, Cisco IP Communicator, and Cisco SCCP Phone

• CTI route point that forwards calls to supported devices

2.4.4.4 Call Diversion

Cisco call manager supports call diversion by rerouting and call diversion by forward switching. When call diversion by rerouting occurs, the originating PINX receives a request from the receiver of the call to divert the call to another user. The system creates a new call between the originator and the diverted-to user. If the receiver of the incoming call and the diverted-to user exist in the same PINX, Cisco call manager uses call diversion by forward switching

2.4.4.5 Call Transfer

Cisco call manager supports call transfer by join only. When a user transfers a call to another user, the QSIG identification service changes the connected name and number that displays on the transferred party phone. Call identification restrictions can impact what displays on the phone. The call transfer supplementary service interacts with the path replacement feature to optimize the trunk connections when a call transfers to a caller in another PINX.

2.4.4.6 Facility Selection and Reservation

The facility selection and reservation feature allows us to make calls by using mixed route lists, which contain route groups that use different protocols. This feature supports mixed route lists that include the following types of facilities:

- E1 or T1 PRI trunks that use the QSIG protocol
- E1 or T1 PRI trunks that use a protocol other than QSIG
- T1-CAS gateways
- FXO ports

2.4.4.7 Identification Services

When a call alerts and connects to a PINX, identification services can display the caller name/ID on a phone in the terminating PINX, and, likewise, the connected party name/ID on a phone in the originating PINX. QSIG identification restrictions allow us to control the presentation or display of this information between Cisco call manager and the connected PINX.

2.4.4.8 Path Replacement

In a QSIG network, after a call is transferred or forwarded to a phone in a third PINX, multiple connections through several PINX(s) can exist for the call. After the call connects, the path replacement feature drops the connection to the transit PINX(s) and creates a new call connection to the terminating PINX.

2.5 Conclusion

The purpose of this chapter is to understand the working and benefits of IP telephony protocols. We covered 3 types of protocols which are used to complete the call of VoIP.

First category is IP protocols, the protocols which are used to gives service providers the capability to seamlessly route voice and data calls between the PSTN and packet networks.

Second category is analog telephony protocols which are used in PSTN networks. Simply analog protocols provide signaling mechanism provides the method for connecting or disconnecting calls on analog trunks.

Third category is digital telephony protocols which are used to provide special kind of services like ISDN, DSL. They all are used common channel signaling. Simply CCS provide separate channel for voice or data and separate for signaling.

Chapter # 03

The H.323 Protocol

3.1 Overview

The H.323 standard provides a foundation for audio, video, and data communications across IP-based networks, including the Internet. The H.323 standard is specified by the ITU–T.Version 1 of the H.323 standard does not provide guaranteed QoS.

The emergence of voice-over–IP (VoIP) applications and IP telephony has paved the way for a revision of the H.323 specification. The absence of a standard for voice over IP resulted in products that were incompatible. With the development of VoIP, new requirements emerged, such as providing communication between a PC–based phone and a phone on a traditional switched circuit network. Such requirements forced the need for a standard for IP telephony. Version 2 of H.323—packet-based a multimedia communications system—was defined to accommodate these additional requirements and was accepted in January 1998.

New features are being added to the H.323 standard, which will evolve to Version 3 shortly. The features being added include fax-over-packet networks, gatekeeper-gatekeeper communications, and fast-connection mechanisms. The H.323 standard is designed to allow clients on H.323 networks to communicate with clients on the other videoconferencing networks [11].

3.2 H.323 Architecture-Components:

The H.323 standard specifies four kinds of components, which, when networked together, provide the point-to-point and point-to-multipoint multimedia-communication services:[12]

1. Terminals
2. Gateways
3. Multipoint control units (mcus)
4. Gatekeepers

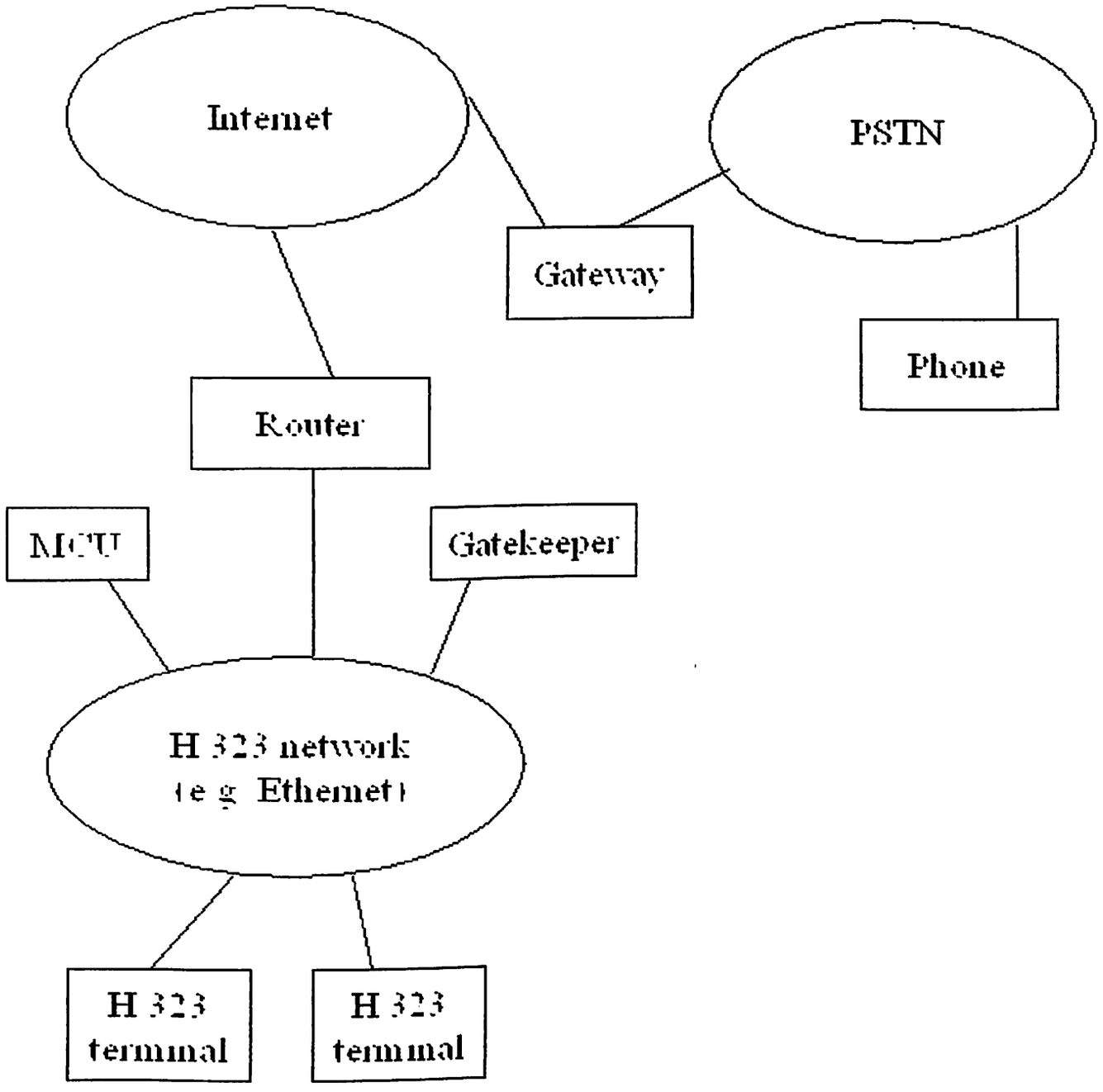


Figure 3 -1 Layout of H.323

3.2.1 Terminals

Used for real-time bi-directional multimedia communications, an H.323 terminal can either be a personal computer (PC) or a stand-alone device, running an H.323 and the multimedia applications. It supports audio communications and can optionally support video or data communications. Because the basic service provided by an H.323 terminal is audio communications, an H.323 terminal plays a key role in IP–telephony services.

The primary goal of H.323 is to work with other multimedia terminals. H.323 terminals are compatible with H.324 terminals on SCN and wireless networks.

H.323 terminals must support the following:

- H.245 for exchanging terminal capabilities and creation of media channels
- H.225 for call signaling and call setup
- RAS for registration and other admission control with a gatekeeper
- RTP/RTCP for sequencing audio and video packets
- H.323 terminals must also support the G.711 audio CODEC. Optional components in an H.323 terminal are video Codec's, T.120 data-conferencing protocols, and MCU capabilities.

3.2.2 Gateway

A gateway provides translation of protocols for call setup and release, conversion of media formats between different networks, and the transfer of information between H.323 and non–H.323 networks. An application of the H.323 gateway is in IP telephony, where the H.323 gateway connects an IP network and SCN network.

On the H.323 side, a gateway runs H.245 control signaling for exchanging capabilities, H.225 call signaling for call setup and release, and H.225 RAS for registration with the gatekeeper. On the SCN side, a gateway runs SCN–specific protocols (e.g., ISDN and SS7 protocols).

Terminals communicate with gateways using the H.245 control-signaling protocol and H.225 call-signaling protocol. The gateway translates these protocols in a transparent fashion to the respective counterparts on the non–H.323 network and vice versa. The gateway also performs call setup and clearing on both the H.323–network side and the non–H.323–network side. Translation between audio, video, and data formats may also be performed by the gateway. Audio and video translation may not be required if both terminal types find a common communications mode. The gateway has the characteristics

of both an H.323 terminal on the H.323 network and the other terminal on the non–H.323 network it connects.

Gatekeepers are aware of which endpoints are gateways because this is indicated when the terminals and gateways register with the gatekeeper. A gateway may be able to support several simultaneous calls between the H.323 and non–H.323 networks. In addition, a gateway may connect an H.323 network to a non–H.323 network. A gateway is a logical component of H.323 and can be implemented as part of a gatekeeper or an MCU.

3.2.3 Multipoint control unit

A multipoint control unit enables conferencing between three or more endpoints. It consists of a mandatory multipoint controller (MC) and zero or more multipoint processors (MP). Although the MCU is a separate logical unit it may be combined into a terminal, gateway, or gatekeeper. The MCU is an optional component of an H.323-enabled network. A MC may also be used in a point-to-point call, which can later be extended into a multipoint conference. Another useful job of the MC is to determine whether to unicast or multicast the audio and video streams depending on the capability of the underlying network and the topology of the multipoint conference. The multipoint processor handles the mixing, switching, and processing of the audio, video, and data streams among the conference endpoints. The MCU is required in a centralized multipoint conference where each terminal establishes a point-to-point connection with the MCU. The MCU determines the capabilities of each terminal and sends each a mixed media stream. In the decentralized model of multipoint conferencing, a MC ensures communication compatibility but the media streams are multicast and the mixing is performed at each terminal.

3.2.4 Gatekeeper

A gatekeeper [13] is an H.323 entity on the network that provides services such as address translation and network access control for H.323 terminals, gateways, and mcus. Also, they can provide other services such as bandwidth management, accounting, and dial plans that you can centralize in order to provide salability.

Gatekeepers are logically separated from H.323 endpoints such as terminals and gateways. They are optional in an H.323 network. But if a gatekeeper is present, endpoints must use the services provided.

3.2.4.1 Gatekeeper Zones and Subnets

A zone is the collection of H.323 nodes such as gateways, terminals, and mcus registered with the gatekeeper. There can only be one active gatekeeper per zone. These zones can overlay subnets and one gatekeeper can manage gateways in one or more of these subnets.

3.2.4.2 Gatekeeper Functionality

The H.323 standard defines mandatory and optional gatekeeper functions:

a) Mandatory Gatekeeper Functions

- **Address Translation**—Translates H.323 ids (such as gwy1@domain.com) and E.164 numbers (standard telephone numbers) to endpoint IP addresses.
- **Admission Control**— Provides authorized access to H.323 using the Admission Request/Admission Confirm/Admission Reject (ARQ/ACF/ARJ) messages.
- **Bandwidth Control**— consists of managing endpoint bandwidth requirements using Bandwidth Request/Bandwidth Confirm/Bandwidth Reject (BRQ/BCF/BRJ) messages.
- **Zone Management**—The gatekeeper provides zone management for all registered endpoints in the zone. For example, controlling the endpoint registration process.

b) Optional Gatekeeper Functions

- **Call Authorization**—With this option, the gatekeeper can restrict access to certain terminals or gateways and/or have time-of-day policies restrict access.
- **Call Management**—With this option, the gatekeeper maintains active call information and uses it to indicate busy endpoints or redirect calls.

- **Bandwidth Management**—With this option, the gatekeeper can reject admission when the required bandwidth is not available.
- **Call Control Signaling**—With this option, the gatekeeper can route call-signaling messages between H.323 endpoints with the use of the Gatekeeper-Routed Call Signaling (GKRCS) model. Alternatively, it allows endpoints to send H.225 call-signaling messages directly to each other.

3.3 H.323 Architecture Protocols

The protocols specified by H.323 are listed below. H.323 is independent of the packet network and the transport protocols over which it runs and does not specify them. [14]

- Audio codec
- Video codec
- H.225 registration, admission, and status (RAS)
- H.225 call signaling
- H.245 control signaling
- T.120 for multimedia conferencing
- H.450 Series for supplementary services
- H.235 for Security and encryption for H-Series
- Real-time transfer protocol (RTP)
- Real-time control protocol (RTCP)

3.3.1 Audio Codec:

An audio CODEC encodes the audio signal from the microphone for transmission on the transmitting H.323 terminal and decodes the received audio code that is sent to the speaker on the receiving H.323 terminal. Because audio is the minimum service provided by the H.323 standard, all H.323 terminals must have at least one audio CODEC support, as specified in the ITU–T G.711 recommendation (audio coding at 64 kbps). Additional audio CODEC recommendations such as G.722 (64, 56, and 48 kbps), G.723.1 (5.3 and 6.3 kbps), G.728 (16 kbps), and G.729 (8 kbps) may also be supported.

3.3.2 Video Codec:

Video communication is bandwidth intensive in nature. Therefore, efficient compression and decompression techniques are essential for good performance. Recommendation H.323 specifies two video codecs: H.261 and H.263. However, H.323 clients are not limited to these codecs only. Other codecs can be used provided both terminals agree on and support it. Video support in H.323 terminals and mcus is optional. The H.261 codec produces video transmission for channels with bandwidths p x 64 kb/s where p can range from 1 to 30. The discrete cosine transform (DCT) is used for compression together with quantization and motion compensation. H.261 supports two video formats. The common intermediate format (CIF) has a resolution of 352 x 288 pixels while the quarter common intermediate format (QCIF) has a resolution of 176 x 144 pixels. The CIF format support is optional. The H.263 codec is designed for low bit rate transmission without loss of quality. It uses the same DCT coding with quantization for compression but this is accompanied by both motion estimation and prediction. The result is better quality at a lower bit rate. The quality of video transmission strongly depends on compression techniques. Active work is on-going in the development of more efficient codecs like MPEG-4 and MPEG-7. The architecture of H.323 is designed to allow the incorporation of new codec's as they become available.

3.3.3 H.225 RAS:

The H.225 RAS is used between H.323 endpoints (terminals and gateways) and gatekeepers for the following:

- Gatekeeper discovery (GRQ)
- Endpoint registration
- Endpoint location
- Admission control
- Access tokens

The RAS messages are carried on a RAS channel that is unreliable. Hence, RAS message exchange may be associated with timeouts and retry counts.

• Gatekeeper discovery: The gatekeeper discovery process is used by the H.323 endpoints to determine the gatekeeper with which the endpoint must register. The gatekeeper discovery can be done statically or dynamically. In static discovery, the endpoint knows the transport address of its gatekeeper a priori. In the dynamic method of gatekeeper discovery, the endpoint multicasts a GRQ message on the gatekeeper's discovery multicast address. One or more gatekeepers may respond with a GCF message.

• Endpoint Registration: Registration is a process used by the endpoints to join a zone and inform the gatekeeper of the zone's transport and alias addresses. All endpoints register with a gatekeeper as part of their configuration.

• Endpoint Location: Endpoint location is a process by which the transport address of an endpoint is determined and given its alias name or E.164 address (telephone number).

• Other Control: The RAS channel is used for other kinds of control mechanisms, such as admission control, to restrict the entry of an endpoint into a zone, bandwidth control, and disengagement control, where an endpoint is disassociated from a gatekeeper and its zone.

3.3.4 H.225 Calla Signaling:

H.225 call signaling is used to set up connections between H.323 endpoints (terminals and gateways), over which the real-time data can be transported. Call signaling involves the exchange of H.225 protocol messages over a reliable call-signaling channel. For example, H.225 protocol messages are carried over TCP in an IP–based H.323 network. H.225 messages are exchanged between the endpoints if there is no gatekeeper in the H.323 network. When a gatekeeper exists in the network, the H.225 messages are exchanged either directly between the endpoints or between the endpoints after being routed through the gatekeeper. The first case is direct call signaling. The second case is called gatekeeper-routed call signaling. The method chosen is decided by the gatekeeper during RAS–admission message exchange.

• Gatekeeper-Routed Call Signaling: The admission messages are exchanged between endpoints and the gatekeeper on RAS channels. The gatekeeper receives the call-

signaling messages on the call-signaling channel from one endpoint and routes them to the other endpoint on the call-signaling channel of the other endpoint.

• *Direct Call Signaling:* During the admission confirmation, the gatekeeper indicates that the endpoints can exchange call-signaling messages directly. The endpoints exchange the call signaling on the call-signaling channel.

3.3.5 H.245 Media Control Signaling:

The flexibility of H.323 requires that endpoints negotiate to determine compatible settings before audio, video, and/or data communication links can be established. H.245 uses control messages and commands that are exchanged during the call to inform and instruct. The implementation of H.245 control is mandatory in all endpoints. H.245 provides the following media control functionalities:

• Capability exchange: H.323 allows endpoints to have different receive and send capabilities. Each endpoint records its receiving and sending capabilities (e.g. Media types, codecs, bit rates, etc) in a message and sends it to the other endpoint(s).

• Opening and closing of logical channels: H.323 audio and video logical channels are uni-directional end-to-end links (or multipoint links in the case of multipoint conferencing). Data channels are bi-directional. A separate channel is needed for audio, video, and data communication. H.245 messages control the opening and closing of such channels. H.245 control messages use logical channel 0 which is always open.

• Flow control messages: These messages provide feedback to the endpoints when communication problems are encountered.

• Other commands and messages: Several other commands and messages may be used during a call like a command to set the codec at the receiving endpoint when the sending endpoint switches its codec.

H.245 control messages may also be routed through a gatekeeper if one exists.

3.3.6 T.120 Data Conferencing:

Real-time data conferencing capability is required for activities such as application sharing, whiteboard sharing, file transfer, fax transmission, and instant messaging. Recommendation T.120 provides this optional capability to H.323. T.120 is a real-time data communication protocol designed specifically for conferencing needs. Like H.323, Recommendation T.120 is an umbrella for a set of standards that enable the real-time sharing of specific applications data among several clients across different networks. T.120 provides several advantages over regular data transmission such as:

- Multipoint conferencing support: T.120 supports multipoint data delivery, which enables group collaboration activities. The MCU handles the mixing and switching of data in a similar manner to that used for video and audio.

- Network and platform independence: T.120 operates on top of the transport layer of the underlying network. As such, it is transparent and independent of the network hardware and software.

- Interoperability: T.120 is referenced in all the H.32X conferencing standards. This cross-referencing, together with the network and platform independence, ensures a high degree of interoperability at the application level.

- Multicast support: T.120 supports multicast of data streams in multicast-capable networks. This support is flexible with mixed unicast and multicast also possible during a conference.

- Other benefits: T.120 provides error correction capability on top of the network transport ensuring reliable delivery. In general, T.120 has a scalable and extendible architecture with provisions for the addition of new applications that take advantage of real-time reliable and efficient data delivery among a group of collaborators.

3.4 Features of H.323:

- Inter-network interoperability: H.323 clients are interoperable with switched circuit network (SCN) conferencing clients such as those based on Recommendations H.320 (ISDN), H.321 (ATM), and H.324 (PSTN/Wireless).

- Heterogeneous client capabilities: A H.323 client must support audio communication; video and data support is optional. This heterogeneity and flexibility does not make the clients incompatible. During call set-up capabilities are exchanged and communication established based on the lowest common denominator.

- Audio and video codecs: H.323 specifies a required audio and video codec. However, there is no restriction on the use of other codecs and two clients can agree on any codec.

- Management and accounting support: H.323 calls can be restricted on a network based on the number of calls already in progress, bandwidth limitations, or time restrictions. Using these policies the network manager can manage H.323 traffic. Further, H.323 also provides accounting facilities that can be used for billing purposes.

- Security: H.323 provides authentication, integrity, privacy, and non-repudiation support.

- Supplementary services: Recommendation H.323 recognizes the huge potential for applications based on IP telephony and multimedia. It provides a basic framework for development of such services. In version 2.0 of H.323, two services -- call transfer and call forwarding -- have been specified.

- Platform and Application Independence: H.323 is not tied to any hardware or operating system. H.323-compliant platforms will be available in many sizes and

shapes, including video-enabled personal computers, dedicated platforms, IP-enabled telephone handsets, cable TV set-top boxes and turnkey boxes.

- *Multipoint Support: Although H.323 can support conferences of three or more* endpoints without requiring a specialized multipoint control unit; *MCUs provide a* more powerful and flexible architecture for hosting multipoint conferences. Multipoint capabilities can be included in other components of an H.323 system.

- Multicast Support: H.323 supports multicast transport in multipoint conferences. Multicast sends a single packet to a subset of destinations on the network without replication. In contrast, unicast sends multiple point-to-point transmissions, while broadcast sends to all destinations. In unicast or broadcast, the network is used inefficiently as packets are replicated throughout the network. Multicast transmission uses bandwidth more efficiently since all stations in the multicast group read a single data stream.

3.5 Scope of H.323

H.323 is a broad and flexible recommendation. As a minimum, H.323 specifies protocols for real-time point-to-point audio communication between two terminals on a packet-based network that do not provide a guaranteed quality of service. The scope of H.323, however, is much broader and encompasses inter-network multipoint conferencing among terminals that support not only audio but also video and data communication.

The scope of Recommendation H.323 can be summarized in the following broad categories:

- **Point-to-point and multipoint conferencing support**: H.323 conferences may be set up between two or more clients without any specialized multipoint control software or hardware. However, when a multipoint control unit (MCU) is used H.323 supports a flexible topology for multipoint conferences. A multipoint conference may be centralized where new participants can join all the others in the

conference. This is the so-called hub-and-spoke topology. Or, a multipoint conference may be decentralized where new participants can elect to join one or more participants in the conference but not all. This approach will produce a flexible tree topology.

- **Inter-network interoperability**: H.323 clients are interoperable with switched-circuit network (SCN) conferencing clients such as those based on Recommendations H.320 (ISDN), H.321 (ATM), and H.324 (PSTN/Wireless).
- **Heterogeneous client capabilities**: A H.323 client must support audio communication; video and data support is optional. This heterogeneity and flexibility does not make the clients incompatible. During call set-up capabilities are exchanged and communication established based on the lowest common denominator.
- **Audio and video codecs**: H.323 specifies a required audio and video codec. However, there is no restriction on the use of other codecs and two clients can agree on any codec which is supported by both of them.
- **Management and accounting support**: H.323 calls can be restricted on a network based on the number of calls already in progress, bandwidth limitations, or time restrictions. Using these policies the network manager can manage H.323 traffic. Further, H.323 also provides accounting facilities that can be used for billing purposes.
- **Security**: H.323 provides authentication, integrity, privacy, and non-repudiation support.
- **Supplementary services**: Recommendation H.323 recognizes the huge potential for applications based on IP telephony and multimedia. It provides a basic framework for development of such services. In version 2.0 of H.323, two services -- call transfer and call forwarding -- have been specified.

3.6 Why is H.323 Important?

Trend: The explosive growth of the Internet and the almost universal deployment of corporate LANS have made packet-based networks ubiquitous. It is therefore natural for

individuals and enterprises to use this resource for audio and video communication to offset some of the tariffs of public switched telephone networks (PSTN). Multimedia over packet-based networks (primarily IP networks) has grown rapidly in the last few years. Industry research put the growth at 37% annually and is expected to reach $39 billion by the year 2002 (In a similar forecast, Probe Research estimates that by the year 2002 18.5% of all U.S. phone traffic will be carried over data networks. This rapid expansion and potential underlies the importance of an enabling and unifying standard such as H.323.

Standardization: In the 1995-1997 time period, several vendors developed products and services to cater for the emerging IP telephony market. These products and services, however, were based on proprietary protocols that prevented widespread interoperability. H.323 is a standard protocol that has been widely accepted. This will promote greater awareness, availability, and acceptability of multimedia conferencing over packet-based networks.

Internetworking: H.323 bridges multimedia communications between packet-based and switched-circuit networks SCN). Existing clients based on SCN conferencing standards like H.320 (ISDN), H.321 (ATM), and H.324 (PSTN) can inter-operate with H.323 clients. For example, it is possible to call from a H.323 client to a regular telephone on a PSTN. At the corporate level this internetworking capability allows enterprises to migrate voice and video from existing networks to their data network.

Integrated services: H.323 makes possible the development of additional services such as e-mail, voice mail, fax, call center functionality, and videoconferencing in an integrated environment. For example, an e-commerce business can provide a direct voice link from their web site to a sales representative to answer customers' questions. A few services have been standardized in H.450.x (e.g. Call transfer, call forwarding).

3.7 Conclusion

In this chapter we have discussed H.323 protocol. These protocols provide the point-to-point and point-to-multipoint multimedia-communication services.

There are four basics components of H.323 that are Terminals, Gateways, Gatekeepers and MCU.

To use components of H.323 required their components protocols that perform different functions to enable communication. The Audio CODECs , Video CODECs , H.225 registration, admission, and status (RAS) ,H.225 call signaling , H.245 control signaling T.120 for multimedia conferencing ,H.450 Series for supplementary services, H.235 for Security and encryption for H-Series , Real-time transfer protocol (RTP) , Real-time control protocol (RTCP) .

Chapter # 04

Quality of Service

4.1 Overview

Now days, it's a trend for network designers to build multi-service networks carrying all types of communications (voice, data, and video) over a packet-based architecture. Bandwidth demand is ever increasing. However, the increased need for bandwidth can cause quality problems, especially degradation in time-sensitive traffic such as voice. Voice doesn't have the same luxury as data for dropping or re-transmitting its payload packets.

QoS refers to the ability of a network to provide better service to selected network traffic over various underlying technologies, including IP, Frame Relay, ATM, Ethernet and Synchronous Optical Network (SONET). In particular, QoS features provide better and more predictable network service by: [14]

1. Supporting dedicated bandwidth
2. Improving loss characteristics
3. Avoiding and managing network congestion
4. Shaping network traffic
5. Setting traffic priorities across the network

For VoIP, QoS defines limits on problems specific to VoIP such as (delay, jitter, signal loss, and echo)

In the case of voip, this typically means prioritizing voice traffic at a higher level than other forms of traffic such as data so that voice traffic will not be delayed or dropped. Most qos solutions focus on either resource reservation or resource prioritization.[15]

The Benefits of providing QoS in VoIP in a data network include:

1. Reduced operational costs
2. Higher performance
3. Greater flexibility
4. Faster application and service deployment

4.2 Definition

"Quality of service (QoS) refers to the measure of service quality provided to the user. QoS is a set of adjustable controls that create selective services for network traffic."

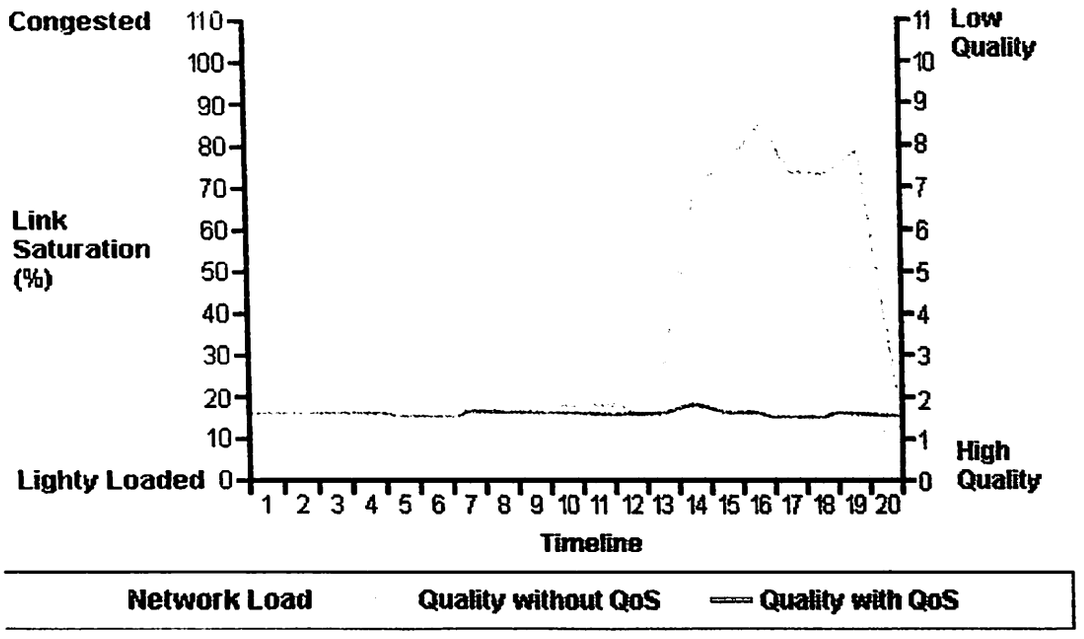


Figure 4-1 QoS vs no QoS

4.3 Factor Affecting QoS in VoIP

Voice quality in a VoIP network is affected by two main factors: [16]

1. Packet loss
2. Packet delay

4.3.1 Packet Loss

VOIP is exceptionally intolerant of packet loss. Packet loss can result from excess latency, where a group of packets arrives late and must be discarded in favor of newer ones. [17]

. There are many reasons for packet loss few are as fallows

1. Congestion caused by queues overflowing their packet limits
2. Network nodes running out of buffer space
3. Memory limitations in network nodes
4. Policing or controls that watch traffic flows, ensuring that they conform to certain bandwidth levels.

Packet loss causes voice "clipping" and "skips".

Packet loss should never exceed 1%. Packet loss of 1% translates into one voice clip or skip every three minutes, while packet loss of .25% translates into one error every 53 minutes.[18]

4.3.2 Packet Delay

Delay is caused when packets of data (voice) take more time than expected to reach their destination. This causes some disruption in the voice quality. However, if it is dealt with properly, its effects can be minimized. [19]

Packet delay can cause either voice-quality degradation, due to the end-to-end voice latency, or packet loss, if the delay is variable.

If the end-to-end voice latency becomes too long (e.g. 250 m-secs,), the conversation will sound like two person talking on a CB radio. (In CB radio communication, each person takes a turn talking and then ending his or her portion of the conversation with a keyword, instead of silence, to show the end).
If the packet delay is variable, there is a risk of jitter buffer overruns at the receiving end. Eliminating drops and delay is even more very important when including fax and modem traffic over IP.

If delay are not minimized, the voice signal received at the other end of the network will have poor quality (long overall delay is called **"unnatural conversation"**, sometimes known as the **"satellite effect"** and sounds like you are talking in an echo chamber), or crackles and missing syllables due to lost or delayed packets.

4.4 Causes of Packet Delay & Loss

There are three factors for packet delay and loss:

1. Poor Network Quality
2. Network Congestion
3. Delay and Jitter

4.4.1 Poor Network Quality

Poor network quality can lead to sessions frequently going out of service because of a loss of physical or logical connections.

4.4.2 Network Congestion

Network congestion may lead to both packet drops and packet delay. Voice packet drops due to network congestion are usually caused by full transmit buffers on the out interfaces somewhere in the network. As links or connections approach 100-percent utilization, the queues servicing that connection will fill. When a queue fills, packets attempting to enter the full queue will be discarded.

Because network congestion is irregular, delays due to congestion is variable in nature. These variable delays stemming from network congestion are caused by out interface queue wait times, or de-jitter buffers, or delays caused by variable packet size.

4.4.3 Delay and Jitter

Delay is the amount of time it takes a packet to reach the receiving endpoint after being transmitted from the sending endpoint. This time period is termed the "end-to-end delay".

In VOIP, **jitter** is the variation in the time between packets arriving, caused by network congestion, timing drift, or route changes. [20]

4.5 Types of delay

The causes of packet loss and delay can be classified into two types:

1. Fixed network delay
2. Variable network delay

4.5.1 Fixed Network Delay

Fixed network delay should be examined during the initial design of the VoIP network. The ITU standard G.114 states that a 150-msec one-way delay budget is acceptable for high voice quality. There is a negligible difference in voice quality scores using networks built with 200-msec delay budgets. There are three components of fixed network delay:

1. Propagation delay
2. Serialization delay
3. Processing delay

4.5.1.1 Propagation Delay

Propagation delay is the delay of signals between the sending and receiving endpoints and is based on the total distance between source and destination. [21]

4.5.1.2 Serialization Delay

Serialization delay is the result of placing bits on the circuit. The higher the circuit speed, the less time it takes to place the bits on the circuit. So, the higher the speed, the less serialization delay. [22]

The larger the packet and the slower the link clocking speed, the greater the serialization delay. Serialization delay is not variable.

4.5.1.3 *Processing Delay*

Processing delays can be defined as follows: [23]

Coding, compression, decompression, and decoding delay will be based on the algorithm used. These functions can be performed on either hardware or software. Using specialized hardware such as DSPs will dramatically improve the quality and reduce the delay associated with different voice-compression schemes.

Packetization delay [24] is the process of holding the digital voice samples for placement into the payload until enough samples are collected to fill the packet or cell payload. To reduce excessive packetization delay associated with some compression schemes, partial packets could be sent.

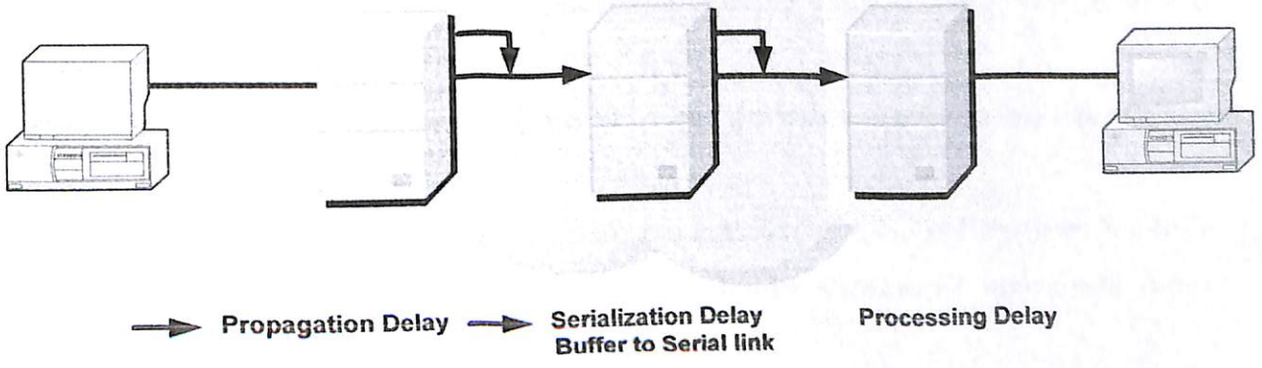


Figure 4-2 Fixed Delay

4.5.2 Variable Network Delay

The following three factors contribute to variable network delay: [25]

1. Queuing delay
2. De-jitter buffers

3. Variable packet size

These are described below. The delay components depicted on this graphic are variable, and they result in higher delay variation; they are also more controllable.

4.5.2.1 Queuing/processing Delay

Congested outlet queues on network interfaces are the biggest source of variable delay. Queuing delay occurs when a packet is waiting for others to be serviced first on the trunk. This waiting time is based on the arrival of traffic. [26]

4.5.2.2 De-jitter Buffers

Jitter, is the variation between when a packet is expected to arrive and when it actually is received.

To compensate for jitter VoIP endpoints (receiving end) use de-jitter buffers (also called *jitter buffers*) to turn these delay variations into a constant value so voice can be played out smoothly.
A jitter buffer is used to temporarily hold the voice packets in order to smooth out the variations in packet-delay values.
Setting these buffers too low will cause overflows and loss of data, setting them too high will cause excessive delay. In effect, a de-jitter buffer reduces or eliminates delay variation by converting it to fixed delay.

4.5.2.3 Variable Packet Size

Bigger packets take longer to transmit than smaller packets. A queue combining both big and small packets will experience varying lengths of delay. Variable packet size is considered part of variable network delay because the random size of packets arriving for

transmission causes a variable length of delay in the transmission time as the packets are queued for sending.

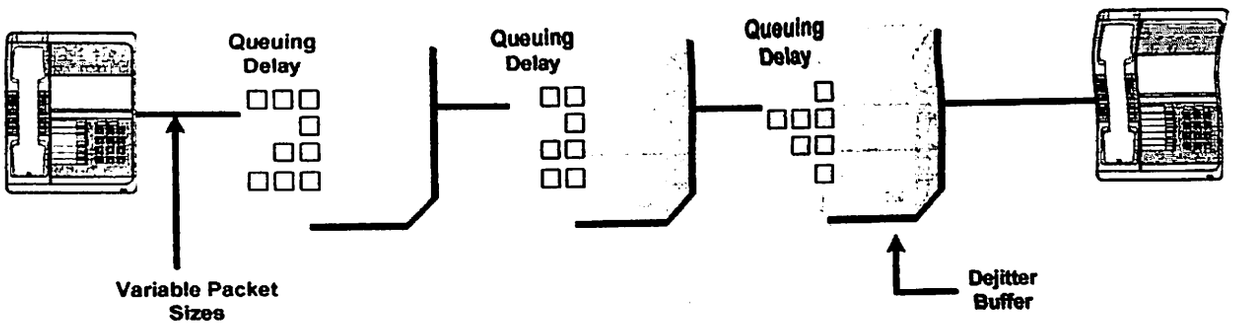


Figure 4-3 Variable Delay

4.6 QoS Categories

VoIP QoS tools can be separated into three categories:

1. Classification
2. Queuing
3. Network provisioning

4.6.1 Classification

Classification is marking a packet or flow with a specific priority. Classification should take place at the network edge, typically in the wiring closet or within the IP phones or voice endpoints themselves.

Packets can be marked as important by using Layer 2 class-of-service (COS) settings in the User Priority bits of the 802.1p portion of the 802.1Q header or the IP Precedence/differentiated-services-code-point (DSCP) bits in the type-of-service (TOS) byte in the ipv4 header.

All IP phone Real-Time Transport Protocol (RTP) packets should be tagged with a values of COS = 5 for the Layer 2 802.1p settings and IP Precedence = 5 for Layer 3 settings. Additionally, all control packets should be tagged with a Layer 2 COS value of 3 and a Layer 3 TOS of 3. [27]

Layer 2 Class of Service	IP Precedence	DSCP
COS 0	Routine (IP Precedence 0)	0–7
COS 1	Priority (IP Precedence 1)	8–15
COS 2	Immediate (IP Precedence 2)	16–23
COS 3	Flash (IP Precedence 3)	24–31
COS 4	Flash-override (IP Precedence 4)	32–39
COS 5	Critical (IP Precedence 5)	40–47
COS 6	Internet (IP Precedence 6)	48–55
COS 7	Network (IP Precedence 7)	56–63

Table 4-1 Relationship between the values for COS, IP Precedence, and DSCP.

This use of IP Precedence to mark traffic is a transitional step until all IP devices support the DSCP. Ideally, all Cisco voice-over-IP (VoIP) endpoints will use a DSCP value of Expedited Forwarding (EF) for the RTP voice-bearer flows and DSCP = Assured Forwarding 31 (AF31) for VoIP control traffic.

4.6.2 Queuing

Queuing assigns a packet or flow to one of multiple queues, (based on classification), for appropriate treatment in the network. When data, voice, and video are placed in the same queue, packet loss and variable delay are much more likely to occur. Using multiple queues on out interfaces (instead of placing voice, data, and video in the same queue) and separating voice into a different queue from data makes network behavior much more predictable.

Serialization delay is considered part of an overall queuing solution. Serialization delay is a factor only on slow speed links (link speeds of less than 1 megabit per second [Mbps]).

4.6.3 Network Provisioning

Network provisioning consists of accurately calculating the required amount of bandwidth for running voice over the wide-area network (WAN), all data traffic, any video applications, and necessary link management overhead, such as routing protocols.
It is important to remember that all the application traffic, (that is, voice, video, and data traffic), when added together should equal only 75 percent of the provisioned bandwidth. The remaining 25 percent is used for overflow and administrative overhead, such as routing protocols.

4.7 QoS Techniques

Various quality-of-service (QoS) techniques are:

1. Compression
2. Call Admission Control (CAC)
3. Tagging
4. Queuing
5. Traffic Shaping
6. Fragmentation
7. Media

The following sections of this document explain the different characteristics of the techniques listed above, with advice on what combinations may be appropriate in different types of network.

4.7.1 Compression

Uncompressed voice, (G.711 pulse code modulation [PCM]), has a bandwidth of 64 kilobits per second (kbps) (this is before adding header overhead). In VoIP networks, a full-rate voice-over-IP (VoIP) call takes about 84 kbps of network bandwidth. Cisco VoIP gateways can use G.711, or can compress the voice signal to 8 kbps using the G.729 coder/decoder (codec).

Compression reduces the bandwidth necessary for each voice call and hence increases the number of calls that can fit on a given trunk. The table below shows the various codecs and their bandwidths:

Codec	Type	Bandwidth
G.711	PCM	64K
G.726	ADPCM	32K
G.726	ADPCM	24K
G.726	ADPCM	16K
G.728	LD-CELP	16K
G.729	CS-ACELP	8K
G.729A	CS-ACELP	8K
G.729B	CS-ACELP	8K (with built- in VAD)
G.729AB	CS-ACELP	8K (with built- in VAD)
G.723.1	MP-MLQ	6.3K
G.723.1	ACELP	5.3K
G.723.1A	MP-MLQ	6.3K(with built- in VAD)
G.723.1A	ACELP	5.3K(with built- in VAD)

Table 4-2 various codecs and their bandwidths

4.7.1.1 Additional QoS Features for Compression

In addition to codec, there are other QoS features that can be compression:

- **Voice activity detection (VAD):** When VAD is enabled on a voice port or a dial peer, silence is not transmitted over the network; only audible speech is transmitted. The sound quality is slightly degraded, but the connection uses much less bandwidth. VAD can reduce the transmitted signal by a factor depending on the silence content of the original signal: for normal conversation, 35-percent reduction is a reasonable.

- **Size of the voice payload:** Another important factor in determining the bandwidth per voice channel is the size of the voice payload in each packet. Here the number of bytes of payload can be specified as part of the codec selection, and it can have a dramatic effect on bandwidth utilization.

- **Compressed RTP (CRTP):** It is configured on a link-by-link basis. It takes the Real-Time Transport Protocol/User Datagram Protocol (RTP/UDP)/IP header on the voice packet and compresses it from 20 bytes to about 2 bytes. This compression technique is applied per link.

Table 4.3 below shows the total bandwidth requirements for two specific codecs and more importantly, the overhead associated with the codecs and Layer 2 overhead. Note the header reduction by using CRTP.

Coding Algorithm	Voice Bandwidth (kbps)	Coded Frame Size (bytes)	Frame in VoIP Packet	IP Header Size (bytes)	Layer 2 Technology Used	Layer 2 Header Size (bytes)	Total Bandwidth Required
G.711	64	80	2	40	Ethernet	14	85.6
G.711	64	80	2	40	MLP	6	82.4
G.711	64	80	2	2(CRTP)	MLP	6	67.2
G.729	8	10	2	40	Ethernet	14	29.6
G.729	8	10	2	40	MLP	6	26.4
G.729	8	10	2	2(CRTP)	MLP	6	11.2

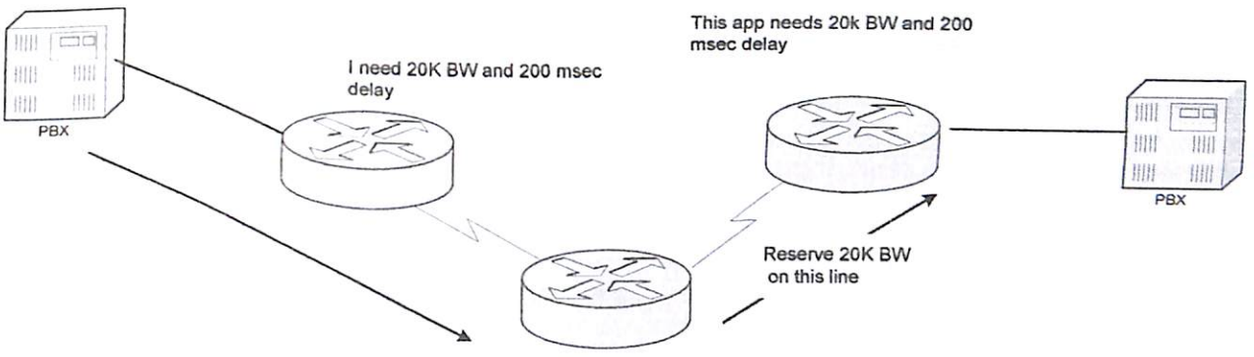
Table 4-3 total bandwidth requirements for two specific codecs, and overheads

4.7.2 Call Admission Control (CAC)

CAC ensures that the network does not accept or set up voice calls for which it has insufficient bandwidth (or other resources). The primary component of CAC is RSVP.

4.7.2.1 Resource Reservation Protocol (RSVP)

RSVP is the only form of end-to-end CAC currently available for VoIP. When a call is set up, messages are passed over the network and back again to the originator requesting a given bandwidth (among other parameters). If that bandwidth is available over the entire path, the call will progress, and queues along the path will be modified to provide a bandwidth reservation. If the requested bandwidth is not available across the network, the call setup should fail (however, in the current software implementation, calls will still progress; thus RSVP could be seen as a way of adjusting queue weights but not really of achieving CAC). The following figure describes the operation of RSVP:



- *RSVP QoS Services*
  - *Guaranteed Service - Mathematically Provable Bounds on End-to-End Datagram Queuing Delay/Bandwidth*
  - **Controlled Service** - Approximate QoS From an Unloaded Network for Delay/Bandwidth
- **RSVP Provides the Policy to WFQ**

Figure 4-4 Operation of RSVP

RSVP works in conjunction with WFQ in this way: when an intermediate router agrees to an RSVP request, it sets up a queue for traffic with the requester's source address and allocates enough weight to the queue to give absolute priority over the interface to the bandwidth requested. In the same way as with RTP reservation, if the reserved traffic does not use all its bandwidth, other traffic on the interface will be able to use it.

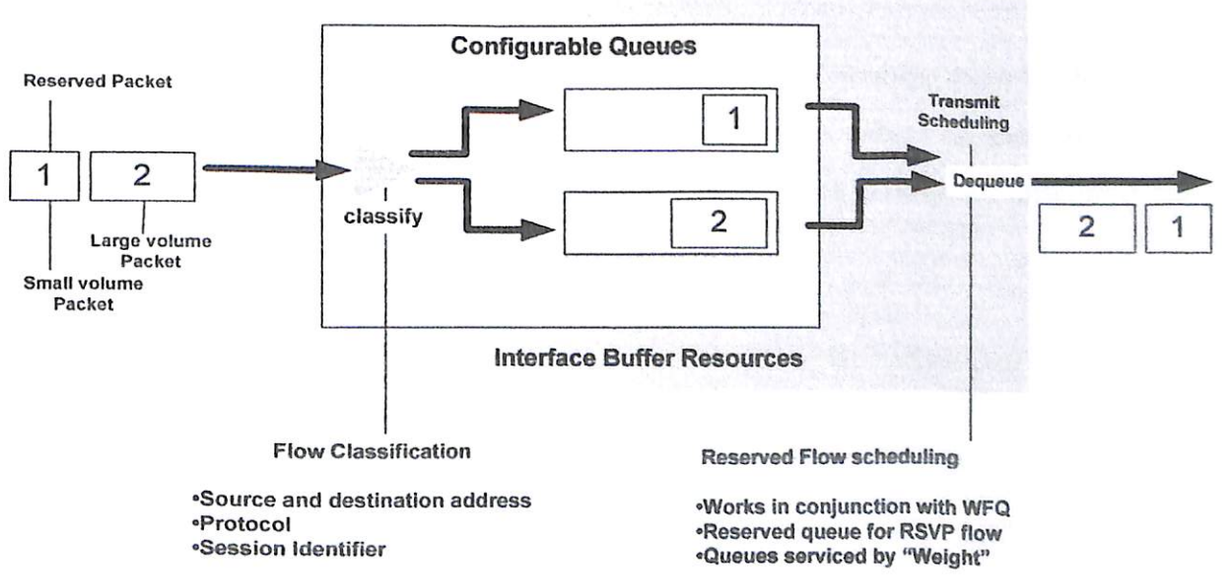


Figure 4-5 RSVP in Conjunction with WFQ

For the purposes of VoIP networks, RSVP can be a useful technique. It ensures that appropriate priority is given in queues so that the traffic will not be discarded, and *that the* traffic will suffer the minimum delay and jitter.

4.7.2.2 *Alternatives to RSVP for CAC*

RSVP is not at present an effective CAC mechanism. So whether or not RSVP is used, the network designer must ensure that voice-call patterns will not overbook the available network bandwidth. The usual way to do this is to overlay the worst-case (or busy-hour) calling pattern over the network and ensure that the required traffic can be accommodated on network trunks. If this is the case (as it probably would be if voice traffic is only 25 percent of total traffic on the network), then CAC is unnecessary and RSVP does not need to be used.

A simple way of designing networks to avoid CAC problems is to limit the number of voice ports, so there can never be too many voice calls for the carrying capacity of any particular trunk.

4.7.3 Tagging

Tagging actually work in pair with queuing. We separated queuing from the tagging. The reason for this is that queuing provides the structure, while tagging provides the indication that a particular packet contains voice, so that it will be processed by the correct queue.

One simple way to tag a packet as voice is to use an implicit characteristic:

currently voice is the only type of traffic to use UDP/RTP packets, which in Cisco router networks have UDP port numbers numbered from 16384 (other vendors use different ranges). When this characteristic is used, queuing on each network interface carrying voice should be configured with an access list to give priority to UDP traffic with these

port numbers. This can be achieved with Priority Queuing (PQ), WFQ, CBWFQ, and LLQ.

Another way is to use IP Precedence.

Configure a dial peer to assign an IP Precedence tag to all voice traffic it generates. Usually an IP Precedence of 5 is used to give voice priority (the range is 0 to 7, with 7 being the highest priority, for network control traffic). Refer to the graphic of IP Precedence below. WFQ will automatically give such traffic the highest priority, without special configuration.

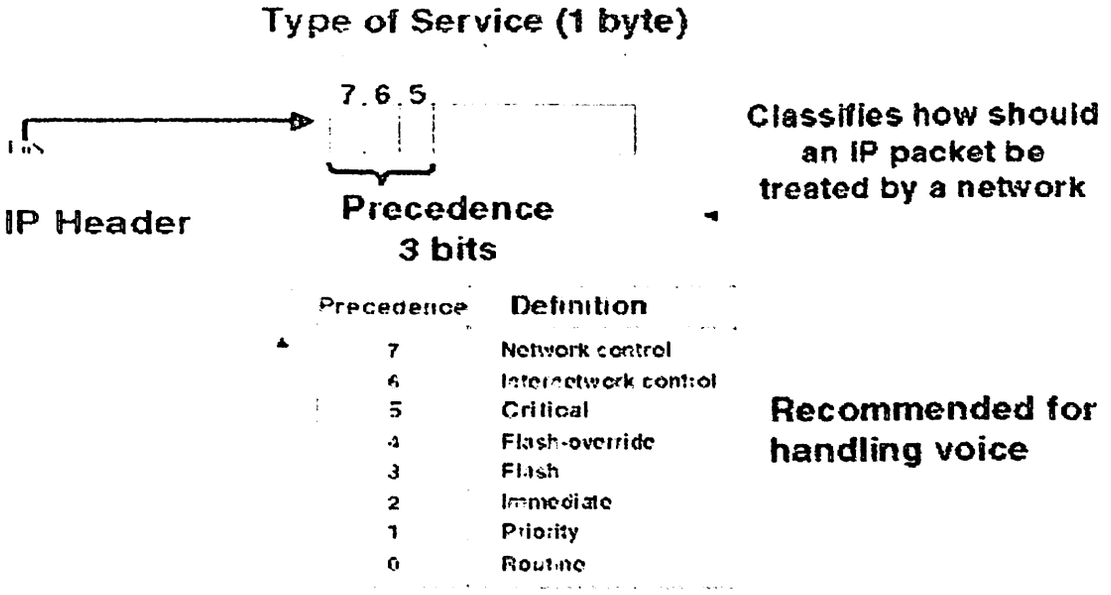


Figure 4-6 Tagging Operation

In the above figure, the Type of Service (TOS) field is 1 byte wide. The 3 most significant bits define the IP Precedence level while the remaining bits are set to 0. For example, to have an IP Precedence level of 5, the value of the TOS field would be 160.

4.7.4 Queuing

Queuing is a QoS function that buffers packets waiting for transmission. It is a process in which packet networks absorb bursts of traffic in excess of trunk bandwidth. If packets arrive at the back of a queue (or if higher precedence traffic arriving later takes a higher position in the queue), then they will be delayed. If the queue becomes full, then the lowest-priority packets will be dropped.

When voice traffic is to be carried on packet networks, queuing generally functions to give voice priority over data traffic. If any traffic is to be unduly delayed or discarded, it should not be voice traffic. Thus any queuing technique used must be able to distinguish the voice packets from data packets, and to give the packets priority through the queues. The queuing techniques covered in this course include:

- Weighted Fair Queuing (WFQ)
- Priority Queuing WFQ (PQWFQ) (also known as IP RTP priority)
- Class-Based WFQ (CBWFQ)
- Low-latency queuing (LLQ) (also known as Priority Queue, Class-Based Weighted Fair Queuing [PQCBWFQ])

The latest advanced technique is LLQ. This queuing technique supersedes all previous queuing methods [28].

4.7.4.1 Weighted Fair Queuing (WFQ)

WFQ in its default mode separates traffic on an interface into flows, determines the transmission rate of each flow, and then weights the priority of each flow. From the perspective of WFQ, there are two categories of data streams: high-bandwidth sessions and low-bandwidth sessions. Low-bandwidth traffic is given effective priority over high-bandwidth traffic, and high-bandwidth traffic shares the transmission service proportionally according to assigned weights.

WFQ in its default state is self-discovering; that is, the queuing mechanism separates and measures the different traffic flows, and then assigns weights. In networks carrying VoIP, it is normally expected that the voice traffic would be a low-bandwidth flow, and that it would thus be given higher weighting, but this is not a recommended configuration. For a stronger prioritization, WFQ can be configured to give absolute priority to traffic based on IP Precedence, or on RTP/UDP port number. The graphic below shows conceptually how WFQ works to handle traffic.

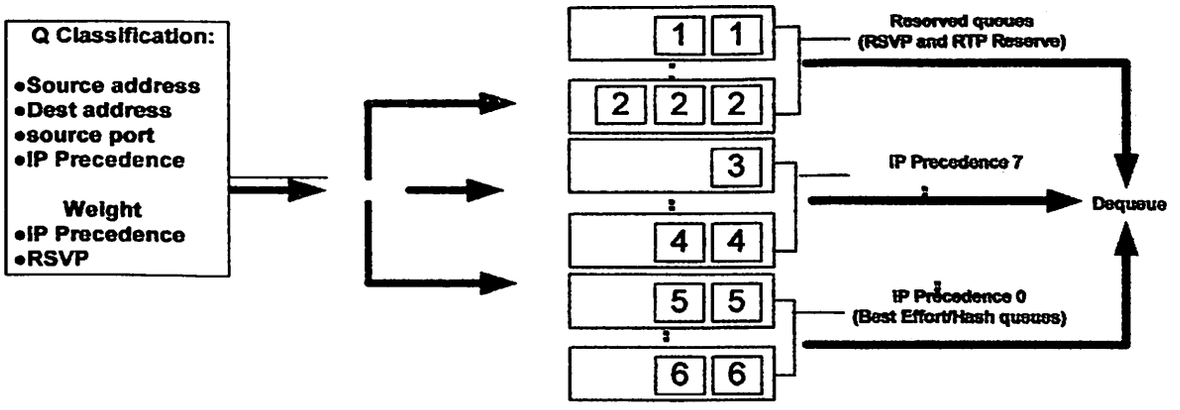


Figure 4-7 Operation of Weighted Fair Queuing (WFQ)

4.7.4.2 IP RTP Priority

IP RTP priority queuing (also known as Priority Queue, Weighted Fair Queuing or PQWFQ) is a queuing feature that provides a strict priority queuing scheme for delay-sensitive data such as voice. Using this feature, voice traffic is identified by its RTP port numbers and classified into a priority queue configured by the IP RTP priority command. This enables voice to receive absolute priority in service handling over non-voice traffic. When the priority queue is empty, the other queues are serviced using standard WFQ. The figure is on the next page that shows how PQWFQ operation:

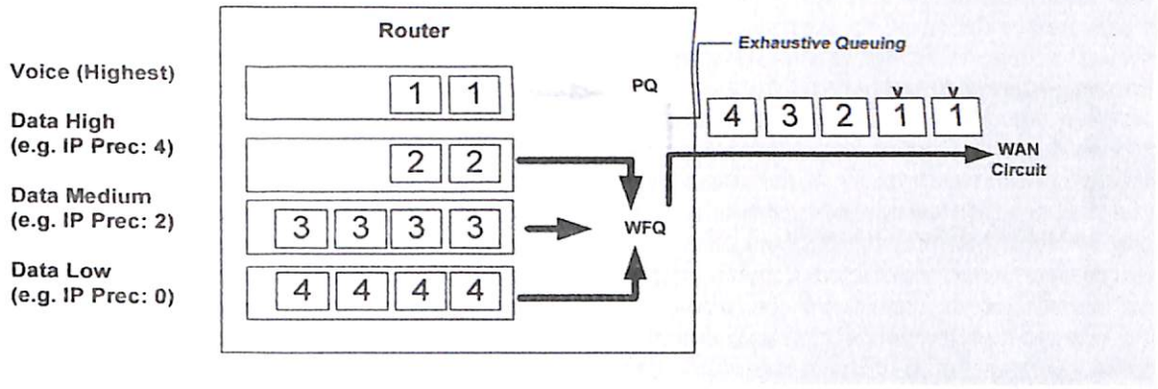


Figure 4-8 Operation of Priority Queuing WFQ (PQWFQ)

With PQWFQ, customers can specify the exact amount of bandwidth allocated for priority voice traffic. PQWFQ closely polices use of the bandwidth and drops voice packets only when exceeded. This feature helps guarantee voice performance by assigning strict VIP status to voice traffic.

4.7.4.3 Class-Based Weighted Fair Queuing

CBWFQ allows customers to define a specific class for voice traffic using standard and extended numbered access control lists (ACLs) and protocols, as well as interface names. Packets satisfying the match criteria for a class become the traffic for that class.

Each class defined using CQWFQ is associated with a specific queue. Customers can allocate the minimum amount of bandwidth guaranteed to the class as a percentage of the link or in kbps. Unused bandwidth can be shared by other classes in proportion to their assigned weights.

Although CBWFQ does not assign voice traffic absolute priority as does PQWFQ, weights can be configured to simulate priority queuing. Because the weight for a packet belonging to a specific class is derived from the bandwidth assigned to the class when created, weights are, in effect, user configurable.Below figure is showing how CBWFQ works

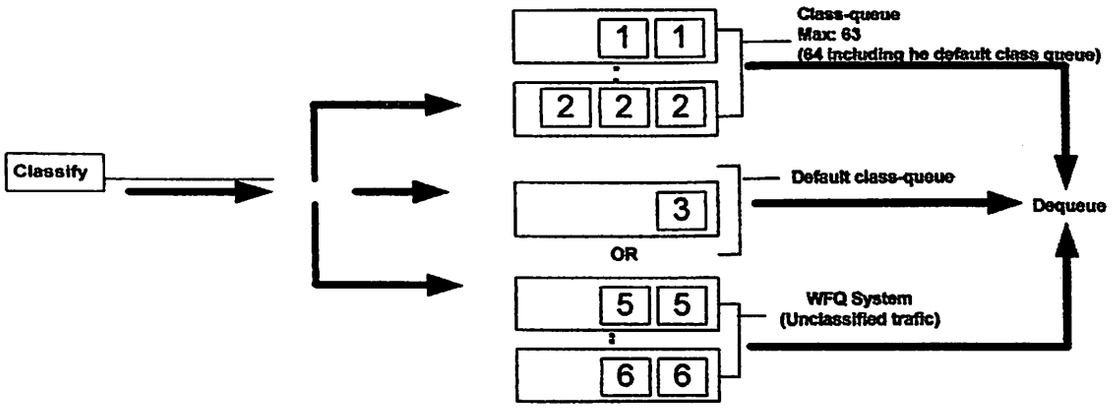


Figure 4-9 Operation of Class-Based WFQ (CBWFQ)

4.7.4.4 Low Latency Queuing (also known as PQCBWFQ)

Low-latency queuing (LLQ), also known as Priority Queuing, Class-Based Weighted Fair Queuing (PQCBWFQ), is a combination of the two queuing techniques listed above (PQWFQ and CBWFQ). In this technique, the voice classes can be configured for priority x instead of for bandwidth x to guarantee x bandwidth for voice and give it priority. The rest of the traffic can also be classified and CBWFQ will be applied to it.

LLQ offers the same voice-packet treatment as PQWFQ but offers myriad ways of configuring traffic and a much more granular control over what can go into the priority queue (PQ). LLQ is a superset of CBWFQ, so everything you can do with CBWFQ you can also do with LLQ. Below is a figure showing LLQ operation

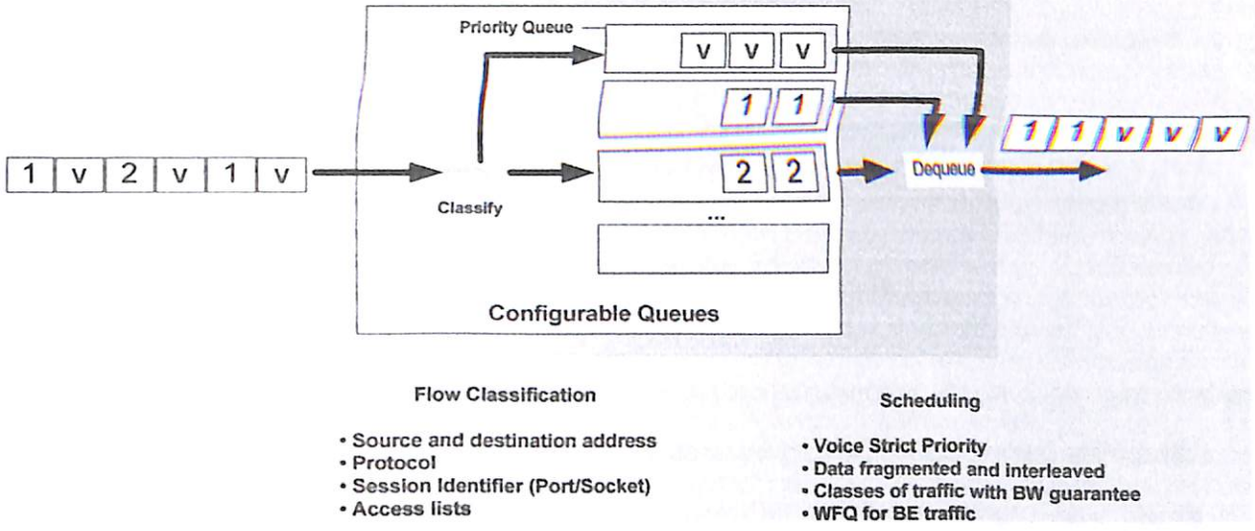


Figure 4-10 Operation of Low-latency queuing (LLQ)

4.7.5 Traffic Shaping

Traffic shaping is applicable only to Frame Relay or ATM networks. Basically, if the media used in the network is a leased line, traffic shaping does not apply. In general, it is important to configure the router traffic shaping such that the Frame Relay network committed information rate (CIR) is not exceeded; this could result in excessive delay or discard of voice packets, either of which would affect voice quality.

Traffic shaping can be a complex problem, especially in Frame Relay networks. As our concerning is VoIP not VoATM or VoFR .So we will not discuss it in our thesis.

4.7.6 Fragmentation

.

Fragmentation may be necessary if a trunk over which voice and data traffic will pass is a low-bandwidth trunk.

Trunk Bandwidth (kbps)	56	64	128	256	512	728	1000+
Fragment Size (bytes)	70	80	160	320	640	1000	Not necessary*
*Assumes maximum packet size of 1500 bytes							

Table 4-4 BW and Fragment Size

The table above is a general rule. The fragmentation is to ensure that a voice packet does not get delayed if it has to wait for a data packet to be played out on a trunk ahead of it.

Clearly, the delay depends on the bandwidth of the trunk and the size of the data packet. The figures above will ensure that the delay seen by the voice packet is no greater than 10 milliseconds (m-sec), which is a very conservative figure.

To identify the allowable delay for a voice packet, it is first necessary to:

- Determine the worst-case route for a voice packet through the network.
- Add the worst-case delays due to queuing and propagation delays and dejitter buffers.
- Finally, subtract that sum from the budgeted delay for voice across the network, usually in the order of 150 to 200 m-sec. The resulting figure will indicate the allowable delay due to fragmentation.

In simple networks, this amount may be in the order of 20 to 50 m-sec; in more complicated networks it will be closer to the 10 m-sec assumed in the table above. In large networks it will be worthwhile to calculate the voice delay budget, because any loosening of the restriction on fragment size will increase the performance of the network (through lower packet-per-second [pps] rates reducing CPU utilization, and lower header overhead)

4.7.6.1 Fragmentation Techniques

The following are fragmentation techniques. Each technique has different characteristics for use in specific situations. These are described below.

- **Multilink PPP (MLP)** with interleaving is a per-link fragmentation technique. This means that long packets are fragmented for transmission onto the trunk, and reassembled on receipt from the trunk; they do not travel through the entire network as fragments. MLP is also transparent to all protocols (for example, TCP/Internetwork Packet Exchange [IPX]/SNA). MLP is generally the preferred fragmentation technique on leased-line trunks, but it cannot be used over Frame Relay trunks (this is a general Cisco IOS Software limitation).

- **FRF.12** is an implementation agreement (standard) of the Frame Relay Forum (www.frforum.com). It may be thought of as an alternative to MLP for fragmentation and interleaving. FRF.12 is often used in conjunction with VoFR (FRF.11). In this case, the fragmentation follows similar architecture but different implementation, so data using the same PVC as VoFR is fragmented according to FRF.11 Annex C. There are some caveats and limitations to mixing VoFR and VoIP on the same interface. However, FRF.12 can be used as a fragmentation technique on Frame Relay trunks, without VoFR present but with VoIP.

FRF.12 has all the advantages of MLP:

  - The overhead per link is low.
  - It can be applied per PVC.
  - The fragmentation threshold can be set.
  - It is protocol transparent.

  It can be applied only on a Frame Relay trunk, but a leased-line trunk can be configured for Frame Relay encapsulation.

When FRF.12 is used on PVC, the queuing technique is automatically LLQ. If for some reason it is necessary to use PQ or Custom Queuing (CQ), then MLP is a better choice (but MLP cannot be used on Frame Relay trunks).

4.7.6.2 IP MTU Size Restriction

An alternative fragmentation technique is IP maximum-transmission-unit (MTU) size restriction. IP MTU fragments IP packets processed by the router when those packets are longer than the configured limit. The fragments then continue through the network to the receiving IP device (the terminating host rather than the last router), which reassembles the long IP packets. There are many drawbacks to using IP MTU size restriction: Some IP endpoints cannot reassemble fragments shorter than 256 bytes. In any case, it may be necessary to reconfigure applications that terminate fragmented traffic.
The use of IP MTU fragmentation places a real-time load on the router processor, both because fragmentation increases the PPS load on the CPU and because MTU size restriction forces process switching.

Fragments are not reassembled at the far end of the link, but continue through the network. This places a greater pps load on the entire network.
IP packets marked "do not fragment" will be passed as long as they are shorter than the configured MTU. If they are longer than the configured MTU, they will be dropped. Other (non-IP) protocol packets longer than the MTU threshold may be passed un-fragmented, interrupting voice conversations, or may be dropped, depending on the configuration.

4.7.7 Media

This function distinguishes between leased-line trunks, Frame Relay, and ATM networks. With a leased-line trunk, the router does not use any traffic shaping after the queuing technique. If the network uses leased-line trunks, it is not necessary to use a technique from the traffic-shaping function, because bandwidth is guaranteed and fixed.

With a Frame Relay trunk, traffic is queued for particular PVC using the queuing technique of choice, and then shaped so it meets the desired characteristics of the Frame Relay traffic contract. These normally include CIR, Committed Burst (Bc), and Excess Burst (Bc). The bandwidth available on Frame Relay connections should be at least the CIR, but may be more.

4.7.8 Resolving Echo Problems

Echo is a problem for both circuit- and packet-switched networks. There are specific steps to take to identify and correct the source of an echo:

1. Map out your network loss plans.
2. Identify which tail-circuit is the problem.
3. Avoid adding gain on the input side; it amplifies noise.
4. Try to reduce attenuation at the output instead.
5. To raise an output (attenuation) level:
    - Decrease the attenuation at the output side. If you are applying 0 decibels per milliwatt (dbm) of attenuation, and the signal is still too soft, then go to the input side and increase the gain. Working this way avoids overdriving the inputs on the first pass.
6. To lower an output (attenuation) level:
    - Adjust the input side first; then adjust the output side.
7. If the destination telephone is a speakerphone or headset, try replacing the speakerphone or headset with a decent handset, and see if the echo fades away normally.
8. Try performing a Telnet to the destination voice gateway and ensure that the echo canceller is provisioned on and coverage is set to maximum. Test for normal echo canceller behavior, (that, see if echo fades away within a few seconds of speech).
9. To verify that echo canceller is working correctly, perform the following steps:
    - The quickest way to tell if you have a working echo canceller in the circuit is to make a call, and immediately begin to say something like "TAH TAH FISH" repeatedly. The person on the other end of the line should be silent,

so if you are calling a voice-mail system, wait for the announcer to stop talking before starting the experiment.

- If the echo canceller is enabled and the echoes in the system are cancelable, you will hear echo for the first few utterances and will notice the echo dying away. After a few seconds of speech, the echo should be gone or at least more quiet compared to the echo level at the beginning of the call. This is proof of a working echo canceller.
- An echo canceller starts out with no knowledge of the tail circuit that it is looking into. It needs to observe a certain amount of speech and echo flowing through the tail-circuit to form the virtual tail circuit model. This learning period is known as the convergence time of the echo canceller. You should expect convergence within the first few seconds of active speech. If you try this experiment and do not obtain echo reduction with time, there are two possibilities:
  i. The echo canceller is disabled or broken.
  ii. The echo source is uncancelable (either too loud or delayed beyond the tail coverage of the canceller).
- Try making calls to other destinations and check for the standard echo-die-away behavior. If you don't find this behavior, you will have to be a bit more methodical to determine whether your echo canceller is working. The first thing to do is to check your provisioning. Remember, the echo canceller we are interested in is the echo canceller in the destination voice gateway, so perform a Telnet to that voice gateway and check the provisioning of your voice ports.
- Determine if your echo canceller is working by doing an A/B comparison between Canceller = OFF and Canceller = ON, over a call that exhibits echo in the Canceller = OFF case. Disable the echo canceller (no echo-cancel enable), then shut (shutdown) and reopen (no shutdown) the voice port. Make a call to a destination telephone and listen for echo by saying something like "TAH, TAH, TAH." If you don't hear any echo, try different destination phones until you do. When you've found an echo (the echo should persist throughout the call), save the destination number. Now re enable the echo canceller, set coverage to max, then shut and reopen (no shut) the voice port. You should hear the echo die away within the first

few seconds of speech. This procedure will have shown that the echo canceller makes a difference when it is turned on, and, therefore, your canceller is working.

10. Finally, if echo is still persistent and the echo canceller has been demonstrated to be working, this implies that the echo is beyond the ability of the echo canceller to fix. The echo is thus either too loud (more likely) or too delayed (much less likely).

4.8 Conclusion

QoS refers to the ability of a network to provide better service to selected network traffic over various underlying technologies, including IP, Frame Relay, ATM, Ethernet and Synchronous Optical Network (SONET).

The Benefits of providing QoS in VoIP in a data network include:

1. Reduced operational costs
2. Higher performance
3. Greater flexibility
4. Faster application and service deployment

Voice quality in a VoIP network is affected by two main factors:

1. Packet loss
2. Packet delay

There are three factors for packet delay and loss:

1. Poor Network Quality
2. Network Congestion
3. Delay and Jitter

VoIP QoS tools can be separated into three categories:

1. Classification
2. Queuing
3. Network provisioning

Various quality-of-service (QoS) techniques are:

1. Compression
2. Call Admission Control (CAC)
3. Tagging
4. Queuing
5. Traffic Shaping
6. Fragmentation
7. Media

Chapter # 05

Proposed Conclusion

Now days, it's a trend for network designers to build multi-service networks carrying all types of communications (voice, data, and video) over a packet-based architecture. Bandwidth demand is ever increasing. However, the increased need for bandwidth can cause quality problems, especially degradation in time-sensitive traffic such as voice. Voice doesn't have the same luxury as data for dropping or re-transmitting its payload packets.

QoS refers to the ability of a network to provide better service to selected network traffic over various underlying technologies, including IP, Frame Relay, ATM, Ethernet and Synchronous Optical Network (SONET). In particular, QoS features provide better and more predictable network service by:

1. Supporting dedicated bandwidth
2. Improving loss characteristics
3. Avoiding and managing network congestion
4. Shaping network traffic
5. Setting traffic priorities across the network

For VoIP, QoS defines limits on problems specific to VoIP such as (delay, jitter, signal loss, and echo).

In the case of voip, this typically means prioritizing voice traffic at a higher level than other forms of traffic such as data so that voice traffic will not be delayed or dropped. Most QoS solutions focus on either resource reservation or resource prioritization.

The Benefits of providing QoS in VoIP in a data network include:

1. Reduced operational costs
2. Higher performance
3. Greater flexibility
4. Faster application and service deployment

Voice quality in a VoIP network is affected by two main factors.

1. Packet loss
2. Packet delay

We'll remove these two factors by applying QoS techniques which are given below:

1. Compression
2. Call Admission Control (CAC)
3. Tagging
4. Queuing

On the basis of study and hands on experience, we are now in a position to implement these QoS techniques which we'll be implementing in FP-2.

5.1 Compression

We'll use the compression technique in order to remove the wastage of bandwidth. We'll enable voice activity detection (VAD) in which only the audible speech will be transmitted. We'll also use cRTP.

5.2 Call Admission Control (CAC)

CAC ensures that the network does not accept or set up voice calls for which it has insufficient bandwidth (or other resources). That is how we can remove wastage of bandwidth. The primary component of CAC is RSVP. We'll use CAC in our project implementation's phase.

5.3 Tagging

Tagging provides the indication that a particular packet contains voice, so that it will be processed by the correct queue. There are two ways to implement tagging.

I) One simple way to tag a packet as voice is to use an implicit characteristic.

II) Another way is to use IP Precedence.

5.4 Queuing

Queuing assigns a packet or flow to one of multiple queues, (based on classification), for appropriate treatment in the network. When data, voice, and video are placed in the same queue, packet loss and variable delay are much more likely to occur. Using multiple queues on out interfaces (instead of placing voice, data, and video in the same queue) and separating voice into a different queue from data makes network behavior much more predictable.

So, therefore, we'll also use queuing in project in order to remove packet loss and variable delay.

5.5 Rules of Thumb for QoS

Before implementing QoS on edges of network, ask following question

1. **Do you have low bandwidth WAN circuit?**

   If yes, use cRTP

   Also choose a fragmentation method, Multilink PPP (MLP) is recommended for all the networks except Frame relay, where FRF.12 is recommended. MTU is not recommended

2. **Does the traffic need Tagging on WAN link?**

   If yes, use LLQ (LLQ is recommended first), secondly use CB-WFQ

3. **Have u chosen CB-WFQ?**

   If yes, select a method for classify your traffic by tagging.

Final Project-II

Chapter # 06

Implementation

6.1 Project Topology

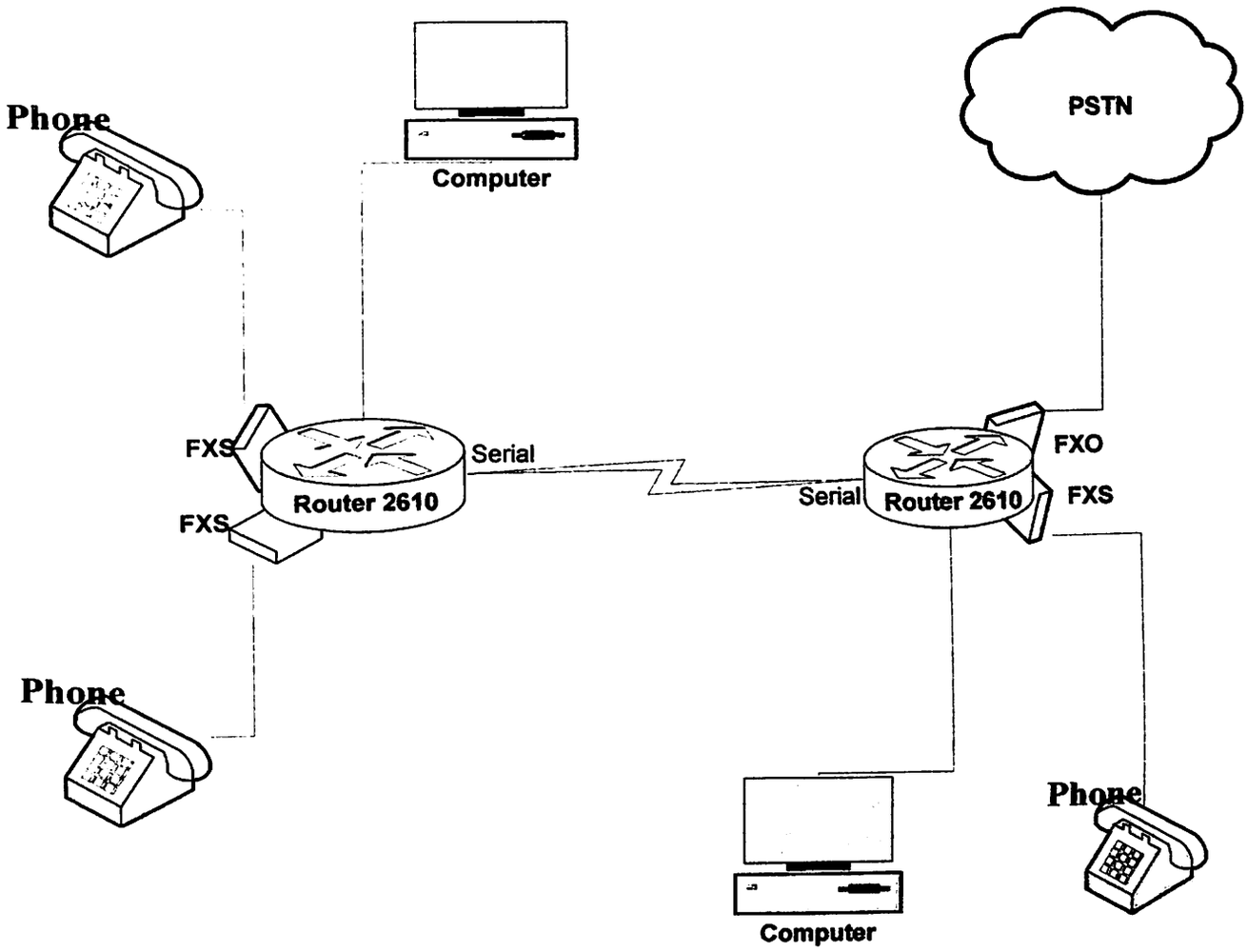


Figure 1.1 Topology of the VoIP

a. Component Detail

6.2.1 FXO and FXS

The terms "FXO" and "FXS" have their origins in an old telephone service called Foreign eXchange (FX) [29]. The original purpose of an FX circuit was to allow an analog phone at a remote location to be connected to a PBX somewhere else. An FX circuit has two ends (the Station end, where the telephone is, and the Office end, where the PBX is). The confusing part about understanding FXO and FXS is that FX cards are not named by what they are, but rather by what is connected to them. An FXS card, therefore, is a card that you connect a Station to. Since that is so, you can see that in order to do its job an FXS card must behave like a central office. Similarly, an FXO card connects to a Central Office, which means it will need to behave like a telephone (a modem is a classic example of an FXO device).

6.2.2 FXS

Foreign eXchange Station (FXS) channels provide the same interface as the traditional analog line your phone company provides to most homes or small businesses. Among other things, FXS channels would normally provide:

- Dial tone
- Ringing voltage
- DTMF (touch tone) detection
- Message waiting
- Calling Line ID

6.2.3 FXO

An FXO card is a card that connects to a central office. A modem is a classic example of an FXO card (in fact, if you have one of Digium's old FXO cards, the X100P, it is in actual fact a modem). An FXO device must be able to:

- Generate DTMF (touch tones)
- Detect dial tone
- Detect ringing
- Detect message waiting
- Interpret Caller ID
- Signal On Hook/Off Hook condition to the far end, as well as Flash

The main difference in the settings that are provided above when configuring an FXO channel instead of an FXS channel is signaling.

6.2.4 What's the difference between FXS and FXO?

These terms originated in traditional telephony, wherein equipment is divided into two major classes: **Office** and **Station**. [30]

In the public telephone system, telephone lines radiate from the public telephone exchange or "Central Office," to "stations" i.e. subscriber telephones.

The office supplies operating voltage, dial tone and ring signals to the station (telephone) while the station sends off-hook indication and dialing signals to the office.

The S in FXS stands for **Station**. An FXS gateway provides analog ports for connection to station equipment such as telephones, fax machines, or the trunk side of a PBX.

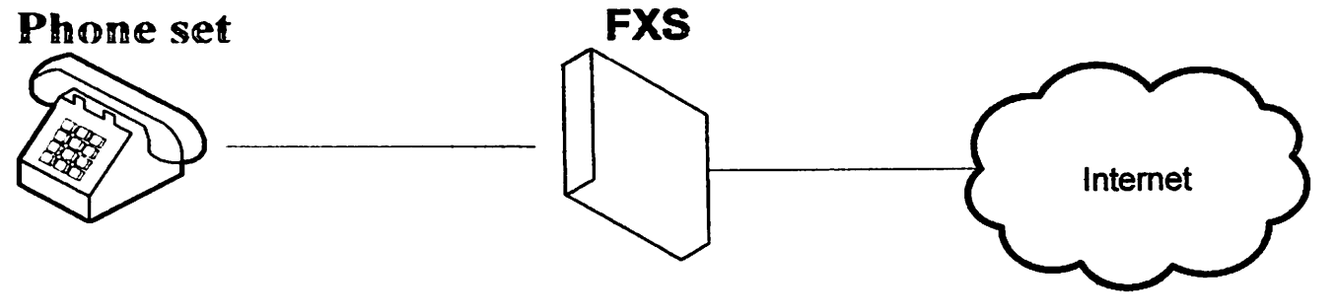


Figure 1.2 FXS card

The **O** in FXO stands for **Office. An FXO gateway provides analog ports for connection to Office equipment, such as lines from the Public Telco system or the line side of a switchboard (PBX).**

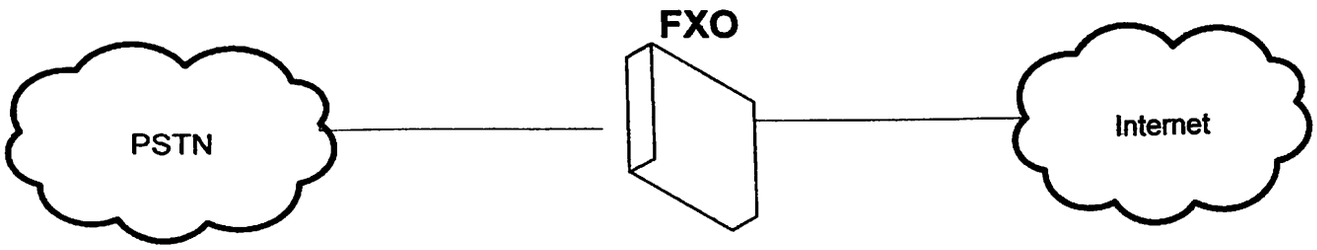


Figure 1.3 FXO card

6.2.5 Router 2610 Series:

Routers are devices that forward data packets along networks. A router is connected to at least two networks, commonly two LANs or WANs or a LAN and its ISP's network.

Routers are located at gateways, the places where two or more networks connect. Routers use headers and forwarding tables to determine the best path for forwarding the packets, and they use protocols such as ICMP to communicate with each other and configure the best route between any two hosts. Very little filtering of data is done through routers.

6.3 Working of Topology:

In this part of the project we are going to discuss the topology of the project, how a call from one end to another end will be routed.
There are two parts of every end station of the network, POTS and VoIP.

6.3.1 POTS

A POTS (Public Old Telephony System) is the part we configure for internal use of one end station. On the POTS end we connect telephone sets to the router using FXS cards. FXS cards generate the tone for the phone sets. After connecting phone sets to the router we do some configuration (which is given in the manual part of the project) to define the destination of the phone sets.

6.3.2 VoIP

VoIP is the part in which we configure the destination of the call of another end station. On the VoIP end we connect the router with the internet. After connecting to the internet we define the destination router using configuration where we want to route our calls.

6.3.3 Route of the call

When a phone set connected (using FXS card) with the router will dial a number, according to the configuration, router will check the outer port of the dialed number. For the destination of the dialed number, there can be two possibilities. It can be inside the end station or it can be outside the end station. If the destination of the dialed number is inside the end station its outer port will be port of may be another FXS card on the same router. So the router will route the call to that port. If the destination of that dialed number

is outside the end station the router will route the call to the configured serial port with the target of specific router where that dialed number is configured locally.

6.4 QoS Techniques:

Various quality-of-service (QoS) techniques we have discussed in earlier phase of the project are:

1. Compression
2. Call Admission Control (CAC)
3. Tagging
4. Queuing
5. Traffic Shaping
6. Fragmentation
7. Media

The techniques those we have implemented due to their various advantages are given below:

6.4.1 Compression

Uncompressed voice, (G.711 pulse code modulation [PCM]), has a bandwidth of 64 kilobits per second (kbps) (this is before adding header overhead). In VoIP networks, a full-rate voice-over-IP (VoIP) call takes about 84 kbps of network bandwidth. Cisco VoIP gateways can use G.711, or can compress the voice signal to 8 kbps using the G.729 coder/decoder (codec). Compression reduces the bandwidth necessary for each voice call and hence increases the number of calls that can fit on a given trunk. So, we have implemented this G.729 coder/decoder (codec).

In addition to codec, we have implemented other QoS features that can be compression:

6.4.1.1 Compressed RTP (cRTP)

Further more to reduce the large percentage of bandwidth consumed by the codec G.729, we are using cRTP. CRTP compresses the 40 byte IP/RTP/UDP header to 2 to 4 bytes most of the time. It is configured on a link by link basis.

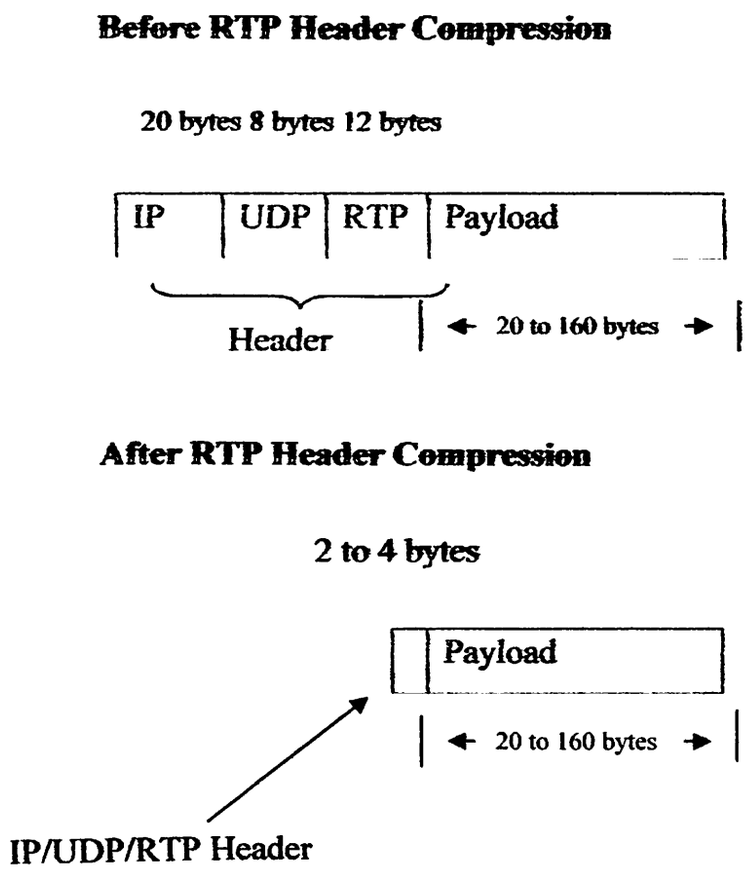


Figure 1.4 RTP Header Compression

6.4.1.2 Voice Activity Detection (VAD):

In normal voice conversations, someone speaks and someone else listens. Today's networks contain a bi-directional, 64,000 bits per second (bps) channel, regardless of whether anyone speaking. This means that in a normal conversation, at least 50 percent of the total bandwidth is wasted. The amount of wasted bandwidth can actually be much

higher if you take a statistical sampling of the breaks and pauses in a person's normal speech patterns.

We have enabled voice activity detection (VAD) to utilize this "wasted" bandwidth for other purposes.

6.4.2 Queuing

Queuing is a QoS function that buffers packets waiting for transmission. It is a process in which packet networks absorb bursts of traffic in excess of trunk bandwidth. If packets arrive at the back of a queue (or if higher precedence traffic arriving later takes a higher position in the queue), then they will be delayed. If the queue becomes full, then the lowest-priority packets will be dropped.

The queuing techniques are given below and we have discussed these in detail in the first part of our project:

- Weighted Fair Queuing (WFQ)
- Priority Queuing WFQ (PQWFQ) (also known as IP RTP priority)
- Class-Based WFQ (CBWFQ)
- Low-latency queuing (LLQ) (also known as Priority Queue, Class-Based Weighted Fair Queuing [PQCBWFQ])

On the basis of study and hands on experience, we observed that the LLQ is the best queuing technique that improves the voice quality. Because we can strictly give priority to voice call. This queuing technique supersedes all previous queuing methods. That is the reason why, we have implemented this technique.

6.5 Results

6.5.1 VoIP without QoS

If we don't follow the QoS techniques, we'll have some problems. For example, if we use G.711 codec. It consumes 96 kbps for a one voice call. If we are using low bandwidth link like 56 kbps. Then, we can't even go for one voice call. If we don't utilize the bandwidth efficiently, then the most of the bandwidth will be wasted. If we don't use

better queuing technique, then we can face these two factors (packet loss, packet delay). Also the voice quality will not good.

6.5.2 VoIP with QoS

Bandwidth is most important factor. If we make utilization of bandwidth better we can do more quality calls using low bandwidth link. To make bandwidth utilization better we have used following techniques:

1) We have used codec G.729 that consumes minimum bandwidth.
2) We have used compression technique cRTP to compress the required bandwidth for the call.
3) We have used VAD technique to enhance the efficiency of bandwidth.

We have used cRTP compression technique for the following reasons:

1. The cRTP reduces the line overhead for multimedia RTP traffic.
2. Reduces the delay.
3. The improvement in bandwidth is gained.
4. RTP header compression feature is used on a link-by-link basis.
5. CRTP compresses the IP/UDP/RTP header in an RTP data packet from 40 bytes to approximately 2 to 5 bytes
6. The de-compressor can reconstruct the original header without any loss of information.
7. CRTP is supported on serial lines using Frame Relay, High-Level Data Link Control (HDLC), or Point-to-Point Protocol (PPP) encapsulation. It is also being supported over ISDN interfaces.
8. Fast-switching with RTP Header Compression

Queuing is another important factor that plays an important part to improve the quality of voice call. Congestion is the main factor that causes packet delay and packet lose. As we have discussed in the earlier part of the project, these are two main factors that affect the

quality of service. By using efficient queuing technique we can control congestion efficiently.

We have used LLQ queuing technique for the following reasons:

a. Combination of CBWFQ and PQWFQ
b. More flexible
c. Not limited to UDP port numbers
d. Simple configuration
e. Ability to provide priority to multiple classes of traffic and give upper bounds on priority bandwidth utilization
f. You can also configure bandwidth guaranteed classes and a default class.
g. Bring strict priority queuing to CBWFQ, reduce jitter give voice packets priority but avoid starving non-priority traffic.

Acronyms

Address Complete Message (ACM)

Answer Message (ANM)

Automatic Number Identification (ANI)

Admission Request (ARQ)

Admission Confirm (ACF)

Admission Reject (ARJ)

Asynchronous Transfer Mode (ATM)

Assured Forwarding 31 (AF31)

Adaptive Differential Pulse Code Modulation (ADPCM)

Basic Rate Interface (BRI)

Bandwidth Request (BRQ)

Bandwidth Confirm (BCF)

Bandwidth Reject (BRJ)

Citizens band (CB)

Central office (CO)

Channel associated signaling (CAS)

Country Code (CC)

Common channel signaling (CCS)

Completion of Calls to Busy Subscribers (CCBS)

Completion of Calls on No Reply (CCNR)

Common intermediate format (CIF)

Class-of-service (COS)

Call Admission Control (CAC)

Coder/decoder (codec)

Compressed RTP (CRTP)

Class-Based WFQ (CBWFQ)

Committed information rate (CIR)

Central Processing Unit (CPU)

Committed Burst (BC)

Decibels per milliwatt (DBM)

Differentiated-services-code-point (DSCP)

Discrete cosine transforms (DCT)

Direct current (DC)

Dual Tone Multi-Frequency (DTMF)

Excess Burst (Be)

Expedited Forwarding (EF)

E1 Primary Rate Interface (E1 PRI)

Foreign Exchange Office (FXO)

Foreign Exchange Station (FXS)

Gatekeeper-Routed Call Signaling (GKRCS)

Gatekeeper Request (GRQ)

Gatekeeper Confirm (GCF)

Hyper Text Transfer Protocol (HTTP)

Internet Protocol (IP)

International Standards for Open Systems (ISO)

Internet Engineering Task Force (IETF)

International Telecommunication Union Telecom (ITU-T).

Initial Address Message (IAM)

Integrated Services Digital Network (ISDN)

Kilobits per second (kbps)

Local Area Networks (LAN)

Low-latency queuing (LLQ)

Multi-Frequency (MF)

Media Gateway Control Protocol (MGCP)

Multipoint Control Units (MCU)

Multipoint controller (MC)

Multipoint processors (MP)

Megabit per second [Mbps]

Multilink PPP (MLP)

Maximum-transmission-unit (MTU)

Milliseconds (m-sec)

MPEG-4 and MPEG-7

North America Numbering Plan (NANP)

Numbering Plan Area (NPA)

National Destination Code (NDC)

Open Systems Interconnection (OSI)

Public Switched Telephone Network (PSTN)

Pulse code modulation (PCM).

Private Branch Exchange (PBX)

Plain Old Telephone Service (POTS)

Private Integrated Services Networks (PISNS)

Personal computer (PC)

Priority Queuing (PQ)

Priority Queuing Weighted Fair Queuing (PQWFQ)

Priority Queue Class-Based Weighted Fair Queuing (PQCBWFQ)

Packet-per-second (pps)

Private Virtual Circuit (PVC)

Quality of Service (QoS)

Q.Signaling (QSIG)

Quarter common intermediate format (QCIF)

Robbed Bit Signaling (RBS)

Registration, admission, and status (RAS)

Real-time transfer protocol (RTP)

Real-time control protocol (RTCP)

Real-Time Transport Protocol (RTP)

Resource Reservation Protocol (RSVP)

Synchronous Optical Network (SONET)

Signaling System 7 (SS7)

Subscriber Number (SN)

Skinny Client Control Protocol (SCCP)

Session Initiation Protocol (SIP)

Simple Mail Transfer Protocol (SMTP)

Switched circuit network (SCN)

Transmission Control Protocol (TCP)

T1 Primary Rate Interface (T1 PRI)

Type-of-service (TOS)

User Datagram Protocol (UDP)

Voice over IP (VoIP)

Virtual private network (VPN)

Voice activity detection (VAD)

Wide-area network (WAN)

Weighted Fair Queuing (WFQ)

References

[1] James Peters, Jonathan Davidson, Brian Gracely Voice over IP Fundamentals Cisco Press, 1st Edition, 2003. (Page # 5)

[2] "Technology backgrounder on Telephony and VoIP Basics" by Lerry Hettick and Steven Taylor July 2004

[3] James Peters, Jonathan Davidson, Brian Gracely Voice over IP Fundamentals Cisco Press, 1st Edition, 2003. (Page # 11)

[4] SS7 services by Syniverse Technologies 2006

[5] James Peters, Jonathan Davidson, Brian Gracely Voice over IP Fundamentals Cisco Press, 1st Edition, 2003. (Page # 17)

[6] From Bradley Mitchell, "Your Guide to Wireless / Networking" Newsletter.

[7] http://www.telecomspace.com/vop-mgcp.html

[8] http://searchvoip.techtarget.com/sDefinition/0,,sid66_gci1077900,00.html

[9] http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci541639,00.html

[10]http://www.cisco.com/univered/cc/td/doc/product/voice/c_callmg/4_0/sys_ad/4_0_1/ccmsys/a08procl.htm

[11] James Peters, Jonathan Davidson, Brian Gracely Voice over IP Fundamentals Cisco Press, 1st Edition, 2003. (page # 229)

[12] http://www.rajjain.com/cis788-99/ftp/h323/index.html

[13] http://www.cisco.com/warp/public/788/voip/understand-gatekeepers.html

[14]http://www.cisco.com/univered/cc/td/doc/product/software/ios120/12cger/qos_c/qcintro.htm

[15] Jan 6,2006 "VoIPNews" newsletter

[16] http://electronics.howstuffworks.com/ip-telephony4.htm

[17] Security Considerations for Voice Over IP Systems

Recommendations of the National Institute of Standards and Technology

By D. Richard Kuhn, Thomas J. Walsh, Steffen Fries

[18] Jan 6.2006 "VoIPNews" Newsletter

[19] From Nadeem Unuth, "Your Guide to Voice Over IP" Newsletter

[20] http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci213534,00.html

[21] http://en.wikipedia.org/wiki/Propagation_delay

[22] Excerpt from a forthcoming Tutorial on Quality of Service (QoS) by Jason Wydra.

[23] http://www.cisco.com/warp/public/788/voip/delay-details.html#codeprocessdelay

[24] http://www.cisco.com/warp/public/788/voip/delay-details.html#packetizationdelay

[25] www.swinog.ch

[26] http://www.cisco.com/warp/public/788/voip/delay-details.html#queuingdelay

[27] http://www.linktionary.com/q/qos.html

[28] http://www.linktionary.com/q/queuing.html

[29] http://www.asteriskdocs.org/modules/tinycontent/content/docbook/current_v1/docs-html/x522.html

[30] http://www.qtelnet.com/telephony-index3.htm#Q0

BOOKS:

- Computer Networks - Andrew.S.Tanenbaum
- Data & Computer Communications - William Stallings
- Voice over IP Fundamental
- Voice over IP, ATM, Frame-Relay

USA Side Configration

A# config t

A# hostname USA_A

Make the dial-peer for both pots and voip sides.

USA_A# config t

For Local Calls within the 2 routers and 2 telephones sets through FXS CARDS.

USA_A (config) # dial-peer voice 1 pots
destination-pattern 101
port 1/0/0

USA_A (config) # dial-peer voice 2 pots
destination-pattern 102
port 1/0/1

USA_A (config) # dial-peer voice 3 voip
destination-pattern 2..
Session target ipv4: 1.1.1.2

(For one city like Lahore Only)

USA_A (config) # dial-peer voice 4 voip
destination-pattern 7.....
Session target ipv4: 1.1.1.2

(For Mobile Calls Only)

USA_A (config) # dial-peer voice 5 voip
destination-pattern 11.....
Session target ipv4: 1.1.1.2

(For Nationwide Calls Only)

USA_A (config) # dial-peer voice 6 voip
destination-pattern 10.....
Session target ipv4: 1.1.1.2

(For International Calls Only)

USA_A (config) # dial-peer voice 7 voip
destination-pattern 13.....

Session target ipv4: 1.1.1.2

How to Change the Codec?

USA A# config t

USA A (config) # dial-peer voice 3 voip

Codec g729 ulaw

How to set the clock rate?

USA A# config t

USA A (config) # clock rate 28800

How to enable the Voice Activity Detection (VAD)?

USA A# config t

USA A (config) # int s0/1

OR

USA A (config) # vad

How to enable cRTP?

USA A# config t

USA A (config) # int s0/1

Ip rtp header-compression

How to Implement the Queuing Technique?

Class Based Weighted Fair Queue with Priority or LLQ:

USA A# config t

USA A (config) # access-list 101 permit udp any any range 16384 32767

USA A (config) # access-list 101 permit tcp any any eq 1720

USA A (config) # class-map match-all data1

USA_A (config) # match input-interface Ethernet 0/0

USA_A (config) # class-map match-all voice1

USA_A (config) # match access-group 101

USA_A (config) # policy-map UMT_Policy

STRICT PRIORITY TO VOICE

USA_A (config) # class voice1

USA_A (config) # priority 14

USA_A (config) # class data1

USA_A (config) # priority 10

APPLYING ON INTERFACE THAT POLICY

USA_A (config) # int s 0/1

service policy output UMT_Policy

Pakistan's side Configration

Make the dial-peer for both pots and voice.

B# config t

B (config) # hostname PAK_B

For Local Calls within the 2 routers and 2 telephones sets through FXS CARDS.

PAK_B (config) # dial-peer voice 1 pots
destination-pattern 201
port 1/1/0

PAK_B (config) # dial-peer voice 2 voip
destination-pattern 1..
Session target ipv4: 1.1.1.1

(For one city like Lahore Only)

PAK_B (config) # dial-peer voice 3 pots
destination-pattern 7.....
Port 1/0/0

(For Mobile Calls Only)

PAK_B (config) # dial-peer voice 4 pots
destination-pattern 11.....
Port 1/0/0

(For Nationwide Calls Only)

PAK_B (config) # dial-peer voice 5 pots
destination-pattern 10.....
Port 1/0/0

(For International Calls Only)

PAK_B (config) # dial-peer voice 6 pots
destination-pattern 13.....
Port 1/0/0

How to Change the Codec?

USA_A# config t

USA_A (config) # dial-peer voice 2 voip

Codec g729 ulaw

The Codec must be same for both sides for the communication.