# Voice Traffic Protection in an IP Network

by

Sanjay Singh

A thesis submitted to

the Faculty of Graduate Studies and Research

in partial fulfillment of

the requirements for the degree of

MSc in Information Systems Science

School of Computer Science

at

CARLETON UNIVERSITY

Ottawa, Ontario

October, 2006

*Your file Votre référence*
*ISBN: 978-0-494-23363-4*
*Our file Notre référence*
*ISBN: 978-0-494-23363-4*

## Abstract

The phenomenal growth of VoIP is attributed to its cost saving and functional advantages over the traditional telephone line. Most of the protection for voice in today's network comes from over-provisioning, which is a viable option for an enterprise network, but may not be suited for public network. In public networks, the inherent unreliability of the router and network congestion often leads to unacceptable performance. Voice traffic from telephone conversation cannot be achieved reliably if the links have large packet loss or the mount-to-ear delays.

The network architecture for voice has major role to play in deciding the kind of protection is needed, e.g. the protection methods for central office based VoIP deployment will be very different from end user based deployment. The former could choose for SONET based protection, while the latter has to rely on IP based protection. For this reason, various VoIP deployments are discussed and general background on voice transportation over the IP is discussed. The review also includes a discussion on various VoIP components, and the factors that affect voice quality.

There are several protection mechanisms to achieve this reliability, and can be used at different protocol layer. This thesis focuses on IP layer protection. Such protection is achieved by MPLS or Differentiated Services. In a multi-homed network, additional protection can be provides by packet duplication. These techniques can be combined together. Unlike MPLS, Differentiated Services are simple to deploy and do not require major network upgrade.

The simulations were conducted for multi-homed networks with voice protected by premium class of Differentiated Services and packet duplication. The results suggest the greatest benefit is achieved when these two are used together. The simulations also show that smaller voice packets have better delay and jitter behaviour. However, in such packet the protocol overheads are large, but they are better choice if used along with the header compression.

## Acknowledgements

I would like to thank Prof. Evangelos Kranakis for providing guidance throughout this thesis. Prof. Kranakis kept me focused on the work and encouraged for in-depth literature survey. I would also like to thank him for being patient to progress in my thesis development.

My thanks also go to my employers Alcatel for providing financial support throughout this work. My special thanks to all my managers at Alcatel over the years for allowing me to take time-off either for attending classes or, preparing the thesis. I would like to mention Mr. Larry Boone, Mr. Piero Sorrini and Mr. Joe Rorai in particular for their understanding.

## Dedication

I dedicate this thesis to my wife Indu. Without her encouragement, I may not have completed this work. She has been very patient with my thesis and spent long hours with our children so than I could get time to work on course assignments or thesis.

# Glossary

| | |
|---|---|
| ADPCM | Codec algorithm used in G726 based encoders |
| AF | Assured Forwarding Per Hop Behaviour of Differentiated Services |
| BF | Best Effort |
| CBR | Constant Bit Rate |
| CE | Customer Edge |
| CO | Central Office of the telephone company |
| CODEC | COder and DECoder of voice |
| CPE | Customer Premises Edge |
| CR-LDP | Constraint-based Routing/Label Distribution Protocol |
| CS-ACELP | Codec algorithm used in G.729a based encoders |
| DOS | Denial of Service |
| DS, DiffServ | Differentiated Services |
| DSCP | Differentiated Services Code Point |
| EF | Expedited Forwarding, a premium class of service of Differentiated Services. |
| FEC | Forward Error Correction, a technique used for correcting lost voice packets. |
| FIFO | First In First Out |
| FTP | File Transfer Protocol |
| H.323 | ITU-T recommended signalling protocol for video telephony |
| HTTP | HyperText Transfer Protocol |
| IETF | Internet Engineering Task Force |
| InterServ | Integrated Services |
| IP | Internet Protocol |
| ITU-T | International Telecommunication Union – Telephony |
| LAN | Local Area Network |
| LSP | Label Switched Path used in MPLS |
| M2E | Mouth to Ear, refers to the instances when source speaks and the destination hears. |

| | |
|---|---|
| MOS | Mean Opinion Score, a measure of voice quality |
| MPLS | Multi-Protocol Label Switching |
| MRED | Multiple Random Early Detection |
| NS-2 | Network Simulator Version 2 |
| OSPF | Open Shortest Path First protocol for routing |
| PCM | Pulse Code Modulation |
| PHB | Per Hop Behaviour in Differentiated Services |
| PLC | Packet Loss Concealment, techniques to hide lost voice packets |
| PSTN | Public Switched Telephone Network |
| QoS | Quality of Service |
| RED | Random Early Detection |
| RFC | Request For Comments, documents issued by IETF |
| RIO | Random Early Detection with "in" (following SLA) and "out" (not following SLA) |
| RIO-C | RIO-Coupled. The out-of-profile packets are dropped based on the weighted average lengths of all virtual queues. In-profile packets are dropped based on the weighted average length of its virtual queue. |
| RTP | Real Time Protocol |
| SLA | Service Level Agreement |
| SIP | Session Initiation Protocol or initiating, modifying, and terminating an interactive user sessions. |
| SMTP | Simple Mail Transfer Protocol |
| SP | Service provider |
| SONET | Synchronous Optical NETwork |
| TCP | Transmission Control Protocol |
| TCP-Reno | Reno redefined RED for slow start, congestion avoidance, fast retransmit and fast recovery. |
| TOS | Type Of Service |
| UDP | User Datagram Protocol |
| UPS | Uninterrupted Power Supply |
| USB | Universal Serial Bus |

| | |
|---|---|
| VAD | Voice Activity Detector, detects silence period following a talkspurt |
| VoIP | Voice over Internet Protocol |
| WLAN | Wireless Local Area Network |
| WRED | Weighted Random Early Detection |

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1
# Introduction

## 1.1 Introduction

Integration of telephone and data networks is increasing among service providers. This is helped by wider deployment of broadband access to home. Services such as voice and streaming video that were limited to academics in the days of narrowband access have broken into the homes and business. Voice over IP enables a service provider to carry digitized voice packets over a data network, similar to an email. At the receiver end, the voice packets are reassembled into the analog voice. Service providers maintain two separate networks, one for voice and the other for data. Voice over IP not only eliminates cost of maintaining two separate networks, but also provides other values added voice services such as caller ID, directory services, call forwarding, call waiting and phone number mobility with little additional cost.

Voice over internet is thriving to replace Public Switched Telephone Networks (PSTN) those provides 99.999% of availability. Essentially, PSTN callers expect to be able to make a call every time they pick up the telephone, and expect high voice quality in telephone conversation. Comments such as "private data networks are only 94% reliable, while carrier data networks are 91% reliable and the public Internet is 61% reliable" (Typhault, 1998) are quite common for data networks. There are several techniques to increase the reliability of voice, such as router redundancy for fault tolerance and quality of service provided by Differentiated Services or MPLS, packet dispersion and multi-homing.

This chapter reviews advantages of VoIP, discusses major issues faced by this service, leading to discussion on the motivation for providing protection to voice.

## 1.2 Advantages of Voice over IP

The phenomenal growth of VoIP is attributed to its advantages over the traditional telephone line. These advantages are both functional, and cost saving as discussed below.

### 1.2.1 Cost Savings

The cost of making a call on a VoIP is much less than those compared to those on a traditional telephone. For long distance calls, VoIP is the primary medium for pre-paid calling cards, and has cut deeply into the profit of traditional telephone companies. The savings have also resulted from efficient bandwidth utilization, especially in long haul networks. VoIP runs on a high-speed network connection, and there is cost associated with the use of this. However, as the most users already have high-speed network service for their data traffic, no additional cost occurs for the network access.

### 1.2.2 Functional Advantages

Adding additional services to a VoIP system incurs minimal charge compared to the traditional phone service. Some of these services include:

- Portability - A VoIP user can travel anywhere in world and can make and receive call on the same phone number.

- Value added services – Features such as caller ID, call forwarding, call waiting, call screening, voice mail, and directory service can be added with little cost. PSTN can offer some of these services, but they are costly.

- Integration with IP services - The voice service on internet can be integrated with other service available on the Internet, e.g., file exchange, messaging and video. These services cannot converge on the PSTN across a 56 kbps modem, and high-speed broadband access is needed. As the last bandwidth issues are resolved with most homes connected with high-speed services, the convergence has already started.

### 1.3 Issues in VoIP

There are several challenges in VoIP deployment. Voice packets have to compete against data packet in the bottleneck links of an IP network. There are several factors that contribute to the reliability and deployment of a VoIP phone system; some of those are discussed below:

- VoIP architecture – IP does not guaranty delivery of packet in order, and voice decoder normally cannot use out of order packets and drops them. When several packets are dropped in a row, voice quality suffers. Further, because of the dynamics of the networks, packets arrival rate varies at the destination. Buffering at destination overcome some of these issues at the cost of additional delay.

- Network availability – This includes backup routes to the destination in case of failure of one or more routers in the path. Data traffic is tolerant to latency in

switches to the alternative path. However, voice traffic is disrupted when switch occurs.

- Network congestion – Voice traffic must be protected from other traffic in the network during congestions. This includes over-provisioning of the network as well as special considerations for voice traffic.

- Low speed access links – On low speed access links, VoIP service cannot be offered. This is because the serialization delay, that is time to push a packet onto a link, can be significant at low speed links. VoIP service providers usually don't offer VoIP below a defined minimum link speed, such as 256 Kbps.

- Reliability – Traditional phones are powered by Telephone Company's central office, and in case of power failure back-up generator continue to provide services for days. The VoIP phones (the modem, voice adapter etc.) are powered by users, and in case of power failure, they have to provide backup power to these equipments. The end users can augment their VoIP phones with a UPS. However, in several emergency cases, a backup power is needed to provide protection for several days. This leads to higher cost, and could make VoIP less economical compared to traditional phones (Chong and Mathews, 2004). Further, the broadband access could fail too during power outages, leading to no phone service.

- Privacy and security – The majority of VoIP customer do not encrypt the voice packets. In certain cases, voice packets may be encrypted to prevent eaves

dropping. However, this must be balanced by need for law enforcement agencies to wiretap a conversation.

- 911 Service – One of the major advantages, as noted earlier, of VoIP service that it is not tied geographically. In case of emergency, this makes it difficult to locate a caller who fails to identify his or her location. As a workaround, the service providers are registering the physical address of the service, so 911 calls can be routed to the local emergency centre. 911 dialled from a location away from home are still non-routable.

## 1.4 Motivation

While transporting voice packet over internet, several factors affect its quality. These include one-way trip delay, variation of this delay (jitter) and packet loss. These can be improved if voice packets are given priority treatment at transit, as in case of the Expedited Forwarding (EF) per hop behaviour (PHB) of Differentiated Services (DS).

Congestion at the access link is one of the common problems in VoIP systems. For example, when a user is downloading files, voice quality deteriorates, and sometimes even the connection is dropped. The primary reason for this is that there is little protection for voice traffic. Similar problem could also occur in the core of the network. These could be handled by using DS classes. The design bandwidth for the EF class carrying voice traffic is recommended to be at least double of the expected traffic (Evans and Filsfils, 2004). The performance of voice traffic in DS essentially depends on how efficient is capacity planned and monitored so that over-provisioning is respected.

Considering that the bandwidth requirement of the voice calls is only a fraction of total bandwidth in the backbone, the over-provisioning for EF class may not be an issue. However, in spite of all these considerations, voice call is affected by the following major issues among others:

1. Fault in the network - The routers lack carrier grade reliability, and often have single points of failure. To overcome this, the network designers generally plan collocated redundant routers. In case of failure, the reroute is triggered by OSPF. The problem here is that it could takes tens of seconds. After recovery routing instability could follow for minutes (Boutremans et al., 2002). Even in case of a redundant element failure within a router, the switchover can take few seconds. Obviously these are not suited for voice calls.
2. Large data packet - Even if there are no failures, certain traffic conditions could arise that may lead to unacceptable delay in voice calls. One such example is when voice packets are stuck behind large data packets in the FIFO queue of the outgoing hardware line. A larger transmit buffer there introduces large delay for voice calls, while a smaller buffer, although good to control delay in voice, leads to packet loss (Evans and Filsfils, 2004).

Few methods are proposed to overcome these problems. These can be divided into two general categories: one voice packet correction based and another routing based. In packet correction techniques such as forward error correction (Jiang and Schulzrinne, 2002; Sanneck, 2000), a portion of previous and the following packets are encoded in

every packet, and in case of loss they are used to reconstruct the lost packet. The routing based solutions normally use computation of backup path in advance, e.g., MPLS fast reroute (Wu et al., 2002). This will though require MPLS deployment in the network. Packet dispersion across multiple routes or links also falls in this category, e.g., Layer 2 link aggregation of IEEE 803.ad, where fault of a link results in quick redirection of packets to the remaining links. The routing and packet correction techniques may be combined together for added reliability.

### 1.5 Statement of the Problem

An approach similar to packet dispersion is packet duplication where the voice packets are duplicated at the source and are sent through different routes (Karol et al., 2003). This approach does not require major changes in the routers; however it needs connectivity to two separate networks at the customer's edge. At the destination only the first arriving packet is processed, and the late arriving duplicate is discarded. In this thesis, this is extended to voice calls in a DS network. The duplication provides protection against failure of a link and network or line congestion, while DS protects voice traffic against non-priority network traffic.

### 1.6 Contributions

Using the simulations, this thesis addresses the network availability and congestion for voice traffic using Differentiated Services and packet duplication. The work presented here shows that the packet duplication mechanism guards against packet loss and network

failures. This method also improves the delay budget if at least one network is lightly loaded, or has acceptable QoS provided by means of another mechanism such as DS.

### 1.7 Organization of Thesis

This thesis is organized as follows. Chapter 2 gives general background on issues affecting voice over Internet. Chapter 3 provides in-depth review of works related to improving voice traffic protection. Chapter 4 presents a simulation model to evaluate packet duplicated voice traffic in a DS network, and the results are presented in Chapter 5. Finally, a summary of work and conclusions are given in Chapter 6.

### 1.8 Summary

Voice over IP is a serious challenger to the traditional telephone network. The major benefit of VoIP is cost saving over PSTN and ability to offer new services with little cost. However, there are several issues such as network availability, congestion and reliability that must be addressed to make VoIP widely acceptable. A congested network adds large delay and loss to voice traffic. These must be minimized for acceptable voice quality. This thesis addresses the availability and congestion issue using Differentiated Services and packet duplication.

# Chapter 2
# Background on VoIP

## 2.1 Introduction

This chapter provides a general background on voice transportation over the Internet. Various voice over IP network architectures, and major components of VoIP networks are discussed. Various factors that affect voice quality are also discussed.

## 2.2 Voice over IP Networks

A simple voice over IP network is shown in Figure 2-1. IP Phones are connected to a LAN as any other computer. The LAN is connected to a wide area network. A phone digitizes voice and transmits that into the network. They are also capable of decoding the arriving voice packets. The signalling between phones is done by the SIP or H.323. Both ends are connected through an IP network. An IP phone can also make call to a traditional phone through a PSTN or media gateway. Traditional telephone services such as 411 could be obtained through this gateway. The local SIP server communicates to the SIP proxy server to setup the call. This enables IP phones not to be tied a fixed location.

There are several variants of a VoIP network configuration. Selection of a configuration depends primarily on the cost a user is willing to pay, and reliability he or she can live with.

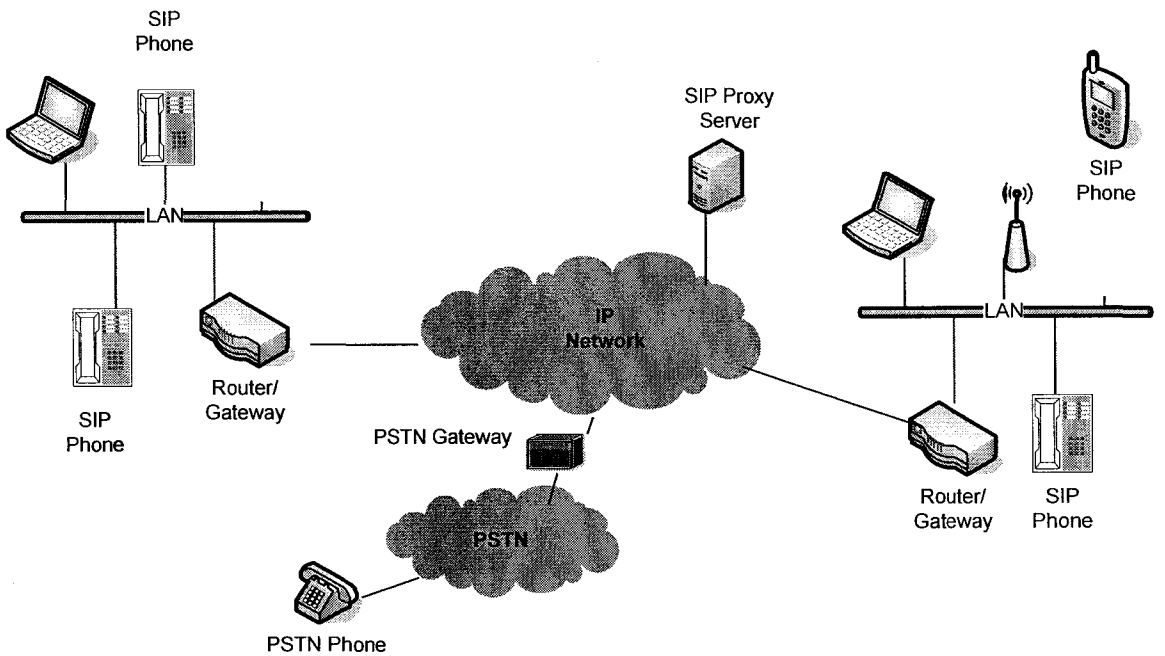


Figure 2-1: Architecture of an end-to-end IP phone network.

The means to provide VoIP services can be broadly divided into three categories, one soft phone based where voice packets are processed over a general purpose computer, the second a telephone adapter placed between modem and a telephone. The third system is central office aggregation based, where the end user continues to use the regular telephone line as before, only the core is IP based. VoIP Info (2006) provides a comprehensive list of resources on voice over internet protocol. In the following sub-sections, some of the common VoIP architectures are discussed.

### 2.2.1 Soft phones

The soft phone is software simulation of telephone on a computer. This technology uses a headset connected to the sound card of the PC or USB phone. Such phones are very popular for computer to computer calls, and also in a call centre environment. Microsoft Windows NetMeeting can provide soft phone service using SIP, and so do several other instant messaging software. There is no extra hardware other than a headset (even computer's speaker and microphones can be used) is required for this service. The voice is digitized and encoded by the phone software running on the computer and sent over the destination, where other computer with compatible phone software decodes and plays the voice. This service is often free as long as calls are made from a computer to another computer, on the Internet (Figure 2-2). The VoIP service offered by Skype (2006) is an example of a soft phone.

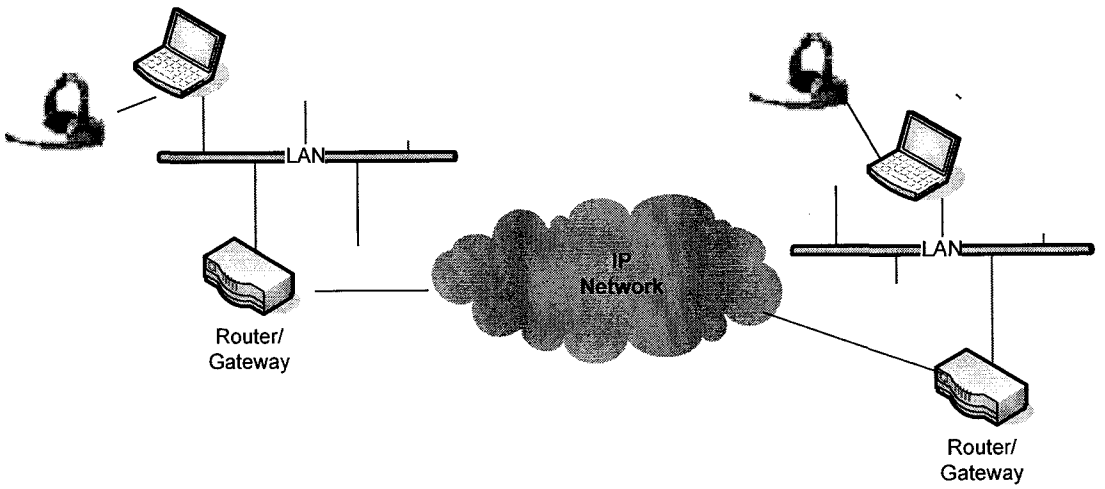


Figure 2-2: A soft phone network

A soft phone service could be extended to a PSTN network as shown in Figure 2-1 using a media gateway. This requires a service provider who can provide connectivity to PSTN, but this has cost associated to it. The gateway resides on the edge on the network and interfaces time division multiplexing on PSTN and packet processing in IP network. This translates between transmission formats and procedures that are used on PSTN on one side to the IP on the other, and vice versa.

The major advantage of soft phone service is its cost, which is often free, however the major drawbacks of the free service is that the parties must have their computer on at the time the call is made.

### 2.2.2 Hard Phones

The incumbent VoIP service providers such as Vonage and cable operators advocate maximum use of IP network to provide phone service, and use the traditional phone network only if IP network is not available at the last mile. This configuration is also used in Bell Canada's (2006) Digital Voice Lite service. This network uses an analog telephone adapter placed between a regular phone and the modem (Figure 2-3). This gateway digitizes, encodes and packetizes the voice for transmit, and buffers, decodes and playback the received voice. The adapter also works as a SIP user agent and communicates to the SIP proxy server located in the service provider's network to setup a call. Another variant of this service is hard phones that integrate the adapter function into a telephone system. These IP phones connect directly in the Ethernet plug as shown in Figure 2-1.

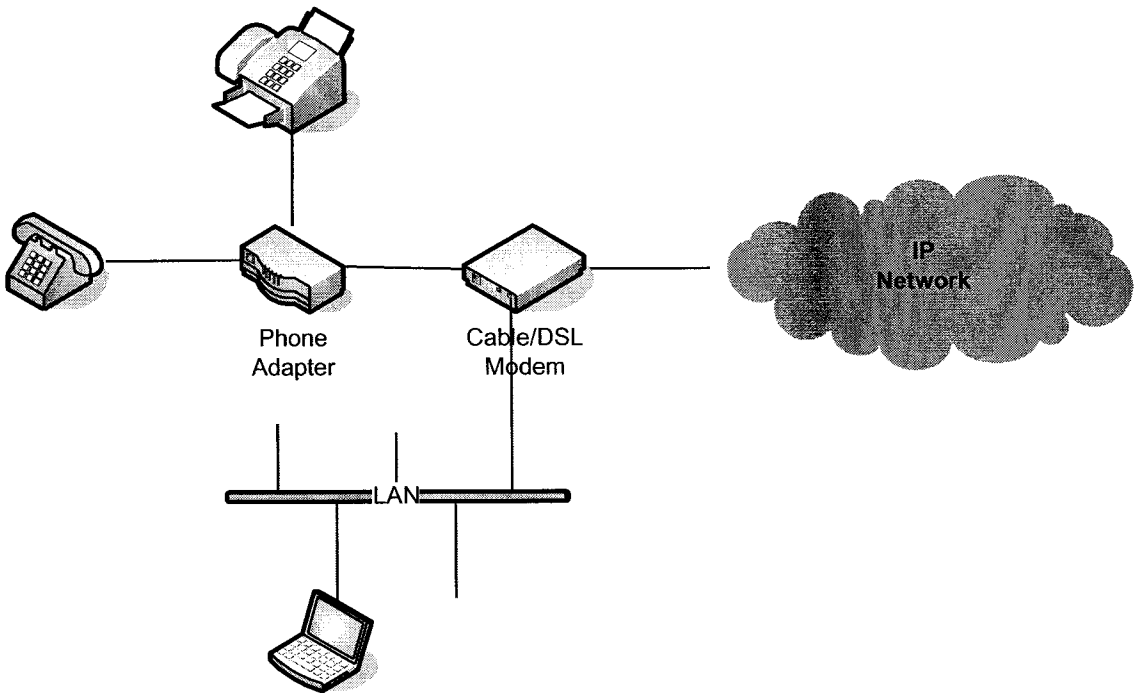


Figure 2-3: Voice-adapter based VoIP configuration (Vonage, 2006).

This cost associated with this service is similar to the regular phone line, but the providers pitch this service for added features such caller ID, call forwarding etc., with little or no additional cost. The service provided in this VoIP model does not require users to be tied up on the computer as in case of the soft phones. Further, this service provides phone number portability. The VoIP packets are normally routed through the private networks of service providers, where network dynamics can be better controlled. The mobility comes from portability of the phone number that is often programmed in the phone adapter. The user can carry the voice adapter while traveling, and can continue to use the phone as at the home location.

There are some important issues that affect this type of service. One, there is little or no control over managing bandwidth for voice packets over the access network. If the access links are congested, voice quality could be severely affected. For example, if the user is downloading a file while on phone, voice quality deteriorates, and sometimes the connection is dropped. The second major problem with this service is that the user has to provide backup power to the modem, the adapter and the telephone system. The 911 service are now offered, but it has limitation when the service is used away from home.

### 2.2.3 Traditional Phones with Central Office Based Aggregation

In the central office (CO) aggregation model, voice is carried, as in traditional telephone, to the phone company's CO. The A/D and D/A conversion, encoding/decoding are performed near the core of the IP network. No new hardware is needed (not even a high

speed connection is needed) for the user, and he or she continues to use the telephone system as before (Figure 2-4), and may not even know that Telco is using VoIP.

A VoIP service based on this configuration overcomes some of the major issues, e.g., the access congestion and the power supply, discussed for the phone adapter model (Section 2.2.2). As a traditional phone has a dedicated line to the CO, the access congestions do not exits. The legacy telephone companies have invested heavily on backup power to power the traditional phones, and can continue to use this for new service. This service is preferable to legacy telephone service providers. Bell Canada (2006) offers such service as Digital Voice.

The end user still has to maintain the voice and data network separately, and the telephone companies the aging access voice switches. For these reasons, although this service provides better migration path, long term future may not be very promising. Further, because the service is not end-to-end IP, additional features offering could be limited.

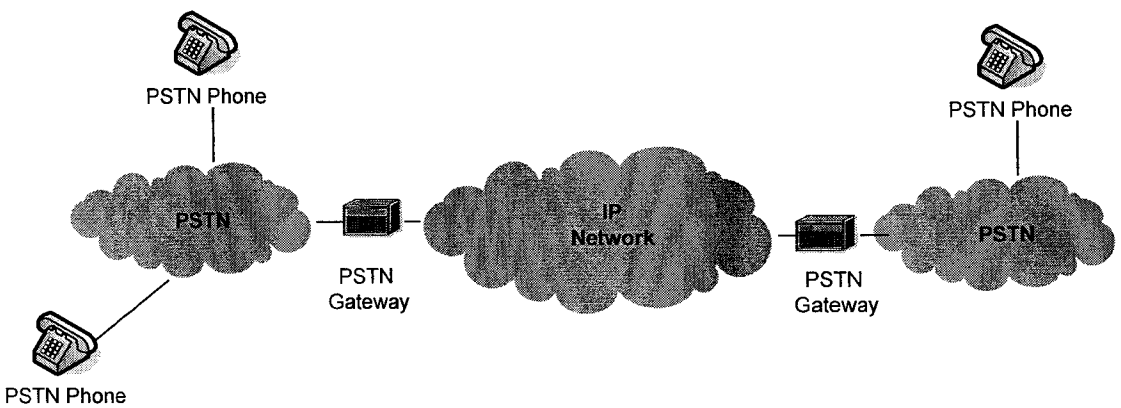


Figure 2-4: PSTN based VoIP network

## 2.3 Signalling and Call Control

A signalling between the parties is used to set a VoIP call. During the call setup, parties exchange encoding parameters. When the parties agree on how to communicate, the voice packets start to flow. There are two common signally protocols: session initiation protocol (SIP) and H.323 (Goode, 2002). Both of these protocols are peer to peer, and run at the gateway to IP network, e.g., used by computers in soft phones, or by voice adapters and hard phones. The signalling protocols performs address translation (translate a phone number to an IP address), admission control, directory services, etc.

Unlike SIP, H.323 was not specifically designed for VoIP, but because of its large installed base, it continued to be used in the core of the network. SIP is very much like HTTP, the Web protocol, or SMTP. Messages consist of headers and a message body. SIP is an application-layer control protocol that can establish, modify, and terminate multimedia sessions including telephony calls. SIP transparently supports name mapping and redirection services, where users can maintain a single externally visible identifier regardless of their network location (Rosenberg, 2002). SIP does not use IP address for identification but it uses URL to identify the logical destination such as an e-mail or phone number.

## 2.4 Components of a VoIP system

The components of a typical VoIP system are shown in Figure 2-5. A phone conversation from the encoder point of view consists of a talk spurt followed by a silence when the user is listening to the other party.

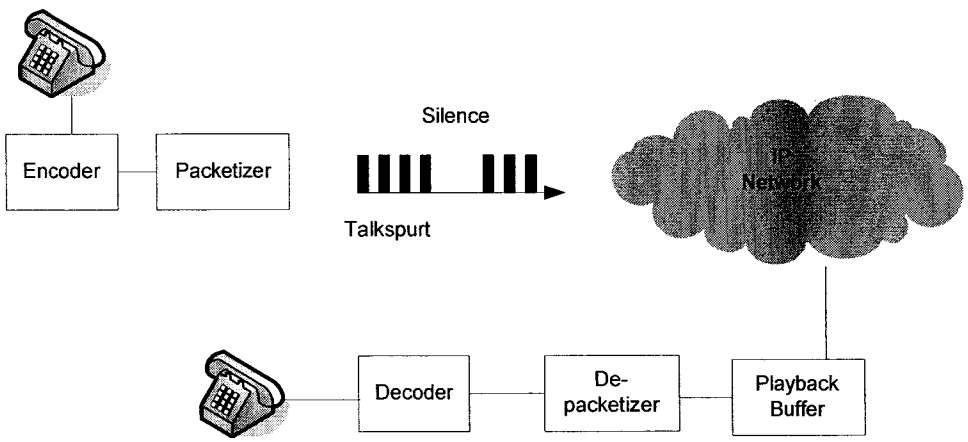


Figure 2-5: Components of a typical VoIP system.

The voice is first sampled by an encoder into a fixed number of bits. The data is compressed to eliminate the silence period, which is added back at the receiver. The voice frame are encapsulated with various headers and transported over an IP network.

At the receiver end, the packets are placed in a jitter buffer. The jitter buffer (also referred layback buffer) removes the variations in the packet arrival time by converting variable delay into a constant delay at the voice destination. Packets are held in the buffer until there are enough packets to play them smoothly.

The received packet is passed to the de-packetizer, which strips of the Layer 3 and 4 headers, and then to the decoder for reconstruction of voice. The decoder adds the silence, if suppressed at the encoder. It can also implements loss concealment algorithms.

### 2.4.1 Voice Codecs

Traditional encoder G.711 creates 8 bit samples at 8000 Hz, creating a data rate of 64 Kbps. The typical data rate varies from 5.3 to 64 kbps depending on the encoding technique. Some of the common encoding schemes are shown in Table 2-1.

In an enterprise network, where a bandwidth between hosts are normally abundant, PCM based encoding is preferable because of its superior voice quality (Takahashi and Yoshino, 2004). However, in a public network, compressed voice, such those produced by G.729a are preferable because of its lower bandwidth requirement.

Table 2-1: Commonly used voice Codec

| *Encoder* | *Data Rate* | *MOS* | *Comments* |
|---|---|---|---|
| G.711 | 64 kbps | 4.1 | Pulse code modulation (PCM) based with voice sampled in 8-bits. Universally used in T-1 carrier system. |
| G.726 | 32 kbps | 3.85 | Referred as Adoptive Differential PCM (ADPCM). 8 bit PCM samples mapped in 4 bits samples. |
| G.729a | 8 kbps | 3.92 | CS-ACELP algorithm. Compresses 10ms of vice and 5ms of look ahead, causing algorithm delay of 15ms. Widespread use in VoIP and cellular network. |

G.729 uses voice activity detection to distinguish between talk-spurt ("on") and silence ("off") as shown in Figure 2-5. The on-off pattern allows for higher bandwidth utilization through multiplexing. The voice activity detection (VAD) can also be used for echo suppression. The "on" and "off" periods are normally exponentially distributed (Jiang and Schulzrinne, 2000). When VAD switches from active (receiving) to inactive (transmitting) mode, hangover frames are sent to enhance the quality of the detection. The distribution of talk-spurt and silence depends on the hangover time. VADs delay the decision for silence by hangover duration to avoid clipping at the end of a speech segment. Voice quality improves with an increase in the hangover time, but again this reduces benefit of multiplexing. If the hangover period is small, the mean talk-spurt is around 200 to 400 ms, and mean silence is around 500 to 700 ms.

### 2.4.2 Voice Packets

The voice bit streams are placed in packets, and encapsulated by Real Time Protocol (RTP) header. The RTP header has a timestamp of the source and a sequence number. This information is used by the intermediate nodes to decide whether the packet needs to be transported further, and by the receiver to construct voice. The late packets cannot be used to re-create voice, and are dropped by the routers. The RTP packet streams are transported over User Datagram Protocol (UDP), and it is placed in IP packets. The overhead of RTP (12 bytes), UDP (8 bytes) and IP (20 bytes) adds 40 bytes to a voice packet. Usually 5 to 48 ms of one or many voice frames are encoded into one packet. A 5 ms voice frame with G.711 encoder is 40 bytes long and a 48 ms frame is 384 byte.

There is trade-off between small voice frames that have large headers, resulting in inefficient bandwidth utilization, and the large frames that cause greater delay in the network. Considering that RTP/UDP/IP header is as large as the payload of a small voice frame, to improve the efficiency, several header compression algorithms are sometimes used to compress these header to 2-4 bytes (Degermark et al., 1999).

## 2.5 Quality of Voice over IP

The quality of voice packet is affected by several factors. These include distortion caused by choice of codec, echo, one-way trip delay, variation of this delay and packet loss. In a congested network, delay or latency of voice packets increases. Packet loss occurs inside the network when queues of switch and routers overflow as a result of congestion. These variables can be controlled by network QoS.

The criteria for assessment of voice are subjective and are measured by mean opinion score, and terms such as excellent, good, fair, poor and bad are used. The factors that normally affect these scores are environmental noise and channel degradation caused by packet loss. ITU-T has proposed more scientific E-Model (G.107, 1998) based on psychological scale called "Mean Opinion Score" (MOS) as shown in Table 2-2. The operational range for PSTN voice corresponds to MOS more than 3.6. Table 2-1 shows MOS for various codecs. G.711 (1999) has best score, but G.729a has better multiplexing benefit for the comparable MOS.

Table 2-2: Voice quality classes (G.107, 1998)

| *User Satisfaction* | *MOS* | *Comments* |
|---|---|---|
| Very satisfied | 4.3 – 4.5 | Desirable |
| Satisfied | 4.0 – 4.3 | |
| Some users dissatisfied | 3.6 – 4.0 | Acceptable |
| Many users dissatisfied | 3.1 – 3.6 | Not acceptable for toll quality |
| Nearly all users dissatisfied | 2.6 – 3.1 | |
| Not recommended | 1.0 – 2.6 | |

### 2.5.1 Delay Budget

The end-to-end delay is a direct measure of VoIP performance. A large delay may not be very crucial for the passive listener (e.g. streaming voice or music), but because in conversation both parties speak, a large delay makes is very difficult to coordinate who will speak next. G.114 (2003) recommends that one way end-to-end delay must be less

than 150 ms to maintain the quality of voice signal as shown in Table 2-3. To put this into prospective, Goode (2002) presented a sample delay budget as shown in Table 2-4.

Table 2-3: Delay requirements as per G.114 (2003)

| ***Delay Range*** | ***Description*** |
|---|---|
| 0 – 150 | Acceptable for most user applications. |
| 150 – 400 | Acceptable provided that administrators are aware of the transmission time and the impact it has on the transmission quality of user applications. |
| > 400 | Unacceptable for general network planning purposes. However, it is recognized that in some exceptional cases this limit is exceeded. |

Table 2-4: A sample delay budget for G.729 encoded voice over internet protocol as presented in Goode (2002).

| ***Delay source (G.729)*** | ***On-net Budget (ms)*** |
|---|---|
| Device sample capture | 0.1 |
| Encoding delay (algorithm delay + processing delay) | 17.5 |
| Packetization /De-packetization delay | 20 |
| Move to output queue/queue delay | 0.5 |
| Access (up) link transmission delay | 10 |
| Backbone network transmission delay | Variable |
| Access (down) link transmission delay | 10 |
| Input queue to application | 0.5 |
| Jitter Buffer | 60 |
| Decoder processing delay | 2 |
| Device play out delay | 0.5 |
| **Total** | 121.1 + variable |

Major components of the delay are:

- Transmission and Forwarding Delays – This is the time the physical signals need to travel across the links along the path taken by the data packets. The forwarding delay occurs in the network layer, and it is the time the router takes to forward a packet from the input to the output port. These delay depend on the outgoing link's speed and congestion in routers.
- Packetization /De-packetization Delays - The time needed to build data packets at the source and as well as to strip of packet headers at the destination. This includes wait for the arrival of sufficient data from the application to form a packet at the source.
- Codec Delay – This is the time needed to perform encoding at the source and decoding at the destination. At source it includes time to digitize speech signals and perform voice encoding. At the destination it includes conversion of digital data into analog signals (see Section 2.4.1).
- Jitter Delay - Although the packets at the sender are periodically generated, they incur random delays while traversing the network. To smooth out such jitter, receiver buffers the arriving packets and delays the playout of received packets until it has enough packets to play. If packets are held for too short a time, variations in delay can potentially cause the buffer to under-run and cause gaps in the speech. If the sample is held for too long, the buffer can overrun, and the dropped packets again cause gaps in the speech.

Figure 2-6 shows a staircase representation of voice generation and playout. The uneven staircases in Figure 2-6 for the receiver are the result of network delays. If the receiver delays the beginning of playout until t2, all packets would have been received by the time their playout is scheduled. However, this would require a larger buffer and will result in large payout delay. However, if the playout begins at time t1, smaller buffer is needed and there is a shorter playout delay, but some packets will be lost at the receiver. This illustrates a trade-off between the jitter delay that voice application is willing to tolerate and the packet loss suffered as a result of the late packets.

The variable backbone delay as shown in Table 2-4 includes propagation delay in the transmission medium and queuing delay at each hop. This delay varies from a few ms in a local or national network to well over 100 ms for overseas networks. The access link transmission delay could be higher than 10 ms depending on the bandwidth available there (Mehmood, 2005). A simple tool such as ping, or trace route can be used to determine the network delays.

Table 2-4 shows that the network transmission delay and delay created by the jitter buffer are two major delay components. These must be controlled for voice traffic to be within ITU-T's recommended limit of 150 ms.

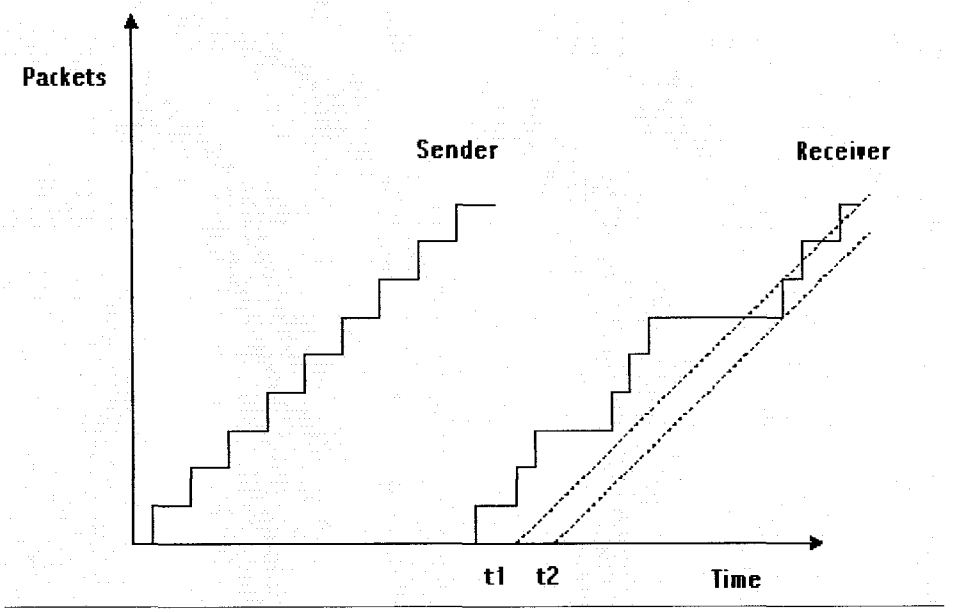


Figure 2-6: Generation and reconstruction of voice

### 2.5.2 Packet Loss

Packet loss causes voice clippings. This is not only irritating to the communicating users, but sometime could be assumed by them that their call is dropped. Some codec algorithms use forward-error-correction, where every voice packet contains portions of the previous and following packet. In this case loss of one packet is concealed, but loss of several packets in a row results in noticeable degradation in voice quality. Backbone engineered for high-quality VoIP services typical target loss is 0.25 percent or less (Evans and Filsfils, 2004). For VoIP systems G.114 (2003) recommends that the packet loss must remain less than 5%.

### 2.5.3 Network QoS

Differentiated Services (DS) is aimed to provide scalable service differentiation in the Internet that can be used to permit differentiated pricing of internet service. This also offers framework within which service providers can offer a range of network services including real-time applications.

Simplicity is the main architecture of DS that sets it apart from other QoS schemes such as IntServ and MPLS. IntServ has serious scalability problem in the core. This is because every router has to have the knowledge of service contract for each flow, and there could be thousands of such flows converging in the core of the network. MPLS handles QoS well, but it is essentially a connection-oriented approach, and will require major network upgrades to handle MPLS labelled packets. DS does not require major

upgrade to the network, as most current routers can handle multiple service queues that are required in a DS network.

## 2.6 Privacy and Security

Voice traffic has similar threat as data traffic. With increasing use of this in businesses, there is a need to protect this traffic from eves dropping and attacks. Voice traffic could be subjected to the attacks such as denial of service, call hijacking and spoofing, call tracking, eavesdropping and snooping, replay attack, man-in-the middle attack and toll fraud (Thaanthry et al, 2005). Voice traffic could be encrypted similar to data traffic. The common security mechanism used in voice traffic is virtual private network, and end-to-end address translation and encryption (Thaanthry et al, 2005). The two parties could be connected through Layer 2 network, which provides basic privacy. In the end-to-end encryption, the parties exchange secret keys which they will be using to encrypt the voice traffic. It should be noted that encryption and decryption adds additional delay.

The law enforcement agencies may need to hear conversation originating or terminated at a particular phone. With encrypted voice, it becomes harder to do this job that is fairly simple in PSTN phones. Because of this voice over IP is subjected to government regulations against encryption.

## 2.7 Wireless VoIP

Wireless LANs are now very popular for data networks. These networks can be used for VoIP calls to provide caller mobility, and can be very cost effective alternative to cellular phone in large enterprise campuses. VoIP runs over a Wireless LANs, which is typically

compliant with the 802.11 standards. As long as callers are within range of a WLAN access point and using a VoIP enabled handset, they can make and receive calls over the wireless network.

Scalability is the one of the major concern, and it can be observed from Table 2-5 (Garg and Kapes, 2003) that with an 802.11b access point, only six simultaneous calls can be placed. Adding more calls results is rapid deterioration of all calls. Considering that 802.11b has 11 Mbps peak rate and fully duplex G.711 voice requires only 128 Kbps, it would suggest that 85 simultaneous calls be made. This low utilization is caused by inherent channel inefficiency of 802.11b for small voice packets. As seen in Table 2-5, by increasing the voice size, more simultaneous calls can be made. However, this leads to larger delay in the multi-hop networks.

Table 2-5: Maximum number of VoIP connections on 802.11b (Garg and Kapes, 2003)

| *Audio (ms)* | *G.711 codec* | *G.729 codec* |
|---|---|---|
| 10 | 6 | 7 |
| 20 | 12 | 14 |
| 30 | 17 | 21 |
| 100 | 39 | 66 |

Other challenges are lack of QoS in wireless networks, and complications in implementing and maintaining a Wireless VoIP.

## 2.8 Summary

This chapter provided a background on VoIP and general description of network configuration and various VoIP components. The factors that affect voice quality were discussed. The next chapter provides a review of voice protection methods.

# Chapter 3
# Voice Packet Protection

## 3.1 Introduction

As discussed in earlier chapters, late voice packets may not be used to reconstruct voice at the destination, and are discarded. For this reason, as long as the voice traffic is below their service level agreement, it must be protected from best-effort traffic. Further, in case of network fault, the voice traffic must be given highest priority for recovery. This chapter discusses various protection mechanisms applicable for voice traffic.

The backbone networks are often over-provisioned because the excess bandwidth may be inexpensive to add than to run complex management schemes on the routers (Goode, 2002). Over-provisioning does solve the delay issues related to voice traffic, but it does not provide protection to this traffic in case of failures such as router fault, or physical damage to the links.

## 3.2 SONET Layer Protection

The protection at SONET layer is a mechanism by which traffic is switched to the available resources when failure occurs. The protection at this layer involves redundant capacity to reroute traffic in case of failure. The protection around the failed link can be done at different points in the network, e.g., at the line layer (between two endpoints of the failed link) or at the path layer (between the source and the destination). The line

layer protection is simpler to implement, but the path layer protection requires less bandwidth (Thiran et al., 2001).

There are two protection mechanisms, one 1+1, and the other 1:1. In 1+1 protection traffic is transmitted simultaneously on two separate fibres. The far end accepts traffic from the primary fibre, and switches to the backup in case of primary failure. In 1:1 protection, traffic is sent only on the primary fibre, and in case of failure, the receiver signals on the backup fibre to the sender to make a switch. This protection can be further extended to 1:N, where one link provides protection to N links.

The biggest advantage of SONET layer protection is fast switch to the protection link; within less than 50ms (Thiran et al., 2001). This makes it an ideal choice for voice packet transportation. The drawback is cost of providing duplicate links and the port in the network.

### 3.3 IP Layer Protection

A link failure typically appears as a period of consecutive packet loss that can last for many seconds, followed by a change in delay after the link is re-established. Despite careful IP route protection, link failures can significantly impact an IP voice service (Boutremans et al., 2002).

IP networks were initially designed to carry best-effort data traffic, which can tolerate short disruptions. Due to the inherent lack of carrier grade reliability, network designers ensure at least dual physically diverse path when constructing the topology. With this configuration, IP traffic is self healing. The routers continuously learn and update their

routing table. In several routing protocols, e.g., Open Shortest Path First (OSPF), router exchange hallo packets with their neighbours to determine the health of the links. If these messages are lost, the backup routes are re-computed by sending discovery messages. All this normally does not require additional hardware because networks are well meshed. The problem here is that it could takes tens of seconds (Johnson, 2004). After recovery routing instability could follow for minutes (Boutremans et al., 2002). Even in case of a redundant element failure within a router, the switchover can take few seconds. Obviously these are not suited for voice calls.

The routers often have single points of failure. The major causes of quality degradation are link and router failures. With redundancy in place, the downtime of routers has greatly improved. However, the problems are still present in the following area (Johnson, 2004):

- The chassis failure of routers still remains as single point of failure causing a downtime up to 10 minutes required to restore the routers.
- Several software and hardware upgrades in year are typical and contribute 10 - to 60 minutes of down time.
- Denial of service (DOS) attacks may cause long outages.

The reliability of routing equipment represents a major obstacle to the introduction of VoIP services. One could use a combination of SONET level and IP level protection. The traffic requiring protection is forwarded over redundant link, and, the regular best effort traffic use IP level protection. There are several variant to this design, e.g. only the high

priority traffic is restored fully in case of failure, while leaving out the best effort IP traffic. This can be implemented using Differentiated Services (DS), or MPLS.

Another drawback for IP networks with respect to voice traffic is that packets are normally routed over shortest path, with little consideration given to link utilization. The main problem in this approach is that some links could be congested by traffic which can be efficiently carried over underutilized links.

One of the simplest approach to achieve protection against best-effort traffic is over-provisioning of the network. In an over-provisioned network, there is little use of sophisticated QoS mechanism. If the occupancy is low, the performance of voice will be good. In the core of network, which has high bandwidth links and where traffic is highly aggregated, over-provisioning is viable solution (Evans and Filsfils, 2005). However, access links that provides connectivity between customer-edge (CE) routers and the provider-edge (PE) router, is likely to be weak section in terms of protection. This is because the customer determines access-link bandwidth, and to minimize networking costs, customers often delay upgrading these links for as long as possible. Consequently, access links are often under-provisioned, resulting in congestion. For these reasons, protection of voice in access links is essential at the network edge (Evans and Filsfils, 2004).

### 3.4 Differentiated Services

As noted earlier, the network transmission delay and delay in jitter buffer can be controlled by network QoS. This can be achieved by the premium class of differentiated

services. In a differentiated services (DS) network, QoS is achieved by policing the classified packets at the edge of the network, while MPLS achieves QoS based on traffic-engineered switch path.

DS supports differentiated and assured delay, jitter, and loss commitments on the same IP network for different service types or classes of traffic. DS is the preferred technology for scalable IP quality-of-service deployments today; it achieves scalability by performing all complex QoS functions, such as per-flow traffic classification, marking, metering, and conditioning, at the network's edge with a relatively simple subset of functionality required in the core. At the edge, DS classifies traffic into a limited number of traffic aggregates or classes and then colors or marks them using the DS code point (DSCP) field in the IP packet header to identify the class (Nichols et al., 1998). These aggregates are checked for conformance against agreed profiles and are conditioned using shaping, in which packets are delayed until they are conformant, or policing, in which non-conformant packets are dropped or re-coloured. To ensure per-class SLA differentiation, DS applies scheduling and queuing control mechanisms to the traffic classes on the basis of these DSCP markings.

In a DS network, the traffic is classified into a premium expedited forwarding (EF) class, several assured forwarding (AF) classes and standard best effort (BF) class. This classification is done at the edge of the network based on service level agreement (SLA) with the end user (Blake et al., 1998). When these packets are inside the DS network, they receive differential treatment. Packets from higher class have preferential treatment

for higher throughput. The service provider (SP) monitors and enforces the SLA at the customer edge (CE) interface.

Working group for DS at IETF has defined Assured Forwarding (AF) as one of the Per Hob Behaviour (PHB) policy. PHB includes policing, shaping, code point marking and enqueueing to provide "differential treatment" at the network output (Blake at al., 1998). AF has four traffic priorities (gold, silver, bronze and best effort) and three-drop precedence.

### 3.4.1 Traffic Conditioning

In order to deliver service agreements, each DS enabled edge router implements Traffic Conditioning function, which performs the functions as shown in Figure 3-1. The classifier measures the temporal property of a traffic stream. The meter monitors the traffic pattern of each flow against the traffic profile. For out-of-profile traffic, the metering function interacts with other components to either re-mark, or drops the traffic for that flow. The marker uses TOS field of IPv4 header to mark the behaviour of the traffic. Customers request a specific performance level on a packet by packet basis, by marking the DSCP field of each packet with a specific value. This value specifies the Per-hop Behaviour (PHB) to be allocated to the packet within the provider's network.

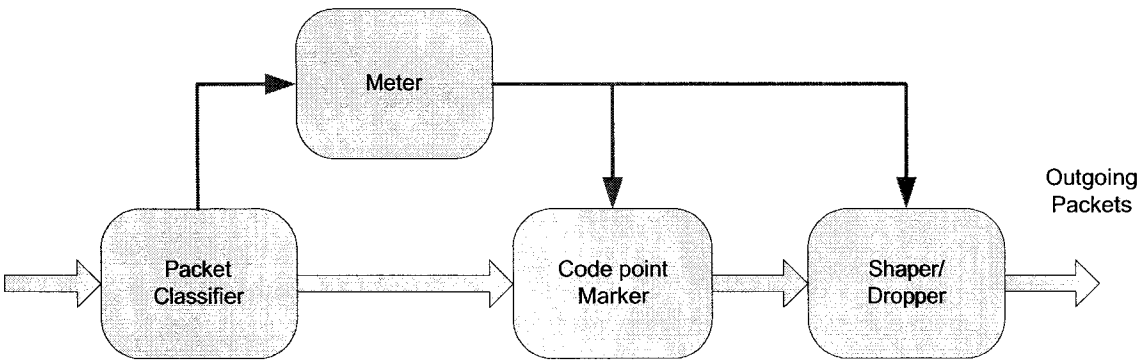


Figure 3-1: Traffic conditioning at an edge router for a DS network

The traffic aggregates are policed according to the SLA. The out-of-profile traffic is either dropped at the edge or is remarked with a different PHB. The routers control the forwarding rate of packets so that flow does not exceed the traffic rate specified by its profile. The shapers ensure fairness between flows that maps to the same class of service, and controls the traffic flow to avoid congestion. When congestion occurs the dropper drops packets based on specific rule. The dropper works on a scheme similar to the RED (Floyd and Jacobson, 1999).

### 3.4.2 RED and RIO Schemes

In case of congestion, IP packets are discarded according to its drop precedence. The most popular mechanism used today to drop packets is Random Early Detection (RED) algorithm (Floyd and Jacobson, 1999). Figure 3-2 shows RED algorithm. RED is an active congestion avoidance mechanism that takes advantage of TCP's congestion control mechanism. RED detects congestion by estimating the average queue size. If the average value exceeds the minimum threshold, RED begins dropping packets based on predefined probability curve. TCP reacts to the packet drops by slowing down the transmission. The drop probability increases as the queue length increases, and once a maximum threshold queue length is reached all packets are dropped.

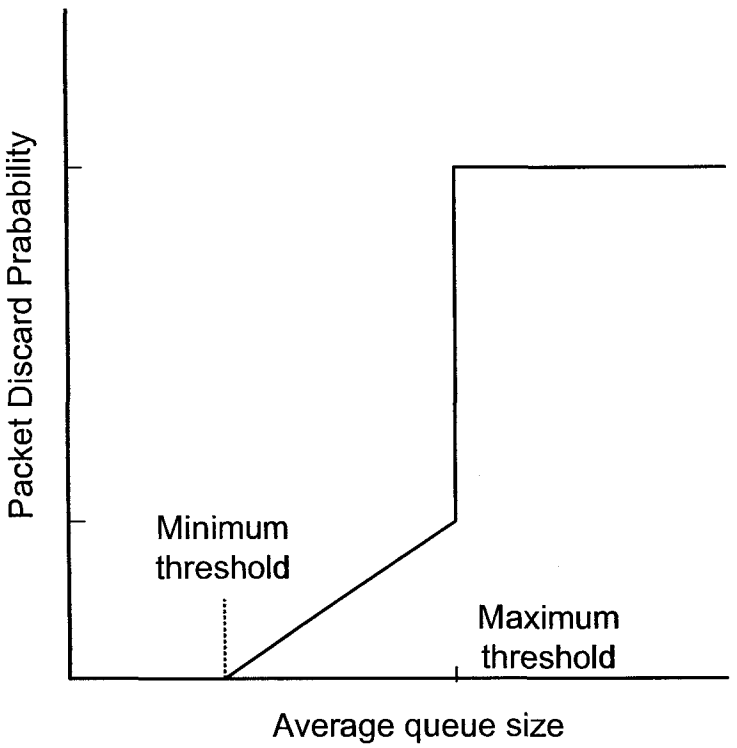


Figure 3-2: Random early detection algorithm

One of the early algorithms for DS is RIO (RED with "in" and "out") (Clark and Fang, 1998). RIO is essentially two RED. RIO has two drop precedence, one marked "in" (packets that follow a SLA) and another marked "out" (not following SLA). The "in" packets can be considered part of reserved flows and "out" packets can be considered part of "best-effort" flows. The "out" curve has lower minimum threshold than "in" curve. Because of this, under low level of congestion, RED will drop only packet marked "out". If congestion is serious and queue length exceeds minimum threshold of "in" packets, RED will start to drop "in" packets as well.

RIO algorithm can be generalized for more than two drop-probability curves for AF forwarding classes. This approach is referred as weighted RED (WRED).

### 3.4.3 Voice over Differentiated Services

Differentiated Services networks are intended to aggregate TCP traffic and real-time traffic. Most real-time applications such as telephony do not require a reliable transport protocol. This traffic runs over UDP and is non-responsive to congestion avoidance support provided by RED type algorithms. Ideally, a non-responsive source should not take more bandwidth than a responsive flow. Consider an aggregate traffic, where both TCP and UDP are given equal treatment for excess bandwidth. The TCP flows will reduce their rate when congestion is detected, whereas the UDP flows will not change and eventually starve the TCP flows (Fang, 2000). Best Effort TCP flows gets much less throughput compared to the unreserved UDP flow. However the reserved UDP flow rate is not affected.

The gold traffic of AF can be used for voice. Alternatively, voice is often recommended to be used with EF class where it gets best treatment and achieves low-delay, low-jitter and low-latency and does not get affected by traffic in other classes. This class is best suited for voice traffic (Fang, 2000; Evans and Filsfils, 2004). The EF class (Jacobson et al., 1999) is served by a priority scheduler, which serves it at line rate till this becomes empty. This ensures no loss, latency, or jitter. The lower priority queues are served only when there are no packets in the EF queue. To prevent starvation of non-EF queues, the traffic in EF class must be shaped. Normally, the bandwidth allocated in EF class is over-provisioned, i.e., the aggregates arrival rate is always less than the departure rate. SLA with the user specifies a peak rate. The SP guarantees the contract as long as the aggregate traffic is below the peak rate; packets violating the SLA are dropped at the edge of the network.

### 3.5 Voice over Label Switch Paths

Voice packets need bandwidth guaranties in the network. Because MPLS is path-oriented it can potentially provide faster and more predictable protection and restoration capabilities in the face of topology changes than conventional hop by hop routed IP systems. MPLS adds traffic engineering capabilities in IP networks, and can be used to protect voice traffic. In an MPLS domain, when a stream of data traverses a common path, a Label Switched Path (LSP) which can be established using MPLS signalling protocols. Some signalling protocol, such as RSVP-TE and CR-LDP create LSPs based on the requested traffic parameters (Awduche et al., 2001; Wu et al., 2002) such as

bandwidth and service category suitable for voice. At the ingress Label Switch Router (LSR), each packet is assigned a label and is transmitted downstream. At each LSR along the LSP, the label is used to forward the packet to the next hop.

The traffic guaranties can be achieved by combining MPLS with differentiated services. The two technologies work in different layers, MPLS is a later 2/3 based while DS is a layer 3 technology. For every DS code point, a separate LSP is established, with voice traffic taking the highest priority class (Wu et al., 2002). The experimental field of MPLS is used to indicate routers that this packet should be given high priority. Because the MPLS label has only 3 experimental bits that came from the old 3-bits precedence in the IP header, it can address up to 8 possible PHBs. The 64 possible code points of DS must be mapped into these 8 PHBs. There are two methods that are used to convey DS information to the LSRs, one E-LSP and the other L-LSP. E-LSP (Wu et al., 2002).

MPLS produces good performance for voice as along as the traffic from voice sources is less than LSP trunk capacity (Al-Irhayim, et al., 2000). An under-provisioned network is not desirable to carry voice. Further best effort traffic has no effect on the steady-state performance of voice traffic protected in LSPs.

Additional protection to voice traffic can be provided by creating a backup LSP. In event that the primary LSP fails, traffic can be forwarded on the backup path.

### 3.6 Multi-Homed Network

Many enterprises have connectivity to multiple service providers to increase reliability in their network. Such multi-homed network can be used to provide protection to voice traffic, along with other high priority traffic. This arrangement, although expensive, provides high availability of service. The secondary networks are used for load balancing to disperse the traffic and provide backup to the primary network.

In traditional IP networks, packets from a session normally follow a single network. In multi-homed networks, the packets for a session are dispersed over multiple networks (Figure 3-3). The dispersion can be implemented by the source application, or by the access node in the network. The parallel paths can be constructed using MPLS LSPs, or using source routing, or constructing parallel static routes in the IP network.

Although best results are obtained when all paths have at least two link, but its benefit is seen by updating a bottleneck link to the destination (Zlatokrilov and Levy, 2004). The dispersion rules account for the congestion. In this case, fewer or no packets are routed on these links, but they are sent over the other link. This acts to balance the load on the links. The packet dispersing devices are located on the path between the sender and the receiver and may take automatic dispersion decisions based on current network conditions or base on a-priori knowledge gathered by network management elements (Zlatokrilov and Levy, 2004).

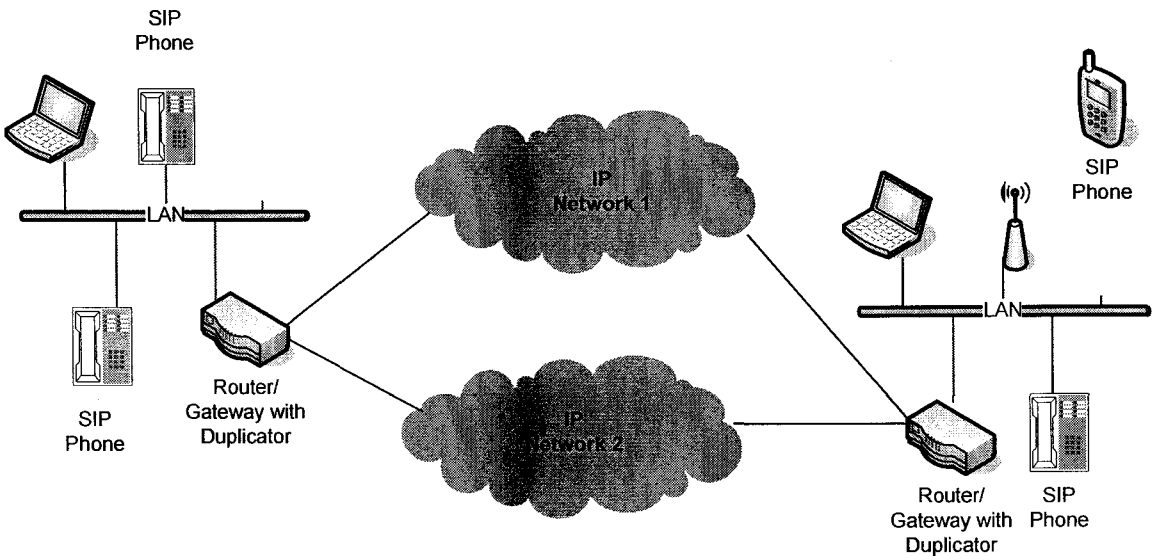


Figure 3-3: Architecture of an IP phone network with voice packet duplication for ingress and selection for egress functions added to the CPE Router.

To preserve the ordering of packets, the load balancing should be at a per-flow rather than per-packet level (Evans and Filsfils, 2004). With packet-level load balancing, different packets within the same flow might follow different paths, which have different delays and can cause re-ordering. Similarly, scheduling algorithms should also ensure that packets from the same flow are serviced in order and from the same queue. For VoIP, the differential delay between the routes have to be larger than inter packet interval (depends on codec, but of the order of 20ms) to cause any concern. Further, packet based load balancing can also increase the jitter experienced in a flow.

Markopoulou et al. (2002) have shown that user dissatisfaction increases with perceptual bursty losses. For voice traffic, traffic dispersion could be used an effective mechanism to reduce losses, thus need for expensive methods such as forward error correction (FEC) can be avoided (Jiang and Schulzrine, 2002).

### 3.7 Voice Packet Protection by Duplication

Despite IP protection of voice traffic by means of MPSL or DS, there could still be a need to protect this traffic against a fault in the network, such a links failure, or router fault. A recovery from these conditions could take several seconds, and may not be acceptable to some customers. The packet duplication method can be used to improve this (Karol et al., 2003).

#### 3.7.1 Packet Duplication Architecture

This method of protection is quite similar to the packet dispersion method used in the multi-homed networks as discussed in section 3.6. Leveraging on such networks, major

improvement in voice quality can be achieved by sending duplicate voice packets over multiple networks (Karol et al., 2003). Figure 3-3 shows a modified network of Figure 2-1 for duplicate voice.

At the ingress of the network, the duplicator sends the voice packets to two separate service provider's networks. Both service providers deliver the packets to the destination router. The selector at the destination simply picks a packet for a session that arrived first, and discard the other packet. This can be done using the packet sequence number set in the RPT header. The duplicator function may be simpler to implement in the encoder. The selector need not be implemented as most decoder will discard the duplicate packet anyway. Alternatively, the selector and the duplicator functions can be integrated into the router as shown in Figure 3-3.

This method not only provides protection for a network fault but also improves upon the performance of the system. If the probability of failure rate in each link is 0.1%, and assuming that their failure rate is statistically independent, the failure rate of this system is only 0.0001%. Similarly if each link has a loss probability of 10%, total loss is only 1%. All these improvement coupled with assurances of expedited forwarding is expected to increase reliability of voice calls over the internet.

One of the disadvantages of this method is that it increases load in the network. However, considering voice traffic is only a fraction of data, this may be acceptable.

### 3.8 Protection with Packet Loss Repair Methods

The methods described in the earlier sections were network based. In this section we outline several mechanisms that are used-based and are used in VoIP systems to correct or conceal the loss of packets in the network. These methods work best when only a few packets are lost in a row, and the losses are scattered in time. Loss of larger number of packets in a row cannot be concealed by these methods.

Forward Error Correction (FEC) is one such mechanism where a part of previous packet and a part of the following packet is encoded in every packet. In case of the loss, destination uses these parts are used to reconstruct the voice frames. Considering that protocol headers take most space in voice packets, FEC adds little overhead. FEC is more efficient if loss is less bursty, and voice packets are larger in size (Jiang and Schulzrine, 2002).

An alternative to FEC is low bit-rate redundancy (LBR). In LBR, a lower quality voice version of same voice is sent simultaneously. When main voice packets are lost, receiver uses the lower quality packets to reconstruct voice. LBR is more complex than FEC, but its performance is not better than FEC (Jiang and Schulzrine, 2002).

Another common method used in VoIP implementation is packet loss concealment (PLC) (G.711, 1999). These methods range from silence substitution, packet repetition, extrapolation and interpolation. These algorithms are used to reduce the effect of packet loss on perceived quality. In case off loss, the decoder derives the data of the lost frame from the previous frames to conceal the losses. In a very simple implementation, the last

frame is duplicated. It is obvious that these extrapolations have their limitations. Loss of 2 or 3 packets in a row could be hidden, but bursty losses are not (Jiang and Schulzrine, 2002).

### 3.9 Summary

In this chapter several methods to protect voice traffic were presented. These include SONET based protection, layer 2/3 based protection using MPLS, and IP based protection such as differentiated services. Other techniques such as multi-homing which includes packet dispersion or duplication, and user based protection to hide losses were also discussed. For voice traffic, selection of a method will depend on its availability and the price a VoIP user is willing to pay. Enterprise user could be willing to pay for the SONET protection, or multi-homed solution, while a private user could rely on IP based protection coupled with user protections. In other cases, the costly but more reliable solutions could be used in the core of the service provider's network, while the IP based protection are used at the edge.

In the following chapter, we will present a simulation to study the performance of voice over a DS network with additional protection provided by packet duplication.

# Chapter 4
# Description of Simulation Model and Network Topology

## 4.1 Introduction

This chapter describes the simulation methodology, and discusses the inputs parameters used in the simulations. A simple model is used to verify the simulation procedure. This is followed by introduction to more complex representative models and discussion of initial results. The detailed simulation results are presented in Chapter 5.

## 4.2 Simulation Procedure

The Network Simulator NS-2 (2005) is used for performance analysis of the voice networks. The DS module that was developed by Pieda et al. (2000) was used to simulate priority processing of the voice packets. This module is briefly described. This follows a description of voice packet, and post processing at the destination to obtain the performance measures. The voice packets in the simulations were duplicated by duplicating various segments of the network; these are discussed latter in this chapter.

### 4.2.1 Differentiated Services

The actions performed by a DS router depend on its location (Section 3.4). The edge router classifies, marks and shapes the packet according to service level. The core router looks at the marking on the packet and schedules them for forwarding according to these markings. The DS module of NS-2 is build on Random Early Discard (RED) (Floyd and Jacobson, 1993) and has two major modules: MRED (multi-RED) and policy

management. The MRED module supports several queue management schemes. For our simulations, we have used RIO-C (RIO Coupled). This model has two separate polices for processing packets, one for IN profile packets and the other for OUT profile packets. Packets are IN profile if they are within service level agreement and OUT profile if they are not. The probability of dropping an OUT profile packet is based on the weighted average lengths of all virtual queues (both IN and OUT); while the probability of dropping an IN profile packet is based solely on the weighted average length of its virtual queue (Figure 4-1).

All flows having the same source and destination are subject to a common policy. We have used time sliding window with 2 colors marking for policy where the lower precedence is used probabilistically when the committed information rate is exceeded.

### 4.2.2 Simulated Voice Packets

Voice is transmitted over Real Time Protocol (RTC, RFC 1889) and UDP. These packets are basically constant bit rate (CBR) with some exceptions. The major exception in voice traffic to CBR, is that in several VoIP implementation empty packets are not sent when there is a speaker silence. This spurt of talk and silence is simulated using an On-Off model (Jiang and Schulzrine, 2000) applied on a CBR traffic generator. Both the On and Off period is often assumed to be exponentially distributed. For further discussion of On-Off model, see Section 2.4.1. In our simulations we have assumed a peak rate of 64 kbps when voice channel is ON. This is based on 8000 Hz voice sampled at 8 bits (G.711). It is also assumed that channel is On for 40% of the time.

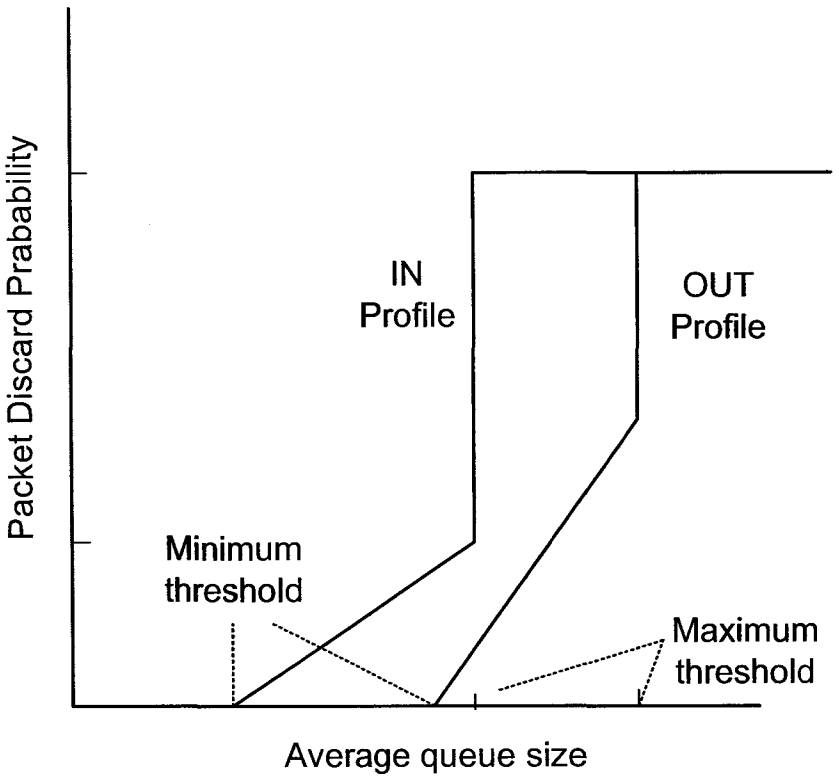


Figure 4-1: Random early detection algorithm for IN profile and OUT profile packets. The average queue size for OUT profile packets is based on average lengths of both IN and OUT queues; while for IN profile packets it is based on the average length of its own queue.

### 4.2.3 Packet Processing at Destination

The RTP header of the packet contains, among other thing, a sequence number and time stamp at the time of creation at the source. Every voice packet is then duplicated across two separate routes to the destination. Information in the RTP header can be used at the receiver to assemble the packets coming from duplicate sources. The sequence number can also be used to calculate losses.

The trace file generated by NS-2 was used to obtain the timestamps arrival and departure at a node while it traverses across various nodes in the network. The trace file also contains a sequence number which is unique between a source and destination pair (Altman and Jimenez, 2003), and for these reasons there was no need to use RTP.

### 4.2.4 Performance Measures

As discussed in earlier sections, the voice packets over internet can be measured by quality factors such as mouth-to-ear delay, variation of this delay (jitter), and packet loss. Measurement methods for these parameters are described below.

**One-way trip delay** – Delay between the sender and the receiver is the most important measure of VoIP performance. A large delay makes a telephone conversation confusing to the users. The delay is the difference in time when the voice was played at the destination and the time packet was digitized at the source. As discussed in Table 2-4, factors such as coding and decoding normally contribute a fixed delay, thus in these simulations we have focused only on the variable delays such as transmission and queuing delays. NS-2 trace file was used to process the delay as discussed in the previous

section. In a real network, this parameter is computed at the destination using the creation time stamp in the RTP header.

**Jitter** – Jitter is the absolute difference between arrival time ($a$) at destination of two consecutive packets ($i$ and $i+1$) and their creation time at source ($d$), or, difference of delay of two consecutive packets, i.e.

$$\text{Jitter}_{i,i+1} = | (a_{i+1} - a_i) - (d_{i+1} - d_i) | = | (a_{i+1} - d_{i+1}) - (a_i - d_i) |$$

The voice decoder at the destination buffers the packets to smooth out the jitter. As discussed in Chapter 2, this adds additional delay.

**Packet loss** - RTP header contains the sequence number of the packet. Out of sequence voice packets were not used in these simulations. In some real VoIP implementations it may be possible to maintain a dynamic list, to buffer and sort the out of sequence voice packets for playback. However, here for simplicity, if a packet with larger sequence number than the expected arrives first, the expected packets are considered to be lost. For the simulations, as described earlier the NS-2 trace file contains the sequence number. The loss is the difference in count of the last packet sequence number assigned at the source and the total number of packets received at the destination node.

## 4.3 Test Model and Testing Procedure

To verify the simulation results, a simple packet duplication network as shown in Figure 4-2 is used. This simple network does not has DS. The parameters used in these simulations are shown in Table 4-1.

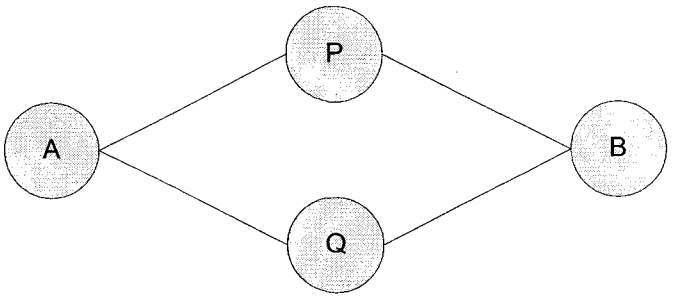


Figure 4-2: Simple non-DS network to verify the simulation results. Both the network data traffic and the voice traffic are traveling from node A to node B.

Table 4-1: Simulation parameters for validation tests.

| Parameter | Description/ Selected value |
|---|---|
| All links | 1 Mbps rate with propagation delay of 10ms |
| Voice Traffic | 0.2 Mbps, 40 bytes packet size, multicast over the two links |
| CBR traffic (other than voice) | 1 Mbps rate with packet size of 200 bytes |
| TCP data traffic | FTP with TCP-Reno with packet size of 552 bytes |
| Queuing parameters | Queue depth of 100, tail drop when full |

All links have a bandwidth of 1Mb/s and have a propagation delay of 10 ms. The voice traffic is CBR and consists of 40 bytes packets and it consumes 0.2 Mb/s bandwidth. The simulation runs for 5 seconds.

Two types of network load are considered here:

I. CBR/UDP traffic at 1 Mb/s with 200 Byte packets on link A-P-B between 1 and 2.5 seconds, and on link A-Q-B between 2 and 3.5 seconds.

II. FTP/TCP traffic with 552 Byte packets on link A-P-B between 1 and 2.5 seconds, and on link A-Q-B between 2 and 3.5 seconds.

Three cases are used in the simulations. These are:

1. Voice traffic on link A-P-B alone,
2. Voice traffic on link A-Q-B, alone and
3. Voice traffic duplicated over both links

The simulation results are presented in Figure 4-3 for network load as described in case I, Figure 4-4 for case II.

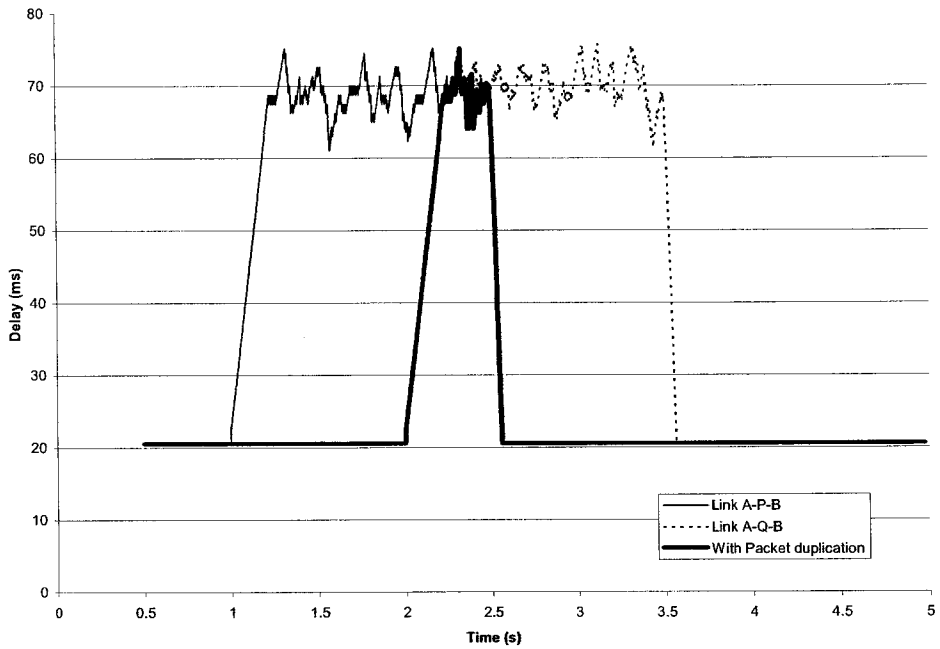


Figure 4-3: Effect of packet duplication algorithm on one-way-delay in voice packets at the receiver in presence of CBR traffic on links.

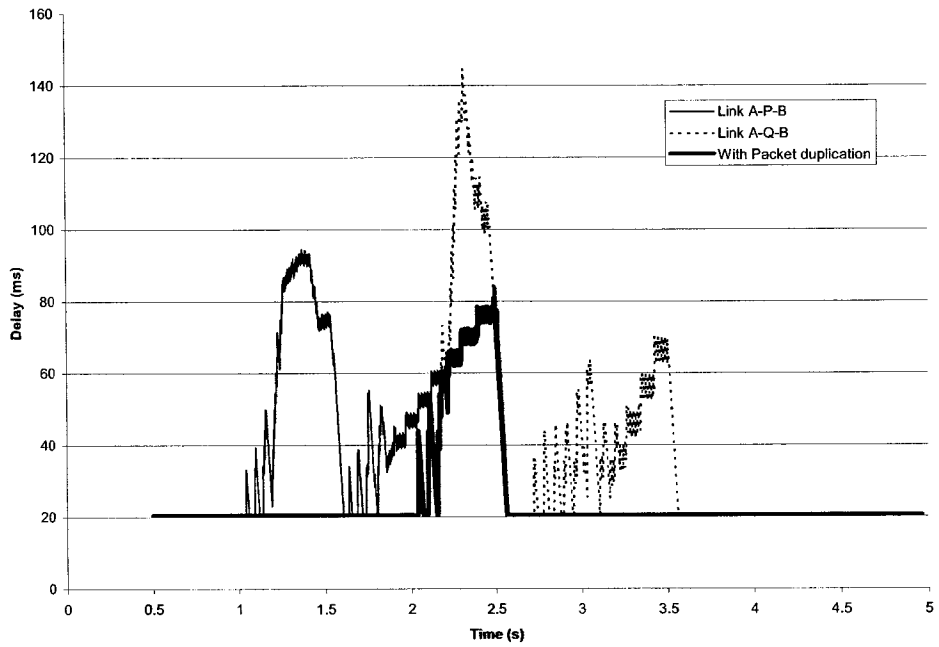


Figure 4-4: Effect of packet duplication algorithm on one-way-delay in voice packets at the receiver in presence of FTP traffic on links.

The results from these simulations can be summarized as follows:

- The packet duplication mechanism achieved least delay and variation of the delay present at any time on two links. The worst delay was seen when both links were heavily loaded.
- The packet duplication mechanism achieved at least an order less packet loss than those on individual links. This depended on the type of other traffic present in the network. For CBR/UDP traffic case, the packet loss for duplicate mechanism is about an order less, while for FTP/TCP traffic case it is negligible compared to individual links. This behaviour is expected because TCP traffic is responsive to network congestion, while UDP traffic is not.

These results show that this simple mechanism has potential to greatly improve the performance of VoIP traffic.

### 4.4 Case Study: Duplication with M/M/1 Queues

While the CBR voice packets are traversing through the network of a service provider, they are expected to randomly reach the destination node. The queuing and service delays in the intermediate nodes introduce this randomness. Further, as the network is also carrying traffic other than voice, and this traffic could have varied packet size, the service times are not constant, but is again randomly distributed. A simple representation of this behaviour can be in the form of two independent M/M/1 queues each representing a service provider as shown in Figure 4-5.

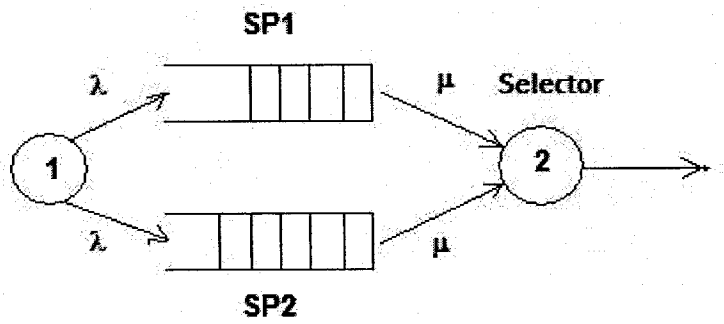


Figure 4-5: Simple duplicated traffic model.

Assume that the networks of the two service providers (SP1 and SP2) are statistically independent. In the actual traffic duplication model, both service providers receive exactly same (deterministic) voice packets. However, in this analysis, to introduce randomness, we assume that both SPs receive statistically similar traffic, with the mean inter-arrival time of packets as $1/\lambda$ with exponential distribution. The service time of the two SPs are assumed to be exponentially distributed with mean as $1/\mu$. The packets egressing from the two service provider's networks are fed into a selector which picks up the first arriving packet.

### 4.4.1 Theoretical Analysis

The network presented in Figure 4-5 is analyzed in this section. The analysis presented below is similar to those of Karol et al. (2003). They assumed that the delay in a SP network is the waiting time in the queue, and ignored the service time. Because the service time also adds to the total delay, the delay is better represented by the response time which is the sum of the waiting time and the service time.

Let $d_1$ be response time in SP1 and $d_2$ be response time in SP2. The response time of the total system is $d$ which is minimum of $d_1$ and $d_2$. For M/M/1 queues (Jain, 1991), the cumulative distribution function $F$ and the mean response time $E$ for each queue is:

$$F(d_i \leq t) = 1 - e^{-\mu(1-\rho)t} \quad (4.1)$$

$$E(d_i) = 1/(\mu-\lambda) \quad (4.2)$$

where $\rho = \lambda/\mu$. Using Equation 4.1 and assuming that the two SPs are independent, the cumulative distribution of the response time of the combined system can be computed to be

$$F(d \leq t) = 1 - F(d_1 > t) * F(d_2 > t) = 1 - e^{-2\mu(1-\rho)t} \quad (4.3)$$

and the probability density function as

$$f(d \leq t) = 2\mu(1-\rho)\, e^{-2\mu(1-\rho)t}. \quad (4.4)$$

The expected response time can be obtained from the probability density function of Equation 4.4 as:

$$E(d) = 0.5/(\mu-\lambda) \quad (4.5)$$

Comparing Equations 4.2 and 4.5, it can be observed that the expected delay is reduced by half. This analysis is over-simplistic, but clearly shows the benefit of packet duplication. Extending this analysis, one can compute that if packets are sent over three independent paths, the expected delay is only third of that when only one route is used. This can be generalised for $n$ independent service providers as:

$$E(d) = (1/n)/(\mu-\lambda). \quad (4.5)$$

In other words, more independent routes will provide better delay performance of the voice packets.

### 4.4.2 Simulation

To verify the simulation procedure, the network presented in Figure 4-5 is simulated using NS-2. Altman and Jimenez (2003) have presented a simulation procedure for

M/M/1 queue using NS-2, we extend that here for duplication with M/M/1 queues. To reduce the effect of truncation at extreme probabilities, the queue sizes are kept very large so that no packet is lost. The packet arrival is exponentially distributed. The exponential distribution of service time is simulated using a constant service time but exponentially distributed packet size. This also results in a limitation that the packet size and the link speed must be adjusted so that the simulated delay while using only one network be same as those computed using Equation 4.2.

The role of the selector at the destination was implemented in a post processing script, where the first arriving packet was used for computing the delay for that packet, and late packet from the other route was dropped.

Using the arrival rate of 20/s and the service rate of 30/s, the simulations were performed for 5000s. Figure 4-6 shows the delays for a segment of simulation. The delay using both SP1 and SP2 is consistently lower than the delay achieved with only SP1. Table 4-2 shows a summary of the computed average delay using Equation 4.2 and 4.5, and the average delay computed in the simulation. The simulated delay for duplicated network is very close to the theoretical value, and verifies the simulation methodology.

Table 4-2: Simulation of M/M/1 queues with packet duplication.

| | *Theoretical Average Delay (ms)* | *Simulated Average Delay (ms)* |
|---|---|---|
| Packet sent over only SP1 | 100 | 100 (normalized) |
| Packet sent over both SP1 and SP2 | 50 | 49.6 |

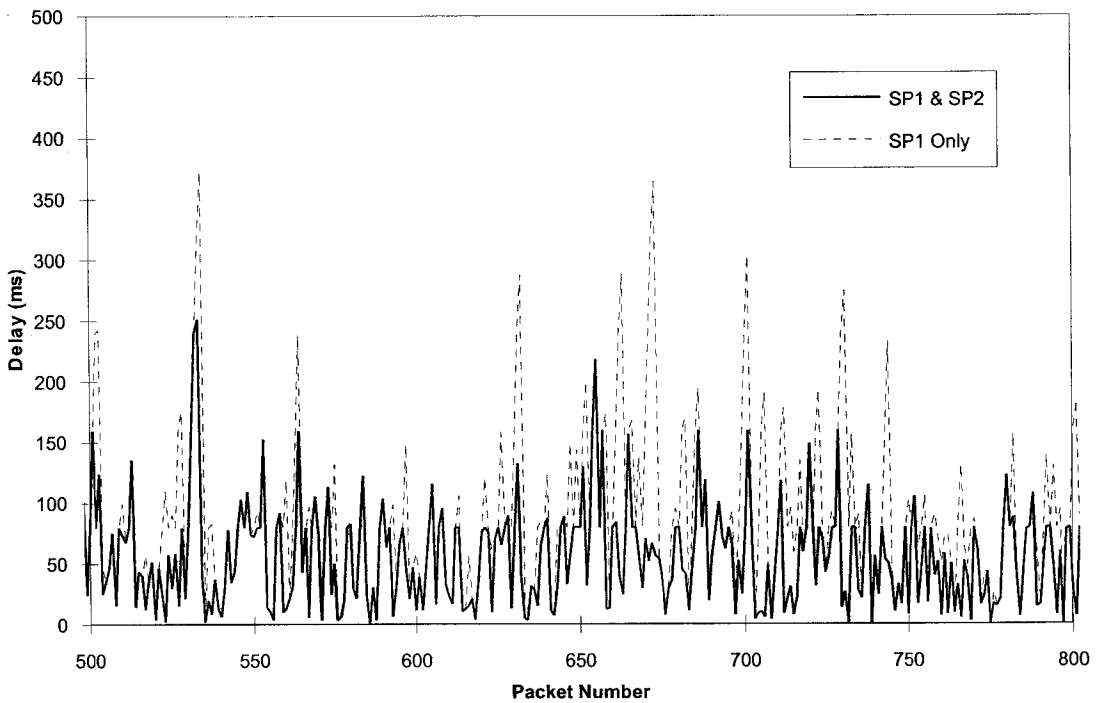


Figure 4-6: Delay behaviour of M/M/1 queues when packets are sent over two independent routes.

### 4.5 Representative Network

A network as shown Figure 4-7 is used to simulate a more realistic scenario. The details of the simulation parameters are shown in Table 4-3. Because we intend to create a congestion in the network core, the edge links are much higher in capacity (150 Mbps bandwidth and 1ms propagation delay) than the core link (P-B) which is 2.048 Mbps bandwidth and 10ms delays. The network is loaded with several combinations of voice traffic with DS class protection and BE traffic.

Table 4-3: Simulation parameters used in the representative networks.

| **Parameter** | **Description/ Selected value** |
|---|---|
| Simulation time | 35 s (data collected over initial 5 s were discarded). Sources were started randomly within first second of simulation. |
| Node queue sizes | 100 |
| Edge Links (V*-A, D*-A, B-D and B-V | 150 Mbps and 1 ms, tail drop |
| Edge link (A-P) | 2.048 Mbps and 10ms, DS RED |
| Core link (P-B) | 2.048 Mbps and 10ms DS CORE |
| Voice traffic | 8 sources (V*-V), 64 kbps with 10% jitter, |
| FTP traffic | 8 sources (D*-D), TCP-Reno with 552 byte packets |
| DS RED queues | RIO-C |
| DS Scheduler | Priority based |
| DS Policier | TSW2CM – RED parameters for IN profile packets are 40/70/0.02, and those for OUT profile packets: 10/30/0.5, the committed information rate of 64kbps. |

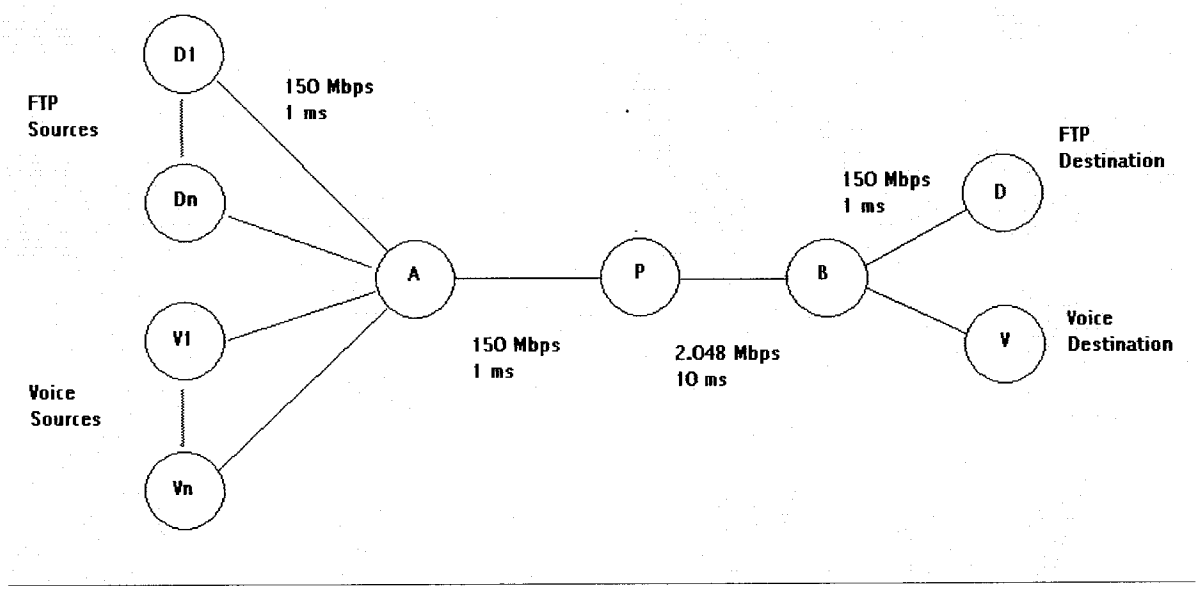


Figure 4-7: DS Network with several DS voice connections (V-V) and BE FTP connections (D-D).

There are 8 CBR voice sources and 8 FTP sources with all CBR sources connected to one destination and the FTP sources attached to another destination. The voice traffic is protected by DS, while FTP traffic is unprotected. The voice sources have a rate of 64 kbps with small jitter around this rate. Ziviani et al. (2002) have noted that the jittered CBR sources provide realistic representation of digitized voice. In the simulations a jitter of 10% was applied to CBR traffic. FTP traffic used TCP-Reno implantation. This provided Slow Start, congestion avoidance, fast retransmit and fast recovery. The data packets are 552 bytes size. The CBR traffic fills up ¼ of the bottleneck link. All sources started randomly within first second of the simulation.

Figure 4-8 shows generalization of Figure 4-7 to include duplication. There are two independent routes between nodes A and B that form the edges the of the bottleneck links, one A-P-B and another A-Q-B. This simulates connectivity of nodes A and B to two service providers. Both of the links have same bandwidth and the delays.

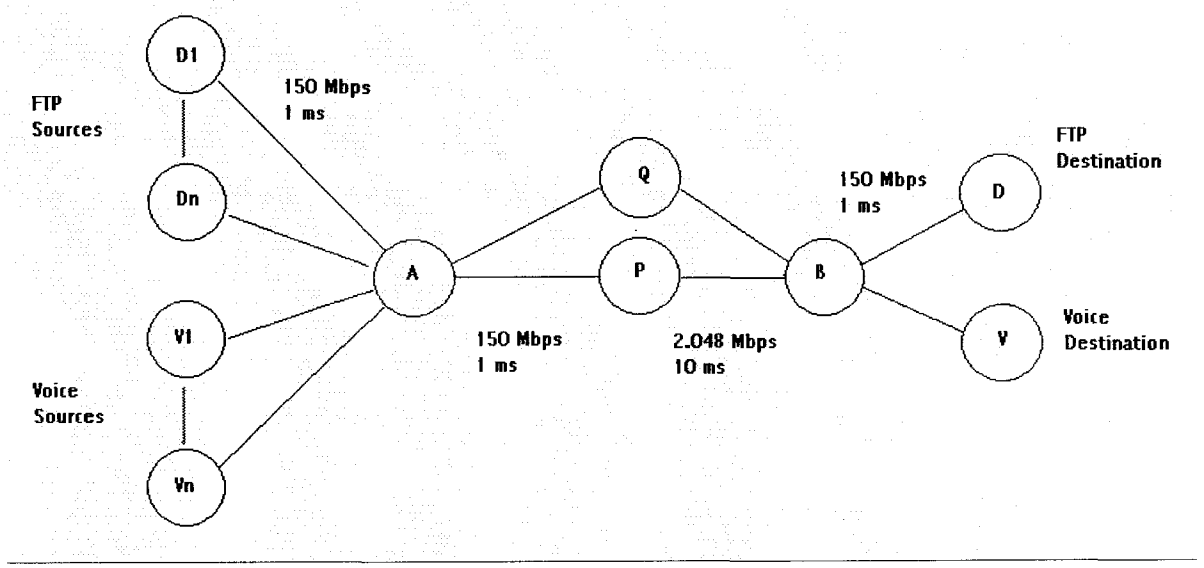


Figure 4-8: A multi-homed DS Network with several DS voice connections (V-V) and BE FTP connections (D-D).

### 4.5.1 Packet Duplication Mechanism

For packet duplication in there simulations multicast could be used. In this method, the voice stream can be duplicated so that they reach the destination taking different route. This procedure did not work because multicast module and the DS modules of NS are not compatible. To get around this problem, several nodes were duplicated. For example, the network of Figure 4-8 is modified as in Figure 4-9. To duplicate the voice, traffic from voice sources (V*) are connected to both destinations (V and V1) separately, and identical traffic is sent over link V*-A-P-B-V and V*-A-Q-B1-V1. The FTP traffic on the two links is not identical, but it is generated based on similar parameters. This assumption is valid because in real life both link will be carrying independent FTP traffic. The edge node B had to be duplicated because of routing constrains.

The node pairs (B, B1), (D, D1) and (V, V1) are virtually attached through post processing. For example, the post processing through the trace file, voice packet for a particular source (V*) arriving first either D or D1 is selected for delay processing.

This network simplification has some limitation, e.g. the dynamics of node B and B1 may be different at any given time because the FTP traffic is not identical. However, considering that this node is at the egress of the network, links coming out of this node will not be congested because large bandwidth there, the simplification is acceptable.

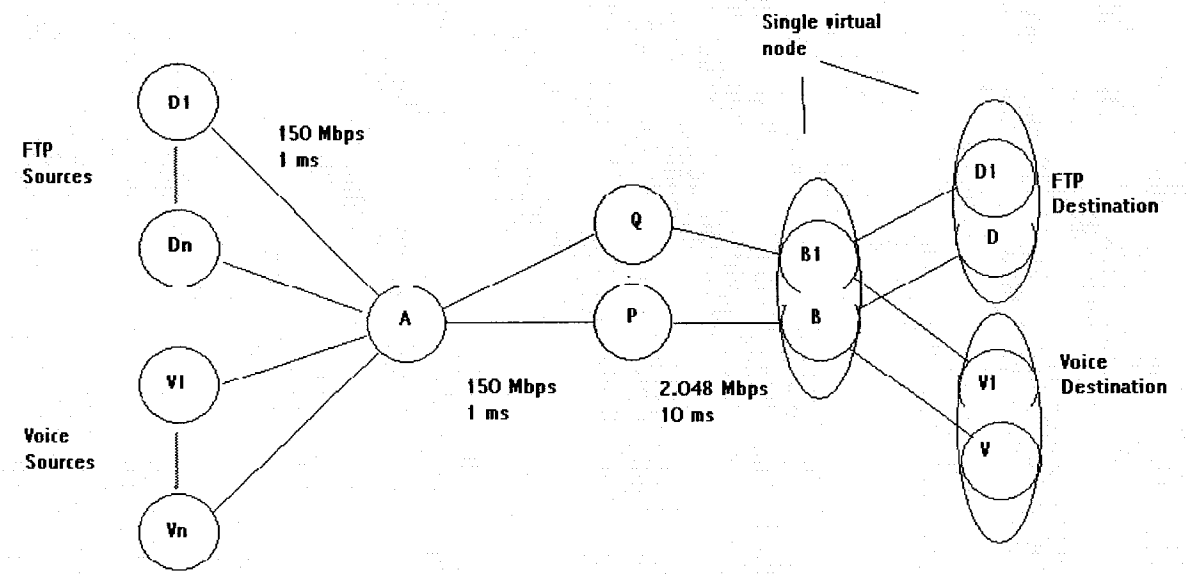


Figure 4-9: Modified network with virtual nodes. Two physical nodes on separate routes form a virtual node.

## 4.6 Steady-State Behaviour and Repeatability

Table 4-4 shows the confident interval estimates for the mean delay obtained by simulating DS and duplication protected voice. The network as shown in Figure 4-9 was used for these simulations. Every simulation was run for 50s, and it was divided into 10 batches of 5s interval each. Batches represent simulation between time 0-5s, 5-10s, 10-15s, 15-20s, 20-25s, 25-30s, 30-35s, 35-40s, 40-45s and 45-50s. The simulations were replicated 10 times.

Table 4-4: 95% confidence interval of the simulated delay (ms)

| Replication | Mean delay (ms) for Replication | | | |
|---|---|---|---|---|
| | Full interval (0s-50s) | Interval (5s-50s) | Interval (10s-50s) | Interval (5s-35s) |
| 1 | 18.6 | 18.4 | 18.6 | 17.8 |
| 2 | 18.1 | 18.2 | 18.2 | 18.4 |
| 3 | 18.9 | 18.6 | 18.6 | 18.7 |
| 4 | 18.5 | 18.2 | 18.1 | 18.5 |
| 5 | 18.0 | 17.8 | 17.9 | 18.9 |
| 6 | 18.2 | 18.2 | 18.0 | |
| 7 | 18.0 | 17.8 | 17.7 | |
| 8 | 18.2 | 18.3 | 18.3 | |
| 9 | 18.5 | 18.6 | 18.7 | |
| 10 | 18.2 | 18.1 | 18.2 | |
| Mean, E | 18.3 | 18.2 | 18.2 | 18.3 |
| Std. Dev. | 0.31 | 0.27 | 0.34 | 0.38 |
| Std. Err (E) | 0.10 | 0.08 | 0.11 | 0.17 |
| 95% min interval | 18.09 | 18.03 | 17.99 | 17.81 |
| 95% max interval | 18.54 | 18.42 | 18.48 | 18.69 |

Table 4-4 shows effect of deleting first one batch (0-5s), and two batches (0-5s and 5-10s). There is little difference in the results showing that steady state behaviour was attained within first 5s of simulations and continues to be in that state. The first 5 second of the simulation data is not used in the analysis. This is because of the obvious dynamic effects at the start up. Further, as various voice and data sources start at different time during the first second in this period including this in analysis does not add any value.

For rest of the analysis an interval between 5s and 35s was chosen with 5 replications. The 95% confidence interval ranges as shown in last column of Table 4-4 suggests that variations in the delay results are still accurate within a ms. The simulated delay distributions are shown in Figure 4-10. As the confidence analysis indicated, there is little variation between 5s-15s, 15s-25s and 25s-35s samples.

Figure 4-11 shows repeatability of the results. The delay distributions were prepared based on data from one replication, three replications and five replications. There is little difference between the distributions of these runs, suggesting that the simulations are repeatable. These results are consistent to those presented in Table 4-4.

Figure 4-12 shows that the distribution delays in voice traffic when passing over the two separate links (A-P-B and A-Q-B) of the network shown in Figure 4-9. Both links are subject to exactly same voice traffic, but independent data traffic. The two delay distributions are identical, showing that there is no bias in the simulation, and the two links are independent.

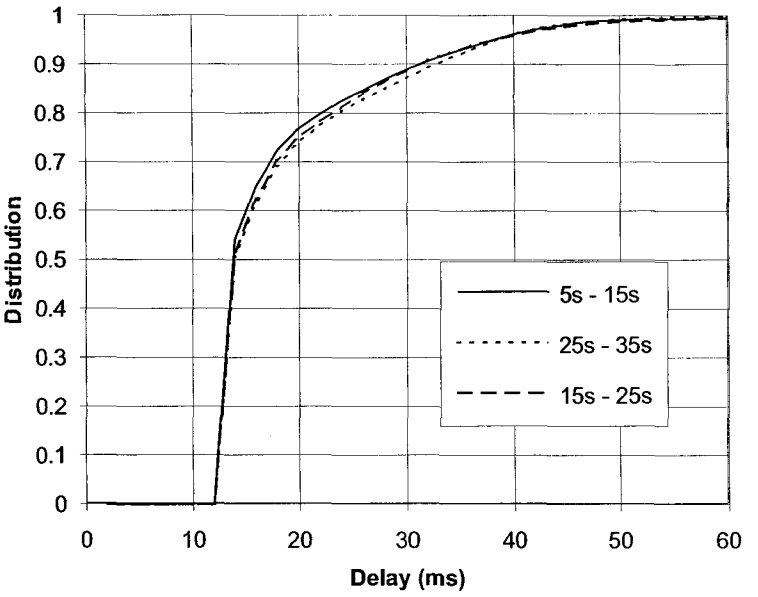


Figure 4-10: Steady state behaviour of the simulations for the network shown in Figure 4-9.

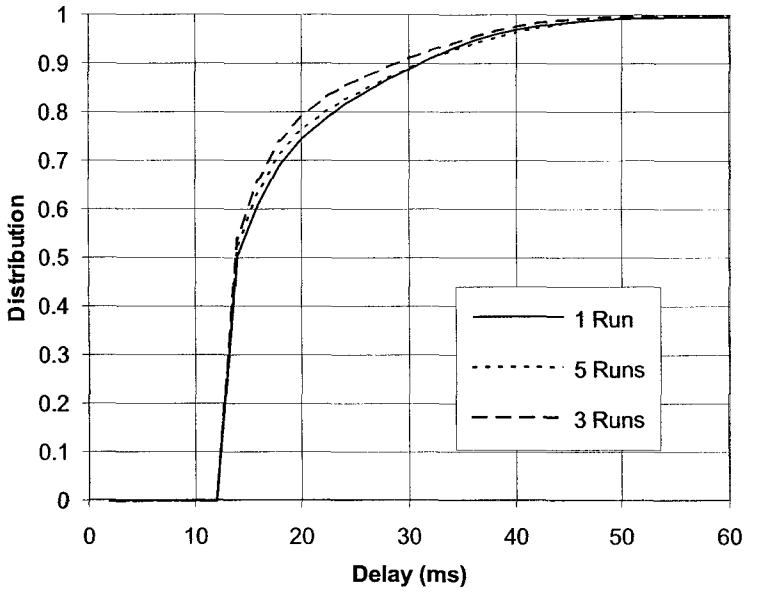


Figure 4-11: Repeatability of the simulation results for the network shown in Figure 4-9.

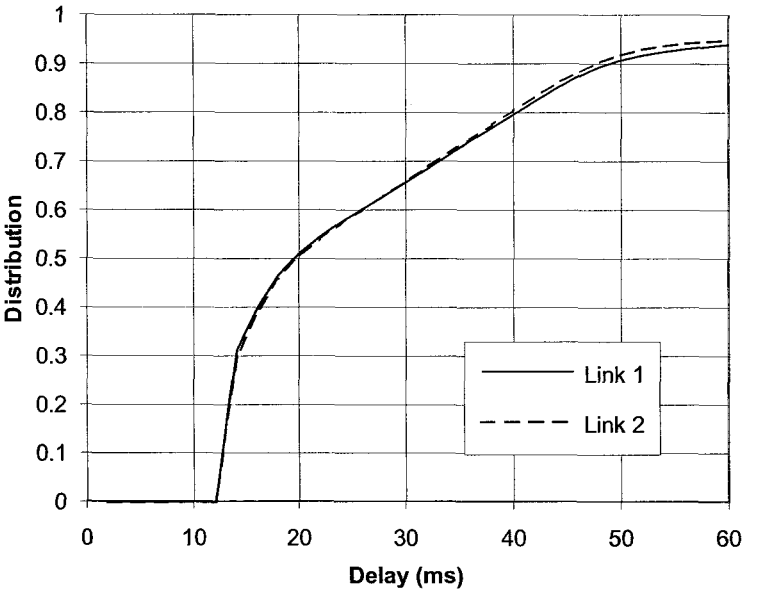


Figure 4-12: Delay distribution of voice on the two independent links of the network shown in Figure 4-9.

## 4.7 Summary

In this chapter the simulation methodology and various input parameters and performance measures were discussed. Simple models were used to verify the procedure. The effect of packet duplication was studied using a simple network based on M/M/1 queues, and found to reduce the average delay by half. Representative simulation models were presented, and the results were verified for their repeatability and steady state behaviour. The results of these models are discussed in the following chapter.

# Chapter 5
# Analysis of Results

## 5.1 Introduction

This chapter presents the results using the simulations models described in Chapter 4. Simulations were performed to study several network cases to protect voice traffic, e.g., effect of traffic duplication and effect of DS protection.

The results are presented in terms of the probability distribution of network delays and the variation in these delays. The distributions are computed from the actual sample size as determined by the sequence numbers of voice packets. This helps to quantify the lost packets.

## 5.2 Duplicated Voice

Using the network of Figure 4-9, the simulated results for 8 CBR voice sources protected by duplication are shown in Figure 5-1 and Figure 5-2. Both independent branches (A-P-B and A-Q-B1) of the network are continuously loaded with FTP traffic. These results show that the delay performance of the duplicated traffic is better than those of unprotected voice using only one route (shown in the figures as regular traffic).

Considering that the fixed delay is 13 ms (10ms for the core link and 3ms for three edge links), the average delay improvements because of duplication is of the order of about 15-20%. This benefit is much less than 50% expected based on the simple analysis based on M/M/1 queues (Section 4.4).

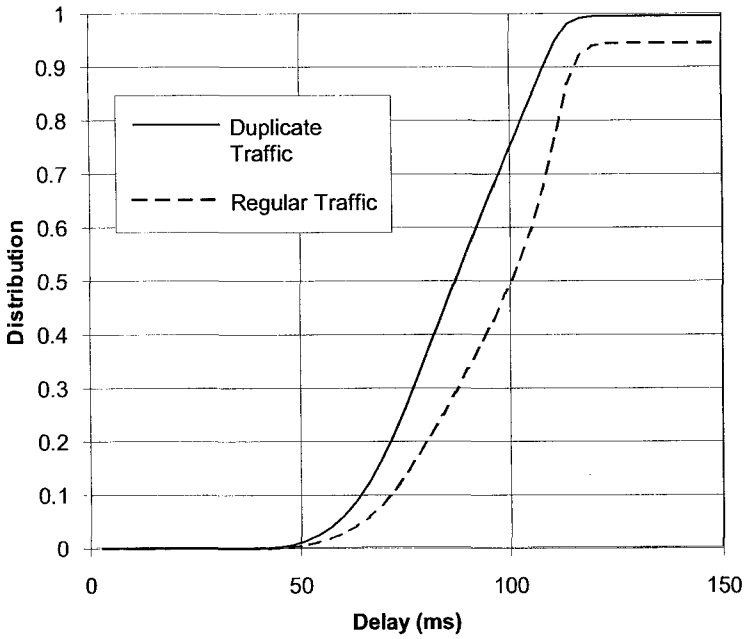


Figure 5-1: Delay distribution of one-way-delay in voice packets in presence of FTP traffic. Voice traffic is duplicated but it is not protected by DS.

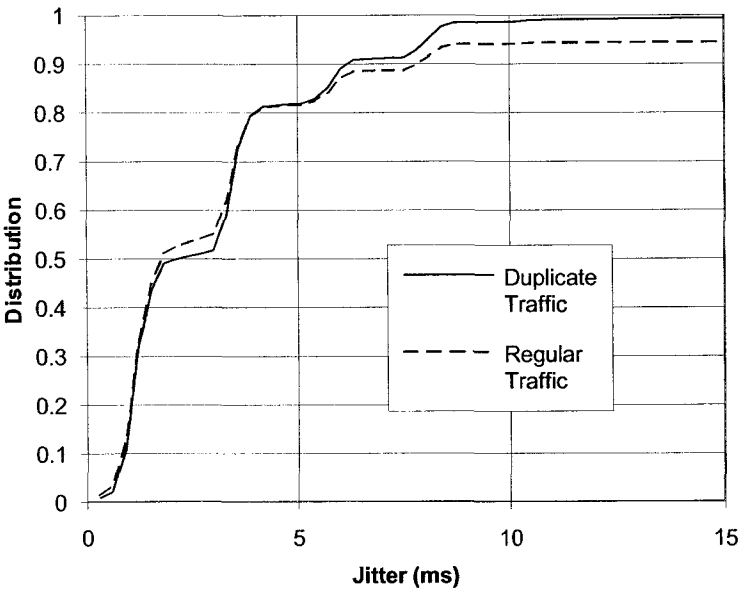


Figure 5-2: Jitter distribution of one-way-delay in voice packets in presence of FTP traffic. Voice traffic is duplicated but it is not protected by DS.

One of the main reasons for this discrepancy is that M/M/1 based model was an oversimplification of the network dynamics. For example, the simple model neither considered the dynamics of TCP, nor discriminated between voice and data traffic.

We should note here that traffic used in the simulations is the worst case, i.e. both links here were heavily loaded with FTP traffic. If one link is less congested than the other, the packet duplication is also expected to provide better performance than those shown here.

A small benefit of 15-20% in delay may not justify the use of duplication mechanism. The real benefit of this method comes from loss protection, and the fault protection. The unprotected voice traffic lost over 5% of the packets, while the duplicated voice had negligible loss.

### 5.3 DS Protected Voice

Figure 5-3 and Figure 5-4 show the performance of DS protected voice. The packets are also protected by duplication. These results show that DS protected and duplicated voice has much better delay behaviour than the regular voice protected only by DS. DS serves its purpose well to deliver packets quickly without any loss. As can be observed from the figures, the DS duplicated voice case achieved delivery of more than 50% of packets within 20ms. Considering that fixed network delay is 13ms as shown in Figure 4-9, this is a great benefit.

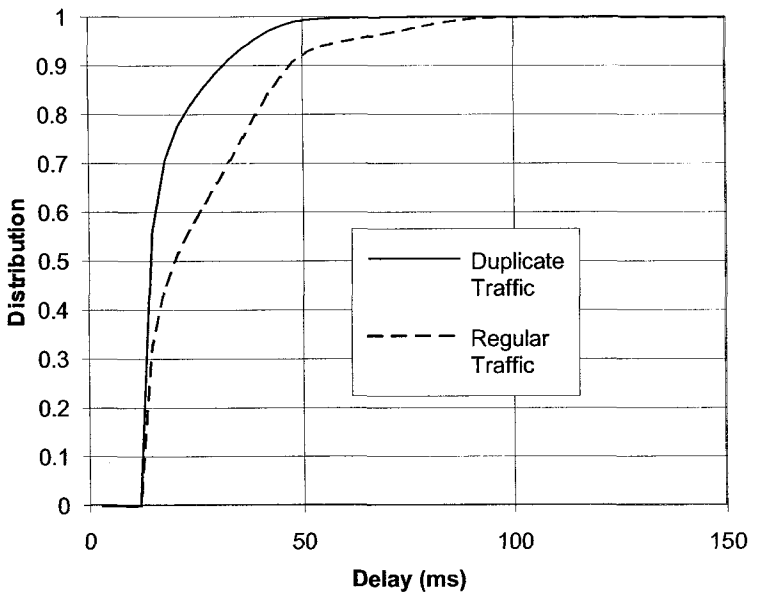


Figure 5-3: Effect of packet duplication algorithm on one-way-delay in voice packets in presence of FTP traffic. Voice packets are duplicated and protected by DS.

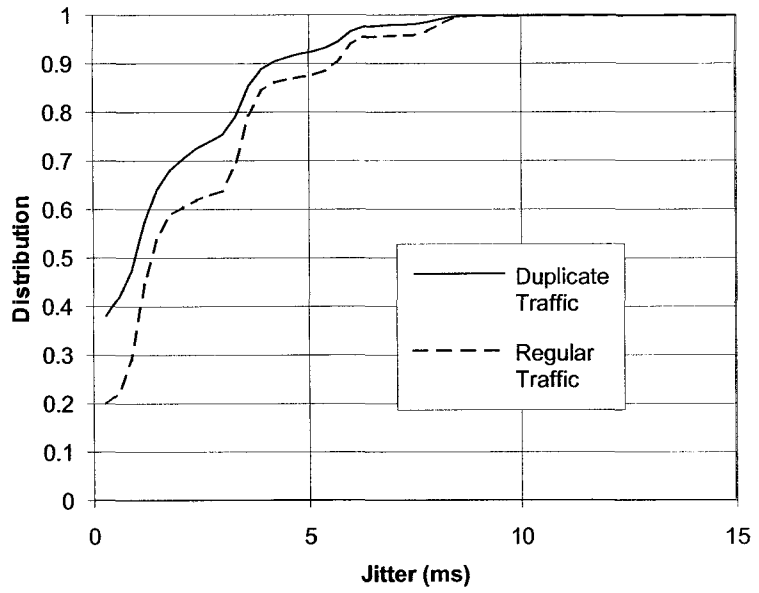


Figure 5-4: Effect of packet duplication algorithm on jitter in voice packets at the receiver in presence of FTP traffic. Voice packets are duplicated and protected by DS.

The benefit of the DS is seen by comparing the results for cases with and without DS protection. Figure 5-5 and Figure 5-6 shows a comparison of duplicated voice with and without DS protection. The voice packets are duplicated in both cases. The loss performance in the two cases are nearly same, but the greatest benefit of DS comes in terms of it ability to control the delay and to jitter.

## 5.4 Performance Measures

Using the results presented in Sections 5.2 and 5.3, a summary of transmission delay, the variation delay (jitter) and packet loss is shown in Table 5-1. If the voice traffic is left unprotected, the network part of delay is very high. Using the delay budget guidance of Table 2-4, this will result is total delay of more than 250ms for 10% of the packet, and is clearly much higher than recommended delay of 150ms. In addition to the delay this network also suffers from larger jitter, and packet loss.

Table 5-1: Effect of network configuration on delay

| **Network Configuration** | **90 Percentile Delay (ms)** | **90 Percentile Jitter (ms)** | **Packet Loss %** |
|---|---|---|---|
| Unprotected voice traffic | 116 | 8.0 | 5.44 |
| Duplicated voice traffic with no DS protection | 108 | 6.3 | 0.43 |
| DS protected voice traffic with no duplication | 48 | 5.7 | 0 |
| DS protected and duplicated voice traffic | 32 | 4.2 | 0 |

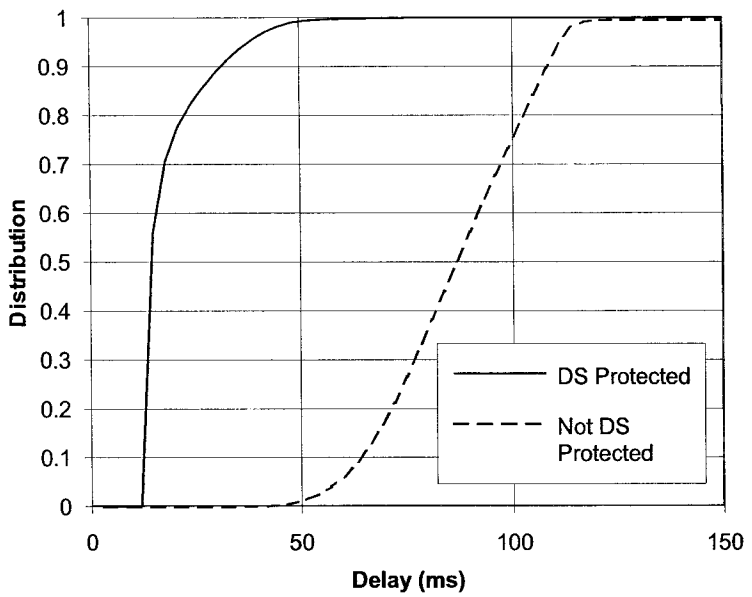


Figure 5-5: Effect of DS protection on one-way-delay in voice packets at the receiver in presence of FTP traffic. Voice packets are duplicated in both cases.

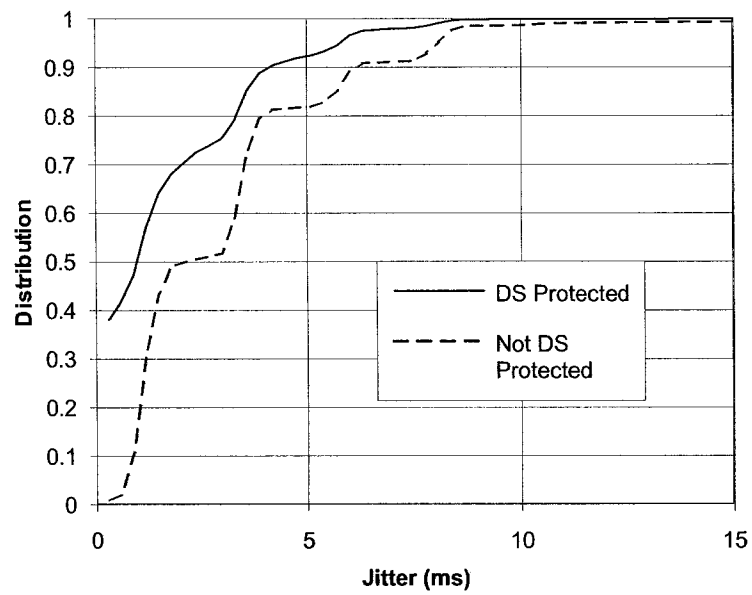


Figure 5-6: Effect of DS protection on variation in delay in voice packets at the receiver in presence of FTP traffic. Voice packets are duplicated in both cases.

When voice traffic is duplicated over the two separate links, there is only marginal improvement in delay and jitter, but the voice packet loss is reduced by an order of magnitude compared to the unprotected network. The measured loss improvement is consistent to those estimated by Karol et al. (2003). As far the delay is concerned, although the packets are duplicated, they are still competing against the data traffic, and in absence of any priority based service, these packets wait much longer in the queues to be serviced.

DS protected voice traffic get priority treatment from the node schedulers, and the wait time is reduced. These configurations are very suited for voice. No packet was lost in simulations. DS without packet duplication is a viable choice for VoIP as have proposed in several studies (Evans and Filsfils, 2004, 2005; Nasser et al., 1998). However, the best performance is achieved by DS with duplication. Note that 10ms delay is introduced by the transmission in the bottleneck link, so the actual delay in the queues is only 22ms.

DS relies on ample over-provisioning of the network, and does little to make efficient use of the network. The advantage in voice is that voice accounts for only small part of total traffic and over-provisioning for this class of traffic may be acceptable. The design bandwidth for the EF class carrying voice traffic is recommended to be at least double of the expected traffic (Evans and Filsfils, 2004).

These results suggest that DS alone in sufficient to provide QoS to VoIP. However, additional mechanism such as duplication is needed to protect against a failure of a link.

## 5.5 Effect of Packet Size

Depending on the length of voice frame packed into a single packet, and the encoding method used, the IP packet size 0.varies. Fugure 5-7 and Figure 5-8 show the effect on delay and jitter on three packet sizes 80 bytes, 160 bytes and 320 bytes. The smaller packets have better delay and jitter behaviour.

Consider that RTP/UDP/IP header could be as much as 40 bytes; the smaller packets have large overhead. Larger packets have less protocol overhead, but there is larger delay in packetization and de-packetization. Essentially there is a trade-off between packet size and protocol overhead.

## 5.6 Talk-Spurt and Silence Simulations

The "On-Off" nature of voice traffic is common in many VoIP implementations. The voice packets are only sent for talk-spurt. This changes the networks dynamics. The TCP traffic can adjust its rate upwards when more bandwidth becomes available. Once the excess bandwidth is occupied by data, another talk-spurt in voice sends non-responsive CBR traffic that has to compete against data traffic. Considering the CBR traffic is of higher priority and protected by DS, routers start dropping data traffic, resulting in TCP to pull back its rate. Figure 5-9 and Figure 5-10 show simulation for "On-Off". Sources are "On" for 400ms and off for 600ms, see Section 4.2.1 for details. Voice is duplicated and protected by DS. These results are comparable to continuous CBR voice traffic case as shown in Figure 5-4 and Figure 5-5.

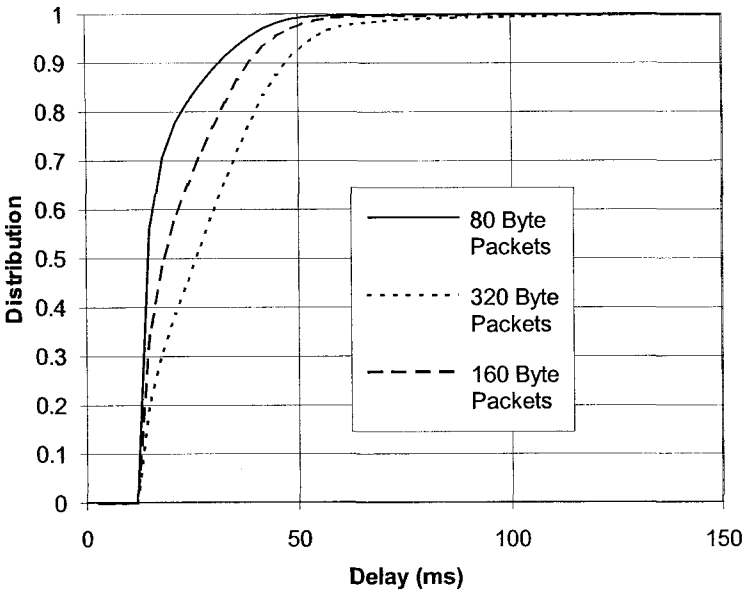


Figure 5-7: Effect of voice packet size on one-way-delay in voice packets at the receiver in presence of FTP traffic.

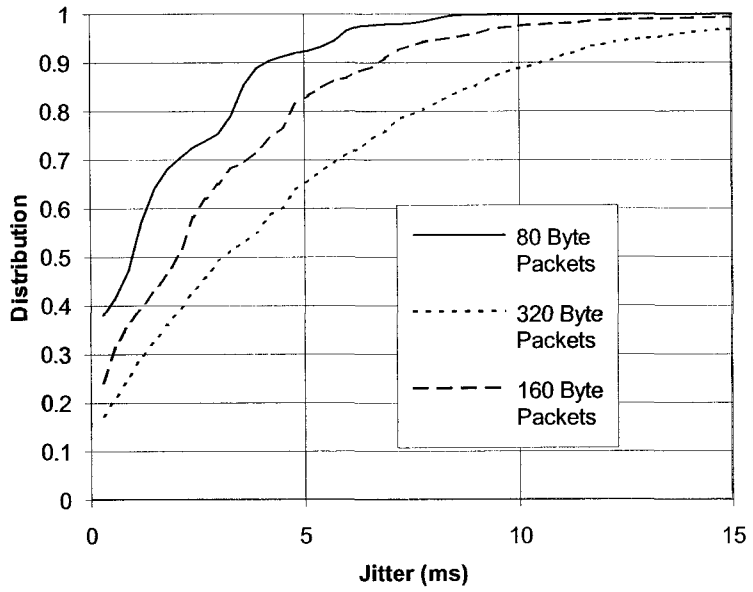


Figure 5-8: Effect of voice packet size on variation in delay in voice packets at the receiver in presence of FTP traffic.

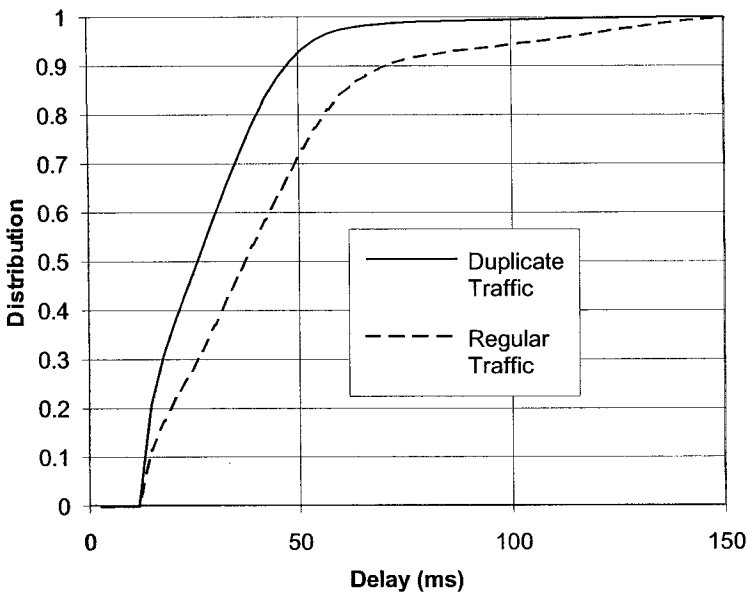


Figure 5-9: Effect of packet duplication algorithm on one-way-delay in On-Off voice packets in presence of FTP traffic. Voice packets are duplicated and protected by DS.

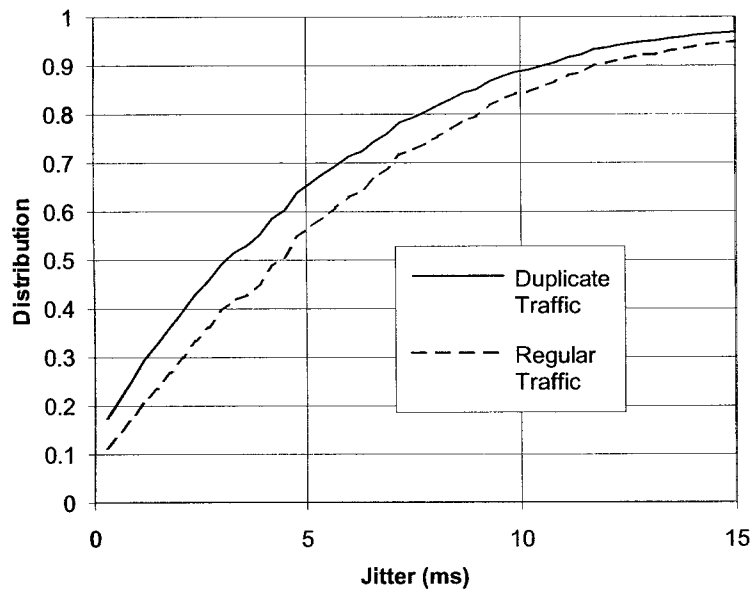


Figure 5-10: Effect of packet duplication algorithm on jitter in On-Off voice packets in presence of FTP traffic. Voice packets are duplicated and protected by DS.

## 5.7 Deploying Packet Duplication

In some networks scenario, independent traffic assumption on the two service provider's network may not be possible, and on some route the duplicated traffic could be forced to take the single path. Further, with the packet duplication, there is a price to pay for increased capacity at the end user, and need for secondary independent service provider. However, considering that the bandwidth requirement of voice calls is only a fraction of total bandwidth of a data network, such refinement may be acceptable. Further, to get the maximum benefit of duplication, it must be done at independent routes, e.g., using two separate SPs (Karol et al., 2003).

## 5.8 Summary

The simulation results of voice protected with DS and packet duplication were presented in this chapter. These results show that DS can provide low delay and low loss voice. The packet duplication guards against packet loss, but makes only small difference in terms of delay if used alone. The duplication however provides protection against fault which DS does not offer. These two mechanisms are complementary, and can provide best protection to voice when used together.

# Chapter 6
# Summary of Work, Conclusions and Future Work

## 6.1 Summary of Work

In this thesis we presented a simulation study of voice over IP, and mechanisms to protect this in case of link congestion and fault. The background information on VoIP was reviewed, and two protection mechanism, DS and packet duplication were evaluated.

IP networks are not as reliable as the PSTN networks. When IP network carries telephone conversations, it must be protected from certain network conditions that may occur. These are links congestions, end-to-end delay, loss and network availability. Voice traffic from telephone conversation cannot be achieved reliably if the links are lossy and delays are large. There are several protection mechanisms to achieve this reliability, and can be used at different layer.

For IP layer protection, DS and packet duplication on a multi-homed network can be used to provide packet protection. The simulations were conducted for these methods and suggest the greatest benefit can be achieved when these two are used together.

The simulations also show that smaller voice packets have better delay and jitter behaviour. However, in such packet the protocol overheads are large.

## 6.2 Conclusions

It was observed that DS can deliver low delay and low loss voice packets. This supports conclusions drawn in other studies (Evans and Filsfils, 2004). The following conclusions can be drawn based on this work:

1. DS can provide the desired level of service to voice over IP in most cases. Voice packets were sent over highest priority queue, and were not affected by FTP data traffic. This method also resulted in low delay and loss.

2. The packet duplication mechanism provides protection against packet loss, but its effect on delay and jitter performances were not significant if this method is used alone. Its real benefit is achieved if both links have QoS provided by DS. One of the disadvantages of packet duplication is the need for additional devices for duplication. The duplication also adds additional traffic in the network, but it may not be significant if only voice traffic is duplicated.

3. It was noted that DS does not protect voice traffic in case of fault of a link. Additional protection is required for this failure. The packet duplication mechanism can protect for fault conditions and could be used along with DS to provide a robust VoIP network.

4. The smaller voice packets have better performance compare to larger packets, but they also carry large protocol overheads. The header compression will certainly handy for small voice packets.

## 6.3 Contributions

The main contribution of this thesis is to show that the packet duplication mechanism guards against packet loss, but makes only small difference to the delay budget if links do not have acceptable QoS. These results are not consistent to those of Karol et al. (2003), who have advocated that using independent routes could provide great QoS benefit in terms of delay behaviour. It is expected that interaction with FTP traffic has resulted in poor performance for voice. This is especially true when both networks are heavily loaded. In duplication only based protection, voice is competing against data traffic on equal footing. There is no addition protection for voice. A major benefit of duplication is fault tolerance and, as is for other packet dispersion methods.

The study also confirmed that differentiated service can provide great relief to voice in networks congestion scenarios. DS is complementary to the duplication method and can be used to provide greater protection to voice when used together.

## 6.4 Future Work

This work can be extended in the following areas:

1. When non-responsive CBR traffic is carried in a high priority class in DS, this traffic needs to be shaped to avoid degradation. Further, uncontrolled voice traffic could starve the best-effort traffic. The effect of a shaper on the performance measures will be valuable addition.

2. For multi-homed network, where there is connectivity to mere than one independent route, packet dispersion is expected to provide similar benefit as packet duplication. This can be verified by simulations.

3. Study the performance of voice when packets are triplicated and sent over three independent routes.

4. Expand the study of On-Off voice sources to include the effect of "Off" period on the overall delay behaviour.

# Bibliography

Al-Irhayim, S., Zubairi, J.A., Qahhar, M. and Abul Latif, S. (2000), "Issues in Voice Over MPLS/Diffserv", Proc. PDCS'2000, Volume II, pp467-472.

Altman, E. and Jimenez (2003), "NS Simulator for Beginners", Lecture Notes.

Awduche, D. Berger, L. Gan, D., Li, T., Srinivasan, V. and Swallow, G. (2001), "RSVP—TE: Extentions to RSVP for LSP", IETF RFC 3209.

Bell Canada (2006), www.bell.ca

Blake et al. (1998), "An Architecture of Differentiated Services", IETF RFC 4275.

Boutremans, C., Iannaccone, G. and Diot, C. (2002), "Impact of Link Failure on VoIP Performace", Int. Workshop on Network and Operating System for Digital Audio and Video (NOSSDAV).

Chong, H. M and Matthews, H.S. (2004), "Comparative Analysis of Traditional Telephone and Voice-over-Internet Protocol (VoIP) Systems", IEEE, pp.106-111.

Cherry, S. (2005), "Seven Myth about Voice over IP", IEEE Spectrum, March 2005.

Chuah, Chen-Nee (2005), A Scalable Framework for IP-Network Resource Provisioning through Aggregation and Hierarchical Control, Ph.D. Dissertation, Department of Electrical Engineering and Computer Science University of California at Berkeley.

Clark, D. and Fang, W. (1998), Explicit Allocation of Best Effort Packet Delivery Service, ACM Transactions on Networking.

Degermark, M., Nordgen, B. and Pink., S. (1999), "IP Header compression", IETF RFC 2507.

Evans, J. and Filsfils, C. (2004), "Deploying Diffserv at the Network Edge for Tight SLAs", Part 1 (Jan-Feb 2004, pp 61-65) and 2 (Mar-Apr 2004, pp 61-69), IEEE Internet Computing.

Evans, J. and Filsfils, C. (2005), "Deploying Diffserv in Backbone Networks for Tight SLA control", IEEE Internet Computing, Jan-Feb 2005, pp 58-65.

Fang, W. (2000), "Differentiated Services: Architecture, Mechanisms and an Evaluation", Princeton Ph.D. thesis, Computer Science, http://www.cs.princeton.edu/~wfang/Research/main.ps.

Floyd, S. and Jacobson, V. (1993), "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking.

Garg, S., Kapes, M. (2003), "Can I Add a VoIP Call", Proc. IEEE International Conference on Communications.

Goyal, M., Durresi, A., Jain, A. and Liu, C. (2000), Performance Analysis of Assured Forwarding, Internet Draft.

Goode, B. (2002), "Voice over Internet Protocol (VoIP)", Proceedings of the IEEE, Vol.90, No.9, 2002, pp.1495-1517.

ITU-T Recommendation G.711 (1999), "A High Quality Low Complexity Algorithm for Packet Loss Concealment with G.711", Appendix I.

ITU-T Recommendation G.107 (1998), "E-Model, a Computational Model for Use in Transmission Planning".

ITU-T Recommendation G.114 (2003), "One Way Transmission Time".

ITU-T Recommendation G.872 (1999), "Optical Transport Networks".

Jacobson, V., Nichols, K., Poduri, K. (1999), "An Expedited Forwarding PHB", RFC 2598, June.

Jain, R. (1991), "The Art of Computer Systems Performance Analysis, John Wiley and Sons.

Jamoussi, B. (2002), "Contrant-Based LSP Setup using LDP", IETF RFC 3212.

Jiang, W., Koguchi, K. and Schulzrinne, H. (2003), "QoS Evaluation of VoIP Endpoints, Proc. IEEE International Conference on Communications, 2003, Vol.3, pp.1917-1921.

Jiang, W. and Schulzrinne, H. (2000), "Analysis of on-off patterns in VoIP and their effect on voice traffic aggregation", 9th IEEE International Conference on Computer Communication Networks.

Jiang, W. and Schulzrine, H. (2002), "Comparison and Optimization of Packet Loss Repair Method on VoIP Perceived Quality under Bursty Loss", NOSSDAV 2002.

Johnson, C. et al. (2004), "VoIP Reliability: A Service Provider's Perspective", IEEE Communications Magazine, July 2004, pp.48-54

Karol, M., Krishnan, P. and Li, J.J. (2003), "VoIP protection and performance improvement", Computer Communications and Networks, ICCCN 2003. Proceedings. The 12th International Conference on 20-22 Oct. 2003, pp.505 - 510.

Markopoulou, A.P., Tobagi, F.A and Karam, M. (2002), "Assessment of VoIP Quality over Internet Backbones", IEEE INFOCOM 2002.

Mehmood, M.A, Jadoon, T.M. and Sheikh, N.M (2005), " Assessment of VoIP Quality over Access Networks", Internet 2005, Proc. The First IEEE and IFIP Int. Conf. in Central Asia, pp.1-5.

Nasser, H., Leon-Garcia, A. and Aboul-Magd, O. (1998)"Voice Over Differentiated Services", Internet Draft, IETF, December.

Nichols, K. et al. (1998), Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers (RFC 2474), IETF.

NS-2 (2005), The Network Simulator, http://www.isi.edu/nsnam/ns/, October 2005.

Pieda, P. Ethridge, J., Bains, M. and Shallwani, F. (2000), "A network Simulator Differentiated Services Implementation", Open IP, Nortel Networks, July 26, 2000.

Rosenberg, J., et al. (2002), "SIP: Session Initiation Protocol", IETF RFC-3261.

Sanneck, D.H. (2000), "Packet Loss Recovery and Control for Voice Transmission over the Internet", PhD. Thesis, Techical University of Berlin.

Skype (2006), www.skype.com.

Takahashi, A. and Yoshino, H. (2004), "Perceptual QoS Assessment Technology for VoIP", IEEE Communications Magazine, July 2004, pp.28-34.

Thaanthry, N., Pendse, R. and Namuduri, K. (2005), "Voice over IP Security and Law Enforcement", CCST05, 39th Annual Int. Carnahan Conf. on Security Technology, 2005, pp. 246-250.

Thiran, P., Taft, N., Diot, C., Zang, H. and McDonold, R. (2001), "A Protection-Based Approach to QoS in Packet over Fibre Networks", IWDC 2001, pp.266-278.

Typhalt, M. (1998), "Voice-Data Integration: Resurgence Of Convergence" Information Week, April 13, 1998.

VoIP Info (2006), http://www.voip-info.org.

Vonage (2006), VoIP setup, www.vonage.com.

Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P. and Heinanen, J. (2002), "Multi- Protocol Label Switching (MPLS) Support of Differentiated Services", IETF RFC-3270.

Zeadally, S. Siddiqui, F., Kubher, P. (2004), "Voice over IP Intranet and Internet Environments", IEE Proc.-Commun, Vol 151, No.3, June 2004.

Ziviani, A,Rezende, J. and Duarte, O. (2002), "Evaluating the Expediated Forwarding of Voice Traffic in a Differentiated Services Network", Int. J. Commun. Systems.

Zlatokrilov, H. and Levy, H. (2004), " Packet Dispersion and the Quality of Voice over IP Applications in IP Networks", Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2004.