VoIP PERFORMANCE IN NGEO SATELLITE IP NETWORKS WITH ON-BOARD PROCESSING CAPABILITY

by

Suzan Bayhan

B.S. in Computer Engineering, Boğaziçi University, 2003

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Department of Computer Engineering
Boğaziçi University
2006

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor Assoc. Prof. Fatih Alagöz, for being not only such a great mentor, but also a great person with his kind, never-ending encouragement and nice talks. I have been really fortunate to have the chance of working with him.

I would like to express my sincere thanks to the faculty staff of Mathematics and Computer Science Department of İstanbul Kültür University and faculty dean Prof. Dursun Koçer for tolerating and encouraging me during my thesis study. I would like to thank my office mate Çiğdem Gencer for both being my academic consultant and close friend. She guided me with her experience and expanded my vision of life.

I am really grateful to Gürkan Gür for being my "sensei" and for his interest, concern, and valuable counsel on both research and writing of this thesis. He spent a great deal of time, which he already lacked, dealing with my LaTEX and other matters. This thesis would certainly not come to an end without him.

This thesis is completed thanks to the encouragements of my beloved friends. I would like to thank Nilay Özok and Azat Zana Gündoğan for opening the great doors of academic enthusiasm and giving very nice lectures on life. Mustafa Çelikkaya was always with me and amused me a lot with his nice and unique comments on everything. Special thanks to M. Gökçen Arslan for her valuable comments on this thesis.

My appreciation for my sister Sezen Bayhan is beyond the words for making me feel the most precious creature on Earth, checking out my papers and listening to my presentations. I would like to thank my flatmates Özlem Kırkım and also Sezen, for tolerating me and for entertaining me with their humorous talks during my coffee breaks. I wish to express my gratitude to the greatest parents ever, Meryem and Fikret Bayhan for embracing me with their endless love and care. Last but not least, special thanks to those anonymous people who once touched my soul.

This work is supported in part by The State Planning Organization of Turkey, Project: Next-generation Satellite Networks and Applications, DPT-03 K120250.

ABSTRACT

VoIP PERFORMANCE IN NGEO SATELLITE IP NETWORKS WITH ON-BOARD PROCESSING CAPABILITY

In this thesis study, an adaptive routing policy utilizing the real-time network information of a two-layered satellite network is introduced. In a satellite network, depending on the requirements and properties of services provided, various kinds of satellites from different orbits can be employed. Geostationary Earth Orbit (GEO) systems are not suitable for Voice over Internet Protocol (VoIP) applications due to long end-to-end delay values about 250-270 ms. Non-Geostationary Earth Orbit (NGEO) systems consisting of Low Earth Orbit (LEO) and Medium Earth Orbit (MEO) satellites can satisfy the performance requirements of VoIP applications. Moreover, a two-layered system of LEOs and MEOs can outperform single plane satellite networks. However, due to the dynamic topology of these networks and nonuniform traffic distribution over the Earth, terrestrial packet based routing algorithms cannot perform well. The proposed routing scheme dubbed as “Adaptive Routing Protocol for Quality of Service” (ARPQ) prevents the congestion on some bottleneck links by distributing the traffic over the entire network. Furthermore, link capacities can be efficiently used. Additionally, delay and jitter sensitive voice traffic is processed in a prioritized way to prevent long queueing delays. By a set of simulations, we showed that proposed mechanism performs better than nonadaptive routing mechanisms and therefore can enable VoIP applications over satellite networks.

ÖZET

ARAÇ ÜSTÜ İŞLEME YAPABİLEN NGEO UYDU SİSTEMLERİNDE VoIP BAŞARIMI

Bu tez çalışmasında, iki katmanlı bir uydu sisteminde, ağın o anki durumu dikkate alınarak yapılan yeni bir yönlendirme mekanizması tanıtılmaktadır. Uydu sistemlerinde, sağlanan servislerin gereksinimlerine ve özelliklerine göre değişik yörüngede uydular kullanılabilir. Yerdurağan uydu (GEO) sistemleri uçtan uca 250-270 ms gecikme değerlerinden dolayı VoIP (Internet Protokolü üzerinden ses aktarımı) servisleri için elverişli değildir. Alçak yörünge uyduları (LEO) ve orta yörünge uydularından (MEO) oluşan yerdurağan olmayan uydu (NGEO) sistemleri VoIP uygulamalarının gereksinimlerini karşılayabilirler. Bununla birlikte, LEO ve MEO uydulardan oluşan iki katmanlı bir sistem, tek katmanlı uydu sistemlerinden daha iyi başarım sonuçları verebilir. Ancak bu sistemlerin dinamik bir topolojiye sahip olmaları ve Dünya üzerinde düzenli bir trafik dağılımının olmaması gibi sebeplerden dolayı, karasal paket tabanlı sistemlerde kullanılan yönlendirme protokollerinin kullanılması uygun değildir. ARPQ olarak adlandırılan önerdiğimiz yönlendirme mekanizması, LEO ve MEO katmanları üzerinde yük dağılımı yaparak, bazı ana noktalarda sıkışmanın önlenmesini ve tüm ağ üzerindeki kanalların verimli bir şekilde kullanılmasını sağlar. Ayrıca, gecikme ve gecikmedeki değişime duyarlı olan trafik (VoIP), uydularda kuyruklama gecikmesini azaltacak şekilde öncelikli olarak işlenir. Çeşitli benzetim çalışmaları ile, önerdiğimiz uyarlamalı yönlendirme mekanizmasının, uyarlamalı olmayan yönlendirme mekanizmalarından daha iyi başarım değerlerine sahip olduğunu ve dolayısı ile VoIP uygulamaları için elverişli olduğunu gösterdik.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3GPP	Third Generation Partnership Project
ADPCM	Adaptive Differential Pulse Code Modulation
ARPQ	Adaptive Routing Protocol for Quality of Service
BWFA	Broadband Wireless Fixed Access
CS-ACELP	Conjugate Structure Algebraic CELP
D_{thrsh}	Threshold delay value to mark the voice packets as either *LDV* or *SDV*
ECC	Error Correcting Code
FEC	Forward Error Correction
FSM	Finite State Machine
GEO	Geostationary Earth Orbit
GM	Group Manager
GPS	Global Positioning System
GS	Gateway Station
HAP	High Altitude Platform
IOL	Inter Orbital Link
ISL	Inter-satellite Link
LD-CELP	Low Delay Code Excited Linear Prediction
LDV	Long Distance Voice packet
LEO	Low Earth Orbit
MEO	Medium Earth Orbit
NGEO	Non-Geostationary Earth Orbit
NSL_i	Neighbor Status List of LEO_i holding the information about the states of ISLs associated with this satellite
OBP	On-Board Processing
PCM	Pulse Code Modulation
QoS	Quality of Service
RIB	Routing Information Base
RTP	Real Time Protocol

RTCP	Real Time Control Protocol
SDV	Short Distance Voice packet
SOS	Satellite over Satellite
VoIP	Voice over Internet Protocol
VSAT	Very Small Aperture Terminal

1. INTRODUCTION

The idea of artificial satellites' being used for communications is firstly introduced by the British science fiction writer Arthur C. Clarke in 1945. Clarke wrote in Wireless World magazine [1] that a satellite with a circular equatorial orbit at a correct altitude of 35786 km would make one revolution every 24 hour; that is, it would rotate at the same angular velocity as the Earth. An observer looking at such a geostationary satellite would see it hanging at a fixed spot in the sky. Clarke showed that three such satellites powered by solar energy could provide worldwide communications for all possible types of services [2]. After his remark on satellites, many research studies began on satellite and space communication systems.

The earliest satellite "Sputnik I" owned by the Soviet Union is launched in 1957. The first artificial communications satellite (SCORE in 1958) did not follow long after the Sputnik launch, and the first commercial geostationary satellite (INTELSAT 1, or "Early Bird") in 1965 ushered in the era of overseas telephony via satellite. In the 1970s and 1980s, both the market for satellite communication services and the technology grew rapidly. Besides providing international telephony and data services between large earth stations owned by national carriers, communication satellites were increasingly used for video (television) distribution. The international organization INMARSAT was founded to provide telephony and data services to maritime customers. Finally, the construction of systems based on Very Small Aperture Terminals (VSATs) for transaction-oriented traffic such as credit card verification and database management was begun in the 1980s. In the 1990s the growth of alternative, cheaper technologies such as high speed fiber optic networks has gradually eliminated much of the international telephony service for non-mobile customers [3].

Satellites can be categorized according to their orbit types. Specifically, there are three types of orbits that need to be considered: Geostationary Earth Orbit, Low Earth Orbit (LEO) and Medium Earth Orbit (MEO) (the latter two also called as NGEO satellites). Satellite systems will continue to be an essential element in the establishment of long-distance telecommunications for many years, and it will have a major role in the implementation of the so-called *global information infrastructure* (GII) in the future. This is because of the particular feature of the satellite that can provide wide coverage independent of the actual land distance between any pair of communicating entities [4]. NGEO satellite networks can provide next generation internet service requirements with high bit rate and much lower latency than GEO systems. However, some characteristics of these systems dissimilar from wired and GEO systems expose some challenges. Dynamic topology of the network and constraints on key system resources such as on-board CPU and memory are essential points to be considered for an efficient system design. From the users' perspective some quality-of-service (QoS) requirements must be satisfied, i.e. bounded delay values, guaranteed minimum bandwidth. From the network providers view, system resources such as link capacities must be effectively used.

In this thesis, a novel adaptive routing mechanism on two-layered satellite network considering the network's real time information is introduced. *Adaptive Routing Protocol for Quality of Service (ARPQ)* avoids congestion by distributing traffic load between the satellites in the two layers. In order to change satellite networks' being used merely as backup systems or merely in niche markets i.e. maritime or rural communications, on-board processing (OBP) is essential. We utilize a kind of round robin queueing policy to satisfy delay-sensitive application QoS requirements while evading non real-time traffic suffer from low performance level. In this thesis, we focus on what can be done on-board the satellite to improve the performance of VoIP services.

2. SATELLITE NETWORKS

The last two decades of the twentieth century marked an explosion in the growth of wireless and mobile communications, fueled by the demand for cellular telephones, pagers and messaging devices. Now, at the beginning of a new century, market growth is being fueled by the promise of multimedia applications and Internet access for wireless laptops, cellular telephones, and personal digital assistants (PDAs) [5]. This growth in demand intensifies interest in satellite technologies, services and networks. In this chapter, beginning with a short history of communication satellites, a brief overview of satellite technical features is presented. Following is a list of early milestones in the history of communications satellites [6].

- Herman Potocnik- describes a space station in geosynchronous orbit - 1928
- Arthur C. Clarke - proposes a station in geosynchronous orbit to relay communications and broadcast television - 1945
- Project SCORE - first communications satellite - 1958
- Echo I - first passive reflector satellite - August 1960
- Courier 1B - first active repeater satellite - October 1960
- Telstar - the first satellite designed to transmit television and high-speed data communications - July 1962
- Syncom - first communications satellite in geosynchronous orbit - 1963
- OSCAR-III - first amateur radio communications satellite - March 1965
- Molniya - first Soviet communication satellite, highly elliptic orbit - October 1965
- Early Bird - INTELSAT's first satellite for commercial service - April 1965
- Orbita - first national TV network based on satellite television - November 1967
- Anik 1 - the first national satellite television system, Canada, - 1973
- Westar 1, the USA's first geosynchronous communications satellite - April 1974
- Ekran - first serial Direct-To-Home TV communication satellite 1976
- TDRSS - first satellite designed to provide communications relay services for other spacecraft - 1983
- Mars Global Surveyor - first communications satellite in orbit around another

planet (Mars) - 1997

- Cassini spacecraft relays to Earth images from the Huygens probe as it lands on Saturn's moon, Titan, the longest relay to date. – January 14, 2005

Satellite systems consist of two segments: *Earth Segment* and *Satellite Segment*. Earth Segment consists of Earth Stations and Ground Control Center. Ground Control Center is the part of the satellite system that acts as the interface point between the satellite user and the Earth stations. It handles the actual connections to the satellite, including the functions of satellite acquisition, tracking, hand-offs, signal modulation and multiplexing. Earth stations can be accepted as users of the satellite services or related terminal equipments such as transmitter and receiver antennas. Satellite segment is the actual satellite repeater consisting of communication payload and other satellite subsystems (e.g. power supply, bus structure).

Satellites are extensively used for a variety of applications as a result of several distinctive characteristics such as global coverage, scalability, broadcast and multicast capability, bandwidth flexibility and reliability. With a proper network design, satellites can serve to any region, even to those areas that are very difficult to serve by terrestrial systems. Installing some Earth stations at the point of application allows users to communicate without any external connections. This can be attractive especially in places where the terrestrial infrastructure is poor or expensive to employ or install [7]. Satellites have also exclusive status in maritime and aeronautical communications. Moreover, since satellites are not affected by natural disasters, they are the only solution for post-disaster management. Installation and maintenance of ground stations are much easier and faster than terrestrial systems. Once the satellites are deployed and system is ready to provide services, users can get service quickly.

2.1. Orbits

Satellites can operate in several types of Earth orbit. The most common orbits for environmental satellites are geostationary and polar, but some instruments also fly in inclined orbits. The inclination (i) determines the tilt of the orbital plane with respect to the equatorial plane of the Earth and is an angle measured in degrees. This element is defined as the angle between the two normal vectors of the equatorial and orbital plane. An orbit with an inclination of zero degrees is equatorial; an orbit with an inclination of 90 degrees is polar. Inclinations are limited to a maximum of 180 degrees [8]. In a polar orbit, the satellite generally flies at a low altitude and passes over the planet's poles on each revolution. The polar orbit remains fixed in space as Earth rotates inside the orbit. As a result, much of Earth passes under a satellite in a polar orbit. Because polar orbits achieve excellent coverage of the planet, they are often used for satellites that do mapping and photography. Satellites can also be classified according to the altitude of their orbits as Geostationary Earth Orbit (GEO) and Non-geostationary Earth Orbit (NGEO) as shown in Figure 2.1.

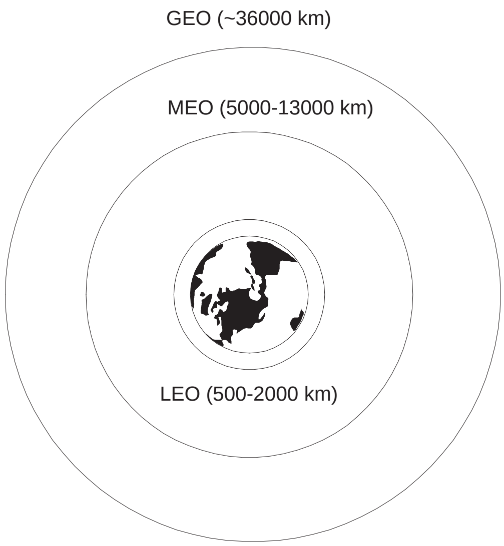


Figure 2.1. Satellite orbits

2.1.1. Geostationary Earth Orbit (GEO) Satellites

Geostationary Earth Orbit satellites -also called Geosynchronous Earth Orbit or synchronous, circle the Earth at the same rate as the Earth spins. The Earth actually takes 23 hours, 56 minutes, and 4.09 seconds to make one full revolution. So based on Kepler's Laws of Planetary Motion, this would put the satellite at approximately 35,790 km above the Earth. The satellites are located near the equator since at this latitude, there is a constant force of gravity from all directions. Geosynchronous orbits allow the satellite to observe almost a full hemisphere of the Earth. These satellites are used to study large scale phenomenon such as hurricanes and also used for communication services. The disadvantage of this type of orbit is that since these satellites are very far away, they have poor resolution. The other disadvantage is that these satellites have trouble monitoring activities near the poles.

2.1.2. Non-geostationary Earth Orbit (NGEO) Satellites

NGEO satellites as opposed to GEOs are not relatively stationary to the Earth. NGEO satellites comprise of Low Earth Orbit (LEO) and Medium Earth Orbit (MEO) satellites, the former being at the altitudes of 500-2000 and the latter at 5000-13000 km. Due to low altitudes of LEO satellites, their orbital period T is short (between 90 and 120 min) and thereby LEOs have low path loss and less bit error rate (BER). Moreover, lower orbit means low deployment costs which make LEO systems economically attractive. LEO systems have small coverage (also called footprint). Hence, in order to provide global service there is a need for a large number of LEO satellites. This, in turn makes the network complex and hard to manage. Typical end-to-end propagation delay for a LEO satellite is about 20-25 ms. LEO systems are suitable for time sensitive applications related to short propagation delay values. On the other hand, at the altitudes of LEO systems, there is severe atmospheric effect which damages LEO satellites. In higher orbits, this effect becomes less severe and therefore leads to longer satellite lifetime. A summary of LEO, MEO and GEO satellites is provided in Table 2.1.

Table 2.1. Summary of GEO, MEO and LEO properties

	LEO	**MEO**	**GEO**
Altitude (km)	500-2000	5000-13000	36000
One-way Propagation Delay (ms)	20-25	80-100	250-280
Orbital Period	90-120 min	about 6h	23h 56 min
Coverage Radius (km)	small	moderate	large
Path Loss	low	moderate	high
Deployment costs	low	moderate	high
Example constellation	Iridium	ICO	Thuraya

2.2. On-Board Processing (OBP)

The satellites are grouped into two according to their payloads: "bent pipe" and on-board processing/switching. A bent-pipe satellite acts as a repeater in the sky by amplifying the incoming signal and forwarding it to the ground station in its footprint. All work such as routing and congestion management is carried on the ground station. Hence, it is the simplest architecture [9].

On-board processing is a general term that refers to signal processing and routing functions implemented on-board the satellite that goes beyond the amplification and frequency conversion performed in conventional, transparent satellite systems. The next generation satellites extensively need to use OBP, including demodulation/remodulation, decoding/recoding, transponder/beam switching, and routing to design cost effective system solutions for the customer needs [10]. The OBP in satellites eliminates the inherent disadvantages of the bent pipe transponders. The main advantages of satellite systems with OBP are: improved link quality with respect to transparent systems due to signal regeneration on board, efficient bandwidth and power level control by multi-beam frequency re-use which increases satellite raw capacity, discarding empty uplink time slots resulting in increased efficiency of downlink transmission, dynamic reallocation of unused bandwidth, asymmetric uplink and downlink bandwidth to take advantage of traffic statistics, on-board management of network traffic, capacity and quality of service (QoS), statistical multiplexing which supports varying degrees of bursty traffic, and direct interconnections between user terminals through on-board switching [11, 12].

OBP can support high-capacity inter-satellite links (ISLs) connecting two satellites within line of sight. Switches in the satellites provide short latency and thus improve the quality of service (QoS) with regard to systems using hub stations on ground. By using a sophisticated constellation with ISLs, connectivity in space without any terrestrial resource is possible. This feature enables far more autonomous satellite networks which may be imperative especially for military purposes and post-disaster-communications situations, where ground facilities may become potential targets or be damaged. These benefits, however, demand payloads with higher complexity [11]. With more advanced and powerful integrated circuitry and microelectronics, OBP has become more feasible and sensible cost-wise. Thus it has the potential for enabling satellite networks to cope with the inherent propagation delay burden [13] and contribute performance of VoIP applications over satellite networks.

2.3. Satellite Services

Satellites support a broad array of applications (Figure 2.2) which can be listed mainly in four groups: communications, Global Positioning System (GPS) and navigation services, remote sensing and direct-to-consumer. Each group and provided services are listed below [14].

- **Voice, video and data communications:** Rural telephony, news gathering and distribution, internet trunking, corporate, VSAT networks (VoIP and multimedia over IP, mobile telephony, videoconferencing, broadcast and cable relay, distance-learning, tele-medicine)
- **GPS and navigation:** Emergency services, search and rescue, security and database access, mapping.
- **Remote sensing:** Forest fire prevention, flood and storm watches, air pollution management, infrastructure planning, urban planning
- **Direct-to-Consumer:** Broadband IP, DTH/DBS television, interactive entertainment and games, video and data to handheld devices

Depending on the type of services offered, satellites usually operate in the part of the radio spectrum ranging between 1 and 30 GHz. Originally, the International Telecommunication Union (ITU) allocated spectrum to mobile-satellite services in the L/S bands. As the range of systems and services on offer have increased, the demand for bandwidth has resulted in a greater range of operating frequencies, from very high frequency (VHF) up to Ka band, and eventually even into the V band of 40-75 GHz as can be seen from Table 2.2. In general, the lower the microwave band, the more efficient it is for mobile applications. This is due to cheaper terminal equipment. For

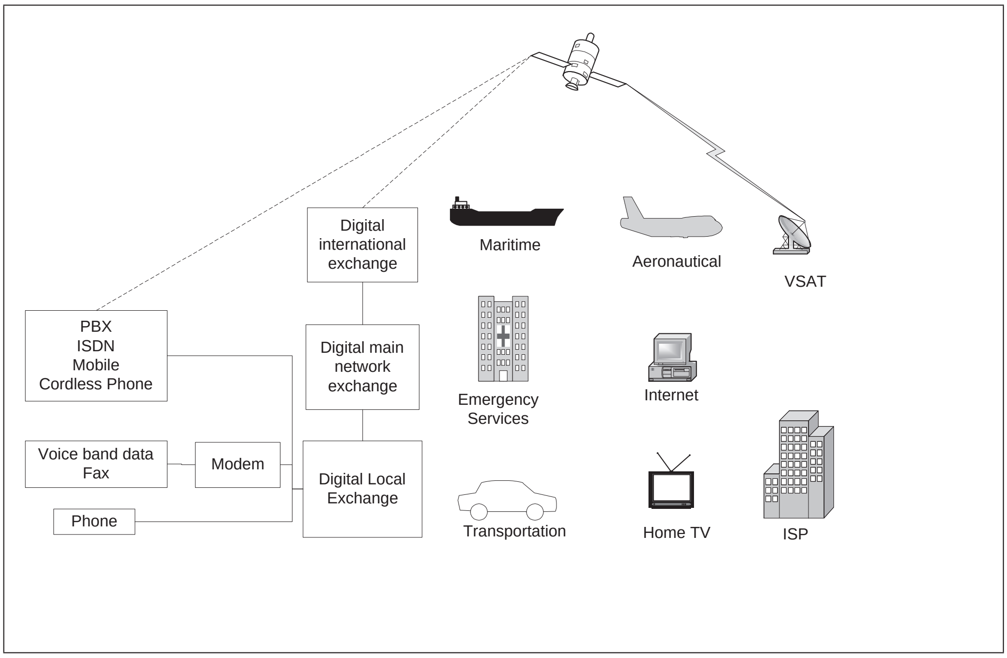


Figure 2.2. Satellite services and coverage

fixed satellite systems the Ku band is the current workhorse for all operators with Ka band gradually being introduced for unicast services [15].

Table 2.2. Satellite Frequency Bands

Band	Frequency Range(GHz)
L	1-2
S	2-4
C	4-8
X	8-12
Ku	12-18
Ka	18-40
V	40-75

2.4. An Example Satellite Constellation: Iridium

The Iridium System is a satellite-based, wireless personal communications network providing a robust suite of voice features to virtually any destination anywhere on earth. The Iridium system comprises three principal components: the satellite network, the ground network and the subscriber products including phones and pagers. The Iridium system requires 66 active satellites in 6 polar orbits as in Figure 2.3. Satellites are LEO satellites at a height of approximately 780 km. Satellites communicate

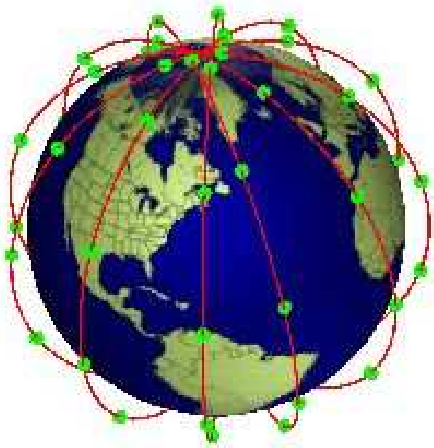


Figure 2.3. Iridium constellation of 66 LEO satellites in 6 planes

with neighboring satellites via inter-satellite links (ISL). Each satellite can have four ISLs: two to neighbors fore and aft in the same orbital plane, and two to satellites in neighboring planes to either side. The satellites orbit from pole to pole with an orbit of roughly 100 minutes.

The design of the Iridium network allows voice and data to be routed virtually anywhere in the world. Voice and data calls are relayed from one satellite to another until they reach the satellite above the Iridium Subscriber Unit (handset) and the signal is relayed back to Earth [16]. Iridium satellites were launched on November 1, 1998 and went into Chapter 11 bankruptcy on August 13, 1999 due to insufficient demand for the service. The increased coverage of terrestrial cellular networks (e.g. GSM) and the rise of roaming agreements between cellular providers proved to be fierce competition. The cost of service was also prohibitive for many users, despite the continuous world-wide coverage of the Iridium service. In addition, the bulkiness and expense of the handheld devices when compared to terrestrial cellular mobile phones discouraged adoption among users. The Iridium satellites were, however, retained in orbit, and their services were re-established in 2001 by the newly founded Iridium Satellite LLC, owned by a group of private investors.

3. VOICE OVER INTERNET PROTOCOL (VoIP)

The Internet is evolving into a universal communication network and it is contemplated that it will carry all types of traffic, including voice, video and data. Among them, telephony is an application of great importance, particularly because of the significant revenue it can generate. VoIP is a technology that uses Internet Protocol (IP) networks to deliver voice services. It takes voice communications and transmits it as packets of data, particularly over broadband networks. In order for the Internet to constitute an attractive alternative to the traditional Public Switched Telephone Network (PSTN), it must provide high quality VoIP services [17]. PSTN uses circuit switching for carrying voice traffic. Since it is a dedicated system for voice communication, it is really very efficient with 99.999% availability. The PSTN has served voice traffic well over the last 100 years, but its success has been paralleled by the rise of separate networks to support data traffic. As more and more PSTN traffic becomes data-oriented, the trend towards voice and data convergence becomes stronger. Convergence has played a major role in the move towards VoIP. For example, there is one network that carries all electronic traffic over the same physical cabling backbone; and one device (the computer) that can handle most transmissions voice calls using softphones, video conferencing, web access, email, faxes, etc.).

VoIP technology enables real-time transmission of voice signals as packetized data over IP networks. IP networks allow each packet to independently find the most efficient path to the intended destination, thereby best using the network resources at any given moment [18]. Internet Protocol (IP) is an attractive choice for voice transport for many reasons, some of which include lower equipment cost, integration of voice and data applications, lower bandwidth requirements and the widespread availability of IP. This chapter presents an overview of VoIP systems and their quality of service requirements.

3.1. VoIP Basics

Figure 3.1 shows the flow of voice packets from one end to other end. Although VoIP involves the transmission of digitized voice in packets, the telephone itself may be analog or digital. The voice may be digitized and encoded either before or concurrently with packetization. The IP phones include codecs that digitize and encode (as well as decode) the speech. The IP phones also packetize and depacketize the encoded speech. Calls between different sites can be made over the wide area IP network. Proxy servers perform IP phone registration and coordinate call signaling, especially between sites. Like all other voice communications VoIP needs two types of protocols: protocol for sending the conversation data in the IP medium and protocol for the signaling.

Since VoIP applications are time-sensitive, User Datagram Protocol (UDP) is used instead of Transmission Control Protocol (TCP). Additionally, for sending the conversation data in the IP medium RTP/RTCP (Real Time Protocol/Real Time Control Protocol) protocol is used over UDP. RTP is responsible to control the voice packet and voice quality. And RTCP is used for exchanging messages between session users regarding the quality of session like lost RTP packets, delay etc. Signaling protocol is needed for Call setup, Monitoring call progress and Call release. There are various protocols available for this purpose like SIP, H.323, MEGACO and H.248. The following subsections explore H.323 and SIP protocol standards.

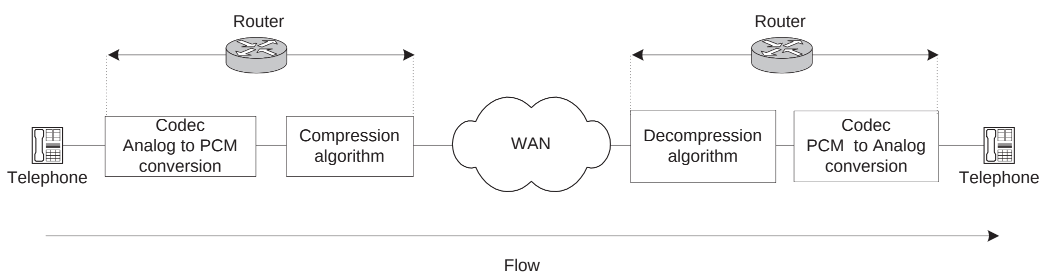


Figure 3.1. End-to-end voice flow

3.1.1. H.323

The H.323 standard is a cornerstone technology for the transmission of real-time audio, video, and data communications over packet-based networks. It specifies the components, protocols, and procedures providing multimedia communication over packet-based networks. H.323 defines how audio and video information is formatted and packaged for transmission over the network.

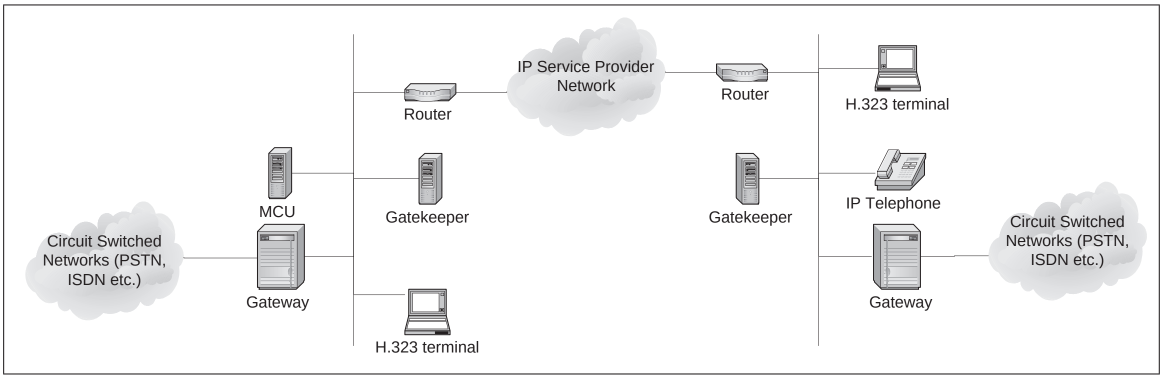


Figure 3.2. H.323 architecture

H.323 defines a number of elements that are required for multimedia transmission. Some elements are mandatory; some are optional. Often, these entities are implemented in software and it can be possible to have more than one entity installed on a single computer. The most important elements (Figure 3.2) are listed below:

- **Gatekeeper:** A gatekeeper is an optional entity which provides network services to H.323 terminals, Multipoint Control Units, and gateways by authorizing (or refusing) communications between other H.323 entities within its zone of control. It also provides an address translation service. H.323 devices register with gatekeepers to send and receive H.323 calls. Gatekeepers give permission to make or accept a call based on a variety of factors.

  Gatekeepers can provide network services such as:

  – Controlling the number and type of connections allowed across the network.
  – Helping to route a call to the correct destination.
  – Determining and maintaining the network address for incoming calls.

- **Multipoint Control Unit (MCU):** An MCU provides services that allow three or more endpoints to take part in a conference call. An MCU comprises a Multipoint Controller for handling call control and optional Multipoint Processors for handling the media exchange (voice, video etc.) in a conference.

- **Gateway:** A gateway provides a protocol conversion service between H.323 terminals and other terminals that do not support H.323. For example, a gateway may route voice over IP calls from an H323 terminal onto the PSTN thus allowing regular telephone calls to be placed from an H323 client such as NetMeeting. Gateways can serve the following purposes:
  - To bridge an H.323 call to another type of call, such as a telephone. Potentially, NetMeeting could call any telephone in the world.
  - To bridge H.323 calls to H.320, which is audio and video transmission over Integrated Services Digital Network (ISDN) .
  - To bridge H.323 calls to H.324, which is audio and video transmission over standard telephone lines.
  - To bridge different networks; an organization could put a bridge on a firewall to connect an internal corporate network with external networks to accept incoming calls.

  In this case, gateway functions are similar to an MCU for connecting people over networks. Typically, though, the gateway is the translation mechanism in a point-to-point connection, where only one endpoint is an H.323 device. On the other hand, an MCU typically connects many H.323 devices in a multipoint conference.
- **H.323 Terminal:** An H.323 Terminal is an endpoint on a network which provides for real-time, two-way communications with another H.323 terminal, Gateway or MCU. A terminal may provide speech only, speech and data, speech and video, or speech, data and video.

3.1.2. Session Initiation Protocol (SIP)

SIP is an Internet Engineering Task Force (IETF), ASCII-based application layer protocol for establishing, manipulating, and tearing down sessions. Like other VoIP protocols, SIP is designed to address the functions of signaling and session management within a packet telephony network. Signaling allows call information to be carried across network boundaries. Session management provides the ability to control the attributes of an end-to-end call. In November 2000, SIP was accepted as a 3rd Generation Partnership Project (3GPP) signaling protocol and permanent element of the IP Multimedia Subsystem (IMS) architecture. It is one of the leading signaling protocols for Voice over IP, along with H.323.

SIP provides various capabilities to:

- Determine the location of the target end point: SIP supports address resolution, name mapping, and call redirection.
- Determine the media capabilities of the target end point via Session Description Protocol (SDP), SIP determines the "lowest level" of common services between the end points. Conferences are established using only the media capabilities that can be supported by all end points.
- Determine the availability of the target end point: If a call cannot be completed because the target end point is unavailable, SIP determines whether the called party is already on the phone or did not answer in the allotted number of rings. It then returns a message indicating why the target end point was unavailable.
- Establish a session between the originating and target end point: If the call can be completed, SIP establishes a session between the end points. SIP also supports mid-call changes, such as the addition of another end point to the conference or the changing of a media characteristic or codec.
- Handle the transfer and termination of calls: SIP supports the transfer of calls from one end point to another. During a call transfer, SIP simply establishes a session between the transferee and a new end point (specified by the transferring party) and terminates the session between the transferee and the transferring party. At the end of a call, SIP terminates the sessions between all parties.

SIP is a peer-to-peer protocol. The peers in a session are called User Agents (UAs). A user agent can function in one of the following roles: User agent client (UAC) or User agent server (UAS). UAC is a client application that initiates the SIP request.

UAS is a server application that contacts the user when a SIP request is received and that returns a response on behalf of the user. SIP clients can include phones and gateways. Phones can act as either a UAS or UAC. Gateways provide many services, the most common being a translation function between SIP conferencing endpoints and other terminal types. This function includes translation between transmission formats and between communications procedures. In addition, the gateway translates between audio and video codecs and performs call setup and clearing on both the LAN side and the switched-circuit network side. SIP servers include Proxy server, Redirect server and Registrar server. Proxy server is an intermediate device that receives SIP requests from a client and then forwards the requests on the client's behalf. Basically, proxy servers receive SIP messages and forward them to the next SIP server in the network. Proxy servers can provide functions such as authentication, authorization, network access control, routing, reliable request retransmission, and security. Redirect server provides the client with information about the next hop or hops that a message should take and then the client contacts the next hop server or UAS directly. Registrar server processes requests from UACs for registration of their current location. Registrar servers are often co-located with a redirect or proxy server.

3.2. QoS Parameters

This section summarizes the main factors determining the voice quality, including the choice of codec, echo control, packet loss, delay, delay variation (jitter), and the design of the network. Figure 3.3 summarizes the factors affecting the voice quality.

3.2.1. Codecs

In order to speed up transmission and save storage space, the size of digital audio samples and video frames need to be reduced. This can be done by codecs which are software or hardware that compresses and decompresses audio and video data streams. However, the compression should not noticeably degrade the audio and video quality. The more complex the compression algorithm, the better the voice quality but higher latency caused by longer processing time. If more bandwidth is used, better

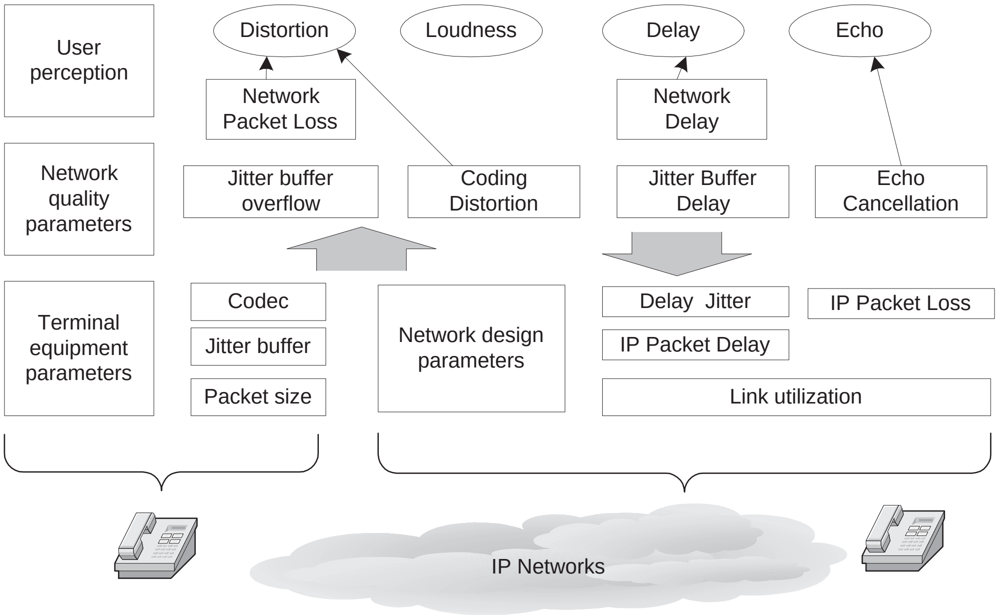


Figure 3.3. Factors that determine the quality of a VoIP service

voice quality can be achieved. There are many codecs available varying in complexity, bandwidth needed and voice quality. Table 3.1 summarizes the properties of some well-known codecs.

Table 3.1. Characteristics of several voice codecs

Codec	**Algorithm**	**Usual rate (Kbps)**
G.711	PCM	64
G.726	ADPCM	32
G.728	LD-CELP	16
G.729(A)	CS-ACELP	8
G.729e	Hybrid CELP	11.8

3.2.2. Delay

There are various sources of packet delay which can be classified into two groups as fixed and variable. As the names specify, fixed delays are constant, and variable delays depend on the system state. Below is a list of delay components.

- Processing (Codec) Delay: The process of encoding the data and combining the data samples to form a packet leads to processing delay which is a fixed delay. The processing delay is influenced by the processor speed and the type of algorithm used.

- Packetization delay is the time spent for accumulating reasonable number of voice samples to place inside a packet. This delay is directly proportional to the size of the packet. On the other hand, IP/UDP/RTP headers add extra 40 bytes (20/8/12 bytes respectively) to the packet size. Thus, to avoid the unnecessary consumption of available bandwidth packet size should not be very small. In other words, there is a trade-off between packetization delay and bandwidth utilization.
- Serialization delay is the fixed delay required to clock a voice or data frame onto the network interface. It is directly related to the clock rate on the trunk. At low clock speeds and small frame sizes, the extra flag needed to separate frames is significant.
- Queueing/Buffering Delay: After the compressed voice payload is built, a header is added and the frame is queued for transmission on the network. Voice needs to have absolute priority in the router/gateway. Therefore, a voice frame must only wait for either a data frame that already plays out, or for other voice frames ahead of it. Essentially the voice frame waits for the serialization delay of any preceding frames in the output queue. Queueing delay is a variable delay.

3.2.3. Jitter Buffers

Delay jitter is the amount of delay variation an application encounters on the network. Since voice packets may encounter different delays from each other on the path from source to destination, there is a need to rearrange the arriving packets. The effects of jitter on VoIP can be counteracted by the use of jitter buffers at the receiver. Jitter buffers are memory areas used to store voice packets arriving with variable delays so that it appears that each voice sample has arrived in the same amount of time. The steady output of the voice samples from the jitter buffer is called playout. The playout is steady and constant, and as long as the jitter buffer receives an ample supply of voice packets, the system appears to have a fixed delay. The delay chosen for the jitter buffer is critical. If the delay is set too low, then the whole scheme will not work. Packets that arrive “late”, i.e., after the jitter buffer delay, must be discarded. This discarding can cause a noticeable gap in the conversation, since a sizable portion of a second of conversation may be discarded. If the delay is set too high, then the jitter buffer may overflow with similar loss of conversation, or the voice is delayed needlessly. Another consideration is how the jitter buffer knows exactly how long it took the voice packets to make their way across the network [19].

3.2.4. Silence Suppression

Silence suppression removes the periods of silence that occur naturally within a voice conversation. The main cause of silence is when one of the parties to the conversation is listening, but shorter periods of silence occur between sentences, phrases, words, and even within longer words. All together, silence accounts for nearly 60 percent of the bits sent during a two-way 64 Kbps PCM voice conversation. The biggest problem with silence suppression is detecting when the speaker begins to talk again after a period of silence.

3.2.5. Comfort Noise

Comfort noise is used to compensate for the loss of sound at the listener when silence suppression is in use. Total silence leads the listener to think the line has failed. Some ambient background noise is generated at the receiver to maintain the illusion of a constant stream of noise across the network. The process of adding comfort noise to a packet voice system is called "comfort noise generation" (CNG). There are many ways to perform CNG. The easiest is just to put a chip in the receiver to generate some random noise called as "white noise". The problem is that listeners quickly become aware of the artificial nature of this comfort noise unless the variation is quite sophisticated.

4. ROUTING IN NGEO SATELLITE NETWORKS

This chapter addresses the routing and related issues in satellite networks. After a short summary of performance issues, previous work on satellite network architectures and routing in these networks is presented. Finally, the proposed algorithm is explained.

NGEO networks are an interesting type of mobile network in that the nodes are moving rapidly with respect to the slow moving or fixed user nodes, causing frequent link handoffs. Despite the highly time-varying nature of the network topology, there are some simplifying properties. First, most of the topology changes of the satellite mesh itself (aside from equipment failures) can be predicted in advance. Second, the graph topology is somewhat regular and dense, leading to a multiplicity of similar routes to most destinations. Both of these simplifying properties can potentially be exploited by routing algorithms. Nevertheless, when compared with routing protocol design for terrestrially-based packet networks, there are several fundamentally different design objectives that complicate the design. First, satellite hardware will continue to be mass and power constrained, thereby limiting the amount of on-board memory and processing. Although it is true that advances in electronics technologies will continue to make memory cheaper and less power-consuming in future years, the satellite payload is still a very power-constrained network node, with as much power as possible allocated to signal transmission. Therefore new routing protocols should be memory efficient and not computationally intensive. Second, conservation of link bandwidth is important because of loss of capacity for user traffic on these expensive links. Third, economic factors limit the number of satellites that can be deployed in a constellation, and consequently cause the coverage footprints of satellites to be stretched thin. For the above reasons, operating traditional distributed routing protocols and using traditional means of hierarchy are not likely to provide the best performance.

Distance vector protocols have well known convergence problems in time-varying topologies. Link state protocols converge much more rapidly upon topology changes, at the expense of a large amount of message traffic, higher protocol complexity, and routing computational overhead. Of course, either distance vector or link state protocols can be made to work in NGEO satellite systems; the point is that because such protocols do not capitalize on the simplifying aspects of NGEO network properties, one is likely to do better with more specialized protocols. In summary, the major challenge in the design of packet routing algorithms for NGEO networks is coping with both a time-varying topology and constraints on key system resources, while trying to capitalize on certain (simplifying) properties of the network topology [3].

4.1. Performance Issues

Various applications can be classified into different groups, based on their latency and error tolerance requirements. ITU-T G.1010 [20] provides a recommendation for various applications considering the packet loss and delay values of those applications. This recommendation defines a model for multimedia Quality of Service (QoS) categories from an end-user viewpoint. By considering user expectations for a range of multimedia applications, eight distinct categories are identified, based on tolerance to information loss and delay. It is intended that these categories form the basis for defining realistic QoS classes for underlying transport networks, and associated QoS control mechanisms. Figure 4.1 summarizes this recommendation.

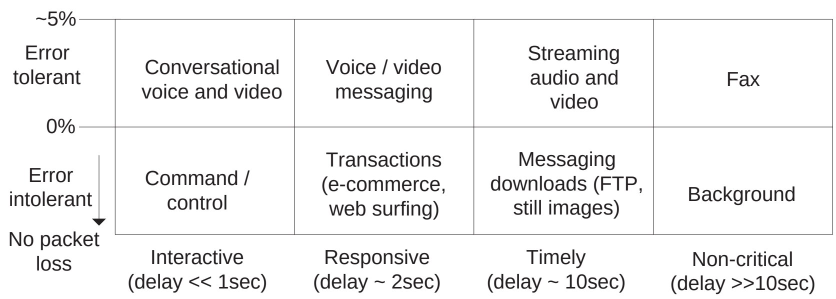


Figure 4.1. Model for user-centric QoS categories (ITU-T G.1010)

Another recommendation by 3GPP [21] groups the applications into four different QoS classes: conversational, streaming, interactive and background. The main distinguishing factor between these QoS classes is how delay sensitive the traffic is. Conversational class is meant for traffic which is very delay sensitive while Background class is the most delay insensitive traffic class. Conversational and streaming classes are mainly intended to be used to carry real-time traffic flows. The main divider between them is how delay sensitive the traffic is. Conversational real-time services, like voice and video telephony, are the most delay sensitive applications and those data streams should be carried in conversational class [21].

The performance metrics of an application provided by a satellite network can be listed as below:

- **Delay:** Time for a packet to be transmitted from sender to receiver. There are many contributing factors to delay, mainly propagation delay, processing delay and queueing delay. The most dominant factor of delay in satellite networks is propagation delay. However, in case of congestion, queueing delay might become more dominant.
- **Delay Jitter:** Variations in voice packet inter-arrival time are known as jitter, generated in satellite system and IP networks due to their multiple access schemes and variations in router queue loading in time. Jitter must be absorbed at the listener side to play out the voice sample in right timing, otherwise it degrades the voice quality and if it is large, a voice conversation is not possible. For this reason dynamic jitter buffering at the listener side is required. The proper selection of buffer size for specific IP network and satellite system characteristics is important to reduce overall delay and at the same time reducing the jitter noise to a minimum.
- **Bandwidth:** Maximal data transfer rate that can be carried. Proper allocation of bandwidth determines the system throughput.
- **Packet Loss:** Packet loss may occur in IP networks due to network congestion causing routers to discard incoming packets. Voice packet loss in satellite links occurs as bits in a packet may be corrupted in the transmission path. Voice packet loss in satellite links can be countered by proper satellite link budget and use of Forward Error Correction (FEC). Proper IP network design and implementation of priority scheme for voice packets can reduce this voice packet loss. Even with proper network design, packet loss still may occur and cause disturbing noise

resulting in the degradation of voice quality. Taking into consideration the human ears characteristics, it is possible to reduce the effect of packet loss by repeating the last voice sample or inserting noise to fill the gap created by the lost packet at the receiving side.

Table 4.1. ITU-T G.114 One-way Delay Recommendations

One-way Delay (ms)	**Description**
<150	Acceptable for most of the applications
150–400	Acceptable but not good quality
$400<$	Unacceptable for most of the cases

The delay requirements for voice communication specified by ITU-T G.114 standard is shown in Table 4.1. In order to satisfy the QoS requirements of applications, management of system resources and network elements is essential. System resources should be efficiently used whilst fairly servicing to all users. Main objectives of traffic management can be listed as : fairness, resource utilization efficiency, bounded queueing delay, stability and scalability. Traffic sources should be treated according to some fairness criteria. System resources i.e. network buffers, link bandwidth, should be efficiently utilized. In order to guarantee low end-to-end delay, some congestion avoidance schemes should be applied to prevent buffer overflows and in turn excessive queueing delays. Additionally, system must converge to a stable state and should not fluctuate in that state and traffic management must be scalable also in case of large number of users [22]. Detailed information will be presented in further sections.

4.2. Related Work

In the following subsections, a literature survey on multi-layered satellite networks and concerning issues is presented.

4.2.1. Multi-layered Architectures

A multi-layered architecture can consist of a combination of satellites at different orbits i.e LEOs, GEOs, MEOs. These networks may also deploy High Altitude

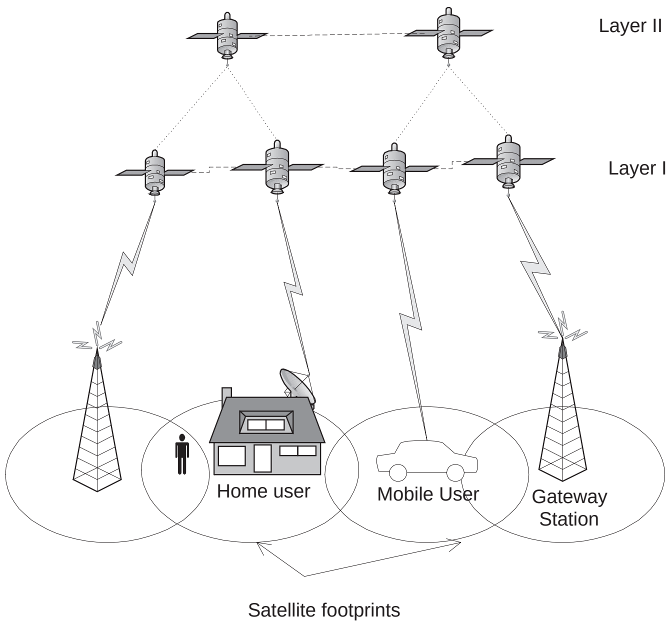


Figure 4.2. A multi-layered satellite network of two layers

Platforms (HAP). HAP systems are airships or aircraft stationed in the stratosphere, at altitudes between 17 and 22 km, to provide wireless communications infrastructure. ITU has assigned two bands of mm-wave frequencies for Broadband Fixed Wireless Access (BFWA) services from HAPs: one at 47-48 GHz (worldwide) and one at 28-31 GHz (40 countries including Russia and most of Asia) [23]. In [24], a three layered architecture consisting of GEOs, LEOs and high altitude platforms (HAPs) is proposed. GEOs act as backbone routers, LEOs as the second layer and HAPs to cover special areas with high and sensitive traffic such as battlefields and disaster areas. In [25], two layered architecture dubbed as "Satellite Over Satellite" (SOS) consists of LEOs in the lower layer and MEOs at the top layer as the system depicted in Figure 4.2. The authors also propose a routing protocol for broadband satellite communication networks - *hierarchial QoS routing protocol* (HQRP) that supports simple routing protocol for long distance multimedia traffic.

In [26], a new *Satellite Grouping and Routing Protocol* (SGRP) on a two-layered satellite network of LEOs and MEOs is developed. LEO satellites are divided into groups according to the the footprint area of the MEO satellites. Based on the delay reports sent by LEO satellites, MEO satellite managers compute the minimum-delay

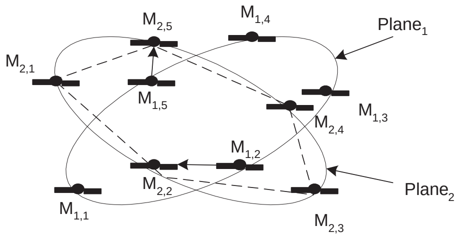


Figure 4.3. Inter-plane message exchange in MEO Layer

paths for their LEO members. The main idea of the SGRP is to transmit packets in minimum-delay path and distribute the routing table calculation for the LEO satellites to multiple MEO satellites. Since routing table calculation is shifted to MEO satellites, power consumption is effectively distributed between the two layers. Routing table calculation is done by a sequence of message exchanges as in Figure 4.3, i.e. inter-plane and intra-plane exchange of delay reports. This routing protocol is also robust to satellite failures and congestion on the links. Every LEO satellite continuously monitors the queue lengths of the output buffers of their adjacent links. If the queue length associated with a link is more than a predefined number, this signals the occurrence of congestion. The delay value associated to this congested link is set to ∞, and MEO manager (or managers) of this satellite is notified of this change. Affected paths are recalculated by MEO managers. Therefore, this link will not be in any path till congestion is overcome. In this scheme, MEO layer is used for managerial purposes and data is actually transmitted via LEO layer. Hence, one drawback is the waste of MEO layer resources such as bandwidth. Authors of this paper have also some other work in the field which can be listed as [27], [28], [29] and [30].

In [25], the system topology is analyzed to estimate the minimum required number of satellites at each orbit to provide global coverage. Minimum elevation angle and orbit types i.e. equatorial or polar, are taken into consideration to determine the number of satellites. Finally, it is concluded that at that specific SOS architecture, for a global coverage at 10 degree minimum elevation, only 10 LEOs and 3 or 4 MEOs are needed. Figure 4.2 shows an example of two-layered satellite network. Last but not least, Wu et al. in [31] models a double layered network of LEO and MEO satellites by using generalized stochastic Petri net (GSPN) model. Other valuable work on multi-layered satellite networks can be found in [31], [32], [33], [34] and [35].

4.2.2. Traffic Load Balancing and Congestion Avoidance / Detection

Due to changing topology of NGEO satellite networks and nonuniform traffic distribution in satellite footprints, some ISLs will be heavily loaded while some ISLs being underutilized. This will in turn lead to congestion on the heavily loaded links and will ultimately result in higher queueing delays and packet loss due to buffer overflows. To prevent from congestion, balancing the traffic load on the links is essential. References [26], [36], [37], [38], [33] and [39] discuss congestion detection and congestion avoidance, and suggest some mechanisms. In [36], a traffic congestion avoidance scheme based on real time traffic information is proposed. In *Explicit Load Balancing* (ELB), a satellite continuously checks its queue size to determine its state which may be *free, fairly-busy and busy.* If ratio of queue size to the total queue length - Q_r is under a threshold value α, its state is marked as *free* meaning that this satellite can be utilized on the path to the destination. If Q_r is between two thresholds α and β, satellite is *fairly-busy.* Finally, in case of Q_r being greater than β, this satellite is accepted as *busy.* A change in the state of satellite is immediately broadcasted to the neighbors of the satellite by Self-State Advertisement packets. Neighbors update Neighbors Status Lists (NSL) and cost of the links between the busy satellite and its neighbors is increased. NSL carries information of queue state of each neighboring satellite. Neighbors forward some portion ($\chi\%$) of traffic to other paths. This scheme therefore alters the traffic sending rate of neighboring nodes of the satellite in question before it gets congested. Since minimum cost links are preferred, packets will be routed on the least loaded links and busy links will therefore have less packets in the queues and will soon become free. Appropriate adjustments of the parameters α, β and χ would result in efficient distribution of traffic over multi-hop satellite constellations. However, this scheme only applies to delay-insensitive applications. In [40], traffic classes are identified by maximum acceptable delay bounds. Delay sensitive traffic has a privilege i.e. a quantum of bandwidth is allocated for high priority traffic. As many other works in the field, time is divided into discrete intervals. Therefore, traffic allocation problem is divided into two sub problems: periodic and incremental.

At beginning of each time interval (periodic allocation) and when a new call arrives (incremental), traffic allocation scheme is applied. The proposed multiservice routing algorithm - GALPEDA, uses different programming facilities like genetic algorithms and linear programming. Poisson and Markov Modulated Poisson Process (MMPP) traffic models are considered, the latter corresponding to bursty traffic. Fair traffic distribution is achieved by forcing low priority traffic to use lightly loaded links.

Jianjun et al. in [39] similarly use queueing information of a satellite to balance the traffic load on each satellite. The proposed algorithm *Compact Explicit Multi-path Routing* (CEMR) consists of three components: route discovery, route maintenance and traffic allocation. Before route discovery phase, special satellites so called "plane speaker" collect link state information of others in the network and build routing information base (RIB). In SOS architecture proposed by Lee et al., each LEO satellite sends its link state information directly to the upper layer MEO. Specialization of some satellites as plane speaker helps decreasing the signalling overhead. The resulting RIB is distributed to all satellites in the system. Link state information is the cost of the link which has two components: propagation delay and expected queueing delay at that node. Propagation delay is deterministic as opposed to queueing delay, therefore can be calculated using predictive ISL length information. Expected queueing delay is calculated according to the number of packets in the outgoing queue of the ISL. Each node calculates two paths: shortest path and minimum cost path. Shortest path can be computed using some well known shortest path algorithms. The sequence of hops and links is encoded using a globally-known unique short path identifier. This scheme also considers the system period as consisting discrete time stamps. CEMR applies some path validation to guarantee loop-free and valid paths. This policy determines whether to forward the packet on the default shortest path or the calculated minimum cost path. Since packets may be routed through multi paths and each path is compactly encoded, this scheme balances the traffic loads with lower signal overhead compared to traditional multi-path routing algorithms. A novel constraint-based routing algorithm on a multi-layered satellite network is introduced in [33]. Delay and jitter sensitive traffic is differentiated from other less sensitive traffic by class based queueing. Moreover, bandwidth availability and bit error rate (BER) of the links are taken into account while calculating route for high priority traffic.

4.2.3. Adaptive Routing Mechanisms

An efficient system should apply a QoS scheme that discriminates packets depending on the traffic classes. To treat each traffic class in a special way, the network structure must be capable of distinguishing between packets by means of classification and scheduling packet queues separately as a result of the classification [22]. In [25], traffic class dependent routing is employed on the two-layered SOS network. A connection request from the user is classified at the LEO layer as either Short Distance Dependent Traffic (SDD) or Long Distance Dependent Traffic (LDD). The first LEO that gets a connection request named as "source satellite" finds a feasible path based on its Global Routing Information (GRI). This path satisfies the delay constraint of the connection i.e the path's expected delay is less than delay bound of this call. If the number of hops of the calculated path from source to destination satellite n_{sd} is smaller than a threshold hop count N_{hop}, then this call is SDD. Otherwise it is LDD. LDD is routed via MEO layer to minimize hop count and in turn transmission delay. SDD is routed via LEO layer satellites. After this layer selection phase, optimal paths are calculated at the chosen layer. Details of optimal path calculation are not given in this study. But it is mentioned that this routing scheme balances the traffic and minimizes the load on the bottleneck link. Call admission messages are sent to all nodes along the calculated path. If admission is not achieved at a node, this call is accepted as *blocked*. The performance of SOS and single layer satellite networks are compared in the first simulation set. Using different values of N_{hop} and changing number of satellites in the LEO layer, system parameters for optimal performance are adjusted. Call blocking rate and average delay values are the performance comparison metrics. Although not explicitly stated in the paper, HQRP differentiates the delay-sensitive and best effort traffic by call admission at each node. On the other hand, real time network conditions are not taken into consideration to route a packet. Because of the connection-oriented nature of the proposed routing scheme, it is not suitable for such a system having dynamic topology.

4.3. Proposed Method : Adaptive Routing Protocol for QoS (ARPQ)

4.3.1. Background and Definitions

In this section, we present some background and relevant definitions for our satellite system.

Source satellite : The satellite which covers a specific user and connection request of this user is firstly received by this satellite.

Destination satellite : The satellite which covers the user who is the target of a communication.

System Period (T_s) : Most of the studies in the field such as [41], [26], [39] and [40], consider the system as a union of states at sufficiently small time intervals. System period T_s is the lowest common multiple of the satellite layer's orbital period and the Earth's period. This period is divided into small time intervals at which system topology is regarded as static. In this way, changing topology of the network is reduced to problem of managing states in T_s that is periodic. We also divide the constellation period into small time slots t.

LEO Group (LG_i) : LEO satellites in the footprint of MEO_i form a group and this LEO group is shown by LG_i. MEO_i is named as *"group manager"* and shown by GM_i. Each group has only one GM and all group members are aware of their GM. Actually a LEO might be covered by more than one satellite, but we assume that the MEO with the longest service time (depending on the satellite calendar) is designated as GM. Figure 4.4 shows a two-layered satellite network and its components.

Link State (LS_i) : LEO_i stores delay information LS_i associated with all its output links.

Plane Manager (PM_i) : A special MEO satellite that is responsible of calculation

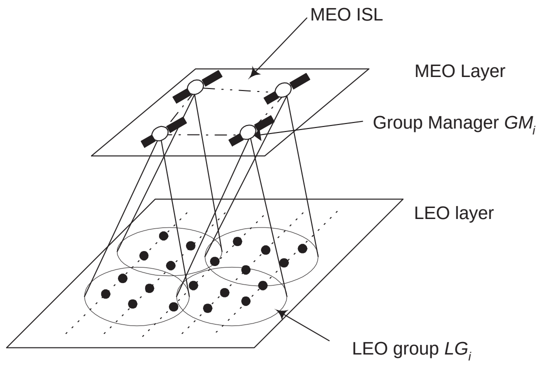


Figure 4.4. Two-layered satellite architecture

and distribution of routing table to the satellites in plane i. Each plane has only one PM and all plane members are aware of their PM.

Neighbor Status List (NSL_i) : A satellite i in the constellation has a NSL_i storing the state of the neighboring satellites. Satellites may be in one of the states: *free (lightly loaded), fairly-busy (fairly loaded)* and *busy (heavily loaded).*

Queue Ratio (Q_r) : The ratio of the current number of bits in a queue at time t denoted by $N(t)$ to the queue capacity in number of bits Q_s.

$$Q_r = \frac{N(t)}{Q_s} \tag{4.1}$$

4.3.2. Routing Table Calculation and Distribution

We apply a virtual node (VN) scheme as in [42]. In VN scheme, satellite positions are assumed to be fixed in the space, and only the actual satellite passing overhead is changing. We consider the system as union of time intervals. At the beginning of each time interval, coverage area of each satellite is updated using the VN topology. Furthermore, all LEO satellites determine delay values of the links associated with each neighboring satellite. A link delay consists of two components : propagation delay (t_p) and queueing delay (t_q). Intra-plane ISL propagation delays are always fixed and therefore can be computed offline. However, the length of inter-plane ISLs are variable and thus, the propagation delay on them is changing all the time in company with the constellation [39]. Furthermore, using periodicity of the satellite topology due to its predetermined motion in its orbit, dynamic inter-plane ISL propagation delays can also be calculated offline and then can be uploaded to the satellites. ISL propagation delays can be calculated using the formulas listed beow [40]. The first equation gives the intra-plane ISL propagation delay and the latter inter-plane ISL propagation delay.

$$\mathbf{ISL_{intra}} = \frac{\sqrt{2}}{c} \times (R_{Earth} + h_{sat}) \times \sqrt{1 - \cos(\frac{2 \times \pi \times n_p}{N})} \tag{4.2}$$

$$\mathbf{ISL_{inter}} = \frac{\sqrt{2}}{c} \times (R_{Earth} + h_{sat}) \times \sqrt{1 - \cos(\frac{\pi}{N})} \times \cos\theta \tag{4.3}$$

where

R_{Earth} : Radius of Earth- 6378.137 km

h_{sat} : Height of satellite above Earth

N : Total number of satellites

n_p : Number of planes

θ : Latitude of the satellite

c : Speed of light (3×10^5 km/s)

At each time interval the routing tables are updated regularly to cope with the satellite mobility and link load changes. Initially at the system start up, routing table calculations are done using only the propagation delay of ISLs. However, processing and queueing delay must be taken into account to make better optimal path calculations. Expected queueing delay at a node $t_q(L)$ can be predicted using the queue size of the outgoing ISLs using Equation 4.4, where L_{av} is the average packet length, C the link capacity, and $N_q(t)$ the number of packets in the queue at time t. A link L between two satellites has total delay $t_{link}(L)$. The queueing delay is integrated over the time interval in [43]. However, in order to mitigate unnecessary processing on-board the satellites, we calculate queueing delay by getting samples at some certain time instants and get the average of all these sample values. Number of samples taken can be adjusted according to the length of a time interval t.

$$\mathbf{t_q}\left(\mathbf{L}\right) = N_q\left(t\right) \times \frac{L_{av}}{C} \tag{4.4}$$

$$\mathbf{t_{link}}\left(\mathbf{L}\right) = t_p\left(L\right) + t_q\left(L\right) \tag{4.5}$$

Similar to other hierarchical routing schemes as [26] and [25], routing table calculation is performed by the higher layer which has better knowledge of the whole network topology. MEO layer collects the link state information of all LEOs and prepares consequent routing tables. More briefly, upon completion of link state information collection, LEO_i directly sends its Link State (LS_i) to its manager satellite GM. After each MEO gets all LS information from the managed LGs, message exchange phase starts to inform other MEOs in this layer about the local topology of the lower LEO group. Each plane manager PM calculates its routing table and distributes this table to the MEOs in its plane. The routing table has entries for each node in the network. Each entry has a destination, next hop and link cost field. Actually link cost is the total delay associated with the link L and is equal to $\mathrm{t}_{link}\left(L\right)$. Upon receipt of the routing table, each MEO distributes it to all LEOs in its footprint. Although some deviations exist, routing table updates and information exchange are mainly based on the method defined in [27]. In our scheme, we do not consider the seams where two ISLs are switched off due to the motion in opposite directions. Thereby, we assume that at any time there are four ISLs belonging to each LEO satellite.

4.3.3. Packet Classification

Since conversational traffic performance is highly dependent on delay and delay jitter values, there is a need for packet type-based routing. In our scheme, voice packets are classified by the source LEO satellites according to the distance between the source satellite and the destination satellite. A path is assigned for each packet by the source LEO using the routing table. Shortest propagation delay path is calculated using Dijkstra's Shortest Path Algorithm. Minimum delay paths are accepted as optimal paths. Total expected delay of a path p consisting of links shown by k is formulated in Equation 4.6. t_{uplink} and $t_{downlink}$ are the propagation delays from source GS to source leo and destination leo to destination GS respectively.

$$\mathbf{t_{path}} = t_{uplink} + \sum_{\forall k \in p} t_{link}(k) + t_{downlink} \tag{4.6}$$

Algorithm 1 Pseudocode of the packet classification algorithm running on LEO_i

Require: Satellites can distinguish voice packets and background packets.

Ensure: Classification of a packet arrived to LEO_i.

```
Arrival of a new packet to source satellite LEO_i
if voicepacket then
    Extract destination address GS_k
    Find the destination satellite LEO_j covering GS_k
    Find the cost of path t_path from LEO_i to LEO_j
    if t_path ≤ D_thrsh then
        Mark the packet as SDV
    else
        Mark as LDV
    end if
end if
```

The pseudocode of packet classification algorithm is given in Algorithm 1. If the calculated path's delay is greater than threshold delay D_{thrsh}, then this voice packet is marked as *Long Distance Voice* (*LDV*). Otherwise, it is *Short Distance Voice* (*SDV*) packet. In [25], packets are classified according to the calculated path's hop count. Since ISL lengths are noticeably different from each other at different parts of the Earth i.e. at polar regions and Equator, hop count does not reflect the real delay values. Hence, we base our marking scheme on ISL delays rather than hop count of the path. Initially, SDV and *Non real-time* (NR) packets are forwarded to the next hop of the calculated path on the LEO layer. In the following sections we refer to NR packets as "background" packet. LDV traffic is forwarded by the source LEO to its GM. After getting the packet, MEO assigns a new path and forwards the packet to the next hop either in MEO layer or LEO layer. Using the MEO layer especially for time-sensitive traffic both balances the link utilization rates and prevents excessive jitter and delay values. The pseudocode of the ARPQ scheme running on a LEO satellite is given in Algorithm 2.

Algorithm 2 Pseudocode of the ARPQ algorithm running on LEO_i

Require: Queueing delay values are sent by each LEO to the corresponding GMs and routing table updates are completed at the beginning of each time interval t

Ensure: Routing of a packet according to the traffic type and link loads.

if $LDVpacket$ **then**
    Forward to GM_i
else
    Get next hop node: LEO_j from Routing Table
    if $SDVpacket$ **then**
        Send to LEO_j
    else
        Check link state of ISL_j from NSL_i
        if $state_j = FREE$ **then**
            Send packet to LEO_j
        else
            Select neighbor node l with Min(Q_r)
            Forward $\lambda\%$ of traffic to $node_l$
        end if
    end if
end if

4.3.4. Link State Assignment and Traffic Load Balancing

Each LEO and MEO satellite continuously checks its outgoing link buffers to detect a sign of congestion. If LEO queue ratio Q_r is under a threshold value α, there is no sign of congestion. If it is between two thresholds α and β, this can be accepted as an indication of forthcoming congestion and thereby some action has to be taken [36]. There is an important point to be considered in link state assignment phase. If α and β are the two thresholds, the queue ratio will oscillate around α. For instance, assuming α=0.5, the link state will become *fairly-busy* if Q_r is 0.51. Upon a change in the link state, the new state and new Q_r will be advertised. The LEO will forward some traffic to other less loaded links till the *fairly-busy* becomes $free$ with a queue ratio of 0.49. Thus, Q_r values will oscillate between 0.49 and 0.51, causing very frequent state advertisements. To tackle this issue, two more threshold values κ and Φ are introduced. The pseudocode for state assignment can be seen in Algorithm 3.

Link state assignment depends on the previous state of a link. Using the previous state information, unnecessary state oscillation is prevented. If a link becomes *fairly-busy*, its queue ratio is decreased till queue ratio is below κ. Similarly, a *busy* link's traffic is spread over other links till its queue ratio is below Φ. If a link advertises its state as *fairly-busy* or *busy*, some portion (λ) of NR traffic is forced through other ISLs. Each LEO has a Neighbor Status List storing the neighbor's status and queue ratio Q_r of the ISL between these nodes. From NSL_i, the outgoing link ISL_l with the smallest Q_r is chosen. Traffic forward rate λ depends on the state of the congested link. If it is busy, then no background traffic is routed over this link. All traffic is split to the other links. Traffic forwarding may cause loops. To prevent loops, a packet is not routed back to its previous hop and there is a hop count threshold to prevent packets traveling too long on the network.

4.3.5. On-board Scheduling

Queueing and scheduling policies are of great importance in order to implement efficient QoS provisions. The default scheduling mechanism for a satellite is First-in

First-Out (FIFO) scheduling policy where the packet entering the queue will leave the queue before the others arriving after it. However, to satisfy the performance requirements of delay and jitter sensitive voice traffic, it must be differentiated from delay tolerant background traffic. Due to satellites' processing limitations, queueing policy must be both simple and fast. Weighted Round Robin Queueing (WRRQ) is quite efficient for on-board processing in that sense. Strict priority may be an alternative policy. However, this policy may lead to suffering of data packets of high delay values and may cause "starvation" anomaly. In our scheme, voice traffic has priority over background traffic. For each traffic class, a portion of satellite's processor is reserved, w_{conv} and w_{bg} respectively. Depending on the values of w_{conv} and w_{bg}, prioritization of one traffic type can be achieved. Since voice traffic is delay and jitter sensitive, it has higher priority and so higher weight w_{conv}. This will yield voice packets be processed faster. In other words, excessive queueing delay values experienced by voice packets due to heavy background traffic will be shortened. Thus, performance of voice communication will be better than usual case of satellites applying FIFO scheduling policy.

Algorithm 3 Pseudocode of state assignment algorithm running on each outgoing queue of satellite LEO_i

Ensure: Notification of LEO_i upon a state change in one of the outgoing queues.

Calculate Q_r
if $Q_r \geq \beta$ **then**
  NewState=$busy$
else
  if $PreviousState = free$ **then**
    if $Q_r \leq \alpha$ **then**
      NewState=$free$
    else
      NewState=$fairly\_busy$
    end if
  end if
  if $PreviousState = busy$ **then**
    if $Q_r \leq \Phi$ and $Q_r \geq \alpha$ **then**
      NewState=$fairly\_busy$
    else
      NewState=$free$
    end if
  end if
  if $PreviousState = fairly\_busy$ **then**
    if $Q_r \leq \kappa$ **then**
      NewState=$free$
    else
      NewState=$fairly\_busy$
    end if
  end if
end if
if $NewState <> PreviousState$ **then**
  Advertise new state and new Q_r to LEO_i
end if

5. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

In this section, we outline the simulation details; tools, models, scenarios and parameters used for performance evaluation of the proposed scheme. In this thesis study, OPNET Modeler 10.5A [44] is used to model and simulate the network. The software runs on a Pentium 333 MHz Processor 800 MB RAM machine with Windows XP operating system. OPNET is a discrete event simulation package which is particularly useful for simulations involving data networks of any kinds, as great scope of abstraction is available: from the low level process models all the way up to end user terminals and cell phones. This work is implemented in the lower abstraction levels of OPNET, mainly in the process model. OPNET Wireless module simulates wireless links using a modular and open framework, called the Transceiver Pipeline^{TM}. The Transceiver Pipeline is fully customizable and designed to efficiently calculate transmission and propagation delays, link closure, bit error rates and error correction and many other satellite link related parameters. OPNET implementation details and models are given in Appendix A.

5.1. Simulation Setup and Scenarios

In the first set of experiments, to validate that multi-layered satellite networks perform better than single plane satellite networks, three simulation scenarios are created. In the first scenario, an Iridium-like satellite constellation is set up. In the second scenario, ICO (Intermediate Circular Orbit) system parameters are utilized to form a MEO constellation. Finally, two-layered satellite constellation of LEOs and MEOs is simulated. The other set of experiments are the ones analyzing the performance of proposed adaptive routing protocol. In the further sections, each scenario components and system properties are analyzed in details. Since the proposed scheme has a number of system parameters, the simulation studies will give an idea on how to adjust these parameters properly. However, it should be noted that these settings are valid for our constellation and parameters. Simulations are run 10 times with different seeds and the average values of results derived from these runs are calculated.

5.1.1. Scenario I: Validation of Multi-layered Architecture Performance

LEO constellation : In this LEO constellation we utilize Iridium system parameters. There are 66 LEO satellites distributed in 6 planes each consisting of 11 satellites. Satellites are identified by their unique numbers between 1-66 and shown as LEO_{id}. Each LEO is connected to two neighbors in the same plane and two other satellites in the neighboring planes by inter-satellite links (ISL). Gateway stations (GS) are directly connected to satellites via user data links (UDL). The world is divided into 6 coverage areas corresponding to the six continents. There are 44 GSs in each region. Each GS is also identified by a unique number between 1-264 and specified as GS_{id}.

Although the world's population, its distribution and communication patterns imply nonuniform traffic density in practice, this nonuniformity is not taken into consideration to keep scenarios simple and easy to manage. Actually, more realistic traffic modeling can be found in [45] and [46]. In [45], Korçak and Alagöz modeled traffic depending on the statistics of the user density levels and host density levels per continent. Additionally, user traffic generation rate changes during the day. This is also considered in [45]. Similarly, Mohorcic et al. in [46] model traffic utilizing the percentage of traffic flows between continents. In the simulations, we use the default satellite link model of OPNET and change some parameters of the channels like frequency and bandwidth. VoIP traffic patterns are created as duplex and symmetric voice communication streams. For a more realistic modeling, GSs generate also background data traffic. All sources are modeled as poisson traffic sources with exponentially distributed packet inter-arrival times. Packet size is also assumed to be exponentially distributed with a mean value of 1 KB.

When the scenario starts running, GS source-destination pairs are uniformly chosen and they generate packets for the entire duration of simulation. A GS sends packets to the corresponding LEO in sight. After LEO receives a packet, it checks the destination address to see if it is in footprint. LEOs are assumed to have knowledge about the network topology, each LEO is aware of GSs in its own footprint and also the other satellites' footprints. This information can be updated to all satellites by some special terrestrial stations or satellites can form overall network topology by some signalling exchange. If GS is in the coverage of the LEO, packets are forwarded directly to the destination GS. If not, LEO knows which is the corresponding LEO that has the destination in coverage. Since a LEO has direct communication links to only four neighboring satellites, it can route the incoming packet through one of these outgoing links. Determining the outgoing link depends on the destination satellite. Optimal shortest paths are determined using Dijkstra's Shortest Path Algorithm. Each LEO plane has intra-plane routing table of 11 elements. Each table entry has two fields showing the destination LEO and next hop LEO to reach that destination. Briefly, each LEO has the simplified route information (destination LEO and next hop LEO) to reach each satellite in its own plane. Finally, if satellites are in different planes, inter plane routing is done using plane routing table. Plane routing table is similar to LEO routing table. It has entries for each plane and destination plane-next plane pair information is stored here. Utilizing this table, LEO knows which plane to forward the packet. According to the destination plane, LEO routes the packet to the corresponding neighbor satellite in the specified plane.

Three different traffic types are modeled: short distance, long distance and random traffic [47]. In short distance communication, each GS is in communication with another GS in close vicinity, in other words both stations are in the footprint of the same LEO. Long distance communication represents long-haul or intercontinental communication. Random communication case allows random pairs of GSs to have voice and data sessions.

MEO constellation : In this scenario, ICO is chosen as reference MEO constellation. There are 10 MEO satellites placed into two planes in ICO's constellation. Actually, ICO satellites are bent pipe satellites, but in our case, they have inter-satellite links with the neighboring satellites [48]. Like LEO satellites considered in the previous case, each MEO satellite has four ISLs- two inter-plane and two intra-plane. Similarly, two routing tables are used for inter-plane and intra-plane routing.

Two-layered Constellation : Multi-layered satellite network, taken as reference constellation in our paper, consists of two satellite layers, LEO layer in the lower part and MEO layer at the top. LEO layer has the same constellation properties defined in the first scenario. MEO layer constellation is slightly different than only MEO satellite constellation case defined in the previous section. There are 6 satellites in MEO layer divided into two MEO planes, achieving global coverage. There are 3 MEOs in each MEO plane. LEO satellites in the footprint of MEO_i form a group and this LEO group is shown by LG_i. MEO_i is named as *"group manager"* and shown by GM. Each group has only one GM and all group members are aware of their GM. Actually a LEO might be covered by more than one satellite, but we assume that the MEO with the longest service time (depending on the satellite calendar) is designated as GM. There are ISLs among each MEO pairs. Previously mentioned LEO ISLs are still valid. Moreover, LEOs are also linked to their MEO group managers by IOLs. There are no direct links between GS and MEO, and GSs can communicate only with LEOs. The network model used in our simulations can be seen in Figure 5.1 and network system parameters are listed in Table 5.1.

Table 5.1. LEO/MEO parameters

Parameters	**LEO**	**MEO**
Altitude (km)	1200	10390
Number of satellites	66	6
Number of planes	6	2
Number of ISLs	4	5
Number of IOLs	1	11

Routing strategy is now different than the previous cases. Voice packets are classified by the source LEO satellites according to the distance between the source satellite and the destination satellite. A path is assigned for each packet by the source LEO using the routing table. Shortest propagation delay path is calculated using Dijkstra's Shortest Path Algorithm. Minimum delay paths are accepted as optimal paths. The packet classification strategy is as defined in the ARPQ algorithm. If the calculated path's delay is greater than threshold delay D_{thrsh}, then this voice packet

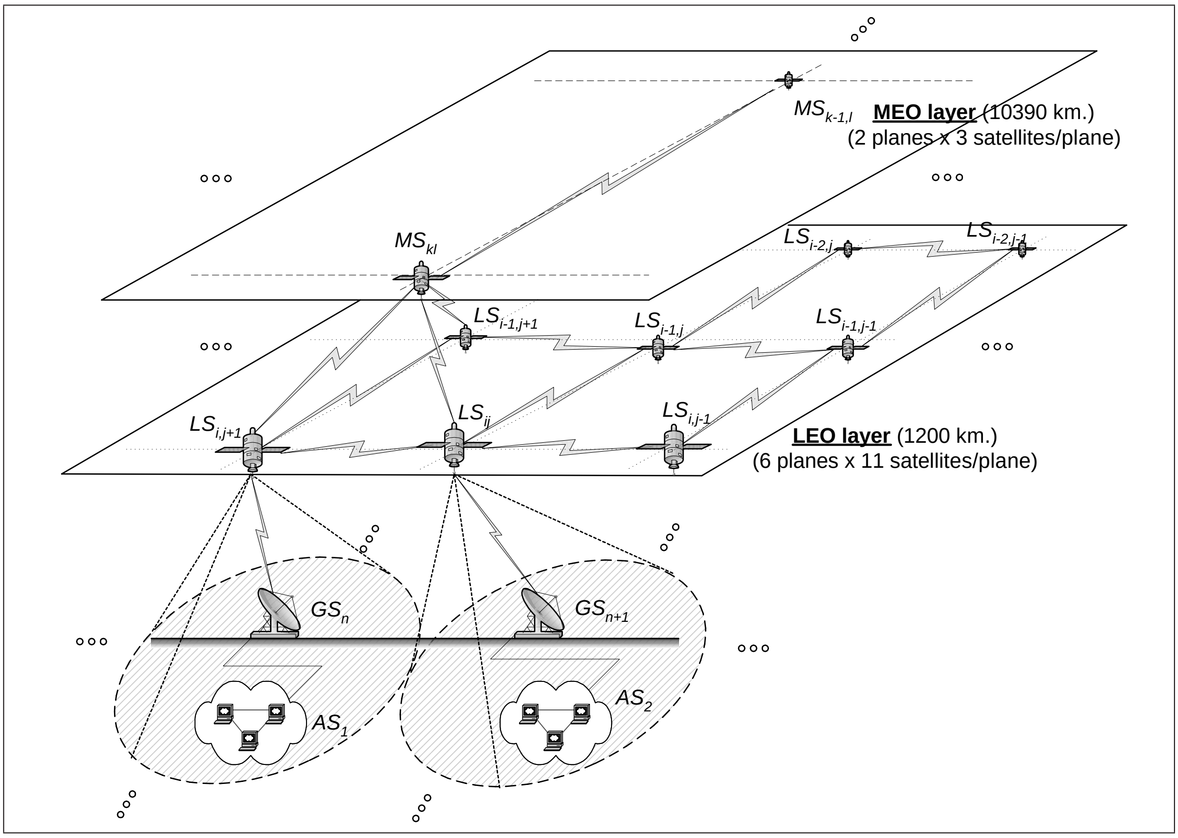


Figure 5.1. Two-layered satellite architecture used in the simulations of this thesis study.

is marked as *Long Distance Voice* (*LDV*). Otherwise, it is a *Short Distance Voice* (*SDV*) packet. However, in this two-layered network, to see the drawbacks of static routing mechanisms, there is no traffic load balancing as explained in previous section.

5.1.2. Scenario II: ARPQ Simulation Studies

In these simulation sets, we analyze the effect of various system parameters on system performance metrics. Three groups of scenarios are defined in this section. In the first group, D_{thrsh} value is changed and system performance metrics are recorded. Next, α and β values are changed. Finally, effect of on-board queueing mechanism is analyzed. The results may give us an idea of how to tune the system parameters for an efficient system. Note that the results are specific to our constellation and depends on the parameters we used, e.g., ISL capacities, buffer spaces and processor speed. These parameters are specified in Table 5.2. Not mentioned in the table, IOL capacity between LEOs and MEOs is 20 Mbps and each IOL has 1 Mb buffer space. Both the uplink (GS to LEO) and downlink (LEO to GS) capacities are 3.2 Mbps corresponding to 400 packets/s.

Table 5.2. Node (LEO/MEO) parameters of the simulation scenarios

Node Type	**ISL capacity**	**ISL Buffersize**	**Processor**	**Processor Buffersize**
LEO	10 Mbps	20 Mb	120 Mbps	2 Mb
MEO	80 Mbps	20 Mb	120 Mbps	4 Mb

The selection of parameters like link capacities is highly dependent on the hardware features of our machine on which the aforementioned OPNET simulations run. However, we believe that changing the parameters to more realistic satellite system parameters do not noticeably affect the simulation results.

5.2. Performance Evaluation

5.2.1. Scenario I

ITU-T G.114 recommendation states that one-way delay for voice transmission should be less than 150 ms. Delay values between 150 ms and 400 ms are acceptable but may result in poor quality of speech. Above 400 ms, it is unacceptable because of very poor speech quality. Therefore, it is important to note that long distance delays should be minimized. As can be seen from the summary of simulation results in Table 5.3, two-layered constellation of LEOs and MEOs can shorten delay values by forwarding packets to MEO layer preventing many LEO hops. Longest path of 9 hops defined in the LEO case is now shortened to 4 hops, LEO-MEO-MEO-LEO, resulting in much shorter delay. Similarly, jitter values also have smaller values changing between 4-16 ms. Forwarding some portion of traffic to MEO layer also facilitates load balancing. Instead of using all ISL capacities in LEO layer, MEO layer is utilized. This approach results in less traffic in LEO layer and less queueing delays.

Simulation results show that in LEO constellation, minimum delay is 30 ms which corresponds to a one-hop communication. The longest path results in a delay of approximately 214 ms in light traffic load. The average delay is 131 ms which corresponds

Table 5.3. Scenario I Simulation Results

Constellation	Long (ms)	Random (ms)	Short (ms)
LEO	210	200	40
MEO	280	220	90
Two-layered	170	160	30

to about 4-5 LEO hops. In MEO constellation, minimum delay - 90 ms - corresponds to one hop MEO path and maximum delay - 286 ms - belongs to the longest route of 5 MEO hops. Jitter value is nearly as high as LEO case because of multiple hops between MEOs. It comes to a conclusion that multi-layered satellite networks can provide better solution for voice services.

5.2.2. Effect of Threshold Delay D_{thrsh}

Changing D_{thrsh} value directly affects the utilization rates of LEO and MEO layer resources. Depending on the values of D_{thrsh}, SDV and LDV traffic percentage will change and therefore load on LEO layer will change. This interaction makes determining D_{thrsh} value a design issue. We run a set of simulations with different D_{thrsh} under changing background traffic load. In the following simulation sets, background traffic load is changed and the effect of D_{thrsh} on system performance metrics is analyzed. The conversational traffic rate is set to 25% of GS uplink capacity in all runs. The packet size is exponentially distributed and has a mean value of 1 KB. We simulated three scenarios: D_{thrsh}=10 ms, 500 ms and 80 ms. The first two cases correspond to two extreme scenarios where nearly all voice packets are marked as LDV in the first scenario and SDV in the latter. Setting D_{thrsh}=80 makes system more balanced where SDV and LDV ratio is nearly equal. The queue ratios are set to $\alpha = 0.4$ and $\beta = 0.8$. Traffic forwarding rate λ is set to 0.9 and 1 in the *fairly-busy* and *busy* states respectively. The following simulations are run for 120 s and system time interval is set to 10 s. Therefore, the scenario simulates one period of the system. In our simulations, we usually consider the case where voice traffic is 25% or less of GS uplink capacity and various background traffic rates. This assumption is quite realistic depending on the statistics of voice and data flow all over the world where background traffic is always more than voice traffic.

OPNET provides a statistics summary of results. This summary includes number of statistic data, minimum, maximum, variance, standard deviation and expected value of the statistic data. Additionally, results are classified according to some confidence interval (CI) values. In the following experiments, we take the results in 95% CI.

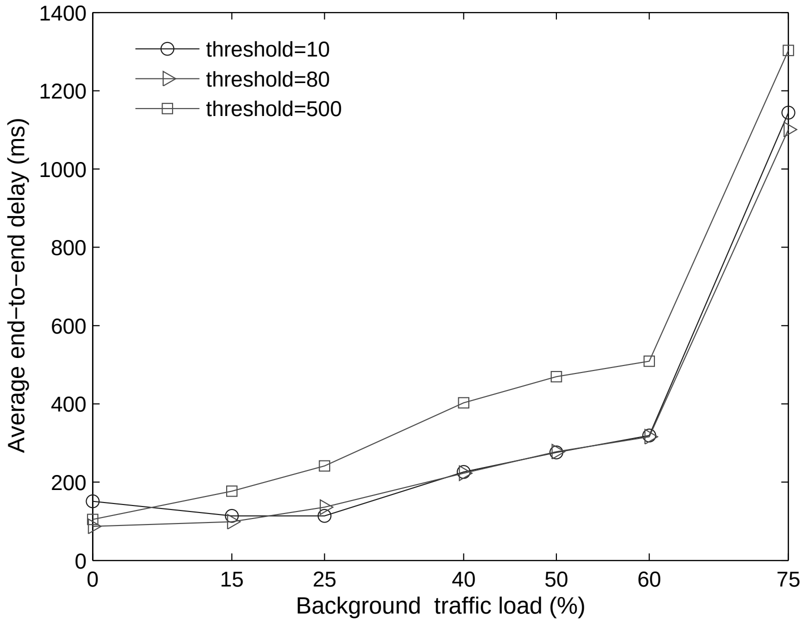
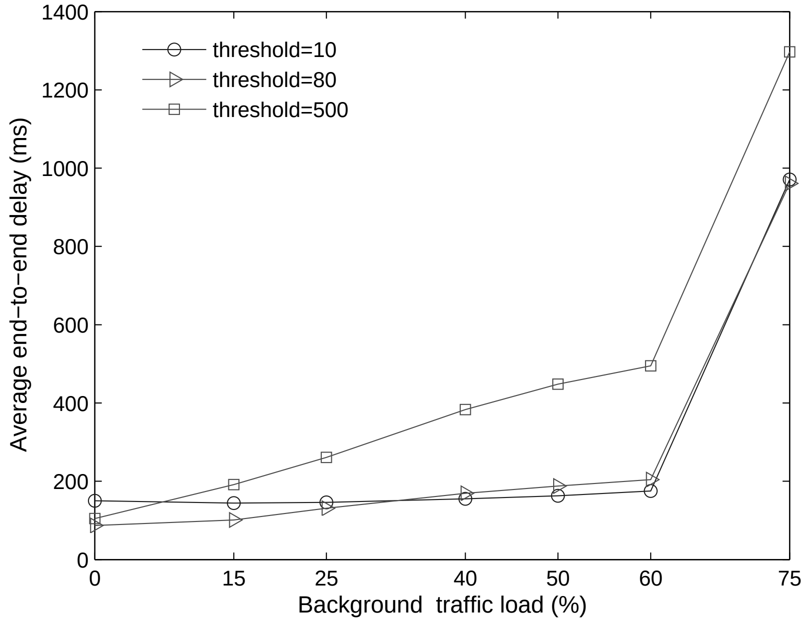
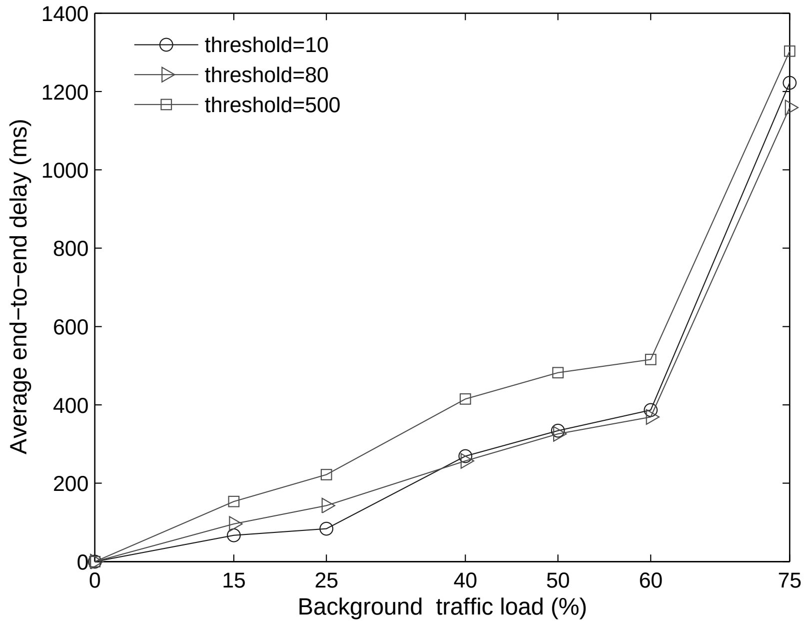


Figure 5.2. Effect of D_{thrsh} on overall average end-to-end delay under changing background traffic load.

The simulation results are plotted in Figures 5.2, 5.3, 5.4 and 5.5. In Figure 5.2, it is seen that average delay is about 150 ms when $D_{thrsh} = 10$ and there is no background traffic. 99% of voice traffic is classified as LDV and the remaining part (1%) corresponding to regional traffic of only one LEO hop is classified as SDV. The packets are forwarded to the MEO layer instead of lower LEO layer. Therefore, the average delay is more than the other two cases where D_{thrsh} is 80 ms and 500 ms with no background traffic. In case of $D_{thrsh} = 500$, all voice packets are marked as SDV and are routed through LEO layer. Under light background traffic this does not affect the system performance, since there is enough capacity for all types of traffic. On the other hand, with the increase in background traffic, voice traffic delay and jitter values also increase due to congested paths in the LEO layer. Since our mechanism forwards background packets to alternate paths in case of congestion (or a sign of congestion), there is not a significant difference in average delay values in scenarios where



(a)



(b)

Figure 5.3. (a) Effect of D_{thrsh} on conversational traffic delay values under changing background traffic load.(b) Effect of D_{thrsh} on background traffic delay values under changing background traffic load.

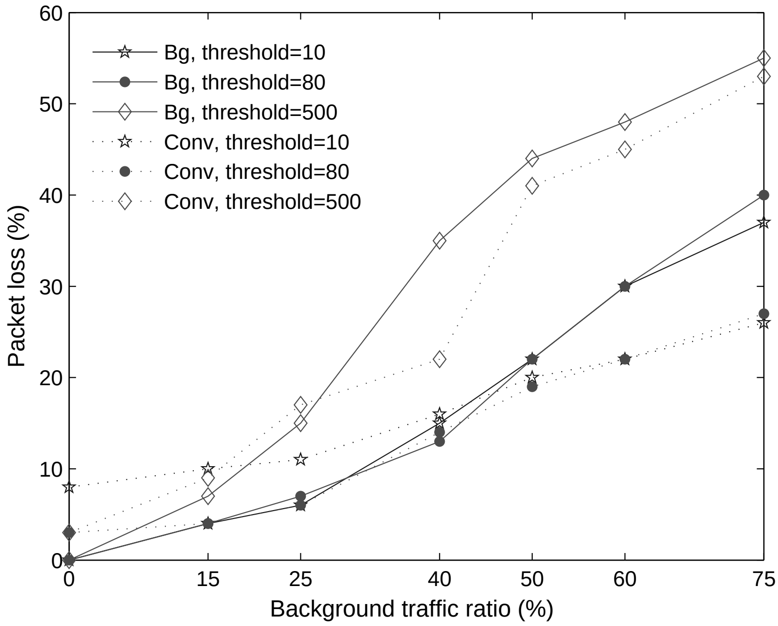
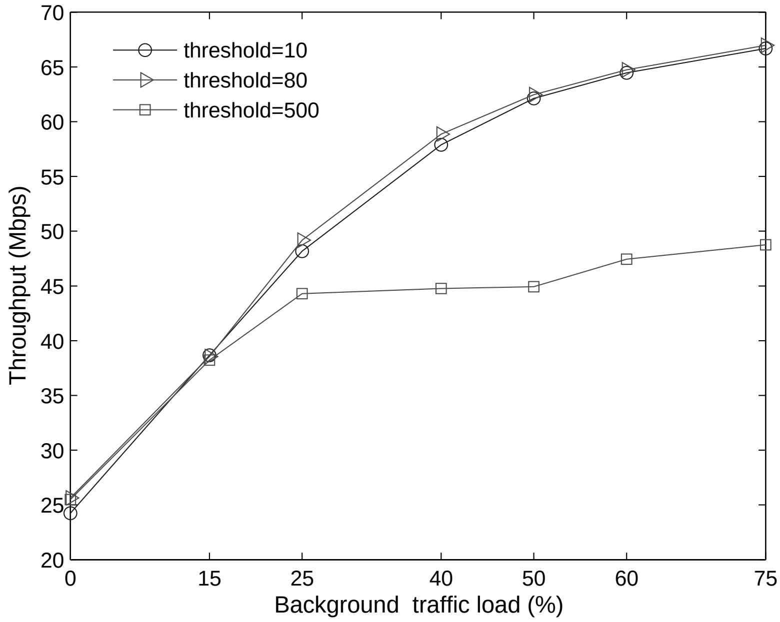


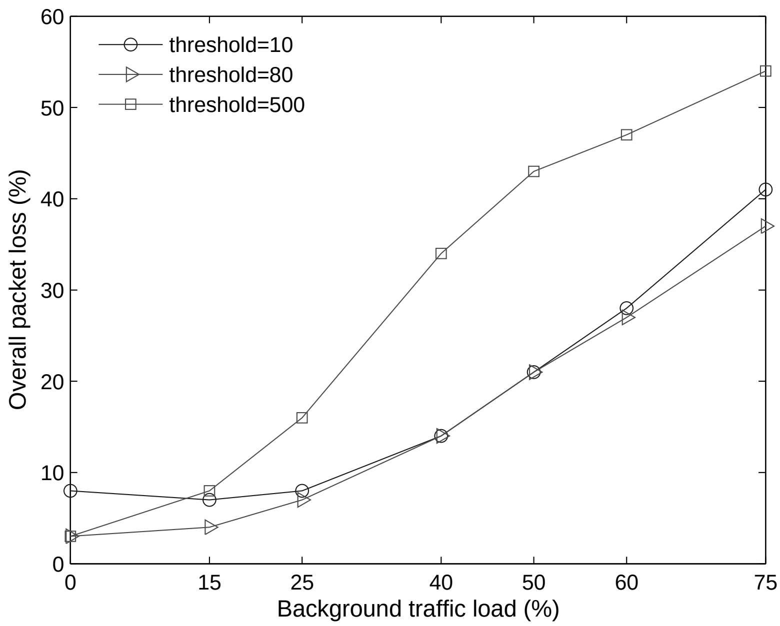
Figure 5.4. Effect of D_{thrsh} on packet loss rate of each traffic type under changing background traffic load.

$D_{thrsh} = 80$ and $D_{thrsh} = 10$. Figure 5.3(a) shows the delay values of each traffic type separately. Conversational delay values similarly increase with the increasing background traffic load. Under light traffic load, delay values are in acceptable region when $D_{thrsh} = 80$ and $D_{thrsh} = 10$. However, when the system uses its entire capacity (when voice traffic is 25% and background traffic is 75%), delay values of conversational traffic are far above the acceptable limits. As can be seen from Figure 5.3(a), average voice delay is nearly 1 s. With such delay values it is impossible to have an intelligible communication. Therefore there is certainly need for voice traffic prioritization on-board the satellite. Background traffic delay values follow a similar pattern as conversational traffic. Moreover, background delay values are higher than conversational delay values. This is due to traffic forwarding applied to background packets in case of queue ratio increase warnings.

When it comes to packet loss, similar to network delay simulation results, scenario with $D_{thrsh} = 500$ has the worst performance results. This is again due to congestion in LEO layer. Additionally, background traffic route is made longer in case of high traffic load. This may also increase the probability of packet loss. Packet loss and



(a)



(b)

Figure 5.5. (a) Effect of D_{thrsh} on network throughput under changing background traffic load.(b) Effect of D_{thrsh} on network packet loss under changing background traffic load.

throughput values of the simulations are depicted in Figures 5.4 and 5.5. Packet loss rates of background and conversational traffic behave similarly. Briefly, with the increasing values of background traffic load, packet loss rates also increase. The most steep change occurs when D_{thrsh} is 500 ms. From Figure 5.5(a), the overall traffic packet loss rate can be seen. Although there is not a significant difference between the cases where D_{thrsh} = 80 ms and 10 ms, it can be seen that it performs better when D_{thrsh} is 80 ms. This is because of the reason that setting D_{thrsh} to 80 ms yields nearly equal number of packets to be marked as LDV and SDV. Hence, voice traffic is split between the LEO and MEO layer thereby causing lower probability of congestion in both layers.

Table 5.4. Simulation Results - Delay Jitter Values

		Background traffic (%)						
D_{thrsh} **(ms)**	**Traffic type**	0	15	25	40	50	60	75
$D_{thrsh} = 10$	Background	-	23	35	184.5	202	228	550
	Conversational	21	20	20	25.5	30	38	444
	LDV	21	20	20	25.5	30	39	444
	SDV	1	1	2	1	1	1	13
$D_{thrsh} = 80$	Background	-	39	72	151	193.5	214	531
	Conversational	27	33	43	59	69.5	82	441
	LDV	13	15	20	28	31	38	382
	SDV	18	33	50	73	84.5	92	349
$D_{thrsh} = 500$	Background	-	81	169	309	404	306	581
	Conversational	65	122	145	209	220	306	533
	LDV	-	-	-	-	-	-	-
	SDV	65	122	145	209	220	306	533

Finally, delay jitter values collected from the simulations are listed in Table 5.4. Background traffic jitter is higher than conversational traffic due to varying route of background packets. LDV traffic has stable jitter values due the paths followed in MEO layer. SDV has usually larger jitter values than LDV because of paths composed of more satellite (LEO) hops. Increase in background load adversely affects jitter values, which shows the necessity of voice traffic prioritization.

To sum up, we can conclude that under heavy traffic load, utilizing MEO layer for voice traffic will be efficient by preventing congested paths on LEO layer. This can be achieved by setting D_{thrsh} to relatively low values, e.g., 20 ms corresponding to 1 LEO hop. Furthermore, voice traffic can be routed through LEO layer under light background traffic load . Setting D_{thrsh} to some higher value, e.g., 80 ms will ensure most of the packets to be routed through LEO layer.

5.2.3. Effect of Queue Threshold Values α and β

Depending on the α and β values, number of packets in the queues therefore queueing delay values change. Setting α to 1 makes the system behave as a nonadaptive system. Therefore we can compare our mechanism ARPQ to the nonadaptive case, ARPQ without load balancing. In the following experiments we set D_{thrsh} to 80 ms, $\kappa = \frac{\alpha}{2}$ and $\lambda = 90\%$. Setting D_{thrsh} to 80 ms, causes almost half of the voice traffic flow through LEO layer. In light traffic case, effect of load balancing will not be apparent. Therefore, the traffic load is 60% of GS uplink utilization rate. We change the first threshold α and κ (κ is set to $\frac{\alpha}{2}$), and analyze their effects on system performance. β is set to 1 or 0.8 (depending on the values of α) and Φ is set to $\frac{\beta}{2}$ in all of the following experiments. It should be noted that these choices are made arbitrarily with no special purpose in mind. Thereby, different values of these parameters can also be set and the effects on system performance can be examined.

Figures 5.6 and 5.7 depict the queue ratios of LEO_{53} with two different α values. Actually the queue ratios correspond to link utilization rates of these related links. In Figure 5.6, IOL is free of congestion and its utilization ratio is very low as opposed to ISL being overloaded. Since there is no load balancing in this case, there is an imbalance in link utilization rates which yields to excessive queueing delay values. Moreover, overflow traffic is dropped rather than being forwarded to less loaded links. On the other hand, in Figure 5.7 it can be seen that all ISLs and IOL utilization rates are similar. This is achieved by splitting the traffic over less loaded links. Moreover,

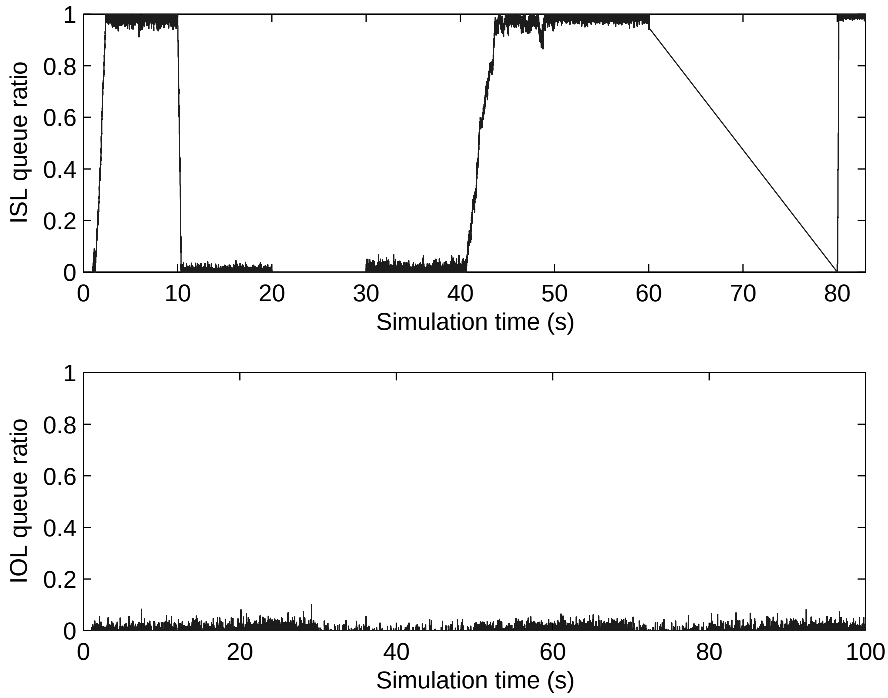


Figure 5.6. The outgoing queues of LEO_{53} with $\alpha = 1$ (nonadaptive case). ISL outgoing queue is congested and experiences overflow as opposed to very low queue ratio of IOL queue.

link utilization rates are more balanced which assures the system resources to be used efficiently. In Figure 5.7, it is seen that the queue ratio Q_r oscillates between 0.8 and 0.4. This is because of the reason that α is set to 0.8 and κ is set to $\frac{\alpha}{2}$. The proposed scheme ensures the background traffic follow the alternate paths in case of Q_r reaching the first threshold value. The packet deflection (forwarding to alternate paths) continues till Q_r reaches the predefined threshold κ, 0.4 in this case.

In Figure 5.8, delay vs. α is plotted. It is seen that with the increase in α, average delay values also increase with the exception of LDV. Since LDV traffic is routed through the upper MEO layer, it is slightly affected by the change in α value. On the other hand, background and SDV traffic are directly affected. Figure 5.8 indicates that increasing α values yields longer queues and therefore longer queueing delays. "Conversational" corresponds to the all conversational traffic class with no classification of LDV or SDV. It should be noted that $\alpha = 1$ corresponds to the case

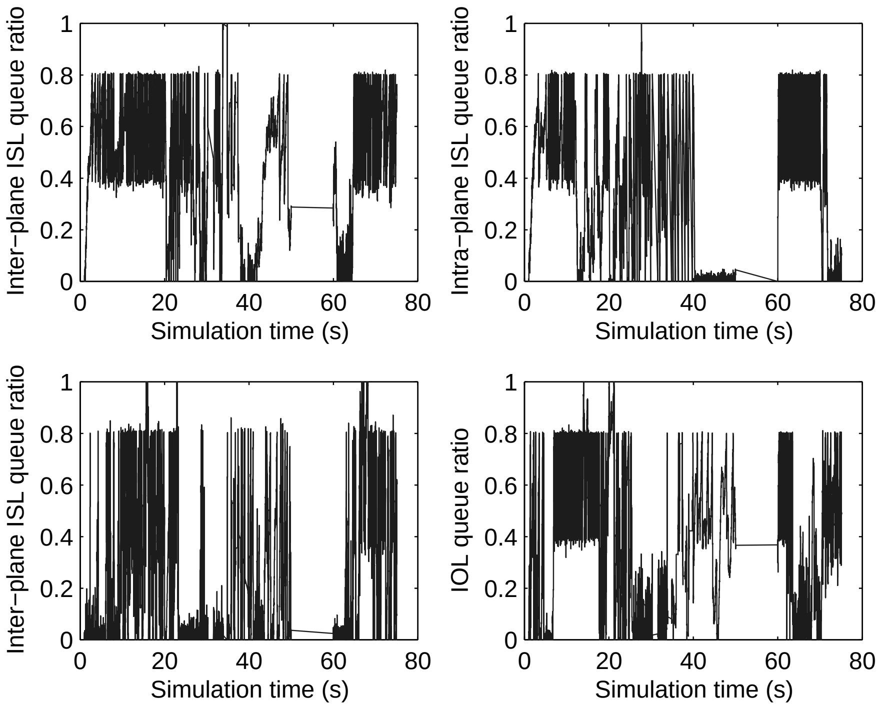


Figure 5.7. Outgoing queue ratios of LEO_{53} with $\alpha = 0.8$ (adaptive case). All queues follow a similar pattern since there is load splitting between the ISLs and IOL.

where no load balancing is applied. It is clear that application of ARPQ improves both the conversational and background traffic performance. With no load balancing, mean conversational delay values are around 240 ms, which causes degradation in the performance of voice communication. The decrease in conversational delay values in case of $\alpha = 1$ is caused by the increase in the background packet drop ratio. Moreover, since overflow traffic is dropped instead of being directed to MEO layer or alternate paths, LDV traffic experiences less queueing delay. Hence LDV, SDV and overall conversational delay decrease. This comes at the expense of many packet drops of background traffic. In our scheme, since no prioritization is done between delay and jitter sensitive traffic and best-effort traffic, LDV and SDV experience as much queueing delay as background traffic. In the following subsections, effect of traffic prioritization will be examined. The exact delay values can be seen in Table 5.5. With $\alpha = 0.2$, the jitter of conversational traffic is around 70 ms which might be acceptable. The increase in queueing delay increases the delay jitter too. Actually, much of background traffic is forwarded to other links in case of low α values and might result in larger jitter values compared to higher α values. But, effect of queueing delay dominates the effect of change in the path of background traffic. Like the decrease in LDV delay, LDV jitter also decreases due to lighter load in MEO layer. By the application of priority queueing, jitter and delay values of conversational traffic can be decreased to more acceptable levels.

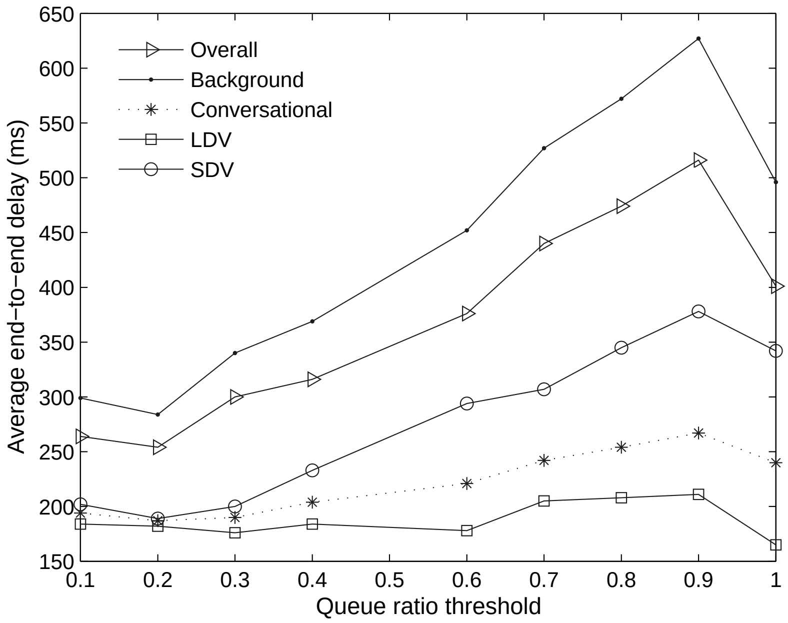


Figure 5.8. Effect of queue ratio threshold α on traffic delay values ($\kappa = \frac{\alpha}{2}$).

The detailed traffic delay values are given in Table 5.5. Increase in α causes increase in delay values. The scenario with $\alpha = 1$ corresponds to nonadaptive case. The delay values seem to be less than the adaptive case with $\alpha = 0.8$. This is due to the reason that in nonadaptive case there is excessive packet loss in background traffic which causes the average delay to decrease. Furthermore, background traffic is not forwarded to MEO layer or other alternate paths. Hence, there is a decrease in MEO layer load which in turn leads to less queueing delay in MEO routes. So, the decrease in LDV traversing the MEO layer can be explained by the decrease in load in this layer.

Figures 5.9 and 5.10 depict the overall packet loss and throughput of system with increasing values of α. With the increase in values of α, packets experience longer queues thereby there is a higher probability of packet loss. This can also be seen from

Table 5.5. Effect of queue ratio threshold α on traffic delay values ($\kappa = \frac{\alpha}{2}$)

	Average delay of all traffic types (ms)				
	Overall	Background	Conversational	*LDV*	*SDV*
$\alpha = 0.1$	264	299	194	184	202
$\alpha = 0.2$	254	284	187	189	182
$\alpha = 0.4$	316	369	204	184	233
$\alpha = 0.6$	376	452	221	178	294
$\alpha = 0.7$	440	527	242	205	307
$\alpha = 0.8$	474	572	254	208	345
$\alpha = 0.9$	515	627	267	211	378
$\alpha = 1$ (Non-adaptive)	401	496	240	165	342

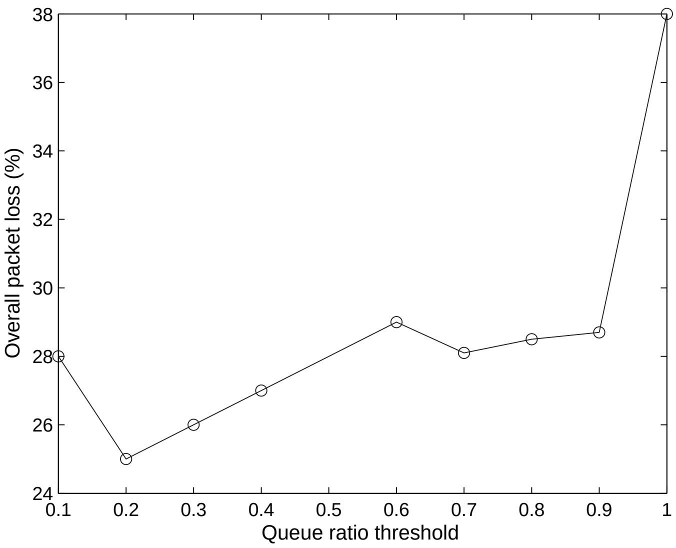


Figure 5.9. Effect of queue ratio threshold α on network packet loss rate ($\kappa = \frac{\alpha}{2}$).

the figures as an increase in packet loss rate and decrease in overall throughput. The loss rate drastically increases when $\alpha = 1$ corresponding to the nonadaptive routing scheme. The loss rate is about 38% which is significantly larger than the loss rate of the closest test point $\alpha = 0.9$. In case of α is 0.9, the loss rate is about 28%. Actually the recorded packet loss rate is usually very large in our simulations, which might be caused by some OPNET related issues. Therefore, we consider these values only to make a comparison between the simulation results. One more point to be considered is that there seems to be a slight difference in packet loss rate with $\alpha = 0.7$, 0.8 and 0.9. This may be due to the reason that the proposed scheme has not enough time to forward the overflow traffic to alternate paths before they are being dropped.

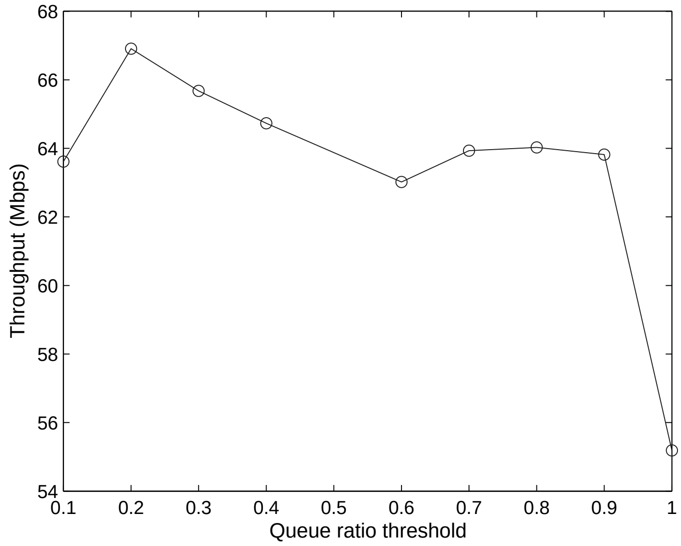


Figure 5.10. Effect of queue ratio threshold α on overall throughput ($\kappa = \frac{\alpha}{2}$).

5.2.4. Effect of Queueing Mechanism

In the following experiments, we apply a non-preemptive, strict priority scheduling mechanism on-board the satellites. In strict priority scheduling policy, the voice packets are always served before the background packets. This will ensure voice traffic not to have long queueing delays due to heavy background traffic. On the other hand, background traffic experiences longer delay and jitter values. However, due to the nature of background traffic, long delay and jitter values do not degrade its performance. The only point to be considered about the background traffic is packet loss rate. In the following scenarios, we set the system parameters as $\alpha = 0.4$, $\beta = 0.8$, $D_{thrsh} = 80$ ms and traffic load is 85% (25% voice and 60% background). Figure 5.11 elucidates the change in average traffic delays by the application of strict priority queueing. Effect of the policy change on LDV and SDV delay values can be seen in Figure 5.12. As the figures show, background traffic delay increases as opposed to decrease in conversational traffic, both LDV and SDV. The main contribution of policy change is the noticeable decrease in SDV traffic. Because SDV traffic is routed on the same route as background packets, it is exposed to congestion in case of heavy background traffic load. Hence, the application of strict priority significantly improves SDV performance. The delay and jitter values which are beyond the acceptable good quality communication limits are now decreased to acceptable levels. By the application of strict priority policy, background delay jitter goes up to 266 ms from 214 ms, while conversational jitter goes down to 28 ms from 82 ms.

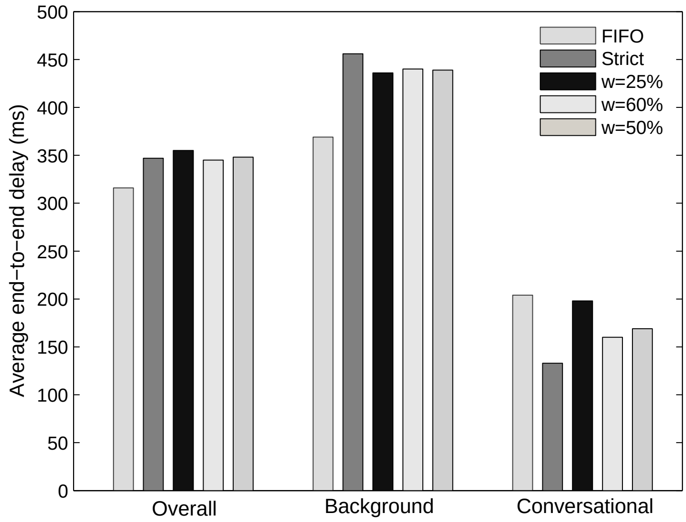


Figure 5.11. Effect of queueing mechanism on end-to-end delay of all traffic types.

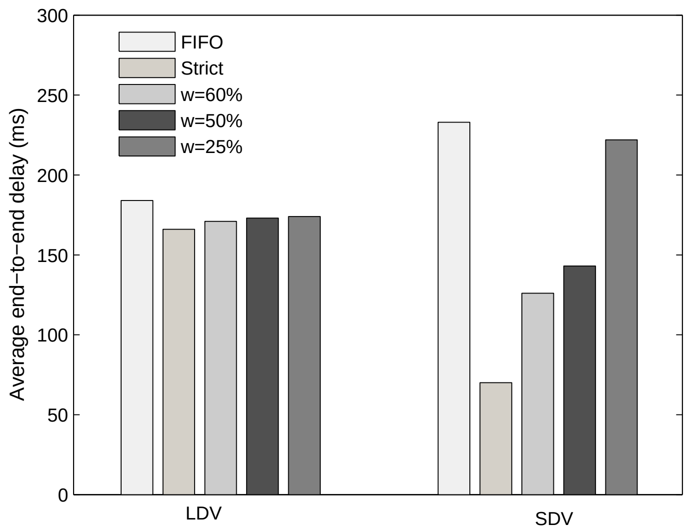


Figure 5.12. Effect of queueing mechanism on end-to-end delay of conversational traffic LDV and SDV.

In order to see the impact of the change in queueing mechanism on individual communications, three GS pairs are taken as reference communication and their corresponding performance metrics are recorded. GS_{58} communicates with GS_{121} which is quite far resulting in 6 LEO hops in the static topology. GS_{107} similarly communi-

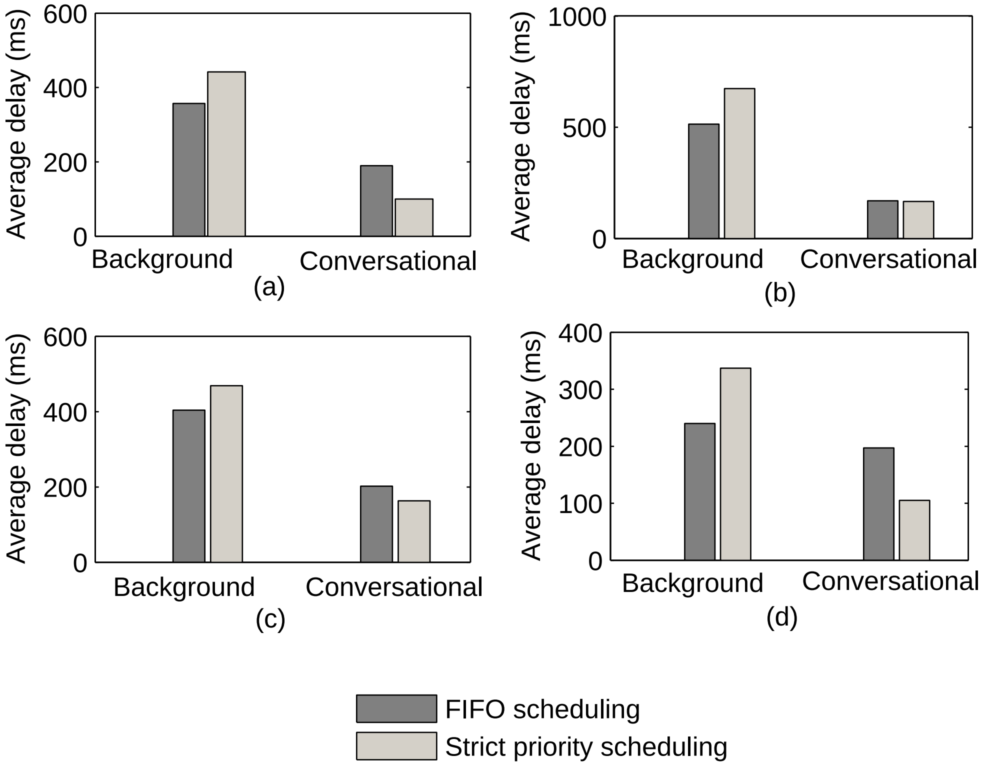


Figure 5.13. (a) The delay values of GS_8 under FIFO scheduling and strict priority scheduling on-board the satellites. (b) The delay values of GS_{58} (c) The delay values of GS_{107} (d) The delay values of GS_{230}

cates with GS_{182} that is 5 LEO hops away from it. GS_8 and GS_{230} have sessions with closer parties in different regions, GS_{261} and GS_{146} respectively. Fig 5.13 elucidates the simulation results. The most prominent decrease in conversational delay is observed at GS_{230}. This is due to the route consisting of many LEO hops of this communication. The performance of GS_{58} and GS_{107} is slightly improved since they are already prevented from congestion by GS_{58} and GS_{107} traffic being forwarded to MEO layer. Note that voice communication will have the desired level of quality, but background traffic will suffer from more packet loss (36%) than the usual case of 30% packet loss in FIFO scheduling. Hence, we should better apply weighted round robin queueing (WRRQ) to satisfy the quality requirements of both traffic type.

We conduct three simulations with w_{conv} set to 25%, 50% and 60%. In Table 5.6, average delay and jitter values of the communications initiated by the predetermined

GSs are listed. The table shows that giving more priority to voice traffic improves the performance by decreasing the delay and jitter values. On the other hand, background traffic delay and jitter values increase with the increasing values of w_{conv}. Background traffic packet loss rate depending on the scheduling policy and weight values of conversational traffic (w_{conv}) is plotted in Figure 5.14. The first and last point in the x-axis correspond to FIFO scheduling and strict priority scheduling respectively. With the increase in voice prioritization (w_{conv}), background traffic experiences more delays in the queues. Queue sizes grow longer with many background traffic packets waiting to be served. Thereby, overflow traffic is dropped. Note that there is a great difference between the strict priority policy and FIFO scheduling policy. The application of WRRQ seems to perform better satisfying the requirement of both traffic types.

Table 5.6. Effect of Queueing Mechanism Simulation Results

		Average delay (ms)		**Delay jitter (ms)**	
GS$_{id}$	**Policy**	Background	Conversational	Background	Conversational
GS_8	FIFO	357	190	244	121
	Strict	442	100	330	46
	$w_{conv} = 25\%$	407	173	334	122
	$w_{conv} = 50\%$	338	124	290	56
	$w_{conv} = 60\%$	356	126	284	56
GS_{58}	FIFO	514	169	266	21
	Strict	674	166	317	22
	$w_{conv} = 25\%$	639	169	335	23
	$w_{conv} = 50\%$	631	167	325	21
	$w_{conv} = 60\%$	718	168	377	22
GS_{107}	FIFO	404	202	222	92
	Strict	469	163	243	33
	$w_{conv} = 25\%$	435	184	267	61
	$w_{conv} = 50\%$	479	177	251	40
	$w_{conv} = 60\%$	482	176	266	37
GS_{230}	FIFO	240	197	161	100
	Strict	337	105	220	45
	$w_{conv} = 25\%$	293	202	217	140
	$w_{conv} = 50\%$	353	150	244	102
	$w_{conv} = 60\%$	320	128	206	83

In our simulations, we do not consider a specific packet dropping mechanism in case of contention. However, the system performance can be improved by the appli-

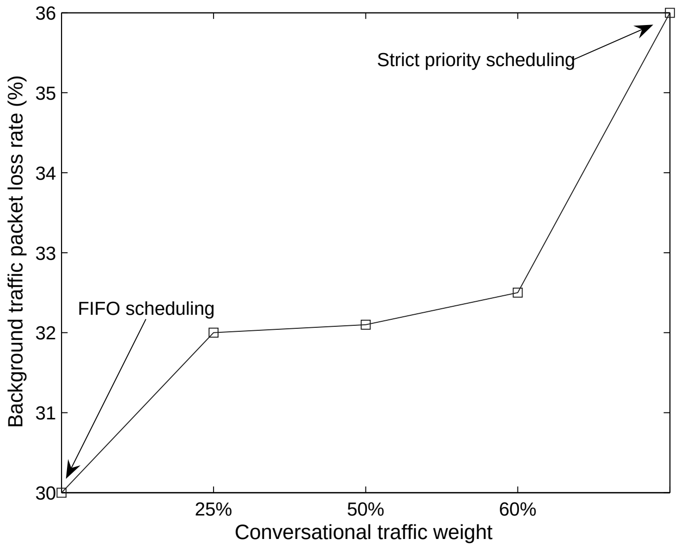


Figure 5.14. Background traffic packet loss rate depending on the values of conversational traffic weight w_{conv}.

cation of some packet dropping mechanisms. In [49] Sun et al. develop a routing and scheduling scheme in a LEO network with limited system resources i.e. transmitters and buffer size. The scheduling mechanism determines which packet to transmit in the next time slot in case of contention and which one to drop in case of buffer overflows. Briefly, three scheduling schemes are employed: random packet win (RPW), oldest packet win (OPW) and shortest hop win (SHW). As the names clearly state, RPW randomly chooses a packet for transmission where OPW transmits the oldest packet that has traveled the longest distance. In SHW scheme, the packet closest to its destination is chosen. The performance comparison of these schemes taking their throughput values as performance metric concludes that SHW outperforms the others in case of no buffer space in satellites. However, having some buffer space, simulation results and theoretical analysis agree that all three schemes achieve nearly same throughput.

6. CONCLUSIONS AND FUTURE DIRECTIONS

Satellites on nongeostationary orbits (NGEOs), or so-called mobile satellites, have had a significant role in the proposal of the first generation of global personal satellite networks such as Iridium and Globalstar. Because NGEO satellites yield smaller coverage areas, a constellation of satellites is required to provide coverage to the Earth. Moreover, because the satellites in NGEO are closer to the Earth, smaller handheld devices than those usual in geostationary satellite systems are practical. By employing LEO and MEO satellites, it is possible to relax the highly restrictive long propagation delay and power loss characteristics of the conventional geostationary orbit (GEO) satellites. Long propagation delay has always been an issue in establishing long-distance real-time communications such as voice and video telephony through satellites. Having satellites in low orbits, it is possible to reduce the transmission delay and power of the transmitters [4].

The routing on satellite networks is of utmost importance to fully utilize the system capacity and provide high quality services. Traditional terrestrial routing mechanisms cannot meet the requirements of NGEO satellite networks with highly time-varying topology. Hence, adaptive routing is an essential requirement to optimize the utilization of satellite payload capacities.

In this thesis, we have studied a new routing scheme called ARPQ on a two-layered satellite architecture. The proposed scheme uses real-time network information to balance the load on the satellite links. Load balancing helps the system resources to be efficiently used and also prevents congestion in some bottleneck links. Hence, applications especially time-sensitive ones have better performance results. Each LEO satellite controls its traffic flow rate to its neighbors and in case of a sign of congestion in one of the links, some portion of background traffic is deflected to other less loaded links. Since this is a kind of self-control mechanism, our proposed scheme achieves load balancing without additional signalling overhead. Furthermore, we have investigated the performance issues and the effects of on-board scheduling mechanisms on VoIP performance in multi-layered satellite networks. We have limited the scope of our interest to the OBP functions. That is, we are only concerned with the mechanisms applied on-board the satellite to improve the performance of VoIP applications. It is of no doubt that the application of previously mentioned mechanisms (e.g. jitter buffers, comfort noise generation etc.) in the terrestrial part of the network amends the performance. The results indicate that OBP, enabling multi-layered satellite and QoS mechanisms, are of great importance for performance enhancement in these networks. With this ability, good-quality VoIP over satellite is feasible. Moreover, OBP reduces the reliance of the satellite systems on the ground. This is especially important in military systems where the ground system might be interfered, and in natural disaster cases where the ground system might also be damaged.

Furthermore, application of some other on-board processing mechanisms like error correction coding helps the system to satisfy the application requirements. The excessive packet loss rate can be decreased to more acceptable levels by the application of some error correction mechanisms as Forward Error Correction (FEC).

Further work includes modeling of the satellite channels with more realistic models and addition of FEC. Moreover, nonuniform traffic distribution through the Earth and during the day needs to be considered to see how our mechanism performs in the case of real traffic generation pattern. We are currently working on switching and routing issues on a satellite network with empirical IP traffic. New features can be added to the scheme, such as a mechanism considering the requirements of all types of traffic (interactive and streaming) other than conversational and background traffic group.

APPENDIX A: IMPLEMENTATION DETAILS

A.1. Network Model

The satellite network model used in the simulation studies of this thesis study is depicted in Figure A.1. There are Gateway Stations (GS) which are placed into subnets, LEO satellites, MEO satellites, Simulation Update node and Global RIB node. Each node and its submodules will be explained in the following sections.

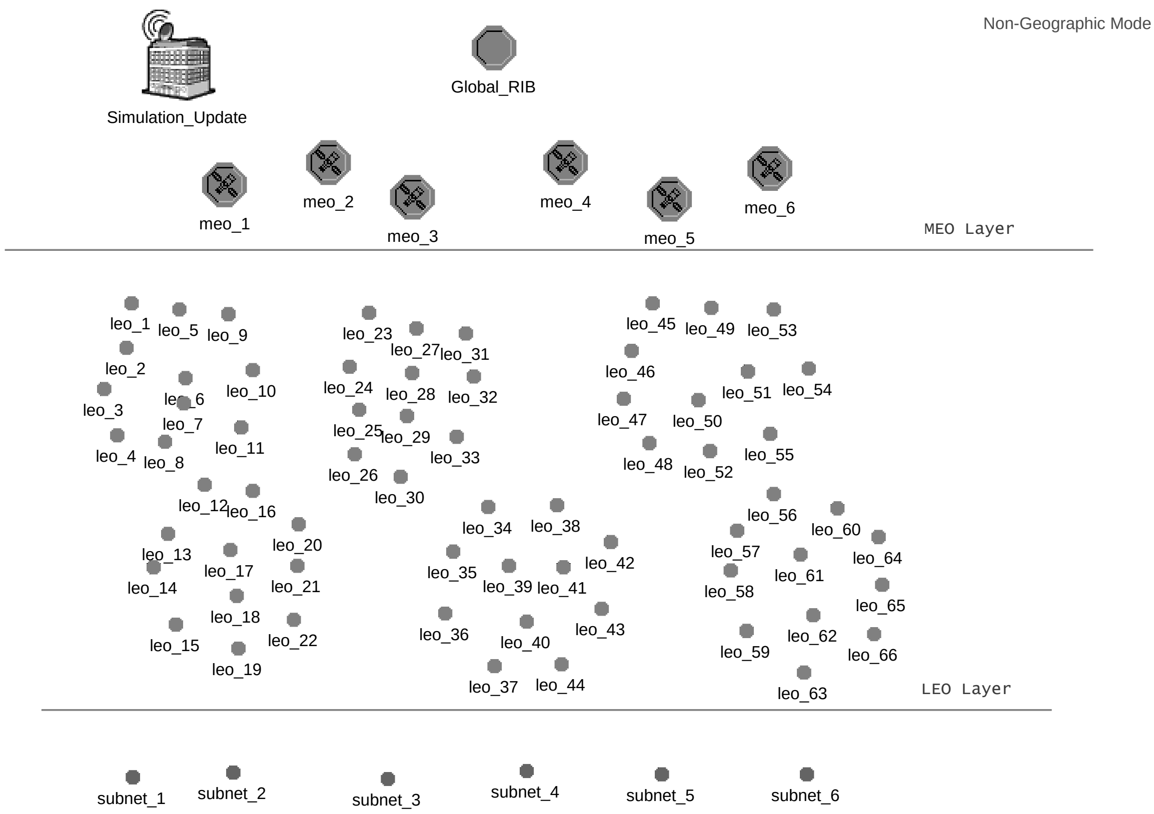


Figure A.1. Two-layered OPNET network model used in the simulations of this thesis study

A.2. Simulation Controller

This module (Figure A.2) deals with the initialization of the system, updating of the system state and collecting the statistics. Initialization includes opening files, assigning source and destination nodes for a communication. Interrupts for a new time interval are generated here, and distributed to all other satellite nodes. Additionally, at the end of the simulation (upon receipt of ENDSIM interrupt), some global statistics are collected in this module.

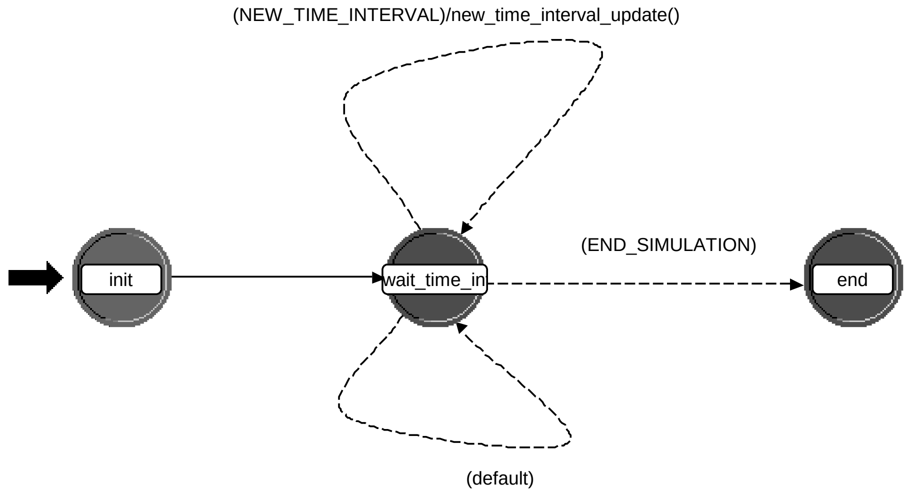


Figure A.2. OPNET state transition diagram of Simulation Update node

A.3. Global Routing Information Base (GRIB)

Rather then choosing a MEO satellite as the routing table updater, we implemented a special node for routing updates. Process model of this node can be seen in Figure A.3. When a scenario starts running, some initializations are done in *init* state. Upon completion of these initializations, process' state changes to *Wait* state. Process stays in this state till **NEW_MEO_INFO_ARRVD** becomes true. At the beginning of each time interval, all LEO satellites report their queueing delays to their group manager MEOs. A MEO sends all this information to the GRIB node which makes **NEW_MEO_INFO_ARRVD** true. When all reports are received, GRIB updates the routing table. After completion of routing table updates, this node goes into *Wait* state and waits for the next time interval updates from MEO nodes.

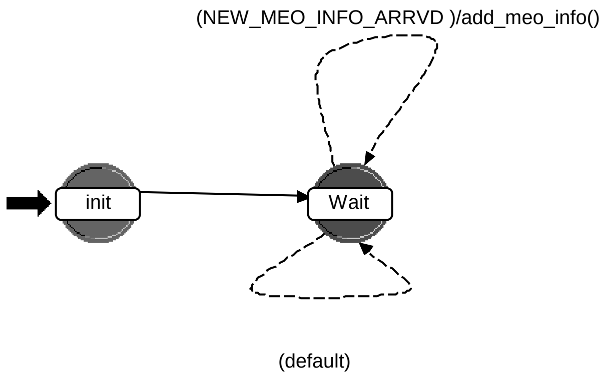


Figure A.3. OPNET state transition diagram of GRIB

A.4. Gateway Stations (GS)

A gateway station is a node that generates traffic and also it is one side of a communication. Each GS has an isotrophic antenna, radio transmitter and receiver, packet source generators, packet send module, a module for recording statistics and a sink. Traffic sources generate packets according to the specified parameters such as packet format, packet interarrival time distribution and packet size. Actually, we have four traffic sources to model four different traffic types corresponding to 3GPP QoS groups: conversational, streaming, interactive and background class. Each class has a different packet format and special traffic parameters. On the other hand, we only considered the two traffic classes (conversational and background) in this thesis. Generated packets are sent to *Send* module and this module decides which satellite to send the packets according to the predetermined satellite calender. *Write_statistics* module records various local and global statistics of an ongoing communication. Finally, packets are destroyed in the sink. Figure A.4 and A.5 show the node model of a GS and process type of *Send* module respectively.

At the very beginning of a scenario, some initializations need to be done in *init* state. When other global initializations are completed by Simulation Controller node, **COMM_UPDATE_COMPLETED** becomes true. Therefore, new state becomes

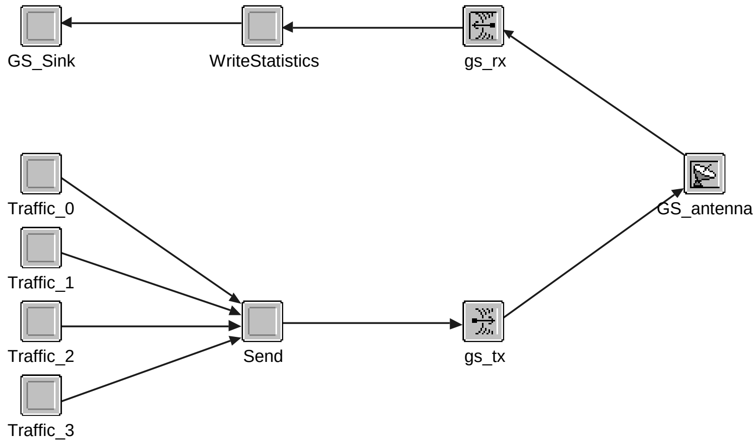


Figure A.4. OPNET model of a GS node. There are traffic sources to generate traffic, a packet send module, a module for recording statistics and a sink module to destroy incoming packets in order to free memory reserved for these packets.

ready. Since this state is a "forced state" (after completion of what has to be done in this state, the state immediately changes to next state without any constraints), it goes into *send_pk* state. If a packet arrives, **PK_ARRVL** becomes true and *Send_to_Leo()* function is called. As the name of the function states, this function sends the incoming packet to the corresponding LEO in sight. Satellite channel assignments are also done in this function. Finally, packets are sent to the radio transmitter *gs_tx* and from there go to the antenna. The radio transmitter has a "receiver group" to specify the nodes which will receive the packets sent by this GS. Since radio links provide a broadcast medium, each transmission can potentially affect multiple receivers throughout the network model therefore we need to specify the target LEO. Furthermore, the propagation delay associated with the link between this GS and corresponding LEO satellite can be set here. OPNET provides a set of functions named as pipeline stages to model wireless transmission of packets.

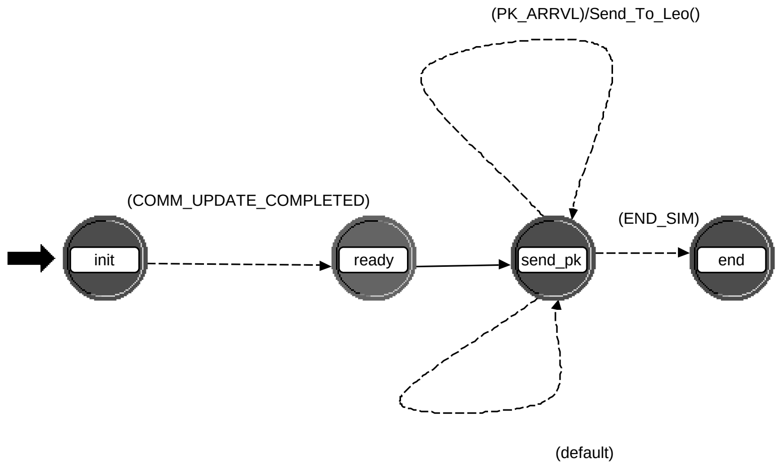


Figure A.5. OPNET state transition diagram of a GS processor.

A.5. LEO Satellites

A LEO satellite node is composed of ISL and IOL transceivers and a processor. Routing and other main satellite functions are implemented in the processor module as in Figure A.7. At the beginning of a new time interval, processor module receives a **NEW_TIME_INTERVAL** remote interrupt and goes to *NewTime* state. In this state, new interrupts are scheduled in order to alert outgoing links to report their previous time-interval queueing information. Each incoming and outgoing link is modeled as a queue. OPNET models of a LEO satellite and its submodules are depicted in Figure A.6, A.7 and A.8.

A.6. MEO Satellites

Figure A.9 shows OPNET model of a MEO satellite. A MEO satellite has one processor, inter and intra-plane ISLs, and IOLs.

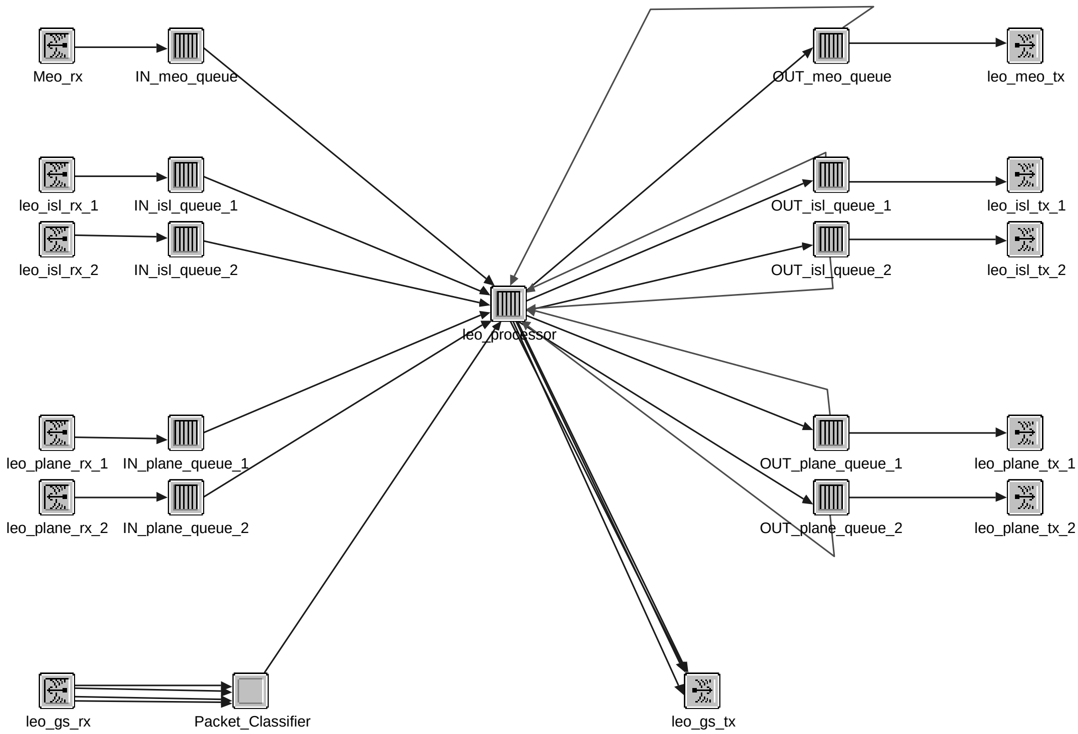


Figure A.6. OPNET model of a LEO satellite. There are two inter-plane ISLs, two intra-plane ISLs and one MEO IOL. Each link is modeled as incoming and outgoing queues.

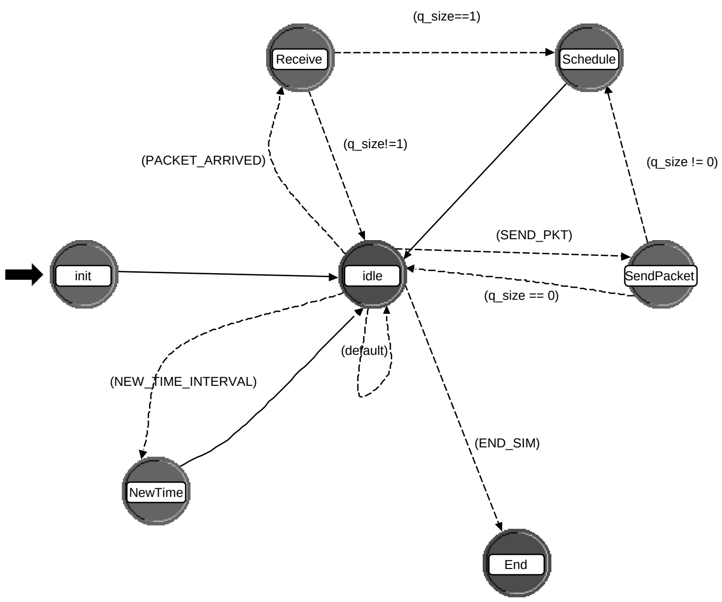


Figure A.7. OPNET state transition diagram of a LEO satellite processor.

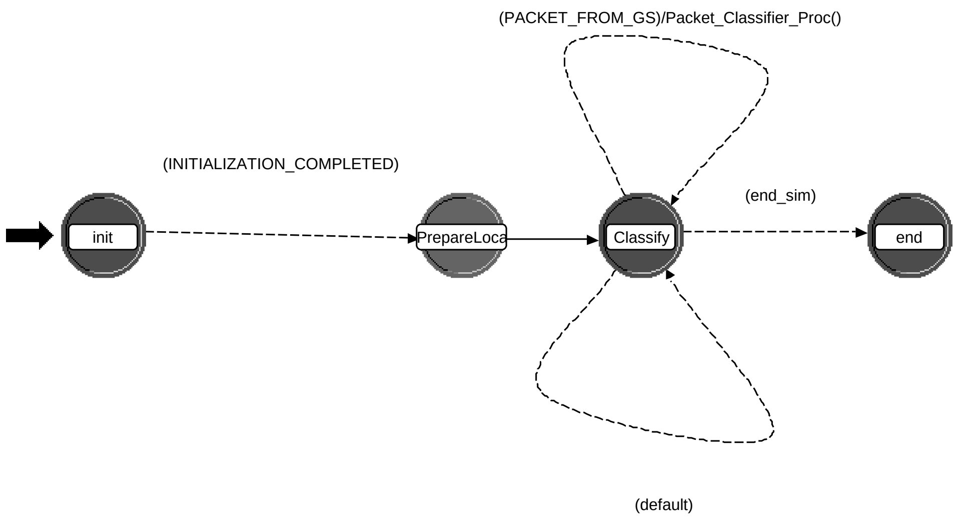


Figure A.8. OPNET state transition diagram of a LEO satellite packet classifier.

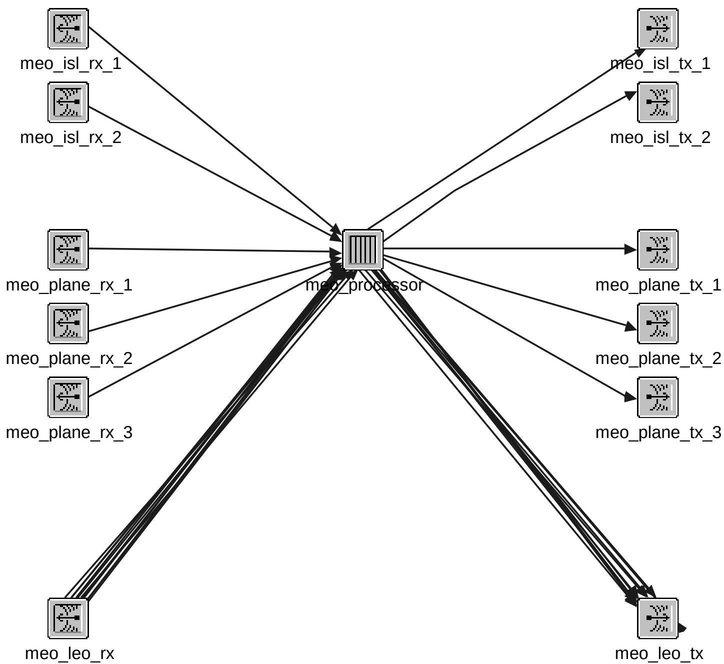


Figure A.9. OPNET model of a MEO satellite. There are three inter-plane ISLs, two intra-plane ISLs and eleven LEO IOLs.

REFERENCES

1. Clarke, A. C., "Extraterrestrial Relays", http://www.sciencemuseum.org.uk.

2. Ha, T. T., *Digital Satellite Communications*, McGraw-Hill, 2nd edition, 1990.

3. Henderson, T. R., *Networking over next-generation satellite systems*, Ph.D. Thesis, University of California at Berkeley, Fall 1999.

4. Jamalipour, A. and T. Tung, "The Role of Satellites in Global IT: Trends and Implications", *IEEE Personal Communications*, June 2001.

5. Dowla, F., *Handbook of RF and Wireless Technologies*, Elsevier, 2004.

6. "Satellite Firsts", http://en.wikipedia.org, 2005.

7. Elbert, B. R., *Introduction to Satellite Communication*, Artech House, 2nd edition, 1999.

8. Dhaou, R., *Modeling of Networks Composed of Satellite Constellations*, Ph.D. Thesis, University of Paris VI, December 2002.

9. Küçükateş, R., *A Minimum Flow Maximum Residual Method for Routing in LEO Satellite Networks*, Ph.D. Thesis, Computer Eng., Boğaziçi University, 2002.

10. Verma, S. and E. Wiswell, "Next Generation Broadband Satellite Communication Systems", *Proc. 20th AIAA International Communication Satellite Systems Conference and Exhibit*, Montreal, Quebec, CA, May 2002.

11. Eckhardt, G., "On-board processing", *Space Communications*, Vol. 17, pp. 1–3, 2001.

12. Gupta, A. R. D. and K. S. D. Gupta, "The Emerging Trends in Satellite and

Wireless Communications Technologies", *Indian Academy of Sciences 69th Annual Meeting*, Guwahati, India, November 2003.

13. Janssen, J., D. D. Vleeschauwer, G. H. Petit, R. Windey, and J.-M. Leroy, "Delay bounds for voice over IP calls transported over satellite access networks", *Mob. Netw. Appl.*, Vol. 7, No. 1, pp. 79–89, 2002.

14. "Satellite Industry Overview", http://www.sia.org, 2005.

15. Evans, B., M. Werner, E. Lutz, M. Bousquet, G. Corazza, and G. Maral, "Integration of Satellite and Terrestrial Systems in Future Multimedia Communications", *IEEE Wireless Communications*, October 2005.

16. "Iridium Satellite", http://www.iridium.com.

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

18. Chong, H. M. and H. S. Matthews, "Comparative Analysis of Traditional Telephone and Voice-over-Internet Protocol (VoIP) Systems", *Proc.of IEEE International Symposium on Electronics and the Environment*, 2004.

19. Goralski, W. J., *IP Telephony*, McGraw-Hill, 1999.

20. "ITU G.1010 End-user multimedia QoS categories", November 2001.

21. "QoS Concept and Architecture", Technical Specification 3GPP TS 23.107, June 2005.

22. Kota, S. and M. Marchese, "Quality of service for satellite ip networks: a survey", *Int. J. Satell. Commun. Network*, 2003.

23. Spillard, C., T. Tozer, B. Gremontand, and D. Grace, "The Performance of High-Altitude Platform Networks in Rainy Conditions", *22nd AIAA International Com-*

munications Satellite Systems Conference, Monterey, USA, May 2004.

24. Dash, D., A. Durresi, and R. Jain, *Routing of VoIP traffic in multi-layered Satellite Networks*, Orlando, Florida, September 2003.

25. Lee, J. and S. Kang, "Satellite over Satellite Network: A Novel Architecture for Satellite Network", *Proc. IEEE INFOCOM*, pp. 315–320, May 2000.

26. Chen, C., E. Ekici, and I. F. Akyildiz, "Satellite Grouping and Routing Protocol for LEO/MEO Satellite IP Networks", *Proceedings of WoWMoM September 2002*, pp. 109–116, Atlanta, Georgia, September 2002.

27. Chen, C. and E. Ekici, "A Routing Protocol for Hierarchical LEO/MEO Satellite IP Networks", *ACM-Springer Journal of Wireless Networks (WINET)*, Vol. 11, No. 4, pp. 507–523, July 2005.

28. Ekici, E., I. F. Akyildiz, and M. Bender, "Datagram Routing Protocol for LEO Satellite Networks", *Proceedings of INFOCOM*, pp. 500–508, Tel Aviv, Israel, March 2000.

29. Akyildiz, I. F., E. Ekici, and M. Bender, "MLSR: A novel routing algorithm for multilayered satellite ip networks", *IEEE/ACM Transactions on Networking*, Vol. 10, No. 3, pp. 411–424, June 2002.

30. Akyildiz, I. F. and E. Ekici, "A Distributed Multicast Routing Scheme for Multi-Layered Satellite IP Networks", *Wireless Networks Kluwer Academic Publishers*, pp. 535–544, 2003.

31. Wu, F. G., F. C. Sun, K. Yu, and C. W. Zheng, "Performance evaluation on a double-layered satellite network", *International Journal of Satellite Communication Networks*, pp. 359–371, 2005.

32. Zhao, Z., J. Zhang, F. Sun, L. Li, and I.-M. Chen, "Routing Strategy for Satellite-Based Networks with a Hierarchial Space Grid Architecture", *CICS-PCM*, Decem-

ber 2003.

33. Durresi, A., D. Dash, B. Anderson, R. Kannan, S. Kota, and R. Jain, "Routing of real-time traffic in a Transformational Communications Architecture", *IEEE Aerospace Conference Proceedings*, 2004.

34. Peng, J. and B. Sikdar, "Multilayer Multicast Congestion Control in Satellite Environments", *IEEE Journal on Selected Areas in Communications*, April 2004.

35. Hu, J. and K. Yeung, "Routing and Re-Routing in a LEO/MEO Two-tier Mobile Satellite Communications System with Inter-Satellite Links", *Proceedings of ICC*, 2000.

36. Taleb, T., N. Kato, and Y. Nemoto, "IP Traffic Load Distribution in NGEO Broadband Satellite Networks", *Proc.20th ISCIS*, Istanbul, Turkey, October 2005.

37. Nanba, S., S. Konishi, and S. Nomoto, "Optimum traffic distribution algorithm for multiple-satellite systems under power constraints", *IEEE Journal on Selected Areas in Communications*, April 2004.

38. Küçükateş, R. and C. Ersoy, "High Performance Routing in a LEO Satellite Network", *Proceedings of the IEEE ISCC*, pp. 1403–1408, Antalya,Turkey, 2003.

39. Jianjuan, B., "Explicit Multi-path Routing for LEO Satellite Networks", *Proc. 2005 IEEE Wksp. High Perf. Switching and Routing*, Hong Kong, China, May 2005.

40. McMahon, G., R. Septiawan, and S. Sugden, "A Multiservice Traffic Allocation Model for LEO Satellite Communication Networks", *IEEE Journal on Selected Areas in Communication*, Vol. 22, No. 3, pp. 501–507, April 2004.

41. Jianjun, B., L. Xicheng, L. Zexin, and P. Wei, "A Distributed Hierarchical Routing Protocol for Non-GEO Satellite Networks", *Proceedings of ICPP Workshop*, pp. 148–154, 2004.

42. Mauger, R. and C. Rosenberg, "QoS Quarantees for Multimedia services on a TDMA-based satellite network", *IEEE Communications Magazine*, Vol. 35, No. 7, pp. 56–65, July 1997.

43. Svigelj, A., M. Mohorcic, G. Kandus, A. Kos, M. Pustisek, and J. Bester, "Routing in ISL Networks Considering Empirical IP Traffic", *IEEE Journal on Selected Areas in Communications*, Vol. 22, pp. 261– 272, February 2004.

44. "Optimizing Network Technologies, OPNET", http://www.opnet.com.

45. Korçak, O. and F. Alagöz, "Deflection Routing over Prioritized Intersatellite Links in LEO Satellite Networks", *2nd IEEE International Conference on Information and Communication Technologies ICTAA*, Syria, April 2006.

46. Mohorcic, M., M. Werner, A. Svigelj, and G. Kandus, "Adaptive routing for packet-oriented intersatellite link networks: performance in various traffic scenarios", *IEEE Transactions on Wireless Communications*, Vol. 1, October 2002.

47. Bayhan, S., G. Gür, and F. Alagöz, "VoIP Performance in Multi-layered Satellite IP Networks with On-Board Processing Capability", *7th International Symposium On Computer Networks (ISCN)*, 2006, accepted.

48. Bayhan, S., G. Gür, and F. Alagöz, "Araç üstü işleme yapabilen NGSO uydu ağlarında VoIP Başarımı", *2. İletişim Teknolojileri Ulusal Sempozyumu*, pp. 31–36, Adana, Türkiye, November 2006.

49. Sun, J. and E. Modian, "Routing Strategies for Maximizing Throughput in LEO Satellite Networks", *IEEE Journal on Selected Areas in Communications*, February 2004.

50. Bayhan, S., G. Gür, and F. Alagöz, "Adaptive Routing Protocol for QoS in Two-Layered Satellite IP Networks", *2nd International Workshop on Satellite and Space Communications (IWSSC)*, Spain, 2006, submitted.