SIP-BASED QOS MANAMGEMENT ARCHITECTURE FOR IP MULTIMEDIA SUBSYSTEMS OVER IMS ACCESS NETWORKS

By

UMBER IQBAL

(2006-NUST-MS PhD-CSE (E)-15)

Submitted to the Department of Computer Engineering
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Software Engineering

Thesis Advisor
Dr. Muhammad Younus Javed

MS-CSE-6

College of Electrical & Mechanical Engineering
National University of Sciences and Technology
2010

Acknowledgements

I thank Almighty Allah for the successful completion of my thesis. I gratefully acknowledge the encouragement and support of my Advisor Dr. Muhammad Younus Javed. He has made available his support in a number of ways and has been a great mentor always providing me with the much needed encouragement and thoughtful direction.

I am heartily thankful to Dr. Shaleeza Sohail for her guidance, time and encouragement from the initial to the mature level, enabled me to develop an understanding of the subject. I was very fortune to have been able to work with her. Her detailed and constructive comments were vital to the development of this thesis.

I owe my deepest gratitude to Dr. Almas Anjum for his help and precious advice. I am also thankful to my advisory committee for their feedback.

I would also like to convey thanks to the National University of Science and Technology (NUST), College of E & ME and Faculty of Department of Computer Engineering for providing me the financial means and laboratory facilities for my research work.

My parents, brothers, sisters and friends are mentioned last to emphasize the special nature of their tremendous support and patience all through my candidature. And also wishes to express love and gratitude for their understanding & endless love, through the duration of my studies and completion of the project.

.

Dedication

I dedicated this thesis in the honour of my parents, my grand parents (Late), my brothers, sisters and my friends who always supported me and prayed for my success

Abstract

True integration of multimedia services over wired or wireless networks increase the productivity and effectiveness in today's networks. IP Multimedia Subsystems are Next Generation Network architecture to provide the multimedia services over fixed or mobile networks. This Research work proposes an extended SIP-based QoS Management architecture for IMS services over underlying IP access networks. To guarantee the end-to-end QoS for IMS services in interconnection backbone, SIP based proxy Modules are introduced to support the QoS provisioning and to reduce the handoff disruption time over IP access networks. In our approach these SIP Modules implement the combination of Diffserv and MPLS QoS mechanisms to assure the guaranteed QoS for real-time multimedia services. To guarantee QoS over access networks, SIP Modules make QoS resource reservations in advance to provide best QoS to IMS users over heterogeneous networks. To obtain more reliable multimedia services, our approach allows the use of SCTP protocol over SIP instead of UDP due to its multi-streaming feature. This architecture enables QoS provisioning for IMS roaming users to differentiate IMS network from other common IP networks for transmission of real-time multimedia services. To validate our approach simulation models are developed on short scale basis. The results show that our approach yields comparable performance for efficient delivery of IMS services over heterogeneous IP access networks.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

The objective of this research is to study existing QoS framework for IMS, investigate the use and feasibility of SIP-Based QoS Management architecture for IMS and Next Generation "all-IP" Networks, and prove the concept model by comparing SIP-Based QoS Management Architecture with an existing QoS framework. The main objective of introducing this architecture is to distinguish IMS networks from other available IP networks to manage the QoS more efficiently and to reduce the hand off disruption time over IMS and underlying access networks. We introduce the SIP-based proxy QoS modules to handle IMS traffic over heterogeneous networks. These QoS modules implement the combination of Diffserv and MPLS mechanisms to support QoS. SIP is a session negotiation protocol provides service continuity in IMS network. But SIP performance is restricted by the performance of TCP and UDP which introduce considerable transmission delay in real-time multimedia services. These modules use SCTP protocol instead of UDP because of its multi-streaming and multi-homing feature.

This Research work proposes an extended SIP-based QoS Management architecture for IMS services over underlying IP access networks. To guarantee QoS over access networks, SIP Modules make QoS resource reservations in advance to provide best QoS to IMS users over heterogeneous networks. To obtain more reliable multimedia services, our approach allows the use of SCTP protocol over SIP instead of UDP due to its multi-streaming feature. This architecture enables QoS provisioning for IMS roaming users to differentiate IMS network from other common IP networks for transmission of real-time multimedia services.

Next Generation Networks, a communication technology capable to provide the QoS-based enrich multimedia applications over heterogeneous networks. IMS is the NGN architecture connected to different access networks to enable the streaming of multimedia contents between various users over the wireless or wire-line networks. IMS was originally introduced by 3GPP (a standard body setup to develop and maintain the technical specifications for the 3G) and adopted by 3GPP2 and TISPAN later.

The promising feature of IMS is its access independence. They are targeted to provide real-time multimedia services with low bit and high error-rate nature over IP communication networks. To support such services, transmission network is conscious about QoS. QoS provisioning assures the reliable delivery of multimedia services with sufficient network resources. IMS Users are able to access and drop variety of IP-based services in a single session and this integration requires more resource reservations and efficient delivery of contents. The IMS policy based QoS mechanism provides the interaction between IMS network and the underlying IP access networks [6]. But when UE roams among IMSs, QoS parameters have to be renegotiated thus introducing significant delay in resource reservation and increase the hand off disruption time in real time multimedia services. So the QoS provisioning and handoff management should be seriously addressed for roaming users in IMS over heterogeneous networks to distinguish IMS from other common IP networks for transmission of real-time multimedia services. The hand off management enables user connectivity with network when point of attachment is changed.

Mobile and fixed operators are deploying NGN based on IMS networks. IMS network should maximize their connectedness with underlying access network by introducing the roaming arrangements and QoS provisioning mechanisms to appreciate the full value of multimedia services. TCP-migrate, Mobile IP and SIP are handoff management protocols proposed for the IP-based Mobile networks [6]. SIP is signaling protocol in IMS and can manage Mobility in IMS networks but at the same time other protocols required significant modifications to work with IMS networking infrastructure.

1.1 Problem Statement

In this research project we will develop a SIP-Based QoS Management Architecture for IP Multimedia System (IMS) over IP Access Networks and Next Generation All-IP Networks, which will complement existing QoS mechanisms over IP Access Networks to deliver the efficient multimedia services. To improve the communication between core and access networks and to reduce the handoff disruption time for multimedia services, this architecture introduces QoS SIP Proxy Modules to monitor the network and to ensure the end-to-end QoS. This architecture introduces the new aspect of handling IMS traffic over access networks.

We will use the Diffserv and MPLS QoS mechanism for our research project which will provide the better utilization of bandwidth and reliable delivery of multimedia services in our proposed framework over IP Access Networks. We have to devise an efficient research methodology to incorporate the results of the research in the final solution in a timely fashion. Our proposed research methodology consists of following important work projects:

a) Analyze the IMS Core QoS Management Policy Based Architecture (Chapter 4)
b) Review the QoS Management Architectures in IP Networks (Chapter 3)
c) Analyze the End-to-End QoS shortcomings in the architecture of IMS framework over Access Networks (Chapter 5)
d) Mapping the shortcomings to IMS Access Networks (Chapter 3, Chapter 5)
e) Propose an Extended QoS Management Architecture for IP Multimedia Subsystems over Access Networks for end-to-end QoS provisioning (Chapter 6)
f) Evaluate and validate the proposed Architecture (Chapter 7, Chapter 8)

1.2 Introduction to Thesis

Chapter 2 explains the Principles, Architecture and Applications of IP Multimedia Subsystem.

Chapter 3 explains the existing QoS methodologies for IP networks and defines the required QoS parameters for multimedia services.

Chapter 4 explains existing Policy Based QoS Management Architecture for IMS core network.

Chapter 5 explains the QoS issues and requirements in IP Access Networks

Chapter 6 describes and explains the proposed SIP Based QoS Management Architecture for IMS over Access Networks.

Chapter 7 this chapter validates the performance of our proposed Architecture using different scenarios with distinctive performance parameters and explains the simulation testbed.

Chapter 8 this chapter concludes the thesis plus recommends the enhancements and future work in this area

Chapter 2

IP Multimedia Subsystems

Introduction

The IP Multimedia Subsystem (IMS) is a series of networking protocols designed to facilitate standard-based fixed/mobile, voice/data and voice/video convergence. It is an architecture based on the Internet Protocol, which fundamentally allows service providers to deliver services to subscribers independent of their device type, network type, or physical location. IMS was designed to fill the gap between the existing traditional telecommunications technology and internet technology. This will allow operators to offer new, innovative services that shareholders and end users are expecting. IMS was specifically architected to enable and enhance real time, multimedia mobile services such as rich voice services, video telephony, messaging, conferencing, and push services. IMS enables these user-to-user communication services via a number of key mechanisms including session negotiation and management, Quality of Service (QoS) and mobility management. However, IMS enables much more than just real time user-to-user services.

2.1 IMS Background and Vision

The need to deploy a new domain may naturally be questioned, especially at a time when network operators are struggling with costs of deploying 3G networks and are also facing reduced voice revenues. To evaluate the need to deploy the IMS, the following discussion examines conventional network domains, the services they can offer, and how they handle various new service offerings. The end users' experiences are also considered.

2.1.1 Shortcomings of Conventional Network Domains

Operators have been looking for ways to differentiate themselves by offering new and creative data services. The PS domain of PS networks or basically 2.5 and 3G networks has helped PS network operators to introduce user-to-server data services, where a user directly addresses a specific server to execute the service in question; furthermore, these services take advantage of IP transport and provide "always on" connectivity.

Unfortunately, PS networks have not been very successful, perhaps due to insufficient bandwidth, lack of enticing applications, confusing charging schemes, long delays in service offerings, etc. Furthermore, increasing the bandwidth alone may not prove sufficient to enable the plethora of new and desirable services that customers may demand.

Also, network operators, in a rush to offer data services, deployed specialized isolated island solutions that often did not integrate well with the other services. These dedicated solutions are typically proprietary and use dedicated components and interfaces that cannot be used for other applications. This is particularly true when the applications are provided by different vendors. Thus, a new platform or domain is needed with unified features; common elements; and open, standardized interfaces that can be used by all existing and future applications and services.

IMS will address these issues by providing:

- A common platform with reusable components, providing quick and easy service development
- Combinations of the functionalities of each solution; e.g. integrating a presence service with conferencing, to produce synergistic gains and avoid redundant development of features.
- Consistent, open interfaces for 3rd party developers
- Consistency to end users: Roaming scenarios are possible, not restricting access to a service to the home network only.

2.1.2 The IMS-Domain of Services

The IMS—the domain of services—meets this need. Desirable Characteristics of the IMS, the desirable features, requirements, and architecture for this new domain are delineated in the 3GPP's technical specification for IMS service requirements [2, 3].

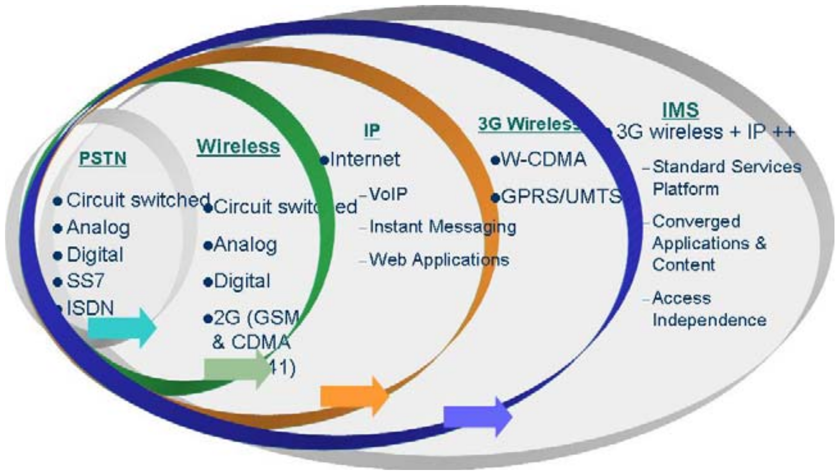

Figure 2.1: Technology Evaluation [16]

The IMS domain supports the following key requirements:

- IP multimedia sessions, i.e., delivery of multimedia sessions over PS networks
- Integration with Internet and CS networks such as public switched telephone networks (PSTNs) and existing cellular networks
- Single sign-on and authentication
- Single converged billing
- Strong operator controls with respect to services delivered to end users
- Rapid service creation without requiring standardization
- Access independence, i.e., allows access technologies other than GPRS and UMTS (e.g., WLANs and x-type digital subscriber line technology [xDSL])

The IMS also provides improved end-user experience over that offered by the other two domains. To fully appreciate these capabilities, a more thorough understanding of SIP and SIP networks is needed.

2.2 IMS Architecture

IMS is being developed by 3GPP [6]; aims to provide converged services over Mobile or fixed networks. The idea behind the IMS evaluation is to provide the capability of using different type of services at single device with best QoS to end users. According to 3GPP

IMS was not intended to standardize applications but rather to aid the access of multimedia applications over network. The basic IMS architecture is the enhancement of UMTS over IP Connectivity adding some different network entities with the QoS provisioning. SIP is the building block in IMS network. The other protocols used in IMS architecture are Diameter, COPS, RTP and RTCP for transmission.

IMS architecture can be functionally divided into three layers.

- Transport and Endpoint Layer.
- Session and Control Layer
- Application or Service Layer

2.2.1 Transport and endpoint Layer

Transport and endpoint layer controls the signaling session and provides the bearer services for the end users. It also supports the media convergence if required. This layer provides the authentication and registration functionality to setup and maintain sessions between end users. Multiple transport services can be merged into a single session. The bearer services used to transport IMS contents over network is provided by the common IP based mechanism and IMS uses the RTP over UDP to transport the media uses the IPv6.

2.2.2 Session and Control Layer

In IMS Core network several SIP servers or proxies are collectively called CSCF and used to process SIP signaling packets in IMS. CSCF (Call Session Control Function) occupies the central position in IMS and are used for setup and control multimedia sessions. There are three different kinds of Call Session Control Functions. They perform specific tasks but all common is that they play a role during the session establishment and registration and form the SIP routing machinery.

2.2.2.1 Call Session Control Functions:

Serving CSCF (S-CSCF)

The S-CSCF (Serving CSCF) is central node of the signaling plane and has the main functionality in user registration. Each SIP message of IMS is inspected by S-CSCF to verify HSS depends upon user Profile. S-CSCF is also used for load balancing and always located in home network.

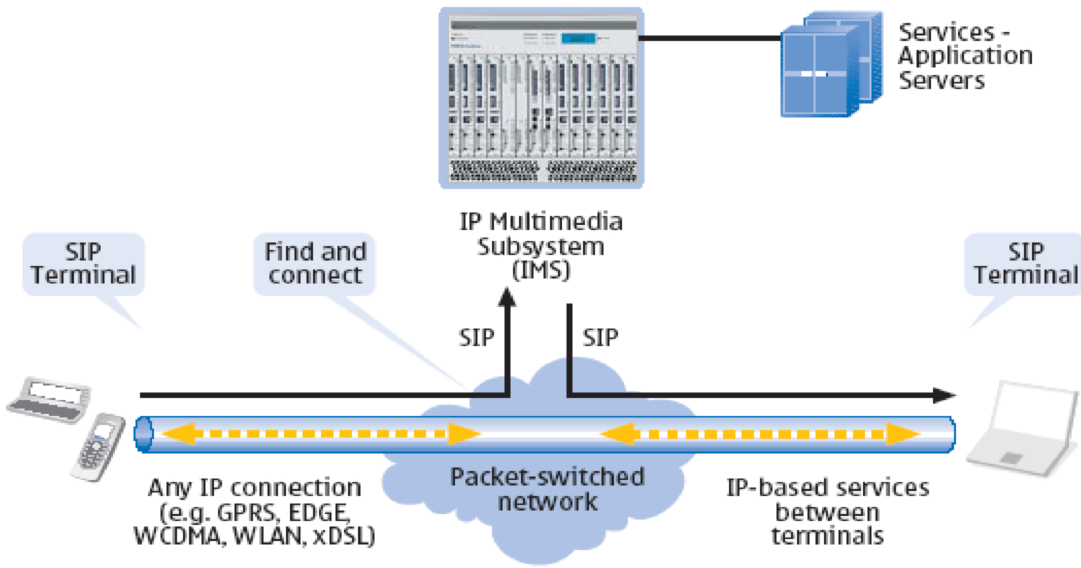

Figure 2.2: SIP based session management []

<u>Proxy CSCF (P-CSCF)</u>

The P-CSCF (Proxy CSCF) is the local contact point of user in the visited IMS network. It acts as a SIP proxy server and forwards all the requests to the directed addresses. This may contain the PDF (Policy Decision Function) for the QoS specification.

<u>Interrogating CSCF (I-CSCF)</u>

The I-CSCF (Interrogating CSCF) is provided at the entry point to the operator's network. It is connected with HSS through Diameter protocol to retrieve the user location for routing purposes. Its IP address in published in the DNS of domain so that the remote servers can find it as a forwarding point.

<u>2.2.2.2 Databases</u>

- HSS (Home Subscriber Server) is the central database containing the subscription related information and the network entities that provide the user's information. It includes subscriber's profile, authentication and authorization information, user physical location, access parameters and security concepts.

- SLF (Subscription Locator Function) SLF is used as resolution mechanism to locate the HSS containing relative information when multiple HSS servers are involved in network. Both HSS and SLF use the DIAMETER protocol for communication. The Figure shows the Complete IMS Architecture.

2.2.2.3 Media Servers

Media Resource Function is used to provide media related information. Each MRF is further divided into two

- MRFP: implement all media-related functions process and mix the media streams.
- MRFC: Signaling plane node to handle the SIP communication to and from S_CSCF.

These both together provide the mechanism for bearer-related services.

2.2.3.4 Gateways

Several types of gateways are supported in the IMS architecture, for example the architecture includes gateways for converting signal from packet switched IMS network to a circuit switched PSTN or vice versa – Signaling gateways performs lower layer protocol conversion from one network to another.

- Gateways to convert the media data - Media Gateway (MG) and Media Gateway Controller Function (MGCF). The MG interfaces the media planes of two networks. Hence, it converts the media over RTP (in IMS network) to the PCM (Pulse Code Modulation) based transport in the PSTN side.

BGCF (Breakout Gateway Controller Functions)

The BGCF is also a SIP server that performs routing functions when the call is addressed to a circuit switched network such as PSTN. It locates the appropriate gateway at the circuit switched destination network for routing the outgoing call.

Media Resource Function (MRF)

An MRF performs several media functions for the home SIP network, such as mix media streams (in a conference bridge), trans-coding functions, playing announcements. An MRF can be further broken into

- MRF Controller (MRFC)
- MRF Processor (MRFP)

The MRFC acts essentially as a SIP UA interfacing with the S-CSCF) to manage resources of the MRFP, while the MRFP performs all the media functions stated above.

2.2.3 Application Layer

Application Layer contains the different Application servers like SIP AS, CAMEL servers to provide the actual services to users [5]. The basic services of IMS are controlled by external or core AS servers. Application Server is a SIP component to host and execute services. AS process and impact the incoming SIP request from IMS and also have the capability to originate SIP request. These servers provide the services like SMS, conference call and presence services according to the server description. The registered sessions are directed towards Application servers to provide the actual services. Multiple sessions can be merged into single application [4]. Media servers (MRF) are used to manipulate the media functions such as voice stream mixing.

2.2.3.1 Application Servers

- SIP-AS native IMS application Server. It is SIP based server that hosts wide range of services such as presence, conference etc.
- IM-SSF: IM-SSF was introduced to support the legacy services developed by CSE (CAMEL Service Environment). It hosts and interface with CAMEL application servers using the CAP.

The IMS architecture follows a functional approach rather than a physical one. The functional architecture consists of separate functional components with standardized interfaces between the components. In other words, it defines the functions that need to exist and not the physical boxes or nodes where each function should reside. The decision as to what functions would reside in which node and how to combine several functions into one node or split functions across several nodes is implementation dependent.

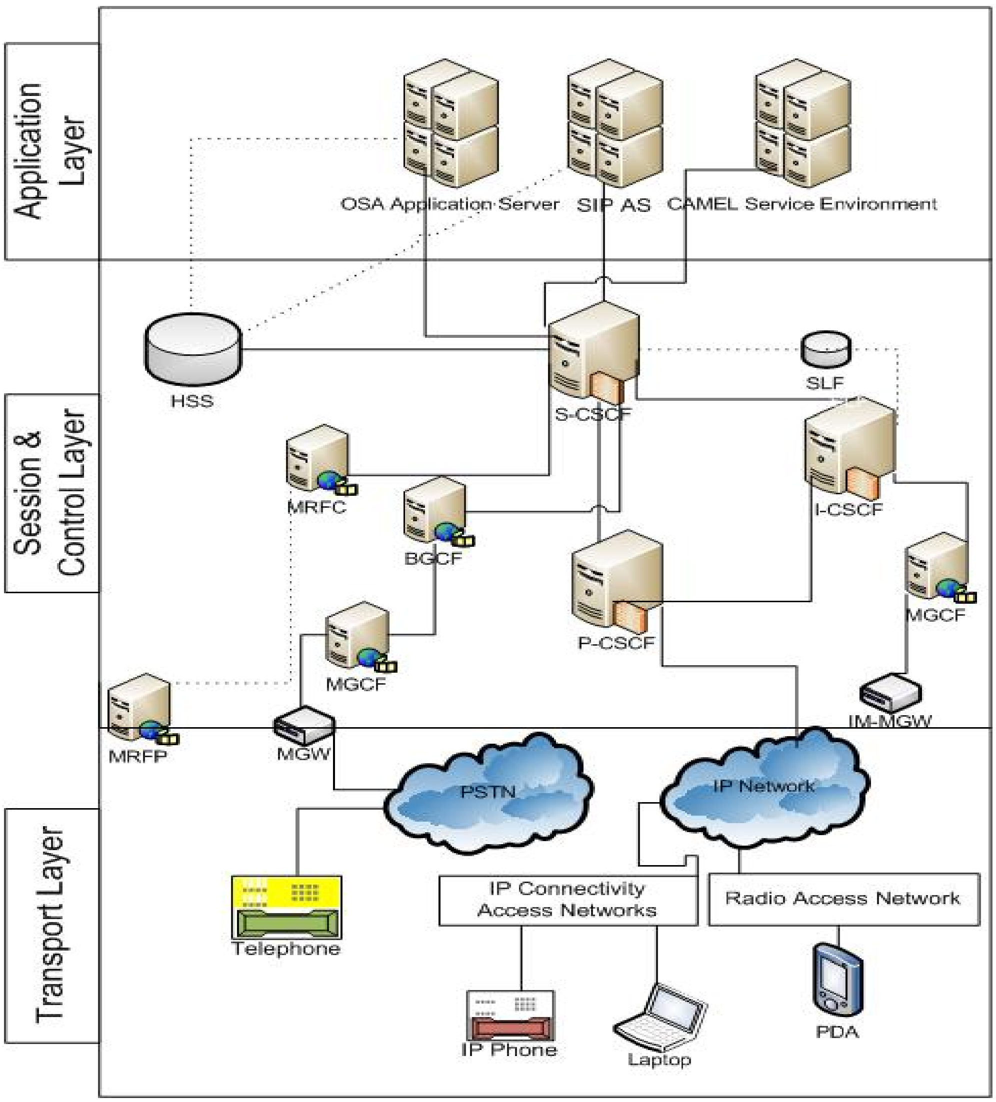

Figure 2.3: IMS Functional Architecture

2.2.4 IMS Terminals

They are typically referred to as User Equipment (UE). Examples of UEs are mobile phones, PDAs (Personal Data Assistant) and computers.

2.2.5 Reference points

Performing functions in IMS network is realized through procedures which define the flows between functional components. The interfaces exposed by the functional components and the control between the components are referred to as "reference points". The following is the description of the reference points for the IP Multimedia Core Network Subsystem.

Table 2.1: IMS Interface Description

Interface Name	IMS Entities	Description	Protocol
Cr	MRFC,AS	Used by MRFC to fetch documents (scripts and other resources) from an AS	HTTP over dedicated TCP/SCTP channels
Cx	I-CSCF, S-CSCF, HSS	Used to communicate between I-CSCF/S-CSCF and HSS	Diameter
Dh	SIP AS, OSA, SCF, IM-SSF, HSS	Used by AS to find a correct HSS in a multi-HSS environment	Diameter
Dx	I-CSCF, S-CSCF, SLF	Used by I-CSCF/S-CSCF to find a correct HSS in a multi-HSS environment	Diameter
Gm	UE, P-CSCF	Used to exchange messages between UE and CSCFs	SIP
Go	PDF, GGSN	Allows operators to control QoS in a user plane and exchange charging correlation informations between IMS and GPRS network	COPS
Gq	P-CSCF, PDF	Used to exchanged policy-decisions related information between P-CSCF and PDF	Diameter
ISC	S-CSCF, I-CSCF, AS	Used to exchanged messages between CSCF and AS	SIP
Ma	I-CSCF -> AS	Used to directly forward SIP requests which are destinated to a Public Service identity hosted by the AS	SIP
Mg	MGCF -> I-CSCF	MGCF converts ISUP signaling to SIP signaling and forward SIP signaling to I-CSCF	SIP
Mi	S-CSCF -> BGCF	Used to exchange messages between S-CSCF and BGCF	SIP
Mj	BGCF -> MGCF	Used to exchange messages between BGCF and MGCF in the same IMS network	SIP

Mk	BGCD -> BGCF	Used to exchange messages between BGCFs in different IMS networks	SIP
Mm	I-CSCF, S-CSCF, external IP network	Used for exchanging messages between IMS and external IP networks	Not Specified
Mn	MGCF, IM-MGW	Allows control of use- plane resources	H.248
Mp	MRFC, MRFP	Used to exchange messages between MRFC and MRFP	H.248
Mr	S-CSCF, MRFC	Used to exchange messages between S-CSCF and MRFC	SIP
Mw	P-CSCF, I-CSCF, S-CSCF	Used to exchange messages between CSCFs	SIP
Rf	P-CSCF, I-CSCF, S-CSCF, BGCF, MRFC, MGCF, AS	Used to exchange offline charging information with CCF	Diameter
Ro	AS, MRFC	Used to exchange offline charging information with ECF	Diameter
Sh	SIP, AS, OSA SCS, HSS	Used to exchange information between SIP AS/OSA SCS and HSS	Diameter
Si	IM-SSF, HSS	Used to exchange information between IM-SSF and HSS	MAP
Sr	MRFC, AS	Used by MRFC to fetch documents (scripts and other resources) from an AS	HTTP
Ut	UE, AS (SIP AS, OSA SCS, IM-SSF)	Enables UE to manage information related to his services	HTTP(s)

2.3 Protocols in the IMS

In any communications network, protocols used fall into two basic categories: signaling and control plane, and media or user plane. The IMS is built based on IP protocols and signaling and control protocols for session initiation and control are based on the SIP and session description protocol (SDP). To transport IMS signaling protocols, the reliable streaming control transmission protocol (SCTP) or transmission control protocol (TCP) is used. Media plane protocols used for media delivery are based on the real-time transport protocol/ real-time control protocol (RTP/RTCP) for transporting real-time media such as audio or video. Near-real-time streaming media are transported using the real-time streaming protocol (RTSP). Both RTP/RTCP and RTSP typically use the user datagram protocol (UDP) as the transport protocol to avoid TCP's setup, teardown, and retransmission delays.

2.3.1 Session Initiation Protocol

The IETF Session Initiation Protocol (SIP) as described in RFC3261 is an application layer protocol for establishing, terminating and modifying multimedia sessions within an IP network. SIP has been embraced as the specified protocol in support of session control protocol for IMS which follows a client server model.

2.3.1.1 Purpose of SIP

SIP is a general-purpose application-layer protocol designed to establish, modify, and terminate multimedia sessions in IP networks. It also allows other participants to be invited to ongoing sessions. The main goal of SIP is to deliver a session description to a user at the user's current location. Once the user has been located and the initial session description delivered, SIP can deliver new session descriptions to modify the characteristics of the ongoing session or to terminate the session. In short, SIP supports the basic aspects of the multimedia session: user location, user availability, user capabilities, session negotiation, and session management.

2.3.1.2 Session Descriptions

A session description contains enough information for a remote user to be able to establish, join, modify, or terminate a session. A session description could include information such as the IP addresses and port numbers to which the media services need to be sent and the coder-decoders (codecs) used to encode the voice, image, and video elements. SIP uses SDP, the most common format to describe a multimedia session.

2.3.2 Session Description Protocol

Session Description Protocol (SDP) is a text based protocol which describes the multimedia session. For example, when initiating a session the caller and callee indicate and exchange their media capabilities as well as receive address and port number. SDP provides an offer/answer model where media attributes such as codec types and capabilities can be offered and negotiated.

2.3.3 Real-Time Protocol

Real-time Protocol (RTP), specified in RFC3550, provides a mechanism to transport real-time multimedia traffic including video and audio over unreliable transport mediums such as User Datagram Protocol (UDP). RTP contains the necessary attributes to ensure correct media buffering and jitter management by providing a timing relationship between source (generate) and sink (terminate) of the media session.
RTCP: RTP Control Protocols are used to monitor QoS of real time data distribution and also convey the session control information.

2.3.4 Diameter

Is used by the network and the user to authenticate and authorize each other. It has a base protocol, complemented by so-called Diameter applications that are customized extensions to the base Diameter to suit a particular application in a given environment. Interacts with SIP during session setup in one application, performs credit control accounting in another application, etc.

2.3.5 Common open policy service (COPS)

This is a request/response protocol used between the policy server (the policy decision function [PDF], also known as the policy decision point [PDP]) and the policy client (the policy enforcement point [PEP]).

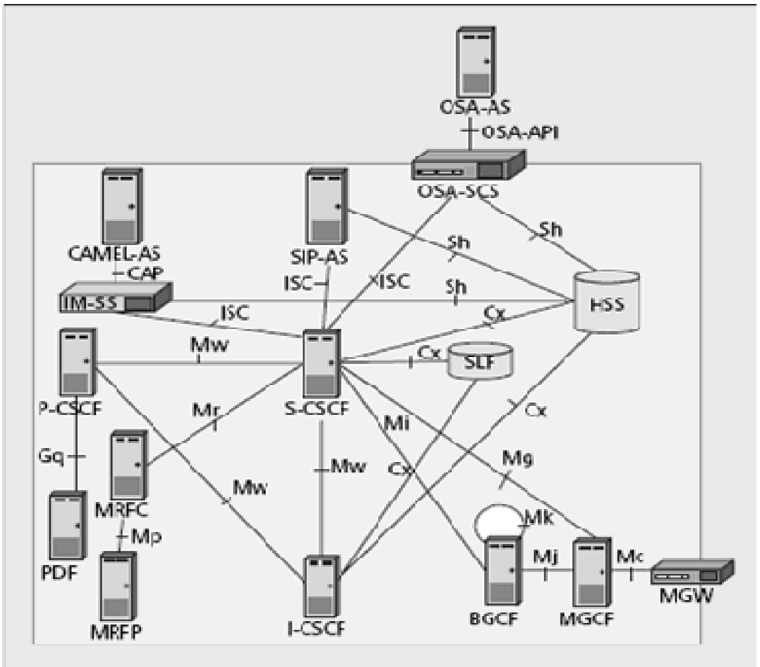

Figure 2.4 Interface Description

2.4 IMS Standards

Different elements and components of the IMS architecture are defined and standardized by different fora. The 3rd Generation Partnership Project (3GPP) and 3GPP2 standardization fora have specified the requirements and defined the overall architecture.

2.4.1 Third Generation Partnership Project (3GPP)

The Global System for Mobile Communication (GSM) during the early 1990s was created by the European Telecommunications Standards Institute (ETSI). The same body was also responsible for the standardization of the General Packet Radio Service (GPRS) known as the first step in the evolution of the GSM network towards a true third generation system. The purpose of 3GPP creation was to develop a third generation

telecommunication system based on the GSM specifications. The IMS was first introduced in 3GPP Release 5. All technical reports and specifications are available to the public through the 3GPP website at:
http://www.3gpp.org/specs/specs.htm

2.4.2 Third Generation Partnership Project 2 (3GPP2)

The 3GPP planned the evolution of the cellular networks based on the GSM (European) specifications into a third generation system. After the successful approach, a similar need was felt to create an organization which would do the same for the North American and Asian cellular networks based on ANSI standards and this led to the creation of 3GPP2. The organization structure of 3GPP2 closely resembles with 3GPP and the technical work being done by Technical Specifications groups whose work is overseen by the Steering Committee (SC). 3GPP2 suggests that the IMS was first introduced in Release A of the specifications.

2.4.3 Internet Engineering Task Force (IETF)

The IETF is known as an organization of operators, vendors, network designers and researchers. Their common goal is to work towards the evolution of the Internet architecture and protocols. The Internet Engineering Task Force (IETF) and more specifically the SIP, SIMPLE and SIPPING working groups have been largely defining the protocols used by the network elements to communicate. These working groups develop protocols related to session and conferencing management (e.g. SIP, SDP), presence and instant messaging (e.g. SIP Message & SIP Events) and list, authorization policies and configuration management (e.g. XCAP). Moreover these groups have specified base formats for presence information and list representation that will be exchanged between the core elements and the service enablers.

2.4.4 Open Mobile Alliance (OMA)

OMA (http://www.openmobilealliance.org) was formed in 2002 by the mobile industry. The Open Mobile Alliance (OMA) is mainly defining applications that are built on top of the IMS infrastructure, such as Push to Talk, Messaging and Presence. The OMA develops the use cases, requirements, architecture and specifications for those applications, such that IMS features are re-used in a network agnostic manner.

2.5 IMS Features

2.5.1 Multimedia Session Negotiation and Management

IMS uses the SIP protocol (Session Initiation Protocol) for multimedia session negotiation and session management. IMS is essentially a mobile SIP network designed to support this functionality, where IMS provides routing, network location, and addressing functionalities.

2.5.2 Mobility Management

The underlying IMS infrastructure enables mobile IP communication services via its ability to find other users in the network and then to establish a session with that user. The key IMS components enabling mobility management are the CSCF (Call Session Control Function) and HSS (Home Subscriber Service). The HSS holds all of the key subscriber data and enables users (or servers) to find and communicate with other end users.

2.5.3 Quality of Service (QoS)

IMS will provide an effective and standardized solution for operators who want to implement real-time IP mobile services without gambling on best effort transmission and the resulting customer dissatisfaction.

The Quality of Service (QoS) mechanisms were developed in order to overcome these issues and provide some type of guaranteed level of transmission instead of ‘best effort’. QoS ensures that critical elements of IP transmission such as transmission rate, gateway delay and error rates can be measured, improved and guaranteed in advance. Users are able to specify the level of quality they require depending on the type of service and the users' circumstances.

2.5.4 Service Execution, Control and Interaction

In a complex mobile service landscape wherein the operator has deployed a large number of services, it is absolutely crucial that the operator is able to control the invocation of services and the interaction between the various service components. In the CS and PS domains, service execution is application-controlled, which makes service interaction increasingly complex and reduces overall service transparency and control. IMS meets this challenge by providing efficient service provisioning functionality.

2.5.5 Third Party Developer Interfaces

Lastly IMS provides the standardized architecture for enabling advanced IP service deployment. Varieties of IMS services can be developed independently and at the same time utilize the common features of the IMS infrastructure. This will facilitate service integration, as well as interoperability (e.g. between mobile and fixed network). Additionally, roaming functionality is automatically supported with little or no additional effort.

2.6 IMS Services

The mobile industry is in a transition phase from traditional voice and short message service centric business to a variety of new and exciting multimedia services and applications [5]. Telephony and messaging is about to be complemented by the next generation of person-to-person applications, making sharing easier - two-way radio sessions (Push-to-Talk), sharing a view, sharing files, shared whiteboards and multiplayer game experiences. It also brings the ability to combine existing services in exciting ways, for example when adding a multiplayer game during a Push-to-Talk session.

IP Multimedia Subsystem (IMS) will also enable new services between mobile and fixed devices. Examples of such services can be content sharing or messaging services between a mobile terminal and PC. A selection of those services is described below.

2.6.1 Presence service

Presence service is a network service which accepts stores and distributes presence information. Presence service may be implemented as a single server or have an internal structure involving multiple servers and proxies. There may be complex patterns of redirection and proxying while retaining logical connectivity to a single presence service. Also presence service may be implemented as direct communication among presentity and watchers, i.e. server is not required.

A user client may publish a presence state to indicate its current communication status. This published state informs others that wish to contact the user of his availability and willingness to communicate. The most common use of presence today is to display an indicator icon on instant messaging clients, typically from a choice of graphic symbol with an easy-to-convey meaning, and a list of corresponding text descriptions of each of the states.

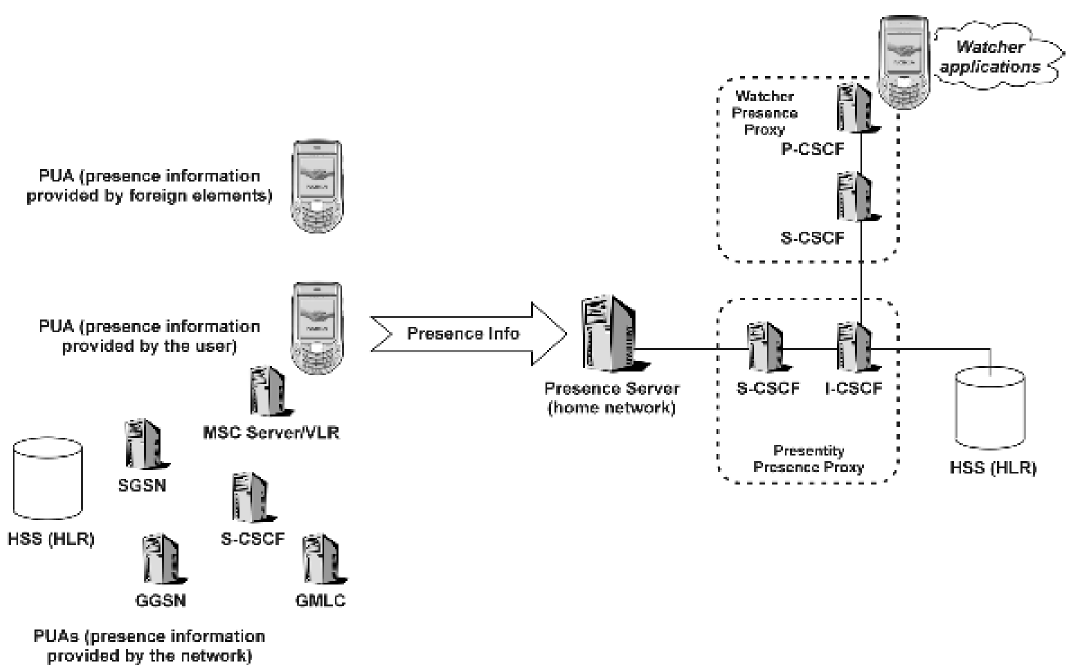

Figure 2.5: Reference architecture to support a presence service [15]

Common states on the user's availability are "free for chat", "busy", "away", "do not disturb", "out to lunch". Such states exist in many variations across different modern instant messaging clients. Current standards support a rich choice of additional presence attributes that can be used for presence information, such as user mood, location, or free text status.

2.6.2 Push to Talk over Cellular (PoC)

Push to talk over Cellular (PoC) introduces a direct one-to-one and one-to-many voice communication service in the cellular network. It makes a popular two-way radio service available through attractive cellular phones, thus enhancing cellular services and bringing new business opportunities to the domain of real-time voice communication.

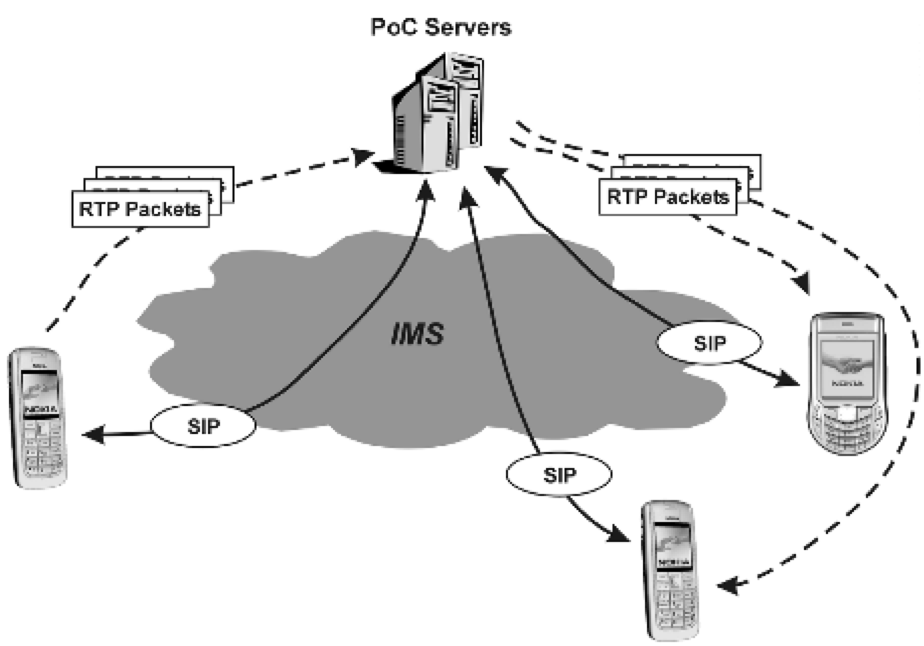

Figure 2.6: Push to talk over cellular [15]

Thanks to GPRS/EDGE technologies, PoC uses cellular access and radio network resources very efficiently. Network resources are thereby used only one-way for the duration of talk spurts instead of two-way for an entire call session. Users can communicate in both one-to-one and one-to-many fashions, with short set-up times.

2.6.3 Real Time Video Sharing

A real time video sharing service is a peer-to-peer, multimedia streaming service that can be offered entirely as a packet switched service or as a "combinational" service, combining the capabilities of the circuit switched and IMS packet switched domains.

In a combinational scenario, the service enriches the user experience during a circuit switched telephony call by, for example, exchanging pictures, video clips or live video over a simultaneous IMS packet-switched connection. This enables operators to leverage upon circuit switched infrastructure, telephony performance, user behavior and experience already in place. Even following the evolution to All-IP networks, CS Telephony enrichment will be maintained for the gradually diminishing CS domain.

2.6.4 Interactive applications

2.6.4.1 Interactive Gaming

In April 2004, mobile users worldwide downloaded already an estimated 14 million Java games per months. Taking the end-user demand for basic gaming with mobile terminals into consideration, it is not difficult to foresee even faster services take off with the capability to establish interactive gaming sessions between players.

2.6.4.2 Shared folders

Content sharing enables users to share files between terminals. A typical use case includes sharing of files or contents such as images, documents, notes, contacts or calendar information, even simultaneously while in a voice-call session.

2.6.5 Instant Messaging services

Instant Messaging is a communication service that allows end-users to send and receive messages instantly. Instant Messaging is well known in today's Internet community. IMS will bring the same service experience to the mobile world, including interoperability without the need for legacy infrastructure.

Messages can contain any MIME type media content such as text, image, audio or video clips, application data or a combination of these. The message is sent through the packet data network to the IP Multimedia Subsystem (IMS), which locates the terminating IP client and routes the message to the recipient.

2.6.6 Voice Messaging

Voice Messaging is a form of instant messaging where the content of the message is an audio file. Using an application in the terminal, users can record the message instantly or use existing audio files stored in the terminal's folders. Voice messages can be sent to one or many recipients. There are many different potential ways to use Voice Messaging, but they are all based on certain common factors:

2.6.7 IMS enabled Voice and Video Telephony

IMS enabled Voice and Video calls are carried over a packet core network (VoIP). Video Telephony is seen as a critical end-user service in Mobile Networks. The Session Initiation Protocol (SIP) enables Voice and Video Telephony Person-to-Person and Multiparty sessions over an IP network.

2.6.8 Video-Conferencing

IP Multimedia Subsystem (IMS) Video-conferencing service extends the point-to-point video call to a multi-point service. Videoconferencing requires an IMS conference bridge service, which links the multiple point-to-point video calls together, and implements the associated service logic. The video telephony connections are made point-to-point from the terminals to the conference bridge, which takes care of joining the point-to-point connections into a conference.

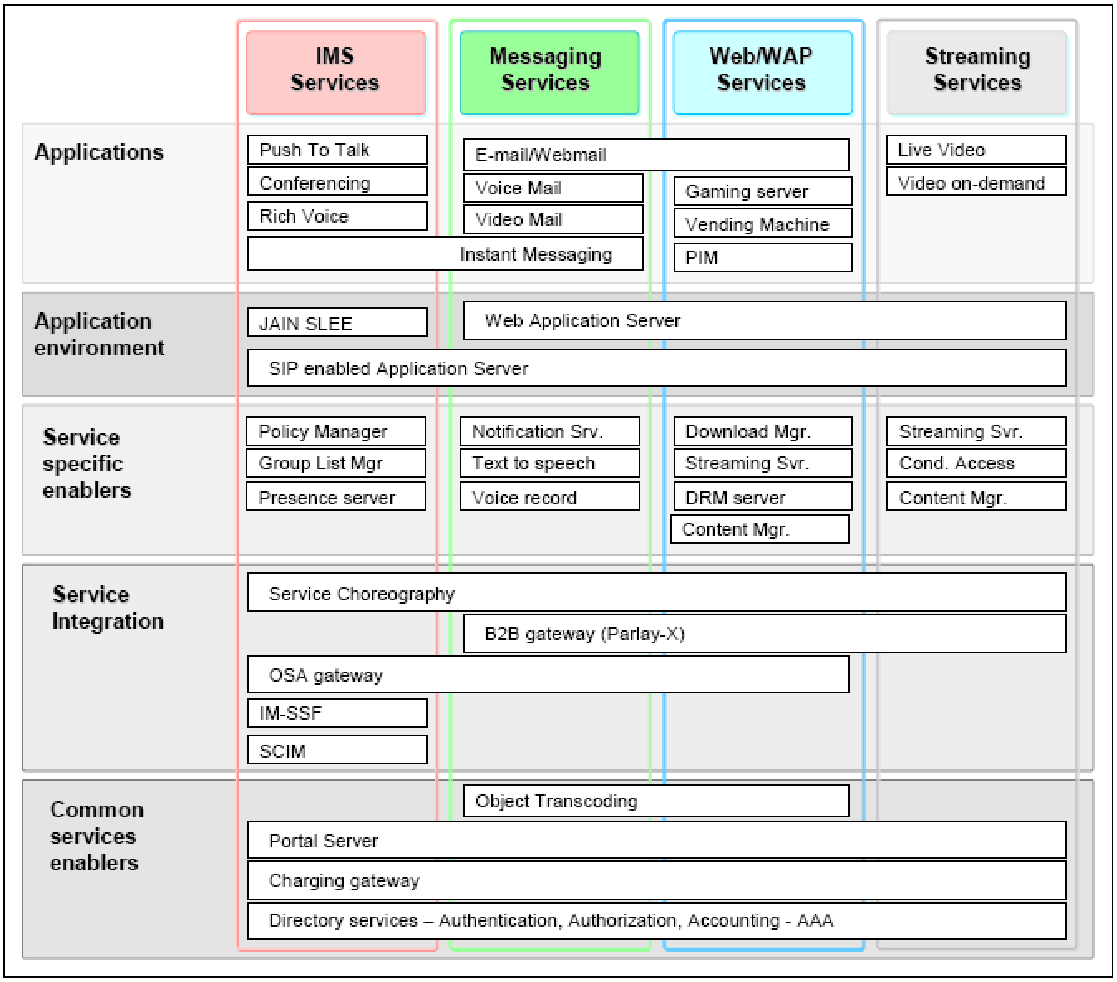

Figure 2.7: IMS Service Architecture [36]

2.7 IMS Security

IMS security can be divided into two areas: access security and network security. Access security involves authentication of users and networks and protection of traffic between IMS terminals and networks. Network security involves protection of traffic between security domains. To provide network security, all traffic entering or leaving a security domain passes through a security gateway (SEG), as shown in Figure. Within a security domain, network nodes use IPSec to exchange traffic with each other and with the SEG. In addition to the mandatory network layer security (IPSec), the IMS also provides optional transport layer security (TLS) and application layer security (e.g., HTTP digest authentication for SIP).

2.7.1 The Seven Domains of IMS Security

To avoid complexity in the IMS security design, the IMS network is divided into six domains which are based on security targets and threats. These six domains include:

- Service Domain

  The service domain contains products that collaborate to provide IMS services such as, CSCF, HSS, MGW, RMS and AS. ed services while, basic services security depends on that of the fundamental services.

- Interconnection Domain

  The interconnection domain provides interconnection between the IMS network and the networks of other carriers, as well as third-party service providers and the Internet.

- Core transmission

  The core transmission domain transmits IP streams between each domain.

- Access Domain

  The access domain provides subscribers access to IMS services via broadband IP networks. The access domain also connects trusted service and distrusted user domains.

- User Domain

  The user domain consists of the equipment and networks of subscribers such as, fixed terminals, mobile terminals, and the user network.

- Management Domain

  The management domain is used in the internal network. It has a higher level of security, which is under strict management.

- DMS Domain

  The DMZ domain is set up for communication among the six security domains. The DMZ domain is the focus of IMS security control.

Chapter 3

Quality of Service over IP Access Networks

Introduction

In recent years Quality of Service for IP networks attain the massive attention which has escort to competing IP QoS Solution being developed. Quality of Service is a method to determine the satisfaction level of service performance athwart networks. It can also be implicit as the capacity of a network to provide a control desired of the traffic that the network complies with the level of service that the final applications to expect. Mechanisms and Quality of Service (QoS), QoS utilizes to handle the congestion in a network for guarantee that the appliances have enough resources of network to be communicated in an efficient and adequate way for the variety of traffic that utilize.

Consequences of QoS procedures on the traffic IP can be studies by analyzing the methods to deliver QoS are necessary for be treaties. The purpose of this chapter is to explain the process of measurement of metric of quality of service as the performance, the latency and jitter for the multimedia traffic and to devise the methodologies of current IP QoS. Configuration of network for constant and expected end-to-end concert can be done by employing QoS, and services of network advertised with the intention that the prices accused are modified to the 'end-users' aspiration service level and price point. In these days, network communication presentation exhortations comprise of following:

- Jitter
- Latency
- Packet Loss
- Bandwidth

3.1 How to deliver QoS?

A QoS-enabled network should be capable of handling different traffic flows in diverse methods [19]. This obliges to classify the traffic in the types or classes and the definition of how each type is handled. Enclosing the scope of QoS subsequent features could be consider [25]

- Admission control
- Differentiation of traffic (Classification)
- Congestion management
- Queuing

3.1.1 Admission Control

A packet can be categorized into two classes, when it appears at the entrance of a network, specifically called ‘in profile’ and ‘out of profile’. It will be categorize as ‘in profile’ if a packet match to a contract with the network owner and will be permitted to penetrate the network with or without any changes to the packet. But if the packet is not adjusted then can fall in the point of entrance in the network. Occasionally due to the shortage of the resources accessible to handle them, packets may be plummeted, apart from a packet is in profile or out. A device can evaluate the accessible resources of the network thorough different ways.

3.1.2 Differentiation of Traffic

IP routers are planed to handle every packet in the similar approach. Packets incoming at the router are regard as new event exclusive of containing any local state memory. So the basic construction chunk here is the aptitude to categorize the traffic into diverse group based on what service they are to collect. Once the packets have been categorized they can be given discrepancy handling based on the categorization. The categorization of a packet can be done into a service class, based on the following constraints:

- Protocol
- Source Protocol port
- Service Mark
- Source host address
- Destination Protocol port
- Source device interface

- Destination host address
- Any grouping of above

3.1.3 Congestion Management

The capability of a network to efficiently pact with weighty traffic volumes is called Congestion management. This feature of QoS is consider beyond a single entity and frequently includes multi-tool to make it exertion. In congestion management, queuing is among those tool which executed for the network and router.

3.1.4 Queuing

In a network each router is a shared resource and any packet receiving a given router can be judged as request for service. Organization the queuing system is the primary phase that establishes the quality of given service which generates the fundamental demarcation among the service levels of different traffic classes. Important methods of queuing are following:

- Priority queuing
- First in First Out (FIFO) queuing
- Weighted Fair Queuing (WFQ)
- Class based queuing (CBQ)

3.2 QoS Measurements

Best example where all packets are handling indistinguishably and share congested conditions are applied, is internet. To guarantee satisfactory service quality in integrating voice and other media, such as data, voice and gaming onto a single network should organized depending on the uniqueness and assortment of each type of traffic.

The data clocking rate of the system is bandwidth, usually articulated in bits per second. Latency passes on to delay, reasoned by the physical traits of the packet reassembly, transport media, queuing delay and processing delay. Jitter is reasoned by latency variations and discrepancy of delay over time.

3.3 QoS Mechanisms

There are a range of methods that can be utilized for provision of quality of service for IP networks and it is not achievable to judge all elucidation here. Therefore it is recommended to scrutinize the most likely applicants for solving the multimedia QoS provisioning problem [35]. Here is the description of IntServ, DiffServ and RSVP MPLS.

3.4 Integrated Services/IntServ

IETF working group present the Integrated Services (IS) model [22] to propose a extensions set to carry the best-effort traffic delivery, currently in place of real time applications, as well as the in the internet. Video broadcasting, video conferencing and audio conferencing software require definite bandwidth to present audio and video of suitable quality. It should essential to revise the Internet infrastructure to grant control over end-to-end packet delay, to sustain these service requirements and bandwidth administration.

3.4.1 Integrated Services model

In agreement with the service model an Internet router should be capable to offer a suitable QoS for each flow to support the Integrated Services model. The function of router that offers diverse characteristics of service is known as traffic control. It comprises of the following elements:

- ✓ The packet classifier
- ✓ The admission control
- ✓ The packet scheduler
- ✓ The reservation setup protocol

On the base of function the router can be separated into two classes: to the background code above the line and to the forwarding path below the thick horizontal line. The background routines manage the forwarding path and grip reservation demands. In the routers, forwarding path is implemented for all packets and it is divided into three sections:

Input driver

Internet forwarder

Output driver

IntServ works by permitting hosts to request per-flow, resources end-to-end along a data path. The network will then provide feedback to the hosts notifying them whether or not the network can assist with such requests. During the development of IntServ the focus primarily was on control mechanism functions:

- Admission and Policy control
- Packet Scheduling

Admission and Policy control

If there are sufficient routing resources to agree the requested QoS for a new flow, the admission controls the decision algorithm that a router applies to establish. If routing resources are not satisfactory free, accepting a new flow would brunt on earlier guarantees and the new flow must be rejected.

To confirm the user authentication for a requested reservation numerous strategies will be employed. Requests of unauthorized reservation can be refused. Consequently, admission control can play a vital task in accounting costs for Internet resources in the prospect. The packet classifier recognized IP packets flow in router and hosts that will obtain a definite level of service. To understand efficient traffic control, every incoming packet is plotted by the classifier into a specific class. All packets that are classified in the same class packets will encompass the same handling from the packet scheduler.

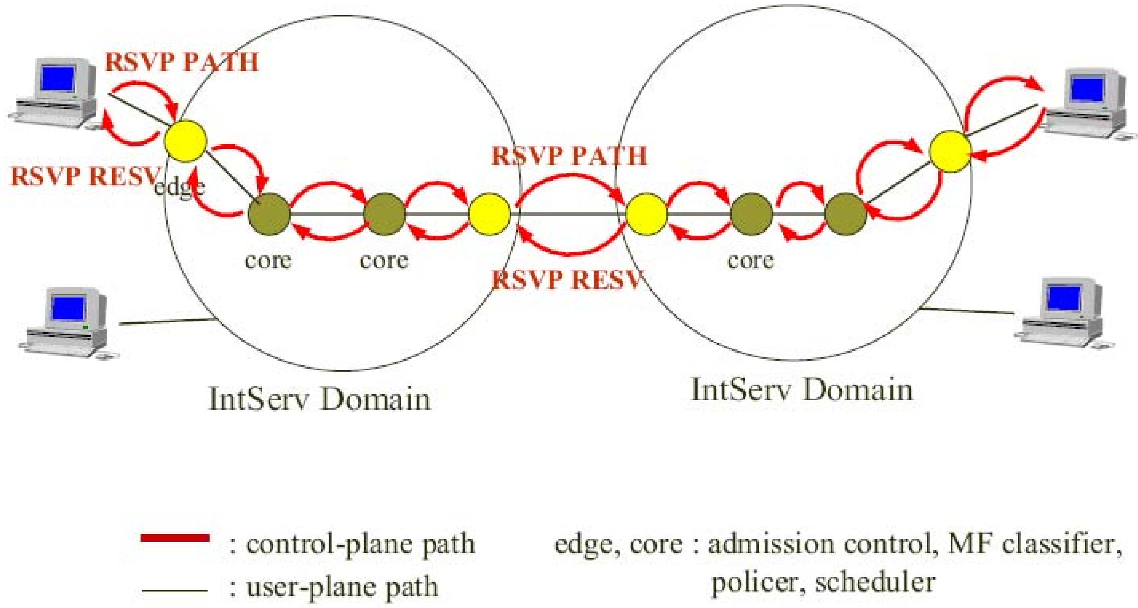

Figure 3.1: IntServ Operational Model [22]

Packet scheduler

The packet scheduler administers the forwarding of diverse packet flows in routers and hosts, depending on their class of service, via queue management and different algorithms of scheduling. The packet scheduler should guarantee that the packet delivery for every flow matches to the QoS parameter. A scheduler is able to shape or police the traffic to ensure a definite level of service. The packet scheduler should execute where packets are queued. This is usually matches to the link layer protocol and the output driver level of an operating system.

3.5 Resource Reservation Protocol

The resource reservation protocol (RSVP) [25] is applies by the IntServ model to handle the prerequisite in IP network of QoS. On a per-flow basis, RSVP and IntServ present unidirectional resource reservations. RSVP may carry IntServ information as a signaling protocol, but is not the only one: RSVP and IntServ are divisible, but the recent established model is based on a combination of both RSVP and IntServ.

In this model the message is updated at each router on the path when the sender of a flow first drives a PATH message to the receiver. The receiver reacts with a RESV message and point outs the resources required at each hop to sustain the forthcoming flow. If resources are insufficient at any router on the end-to-end, path perhaps rejects the flow. RSVP signaling is shown in figure. With the requested QoS the RSVP protocol strive to locate a flow reservation, which will be acknowledged if the routers can handle the requested QoS and the application fulfills the policy restrictions.

RSVP recommends the packet scheduler and packet classifier in every node to process the packets for effectively flow. In the first node if the application distributes the data packets to the classifier, which has plotted this flow into a precise service class submitting with the requested QoS. Depending on their service class the packet scheduler promotes the packets to the subsequently router or lastly to the receiving host.

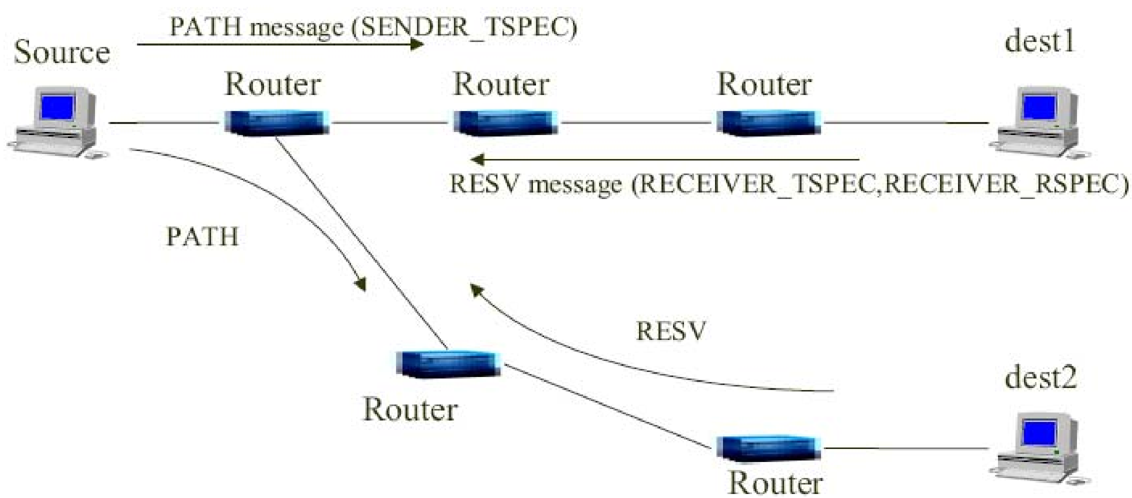

Figure 3.2 RSVP Signaling [21]

In regards to QoS, the host that is receiving sends a reservation-request (RESV) message indicating what service class it is able to support. The reservation messages travel through the network in a reverse path until the sending host receives them. The RSVP protocol makes an effort to make a resource reservation at each network component from which the application flow will pass through. To be aware of QoS requirements particular in each RSVP request the RSVP-enabled routers, as well as, other network components need to maintain "state" on the individual RSVP flows. RSVP institutes a temporary, or soft, state on all RSVP devices between sending and receiving nodes. The protocol will also send intermittent messages to refresh state. State will be deleted if there are no refresh messages received within an allotted.

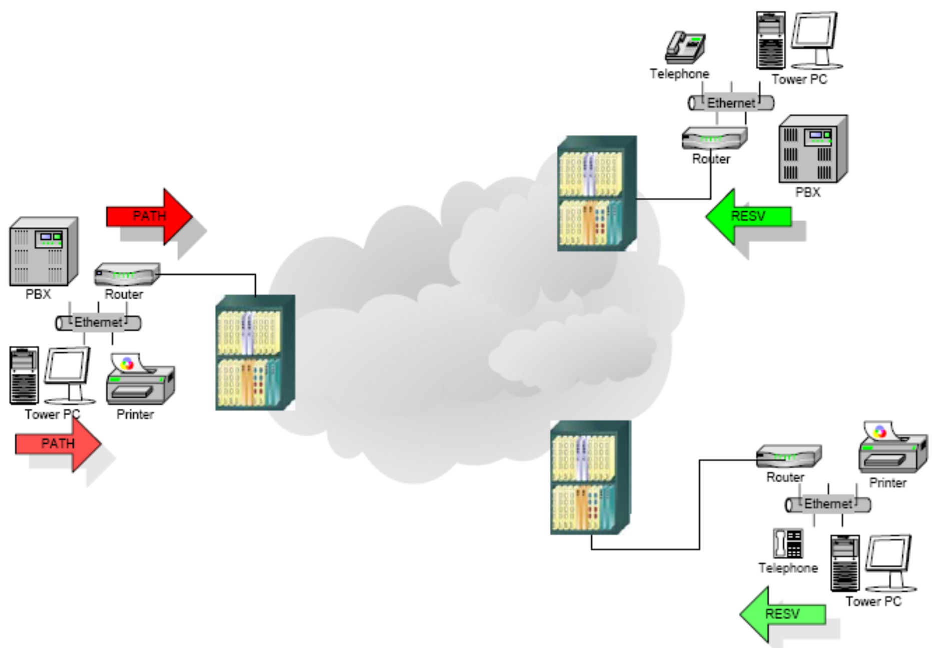

Figure 3.3 Resource Reservation Protocols Architecture [25]

3.6 Differentiated Services

Differentiated Services (DiffServ) [27] is a direct extension of the work carried out by the Services Integrated and RSVP working parties. The objective is basically the same one for both architectures: to provide quality of service and differentiation of services in networks IP. While IntServ give all the guarantees by flow, services differentiated continues the philosophy of the cartography multiple flows levels of service in a few one, a focus at times is called the class of service (CoS). The basic principle of DiffServ is to maintain the architecture of it more simple possible and therefore provides a differentiation of services in a single direction. As opposed to RSVP, DiffServ is oriented sender, what signifies that the sender of the traffic is responsible for the quality of service of the broadcast. The other main difference among DiffServ IntServ and the fact is that IntServ offers classes of service, while DiffServ specifies the blocks of construction of the one that the services can be built. A central component of the architecture DiffServ is the Service Level Agreement (SLA).

3.6.1 DS field

Differentiated Services architecture consists of a small, well-defined set of building blocks, which are deployed in network nodes. One of the most important blocks is the DS field, which is used to select a specific packet-forwarding treatment.

3.6.2 Per-Hop Behavior

The other main building block of DiffServ is Per-Hop Behavior (PHB). A per-hop behavior is defined as "description of the externally observable forwarding behavior of a DS node applied to a particular DS behavior aggregate" [27].

The objective of the Expedited Forwarding PHB is to support a behavior aggregate of low loss, low latency, low jitter, assured bandwidth end-to-end service through a DiffServ network.

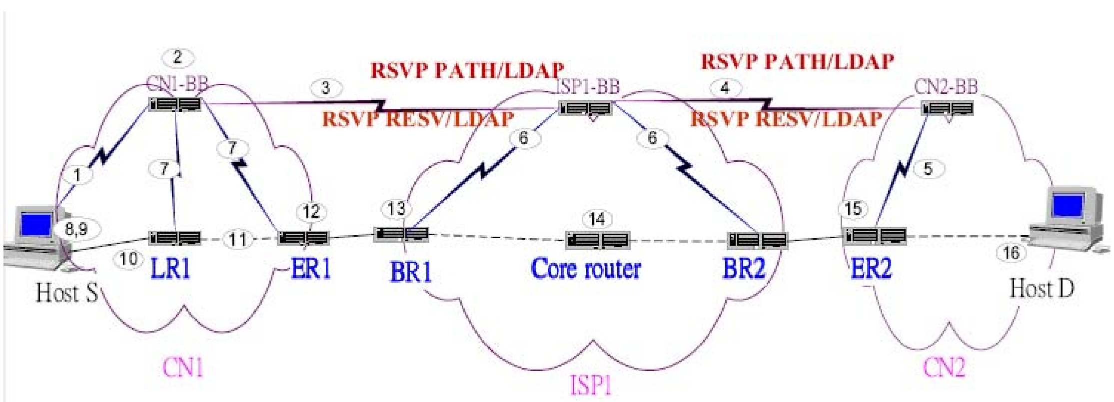

Figure 3.4 Expedited forwarding with dynamic SLA [27]

Thus EF PHB is defined as a forwarding treatment of the traffic aggregate where the configured packet data rate must equal or exceed a real departure data rate. The EF traffic should receive the configured rate independently of the intensity of any other traffic attempting to transit the node. This implies strict bit-rate control in the boundary nodes (i.e. traffic that exceeds the negotiated rate must be discarded) and as quick forwarding in the interior nodes as possible. [31]

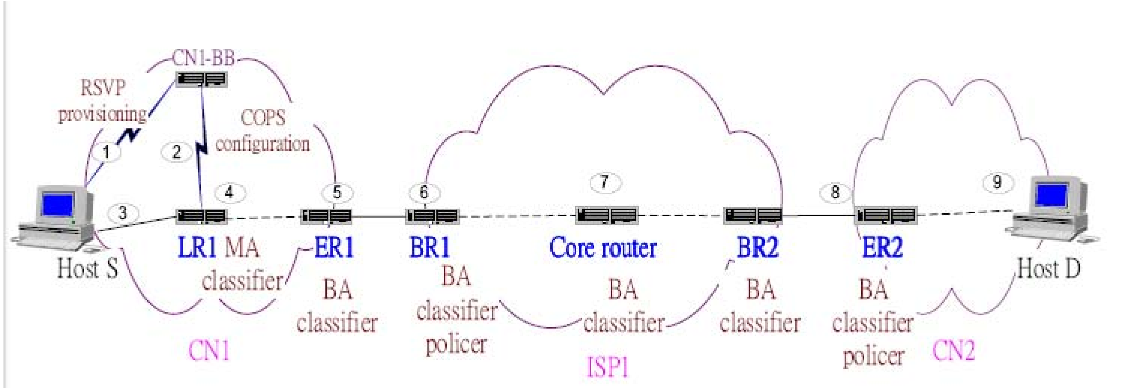

Figure 3.5 Assured Forwarding

3.6.3 Diffserv Domains

The setup of QoS guarantees is not made for specific end-to-end connections but for well defined Differentiated Services domains. The IETF working group defines a Differentiated Services domain as a contiguous portion of the Internet over which a consistent set of Differentiated Services policies are administered in a coordinated fashion. It can represent different administrative domains or autonomous systems and different network technologies. A DS domain consists of boundary components that are used to connect different DS domains to each other and interior components that are only used inside of the domains. The traffic conditioning is done inside a boundary node by a traffic conditioner. It classifies, marks, and possibly conditions packets that enter or leave the DS domain. The most recent conceptual model of a DiffServ router proposed in literature [34] is composed of the following key elements.

Traffic Classification elements Classification is performed by a classifier element. A classifier selects packets based on their packet header and forwards the packets that match the classifier rules for further processing. The DS model specifies substantially two types of packet classifiers:

- Multi-field (MF) classifiers, which can classify on the DS field as well as on any other IP header field.

- Behavior Aggregate (BA) classifiers, which uses only the DiffServ codepoint (DSCP) in a packet's IP header to determine the logical output stream to which the packet should be directed.

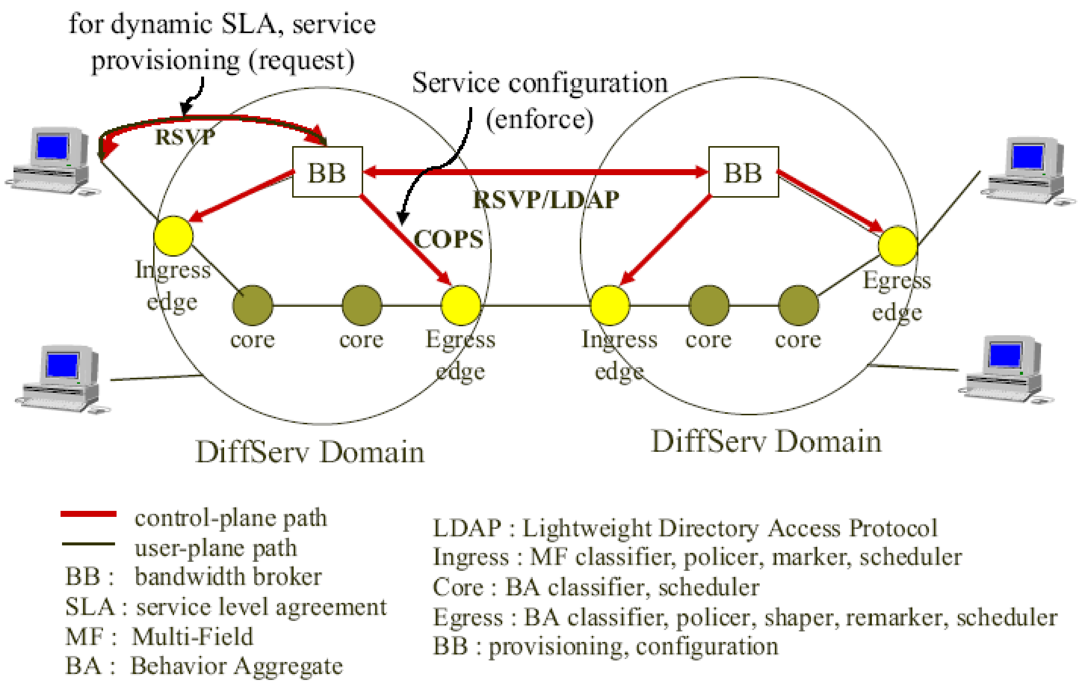

Figure 3.6 Diffserv Operational Model [21]

Metering functions Traffic meters measure whether the forwarding of the packets that are selected by the classifier corresponds to the traffic profile that describes the QoS for the SLA between customer and service provider. A meter passes state information to other conditioning functions to trigger a particular action for each packet, which either does or does not comply with the requested QoS requirements.

Action elements The classifiers and meters described before are generally used to determine the appropriate action to apply to a packet. The set of possible action that can be applied include: Marking, Absolute Dropping, Multiplexing, Counting, Null action.

- *DSCP Marker*
- *Absolute dropper:*
- *Multiplexer*
- *Counter*
- *Null action*

Queuing Elements Queuing disciplines implemented in the queuing elements modulate the transmission of packets belonging to different traffic streams and determine their ordering, possibly storing or discarding them.

3.7 Multi-Protocol Label Switching

The Multi-Protocol Label Switching (MPLS) [32] is specified by the IETF mainly to be used in combination with the DiffServ concept and is an advanced forwarding scheme that extends routing with respect to packet forwarding and path controlling. In a connectionless network layer protocol, a packet traveling from one router to the next causes each router to make an independent forwarding decision for that packet. That is, each router analyzes the packet's header, and each router runs a network layer routing algorithm. Each router independently chooses a next hop for the packet, based on its analysis of the packet's header and the results of running the routing algorithm.

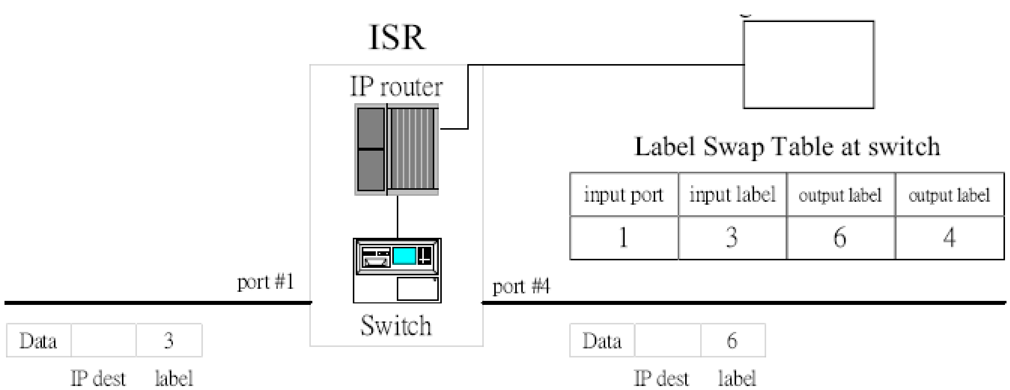

Figure 3.7 MPLS Overview [32]

There are two main approaches in MPLS over the DiffServ implementation field. The first consist on the direct mapping between each DiffServ codepoint and a new MPLS label-switched path. This results in one label-switched path per Forwarding Equivalence Class per DiffServ per- hop- behavior.

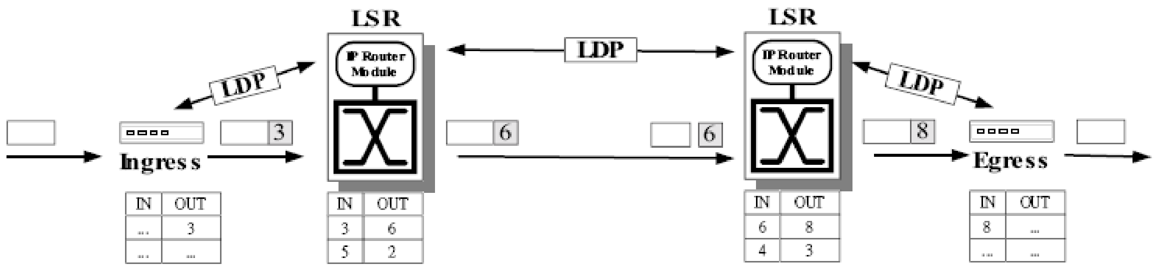

Figure 3.8 MPLS Label Swapping [21]

Chapter 4

Quality of Service in IP Multimedia Subsystems

Introduction

IP multimedia subsystem (IMS) is the future for all IP next-generation converged networks with the potential of enabling service providers to create and provide value added services to users in heterogeneous networks. However, the services provided by the IMS should ensure QoS. The term "quality of service" (QoS) sums up all quality features of a communication network as perceived by a user of a specific service. As any network entity can be a source of network traffic at anytime, it is desirable that QoS management is dynamically adaptable to end users' requirements. The next generation solutions are designed to provide converged multimedia services such as Triple Play and Quadruple Play services, which are becoming the most celebrated services in the near future. This demands a need of end-to-end QoS between the consumer and the service provider.

Applications like Voice over IP have recently gained a large impetus. IMS is able to provide such a service. However, what is the point in using that extra architecture, when a simple IP network may seem enough, and when the good old Circuit Switch network already provides a high quality voice service, with a guaranteed QoS? First, using IP technologies provides extra services, like conferencing, presence service, etc. And secondly, the IMS architecture provides an End to End Quality of Service (QoS) that is critical in such real time applications, QoS that is not guaranteed in a simple data network. The approach used in IMS for providing QoS, is a policy-based architecture.

The idea of transforming a best effort IP network by introducing end-to-end QoS guarantees is an important driver for the development of IMS. This is a key consideration because the level of QoS that the IMS architecture is able to provide determines the services that can be deployed on it, and the value is assumed to reside in real time multimedia services.

4.1 Background

Mobile networks of the 3rd generation are expected to offer wireless Internet services to the mobile subscriber. These networks are designed for multimedia context, high data rates and packet oriented data services. New value added applications have new (high) service requirements which cannot be fulfilled by existing “best effort” IP service that is subject to unpredictable delays and loss.

Introducing 1P Multimedia Subsystem (IMS) into 3rd generation mobile networks gives new dimension to Quality of Service (QoS) support. The IMS is an application environment that uses the packet oriented (PO) domain as a bearer service. For session control IMS uses SIP (Session Initiation Protocol) protocol defined by Internet Engineering Task Force (IETF). Network services are considered end-to-end, this means from a Terminal Equipment (TE) to another TE. An End-to-End Service may have a certain Quality of Service, provided for the users of a network service.

4.2 IMS QoS Support Mechanism

The QoS support mechanism ensures that the critical elements of IP transmission such as transmission rate, gateway delays, and error rates can be measured and guaranteed in advance [12]. This function is performed mainly via the PDF, which interacts with and controls the underlying packet network (i.e., the access network resources) via the Go interface with an element in the PDG called the PEP. The PDG for the GPRS/UMTS is the GGSN, which hosts the PEP. Policy-related information is transmitted between the PDF and the PEP using COPS (RFC 2478). Currently, two basic methods support QoS on the Internet: integrated services (IntServ) (RFC 2215 [22]) and differentiated services (DiffServ) (RFC 3260 [27]). IntServ is designed to provide end-to-end QoS with two classes of services: controlled load and guaranteed. IntServ uses resource reservation protocol (RSVP) to reserve resources with the desired QoS. RSVP also ensures that the routers receiving resource reservation requests are the routers that will actually route the packets. This function is performed via a two-way handshake, in which one endpoint

(endpoint A) sends a PATH message to the other endpoint (endpoint B), recording all the visited intermediate nodes.

Then, in the reverse direction starting at endpoint B, a RES message is sent through all the nodes recorded in the PATH message, this time actually reserving the resources. Note that a router in the path can reject a resource reservation request either because it does not have the required resources or because the requester does not have the permissions to reserve those resources. Thus, RSVP can be considered as not only a resource reservation protocol, but also an admission control protocol. The main drawback of IntServ is that it does not scale well. This is primarily because (1) the network needs to store a large amount of information, and (2) routers need to look up large tables before they can route the packets.

To address these issues, DiffServ architecture was proposed. In DiffServ, routers identify packet treatment without the need for table lookup. Packet treatments, known as per-hop behaviors (PHBs), are identified by 8-bit codes called differentiated services code points (DSCPs). Packets are marked at the edge of the network with a certain DSCP, so that routers in the path apply the correct PHB to them. DSCPs are encoded in the Types of Service field of IPv4 and the Traffic Classes field of IPv6.

Two examples of PHBs are expedited forwarding and assured forwarding. In expedited forwarding PHB, packets never experience congestion in the network. In assured forwarding PHB, packet-drop precedence is determined, allowing low priority packets to be discarded before high priority packets; some packets may in fact be discarded.

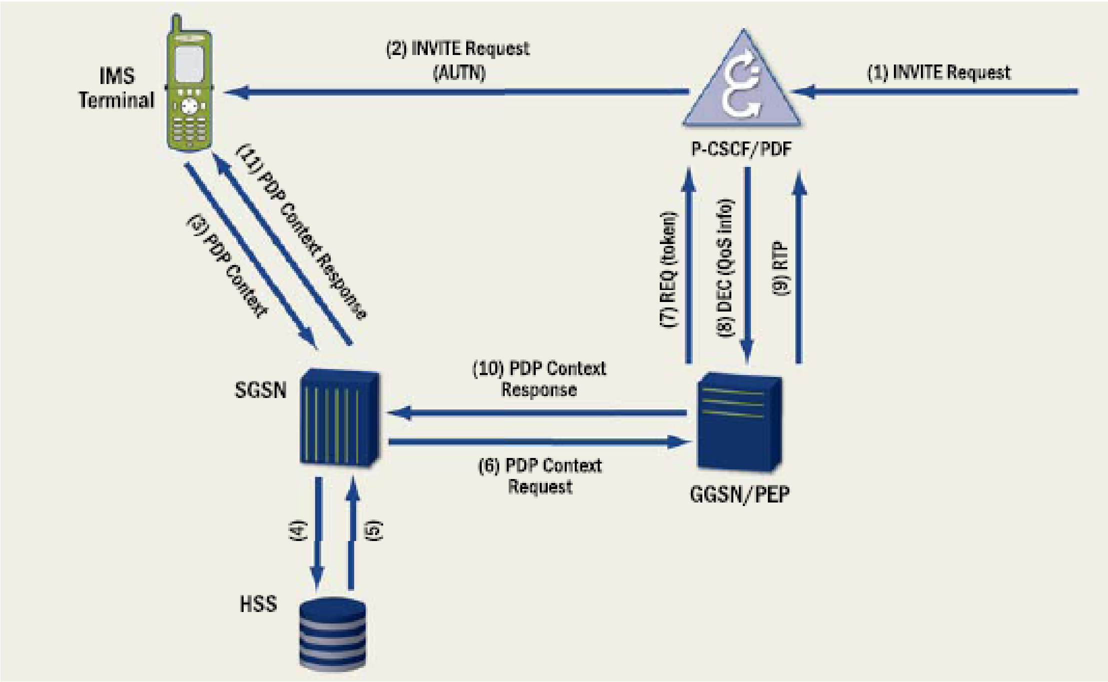

Figure 4.1: QoS Authorization [12]

In the IMS, end-to-end QoS involves both QoS over the access network and QoS in the core network. This implies that QoS-required resources have to be provisioned and enforced on both sides. This can be done by using a link-layer RSVP on the access network side and the DiffServ method (or RSVP) on the network side. On the GPRS/UMTS access network, the link-layer resource reservation is performed via PDP context activation, and the GGSN maps link layer resource reservation flows to DiffServ code points in the network [12].

4.3 Policy-based admission control

The policy-based network is based on two elements,

- Policy Decision Point (PDP)
- Policy Enforcement Point (PEP).

The aim of this policy framework is to transfort sets of policy rules to network devices configuration procedures. These sets of rules are stored in a policy repository (typically an LDAP directory) which is accessed by the PDP when a request is made for example from the PEP. Then, the rules, accordingly chosen given the QoS class requested, are converted into configuration instructions and transmitted to the PEP which will execute them. Two models are defined for the interaction between the PEP and the PDP.

Outsourcing Model

In the outsourcing model, an event, like a resource reservation request, committed to the PEP yields to a request to the PDP for a particular policy. The PDP process the request from the PEP, takes a decision and transfers it to the PEP that enforce that decision.

Provisioning Model

In the provisioning model, the PDP downloads all the policies to the PEP, which will eventually make the decision based on these. In the following, PBN is implemented in IMS to ensure end-to-end QoS communications.

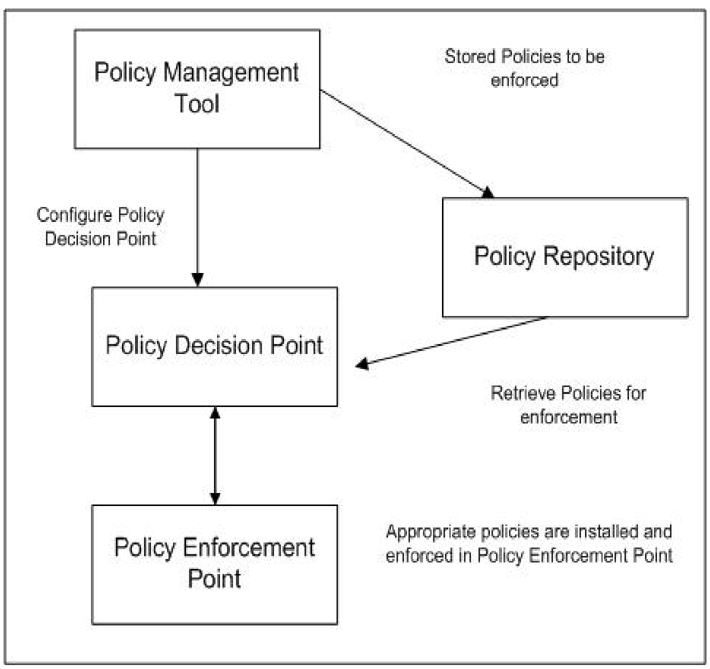

Figure 4.2: Policy-Based Architecture

4.4 IMS Policy-Based QoS Architecture

IMS uses a policy-based mechanism to ensure end-to-end QoS for SIP based multimedia applications. Before elaborating on the IMS QoS provisioning, we briefly describe a simplified 3GPP Release 5 IMS network architecture, where UMTS is assumed to be the underlying IP access network. As illustrated in Figure 4.3, this architecture consists of a radio access network, the General Packet Radio Service (GPRS) core network and the IMS network. The GPRS core network connects to the IMS network through Gateway GPRS Support Nodes (GGSNs; see Fig. 4.3a). The Home Subscriber Server (HSS; see Fig. 4.3b) is the master database containing all user-related subscription information. Both the GPRS and the IMS networks access the HSS for mobility management and session management. A mobile user utilizes a UE (see Fig. 1c) to access IMS services. To provide a data session for the UE, an IP connection between the UE and the GGSN is established. This connection is specified by a Packet Data Protocol (PDP) context. The PDP context must be activated before the UE can access the IMS network. In IMS, the user data traffic and signaling messages are processed separately.

IMS user data is transported through the Media Gateways (MGWs; see Fig. 4.3d), while IMS signaling is carried out by Proxy Call Session Control Function (P-CSCF; see Fig. 4.3e), Interrogating CSCF (I-CSCF; see Fig. 1f), and Serving CSCF (S-CSCF; see Fig. 4.3g). When a UE attaches to the GPRS/IMS network and performs PDP context activation, a P-CSCF is assigned to the UE. The P-CSCF contains limited address translation functions to forward requests (e.g., registration) to the I-CSCF. Authorization for bearer resources in the network (where the UE visits) is also performed by the P-CSCF. The I-CSCF is the contact point for the home IMS network of the destination UE. The role of the I-CSCF is to hide the configuration, capacity, and topology of the IMS network from the external world.

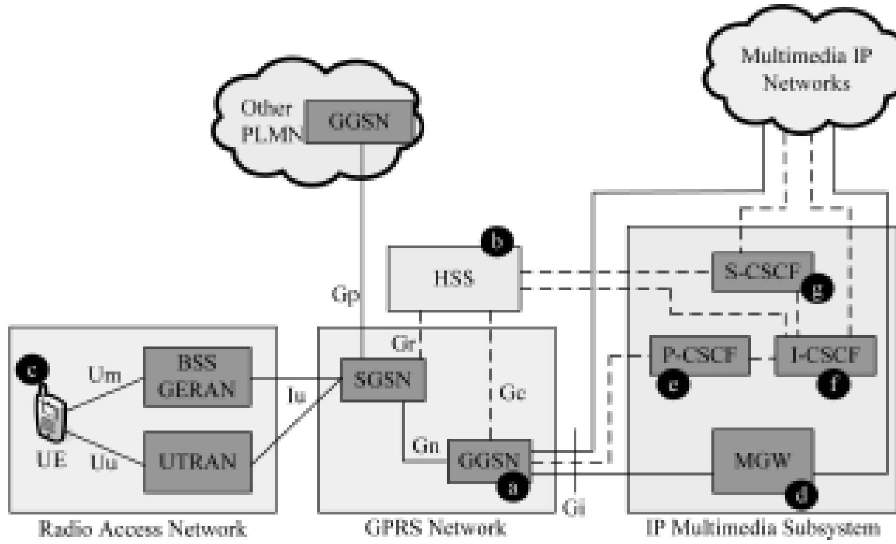

Figure 4.3: A simplified 3GPP Release 5 IMS network architecture [43]

The I-CSCF, based on the capabilities queried from the HSS, determines how to route incoming session requests to the S-CSCF and then to the destination UE. The S-CSCF is assigned to serve the UE during the IMS registration. The S-CSCF supports the signaling interactions with the UE for session setup and supplementary services control.

Figure 4.4 depicts the IETF policy-based QoS architecture [44]. Figure 4.5 depicts the policy-based QoS architecture [44] employed by IMS to ensure that sufficient QoS resources are provided to authorized users. Note that, for the demonstration purpose, we consider the UMTS as the underlying access network. This QoS architecture consists of two main elements: the Policy Decision Function (PDF) and the Policy Enforcement Point (PEP). The PDF is in charge of authorizing usage of network resources such as the bandwidth requested by the UE. The PDF could be integrated with the P-CSCF or be deployed as a stand-alone network entity.

In the stand-alone configuration, the P-CSCF communicates with the PDF through the Gq interface. The PEP is a logical entity that carries out actions stipulated by the policy decisions from the PDF. The PEP resides in the GGSN of the UMTS network and interacts with the PDF via the Go interface. The reader interested in the Gq and Go reference points is referred to [4] for the details. The relationships between the P-CSCF/PDF, the GGSN/PEP, and the UE are briefly described as follows.

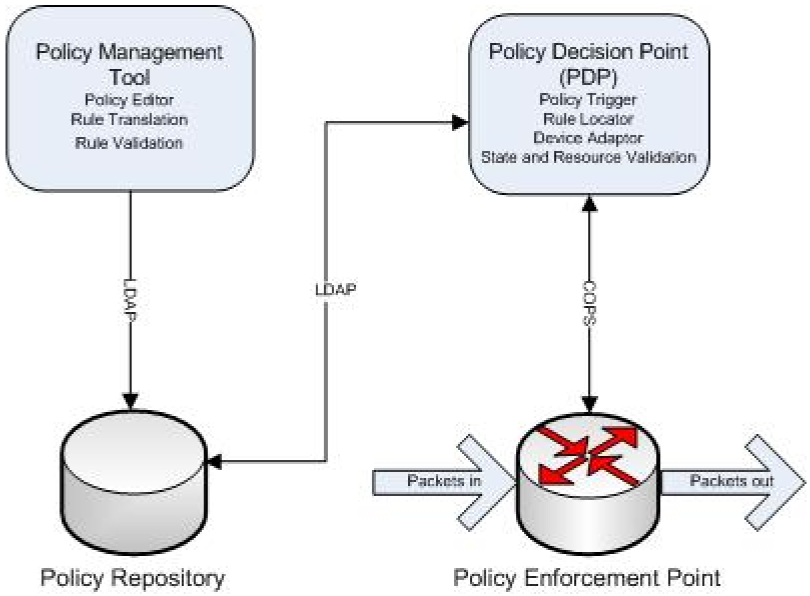

Figure 4.4: IETF Policy Framework [44]

4.4.1 The P-CSCF/PDF Functions

During the establishment of an SIP session, the UE specifies the QoS requirements in the Session Description Protocol (SDP) description within the SIP messages. The P-CSCF processes the SIP messages and sends the relevant SDP information to the PDF together with an indication of the originator. The PDF authorizes the IP flows of the chosen media components by mapping from SDP parameters to authorized IP QoS parameters for transfer to the GGSN via the Go interface. Besides the QoS requirements, the PDF also examines the source and destination IP addresses and port numbers for traffic policing. Then, the PDF generates an authorization token that uniquely identifies the SIP session across multiple PDP contexts terminated by a GGSN. This token is sent to the UE via SIP signaling so that the UE can use it to identify the associated session flows to the PEP in the GGSN in subsequent transmissions of the IP packets.

4.4.2 The UE Functions

The UE obtains the authorization token from the P-CSCF. This token is used in the PDP context activation or modification procedure as the binding information. In addition, the authorization token associates the PDP context bearer with the IP flows. When activating or modifying a PDP context for media, the UE has to perform the mapping from SDP parameters to the requested UMTS QoS parameters. The QoS parameters include traffic class (conversational, streaming, interactive, or background), guaranteed bit rate, maximum bit rate, etc.

4.4.3 The GGSN/PEP Functions

On receiving the PDP context activation or modification request, the GGSN/PEP asks the PDF for the authorization information. The PDF compares the received binding information with the stored authorization information and returns an authorization decision. The PDF communicates details of the media authorization in the decision to the GGSN/PEP. These details contain IP QoS parameters and packet classifiers related to the PDP context. The GGSN/ PEP then maps the authorized IP QoS parameters to the authorized UMTS QoS parameters. Finally, the GGSN/PEP compares the requested UMTS QoS parameters of the PDP context against the authorized UMTS QoS parameters. If the requested UMTS QoS parameters lie between the limits authorized by the PDF, then the PDP context activation or modification request is approved.

The packet classifiers from the PDF are referred to as gates. When an IP flow is authorized by the PDF to use the specified network resources, the PEP opens the gate for the flow and effectively commits the network resources to the flow by allowing it to pass through the policing mechanism. On the contrary, if an IP flow is not permitted by the PDF to use the requested resources, the PEP closes the gate and drops the IP packets of the flow.

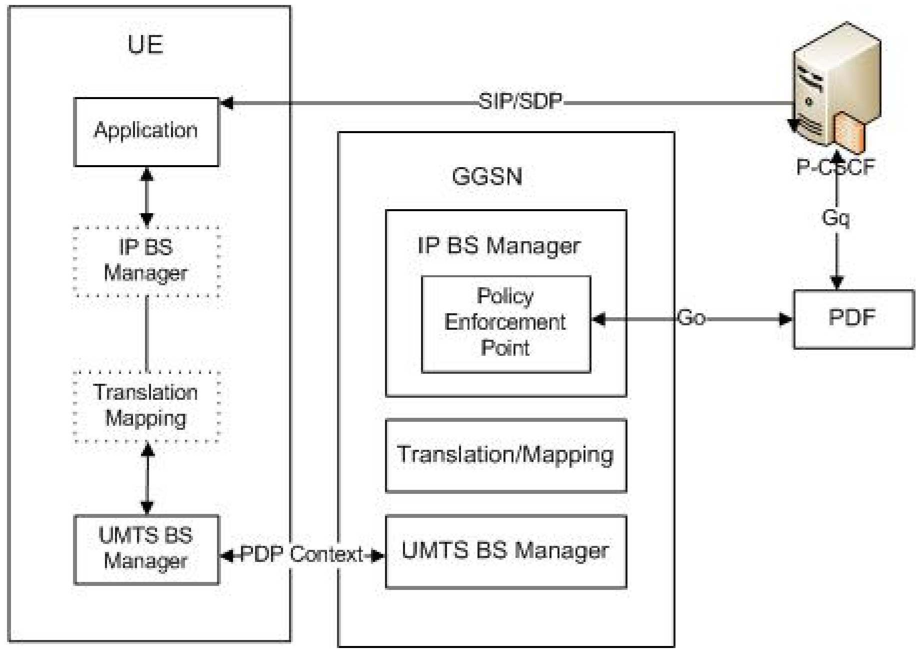

Figure 4.5: Policy Architecture in IMS [15]

Policy Decision Control

During the establishment of new SIP session, GGSN used the information in SDP to map the requirements into Layer2 architecture. P-CSCF processes the user request and communicates PDF through Go interface. GGSN serves as the Edge router in Diffserv network architecture. PDF implements the SBLP in IP bearer layer for media authorization and policy decision [6]. SDP with its QoS parameters activates the PDP context and in result GGSN and PEP negotiates with PDF. PDF makes policy decisions and indicate to PEP that user is allowed to access the request resources for the particular sessions.

Policy Enforcement Point

PEP is the logical entity that enforces policy decisions made by PDF for performing the admission control and authorization. When the PDP context activation request is received PEP asks PDF for the authorization information. Then PDF compares the received information with stored authorization information and if the requested parameters lie in the limit then the request of PDP Context activation is approved [15].

- IP Bearer Service (BS) Manager: IP BS Manager manages the IP BS using the standard IP mechanism. It resides in the GGSN and optionally in UE.

- Translation/Mapping Function: This function provides the internetworking between the mechanisms used by IP BS and UMTS BS. It also resides in GGSN.
- UMTS BS Manager: UMTS BS Manager handles the resource reservation requested by the UE.

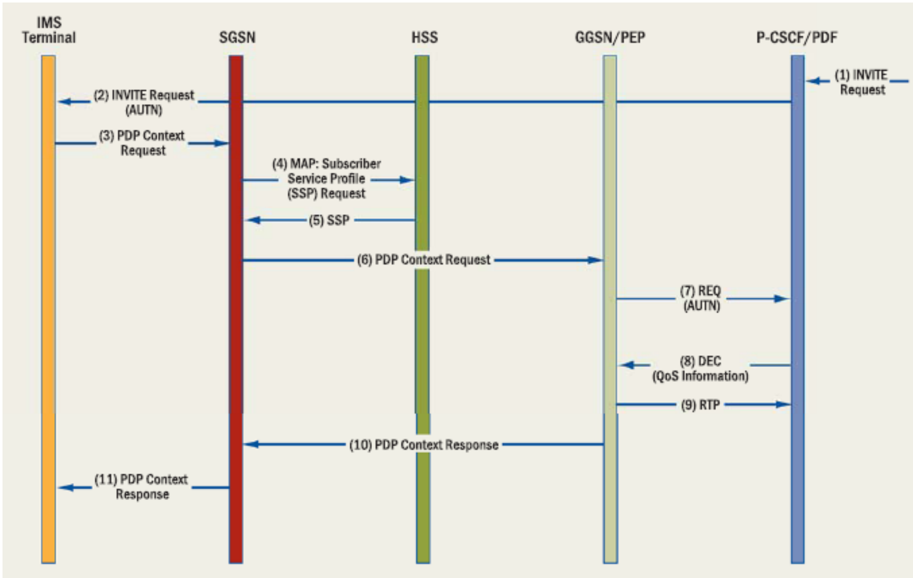

Figure 4.6: QoS Provisioning [12]

4.5 QoS Provisioning Steps in IMS

For inbound sessions:

Step 1: An INVITE request message arrives at the P-CSCF/PDF.

Step 2: The P-CSCF adds a media authorization token to the message and forwards this message to the IMS terminal.

Step 3: The IMS terminal creates a PDP context activation request message and sends it to the SGSN.

Steps 4–6: The SGSN receives this message and checks the user's subscription information stored in the HSS using mobile application part (MAP) protocol. If the IMS terminal requests more resources than it is allowed to use, the SGSN adjusts the requested resources to the appropriate level and sends a PDP context request message to the GGSN, along with the authorization token.

Step 7: The GGSN extracts the token and the packet flow identifier and sends this information to the PDF using the COPS REQ (request) message. The packet flow identifier contains the source address, the destination address, the source port number, the destination port number, and the transport protocol used.

Step 8: The PDF responds to the GGSN with a COPS DEC (decision) message that contains the QoS characteristics of the IMS terminal's authorized session. This is known as service-based local policy (SBLP) information. The GGSN uses this information to install packet filters that allow only authorized packet flows to be transmitted over a given PDP context.

Step 9: The GGSN (actually the PEP residing in the GGSN) sends an RTP message to the PDF indicating that it will comply with the PDF's policy.

Step 10: The GGSN sends a PDP context response message back to the SGSN, authorizing the SGSN for the requested PDP context.

Step 11: The SGSN forwards this response to the IMS terminal.

Figures 4.1 and 4.6 summarize the above process.

For outbound sessions of the QoS provisioning process, only the initial steps differ from those for inbound sessions. Replacing Steps 1 and 2 of inbound sessions, the corresponding steps involved in outbound sessions are:

Step 1a: The IMS terminal sends an INVITE request message to the P-CSCF/PDF.

Step 1b: The P-CSCF/PDF forwards this message to the callee.

Step 2a: The P-CSCF sends a session progress message to the callee.

Step 2b: The P-CSCF also adds the authorization token to this message and forwards it to the IMS terminal.

Steps 3 through 11 of outbound sessions are identical to those of inbound sessions of the QoS provisioning process. As was mentioned earlier, the GGSN/PEP must also play the role of a DiffServ edge router and map the required QoS to the appropriate DSCP to provide the desired QoS on the core network side, assuming DiffServ is used there.

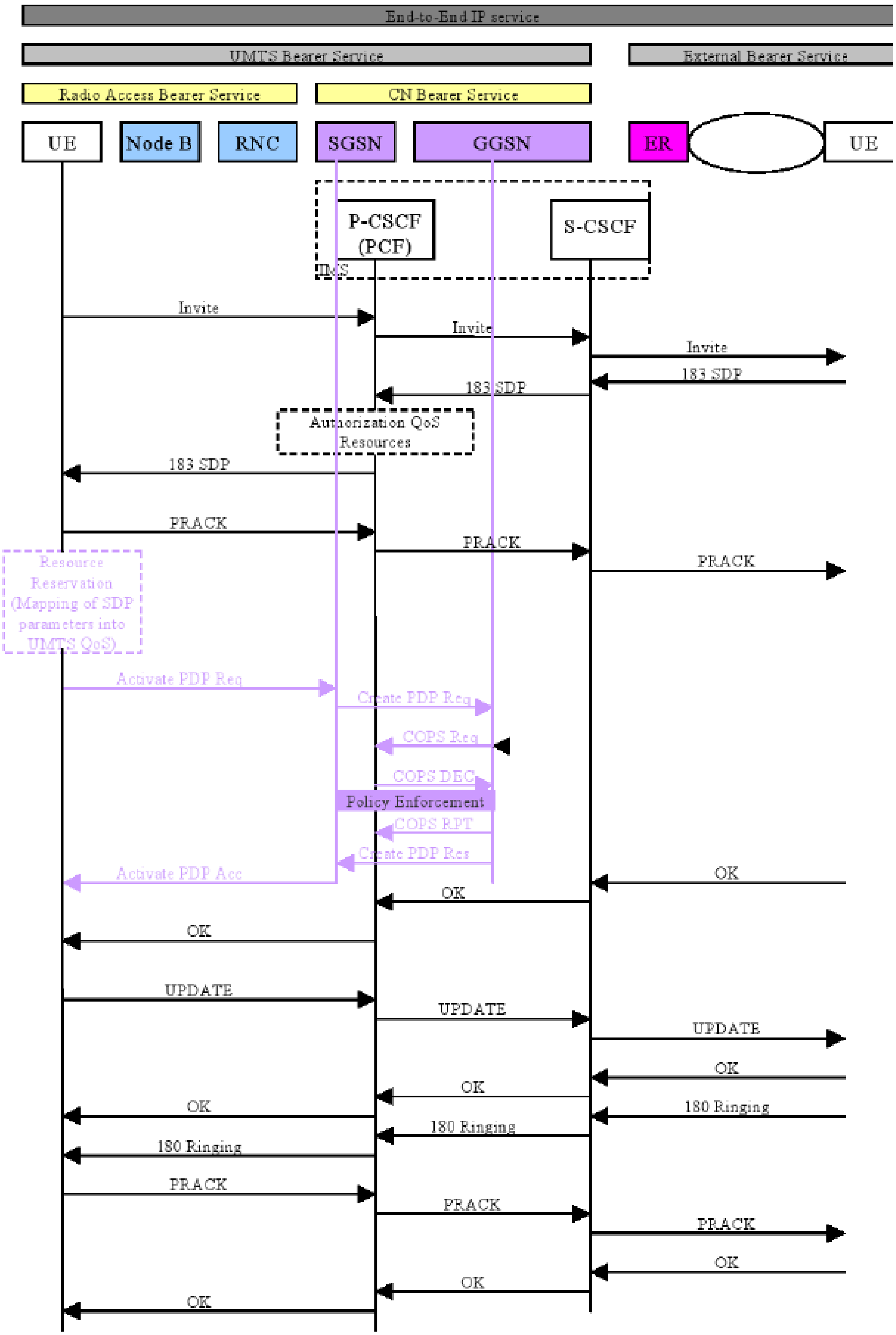

Figure 4.7: QoS Provisioning steps in IMS

Chapter 5

QoS Issues in IMS Access Networks

Introduction

The Internet is composed by different domains, each of them managed by different ISPs (Internet Service Providers)[50]. These Domains have different access networks and capabilities, users and policies that rule their behavior. Each ISP has an SLA (Service Level Agreement) established between its users, which defines what kind of resources and prices was agreed to be offered to each user. However there is no SLA between a visited domain and a mobile user. Thus there is a need to address inter-domain QoS negotiation to ensure that sufficient resources are provided to mobile users. With the new emerging area of IP Multimedia Subsystem (IMS) [4] and Next Generation Networks (NGN), QoS and mobility support are essential for the success of such networks. These NGNs support the provision of integrated information and communication services in face of increasingly more complex value chains, enabled by so-called service delivery platforms (SDPs). Therefore an NGN is an environment of high complexity, in which different actors, such as fixed and mobile network operators, service providers, system integrators and an open set of application providers have to cooperate for the provision of advanced converged services. A converged service is the integration of voice, multimedia, web content and Web Services provided seamlessly over all kinds of access technologies. This heterogeneity increases the complexity and challenges related to the management of such networks. The specific problem addressed by this research is how to ensure necessary levels of service and performance to critical multimedia and data applications, including service guarantees, while automating management.

5.1 Background and Scope

The internet is growing exponentially; every day millions of new users are connected and the internet becomes the backbone to provide various data services. The 4th wireless infrastructure generation propose a fully IP network with packet switching so that traffic processing could be facilitated.

Multimedia applications have very different requirements from applications for which the internet was originally designed. A primary issue is the need for performance assurance. The Datagram model, on which the Internet is based, has few resource management capabilities inside the network, and hence, cannot provide resource guarantees to users. Another issue is the need for differentiation. Because an IP network treats all packets the same way, it can offer only one service. Multimedia applications are real-time applications that require a certain amount of bandwidth to ensure the bit-rate needed by each media stream and strict delay variation needs to avoid buffer underflow at the receivers.

Despite the evolution of computing platforms, computational resources are still scarce because of the increased complexity of applications [52]. Modern networks and operating systems provide predictable behavior through the use of resource reservation mechanisms. However, most applications are still being adapted to make use of these mechanisms in order to support QoS specification and enforcement.

Furthermore, most of the QoS architectures proposed so far target only a particular configuration of processing and communication hardware. This tight dependency on a specific platform constrains their application in open environments, where heterogeneity is always present. Real-time operating systems combined with ATM are the most popular platforms for the development of QoS architectures because of their suitability for the implementation of QoS mechanisms for resource reservation. Some architecture is also targeted at particular application areas, with distributed multimedia being the one where the technology is most mature.

5.2 Key Elements for Providing Access QoS

In this section, we review the key elements that need to be present in an access network to provide QoS. We use the term service as the offering an operator makes to a subscriber. We further distinguish between session-based services and non-session-based services. Session-based services utilize an end-to-end session control protocol such as SIP/SDP or RTSP/SDP. All IMS services are session-based while Internet-Access is an example of a non session- based service. The traffic running between a particular client application and a service can be differentiated into separate service data flows. For example, an IMS-VoIP session can be differentiated into two service data flows, one for the session control signaling, and one for the media. We use the term Traffic Forwarding Policy (TFP) to denote a set of pre-configured traffic handling attributes relevant within a particular user plane network element. For example, a RAN-TFP may include several attributes such as the link layer protocol mode (acknowledged or unacknowledged), the power settings, and a default uplink maximum bit rate; while a GWTFP may only include a default downlink maximum bit rate. Each edge/bottleneck node – potentially including transport network nodes – supports a number of TFPs. Uplink (UL) and Downlink (DL) Guaranteed Bit Rates (GBRs) are not part of a TFP since these traffic handling attributes can not be preconfigured for a QoS class. They must therefore instead be dynamically signaled. TFPs confine traffic handling attributes to those nodes where those attributes are actually needed. TFPs are provided and configurable by the operator from the management plane.

We use the term bearer to refer to an edge-to-edge association between the UE and the GW. Independent of whether it is realized in a connection-oriented or a connectionless way, a bearer is defined through:

- The network to which it connects the UE (referred to as Access Point Name in 3GPP)
- The QoS Class Identifier (QCI) via which it can be associated with a TFP defined within each user plane edge/bottleneck node
- (Optionally) the UL- and DL-GBR. Within an access network the UL-GBR and DL-GBR are only relevant for session-based services, and only if the operator's policy defined for a specific QoS class requires that session admission control (e.g., in the RAN) be triggered when establishing service data flows associated with that QoS class.

Note that the term QCI is not associated with any semantics, e.g., related to traffic characteristics or application layer requirements on end-to-end QoS. That is, a QCI is simply a “pointer” to a TFP. Note further that within a specific node multiple QCIs may be associated with the same TFP. In order to receive a QoS level other than the default QoS level (via the default bearer explained below) a service data flow needs to be bound to what we refer to as a QoS bearer. A QoS bearer is associated with an uplink binding state in the UE for the uplink traffic, and a downlink binding state in the GW for the downlink traffic. The binding state creates the mapping of a service data flow to a QoS bearer. When multiple bearers are established between a UE and a GW then it is the uplink binding states of the QoS bearers that “steer” the aggregate uplink traffic into the right bearers, and likewise the downlink binding states in the GW for the aggregate downlink traffic.

For a specific network that the UE connects to there is at most one bearer without uplink and downlink binding states. This bearer is referred to as the default bearer. It is important that each QoS bearer connecting to a specific network is associated with well defined (“non-overlapping”) uplink and downlink binding states to ensure an unambiguous mapping of packets to the QoS bearers.

5.2.1 Requirements for an Evolved 3GPP QoS Concept

Below we list requirements that we believe should be met by an evolved 3GPP QoS concept followed by a brief discussion of each requirement [45].

a) Operator Controlled Service and User Differentiation
b) Minimize Terminal Involvement in QoS and Policy Control
c) QoS Support for Access Agnostic Client Applications (UE-Based + Non-UE-Based)
d) Fast Session Setup
e) Backwards Compatibility
f) Convergence towards Other Access Types such as IWLAN and fixed broadband
g) Rapid Time To Market (TTM) for the Deployment of New Services

Requirement a)

Service and user differentiation requires a limited set of well defined QoS classes. The number of QoS classes supported within an operator’s network reflects the granularity of differentiation the operator provides. Operators should be free to define the mapping of the service data flow(s) of offered services to the QCI(s). For certain well-known services this mapping could be standardized, or defined as part of roaming agreements. Likewise, operators should be free to define which TFP gets associated with a QCI.

Requirement b)

Operators may regard a UE as a non-trusted device which can be “hacked”, e.g., for the purpose of receiving higher QoS than subscribed and charged for. Therefore, the control over a bearer’s QCI should be located within the network. In principle, there is no reason for a UE to have knowledge of a bearer’s QCI. Another aspect of this requirement is the placement of the exception handling control associated with bearer establishment. To ensure a consistent exception handling across terminals from different vendors, this control should be located within the network.

Requirement c)

Access agnostic client applications do not use any vendor and/ or access-specific QoS-API (Application Programming Interface). A QoS-API can be used to request the establishment of a QoS bearer, and thereby create the UL binding between a service data flow of the requesting client application and the QoS bearer. This requirement basically says that any client application programmed towards the ubiquitous socket-API that is supported by virtually every widely deployed operating system should be able to receive QoS. Note that the socket-API does not support requests for QoS bearers.

Requirement d)

It is widely recognized that low session setup delays are an important factor in user perceived service quality.

Requirement e)

It can be expected that UEs based on the 3GPP Rel. 5 QoS concept will be widely deployed in the coming years. Also, the upgrade of network equipment can not be assumed to be carried out “over night”. Hence, backwards compatibility with Rel. 5 based equipment needs to be ensured by an evolved 3GPP QoS concept.

Requirement f)

An evolved 3GPP QoS concept should be aligned with the QoS concepts of other access types such as I-WLAN and fixed broadband networks. This will facilitate and simplify the provisioning of end-to-end QoS in multi-access networks. An evolved 3GPP QoS concept should allow for deployment of operator-defined services without the need for prior standardization of QoS support in UEs or network elements.

5.3 3GPP IMS

In the late 1990s, 3GPP developed the third-generation mobile system based on the second generation (2G) GSM cellular system. The 2G circuit-switched GSM system supports voice services and text-based short message service (SMS). As the Internet grows, the GPRS, also known as the 2.5G wireless system, has evolved from the GSM to provide packet-switched services. Although limited capabilities have constrained the expansion of GPRS, GPRS laid a foundation for the 3G PS core network. Based on GPRS, the 3GPP Release 4 specifications have specified basic *IP connectivity* for mobile users. However, the *basic IP connectivity* alone is insufficient to support IP multimedia services. Also, the services could be broken while a user roams across network boundaries. In Release 5, IMS is defined to enhance the basic IP connectivity with enhanced signaling and session management.

IMS is an overlay framework over the 3GPP PS domain. It is independent of the 3GPP CS domain. IMS is designed to support real-time voice over IP (VoIP) and other IP-based multimedia services. In this section, we provide a high-level overview of the requirements and architecture of the 3GPP IMS.

5.3.1 Requirements

The IMS service requirements are specified in 3GPP TS 22.228 Release 5. They include:

- Negotiable QoS should be supported. The QoS for IP multimedia sessions or media components should be negotiable, not only at the time of establishment, but also during the session.
- End-to-end QoS for VoIP service should be guaranteed at least equal to or better than the circuit-switched voice service.
- Roaming should be supported. Moreover, QoS and service capabilities are required to be automatically negotiated.
- One or more IP multimedia applications should be supported within each IP multimedia session.
- Inter-working with a circuit-switched network should be supported.
- Services should be rapidly supported without standardization.

In Release 6, 3GPP has added a few extra requirements. The new requirements focus on the support of different radio access networks. The major complementary requirements and modifications are:

- An *access independence principle* should be supported. That is, the services should be offered to users regardless of their underlying access technology, such as GPRS, xDSL, WLAN, and so on.
- Inter-working with other networks including the Internet should be supported. These new requirements reflect a growing trend: IMS will not be limited to cellular networks but is expected to be used to support IP multimedia services in the integrated all-IP networks.

5.4 QoS Issues in Access Networks

Real-time multimedia systems are becoming increasingly popular in a variety of applications. The wide range of applications includes group collaboration, remote medical diagnosis/treatment, conferencing systems, on-demand video services, and distance/remote sensing and monitoring. Applications such as Internet Cellular Phone and access to real-time data through the web require that the network offer quality of service to moving users. There has been especial demand for IP-based networks in which all traffic is delivered in packet-based form. Mobility of hosts has a significant impact on QoS parameters for real-time application.

Note that the absolute amount of signaling effort in order to provide per-session QoS finally depends on the subscribers' session behavior, because the standardized NGN QoS architecture works on a per-session basis [51]. That is, a subscriber establishing and canceling many sessions in a short period of time will require a substantial amount of resources and QoS management traffic.

Based on earlier research work, the following main requirements for the provision of QoS in SIP-based NGN have been identified (Weber *et al.*, 2007).

- Functions and mechanisms, leading to a trustworthy QoS for each established session and, at the same time, do not occupy resources on a per session basis themselves.
- Simple and resource saving QoS control should be preferred. If possible, approaches should rely on already standardized protocols (such as SIP) and architectures (such as NGN according to (ITU-T Y.2012, 2006) and (ETSI ES 282 001, 2005)).
- NGN QoS control has to be aware of a certain amount of traffic that is not session-based (such as TCP web traffic).
- The QoS provision in NGN should be independent of underlying transport technologies such as MPLS, ATM, and VLAN. Arbitrary IP network architectures should be supported; regardless their specific integrated QoS mechanisms.

5.4.1 QoS issues in Wireless NGN

The issue of QoS mapping across different types of networks is being discussed. To ensure satisfactory user experiences wireless network service providers strive to engineer their networking systems to achieve high QoS levels. In general network QoS can be categorized into two types of QoS i.e differentiated QoS and guaranteed QoS. With differentiated QoS, network traffic flows can be categorized into several classes. Network service providers give different priorities to serve packets in different classes. Data packets are classified and delivered with different service classes. Packets with delivery constraints could be sent with high priority class, while delay tolerant packets could be sent with a low priority class. Through a mechanism supporting guaranteed QoS, network resources are reserved for the traffic flows to provide QoS bounds such as bounds on throughput and delay. For example network resources could be reserved for real time traffic (video streaming and voice over IP flows) to ensure service quality. In wireless NGN, network has to support both differentiated and guaranteed QoS because ideally, in a true FMC environment the wireless segment of a network will be indistinguishable from the wired segment. There are three main QoS requirements of FMC for achieving end to end QoS assurance.

- TE and resource reservation: Network controllers need to allocate and reserve resources according to the demand of aggregated traffic.
- Conversion of QoS Requirements: The QoS mechanisms differ in various types of networks. A mapping function is needed to convert the QoS requirements between different networks.
- Fast handover: we need to reduce the handover latency in FMC in terms of forwarding time at relay nodes between networks and in terms of handover time when an end user roams from one network to another one.

5.4.2 Scalability Issues

IMS is designed to support extensive and complex IP multimedia services for a lager number of users. It provides a traffic-independent scalable architecture. That is, CSCFs and application servers can be allocated dynamically to users. The servers are distributed such that the capacity is extensible. However SIP is the major protocol of IMS. SIP is text-based, which is easy to debug, but has a large message size. Moreover, there are numerous nodes with extensive functionality within the IMS. They produce a large number of messages and must process complicated message flows. In this section, we investigate potential scalability issues in IMS.

In addition to the scalability issues incurred by users, the scalability issues of the IMS architecture itself also are indicated in 3GPP TS 23.218. When a myriad of S-CSCFs and ASs are deployed, signaling propagation delay may arise if many of them are involved in the session signaling path. Care must be taken to minimize the transaction delay when intending to share the loads with multiple servers.

In short, a *load-aware routing algorithm* is essential to dynamically route the SIP messages to a proper server in the IMS network. There are challenges as to how to measure and exchange the current information of load among the servers without introducing additional signaling overhead. The decision also should consider the propagation delay which may be incurred by geographical distance.

5.4.3 Business issues

IMS leads networks operators to play a central role in service distribution. This involves that carriers will have to obtain content. The role of the operators in the billing of services provided by third parties has also to be clarified. With IMS a single customer may subscribe to services from several providers. IMS therefore leads network operators into a competition with players of the Internet world. The decision to deploy IMS is strategic. Network operators may choose an early deployment scheme in order to take advantage of the higher prices charged to early adopters. In this pioneering strategy they would open the way for their competitors and take significant risks. Alternatively, a network operator may wait in order to reduce his investment costs and learn from his competitors failures. But he would face several competitors having better market experience.

5.4.4 Technical Issues

Network transmission is liable to errors and data loss. Excessive loss in video transmission affects the performance of video quality [53]. Video quality can be enhanced by proper frame synchronization between the video and audio streams. The transmitting a movie over low bandwidth channels can be enhanced by separation of audio and video streams and routed them independently to offload certain channels. Adaptive and dynamic resource management and reliability issues need to be addressed for continuous media streaming. Emerging continuous media applications have well defined QoS constraints. These applications have stringent resource requirements and can benefit from non-interference to provide forms of progress guarantees. Video stream places high demands for QoS, performance, and reliability on storage servers and communication networks. To support video conferencing, network and application level multicast is a powerful technique. End-to-end congestion control is an important issue for reliable and unreliable multicast transport protocols. Like wired network, resource reservation and rate adaptation can be used to support multimedia services to provide quality of service in wireless networks.

Multimedia networks support real-time data, bulk data, and statistically multiplexed data, which make the traffic management in the network hard. The necessary traffic management components to support QoS are:

- Admission control: The admission control component takes into account resource reservation requests and the available capacity to determine whether to accept a new request with its QoS requirements.
- Scheduling: The scheduling component provides QoS by allocating resources depending on the service requirements. This requires mapping the user-defined QoS requirement to resource allocations for providing the service.
- Resource management: QoS can be provided using over-provisioning of a network, which increases the cost incurred by the provider. Efficient resource management is a cost-effective solution for the provider and it ensures that applications will get the specified QoS during the course of its execution.
- Congestion control: Congestion control is required to avoid anything bad from happening inside a network domain. Some applications may not follow the standard protocol description and try to steal resources, thereby deteriorating the QoS of other applications. Mechanisms are needed to recover from congestion and control flows accordingly.
- Policing/Shaping: Users might send traffic at a rate higher than the agreement. Policing is necessary to monitor these situations, and shaping makes the traffic smooth and reduces its variations over time.

Chapter 6

Proposed SIP-Based QoS Management Architecture

Introduction

IMS provide the convergence between packet-switched and circuit-switched networks. IMS is an umbrella framework for providing enhanced IP-based services developed by the Third Generation Partnership Project (3GPP), collaboration among a number of telecommunications standards bodies. The original scope of 3GPP was to develop specifications for a 3rd Generation Mobile System based on evolved Global System for Mobile communication (GSM) core networks and the radio access technologies that they support. The scope was later amended to include the maintenance and development of the GSM specification and its related radio access technologies. To date, three "phases," known as Releases 5 through 7, have been published for IMS, and provide a framework for an IP/SIP-based network services architecture that is designed to span wireless, wire-line, and cable networks.

6.1 Background

The increasing demand for multimedia applications such as VoIP, video conferencing and video streaming has created new challenges for network service providers. The requirements for delay, throughput and packet loss for such applications are inherently different from those of simple network traffic. The IP service, which is characterized as best effort and performs satisfactorily with data traffic, fails to provide the desired Quality of Service (QoS) guarantees for multimedia applications. It becomes increasingly important to manage network effectively and utilize network resources efficiently to fulfill the requirements of various IMS services. But at present, it is managed rather inefficiently. There is a need for providing such guarantees by differential treatment of the traffic for these multimedia applications over IP. The 3GPP QoS architecture for IMS does not provide the correspondence between different access networks to exchange their policies and introduce network restrictions which cause the network degradation. So it is relatively inflexible to manage QoS in IP access networks.

To improve the communication between core and IP access networks, an extended SIP based QoS Architecture solution is proposed. In this architecture QoS SIP Proxy Modules are introduced to monitor the network and to ensure the end-to-end QoS. The QoS methodology used in this architecture is MPLS and Diffserv. The motivations for DiffServ + MPLS include user demands for consistent QoS guarantees, efficient network resource requirements by network providers, and reliability and adaptation of node and link failures for the multimedia services. DiffServ provides scalable edge-to-edge QoS, while MPLS performs traffic engineering to evenly distribute traffic load on available links and fast rerouting to route around node and link failures. Traffic Engineering aims to facilitate efficient and reliable network operations while simultaneously optimizing network resource utilization and traffic performance. As it turns out, this is indispensable to provide Quality of Service, as it provides a means for network optimization and bandwidth provisioning. Moreover, MPLS can be deployed over a wide variety of link layer technologies such as IP, ATM, and Frame Relay. MPLS support for DiffServ over IP is used in this SIP QoS Modules for IMS. These SIP proxy Modules are assigned to Access networks among core networks.

6.2 Proposed SIP-Based QoS Management Architecture

IMS provide the convergence of multimedia applications over IP-Based networks and implements the 3GPP QoS architecture for QoS provisioning over IP networks. But 3GPP QoS architecture does not provide any mapping mechanism to exchange SLAs between heterogeneous domains and thus introducing hands off disruption time and network degradation.

SIP-Based QoS Management Architecture improves the reliable delivery of IMS services through SIP-Based Proxy QoS Modules and also improves the communication between core and IMS access networks. To manage QoS in underlying access networks these modules implement the Diffserv aware MPLS techniques for the efficient delivery of real-time multimedia services. Diffserv and MPLS enhance the QoS provisioning capabilities on IP Network.

SIP Proxy Modules monitor the network to ensure the end-to-end QoS by using MPLS and Diffserv routing schemes. DiffServ & MPLS fulfill the user demands of consistent QoS provisioning for the multimedia services. MPLS Traffic Engineering alleviates the efficient and reliable network operations while at the same time optimizing network resource utilization and traffic performance.

The SIP-Based QoS management architecture is shown in figure 6.1. The IMS User must have IP connectivity which can be obtained either from home network or visited network to access the IMS services. UE requests required resources from IMS core network using SIP and SDP protocol during session establishment and in access networks SIP proxy Modules provides the desired resources for user applications.

SPQMs communicate with the IMS access networks and transport networks for end-to-end QoS provisioning. These SIP modules treat the SIP traffic received from IMS core network or SIP Module and forward data to next available module or IMS network. MPLS and Diffserv routing mechanisms are used in SPQM. To support DiffServ over MPLS, packets need to get the proper QoS at each LSR in the network. To manage QoS in MPLS aware Diffserv networks, our architecture uses the modification of LDP.

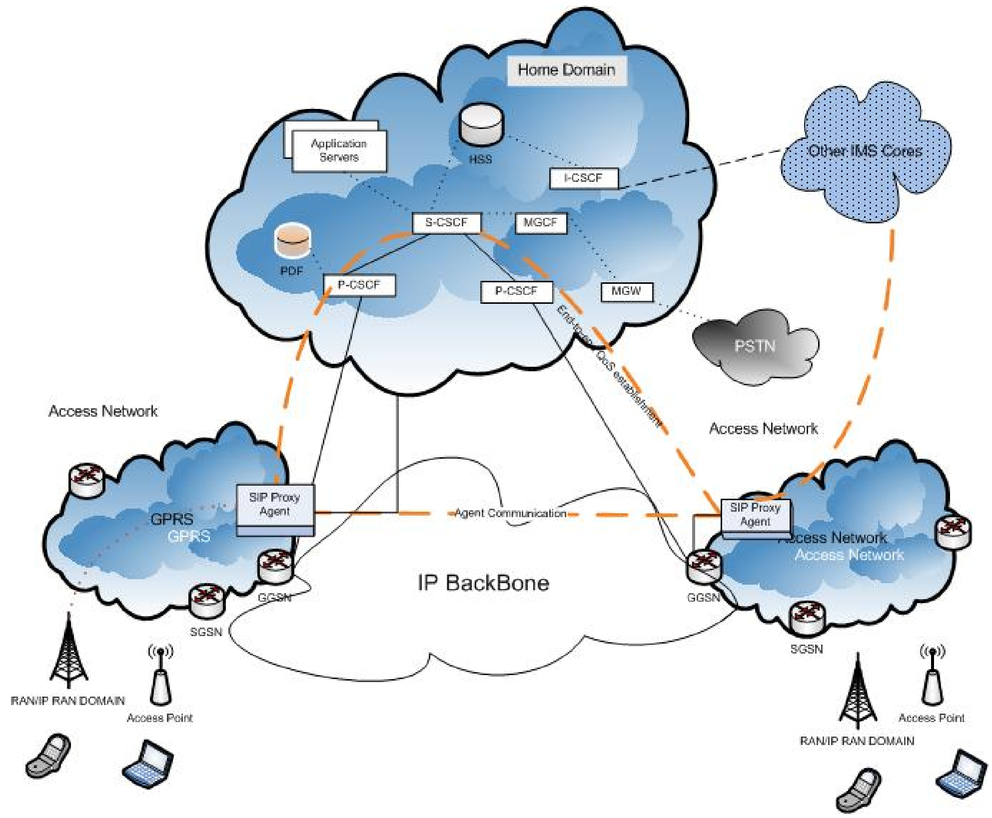

Figure 6.1: Proposed Architecture for QoS Management

6.2.1 SCTP Features

SCTP is reliable, general purpose transport layer protocol used in IP networks. It solves some limitations of TCP while borrowing the some beneficial features of UDP. The main purposes of using SCTP in SPQM are:

- Multi-homing: A Multi-homed host has more than one network interfaces and thus provide the applications with high availability.
- Multi-streaming: SCTP provides multiple streams within an association and each stream in association is independent but related to association.
- Initiation Protection: This feature protects the initiation request from DOS attack by introducing the four way handshake mechanism.

- Graceful shutdown: SCTP provides cleaner termination sequence as compared to TCP. When a peer close its socket both endpoints are required to close and no further data movement is allowed to transfer data.

6.2.2 SIP-Based Proxy QoS Modules (SPQMs)

QoS provisioning tasks for IMS services over access networks are performed by SPQMs. SPQM receives SIP traffic over IP access networks from the IMS network and requests the network resources by tracking the available resource information for particular session. QoS provisioning in SPQMs is multi-party service negotiation mechanism in which SPQM on both neighbor Access Networks collectively decide end-to-end efficient service delivery. Diffserv and MPLS QoS mechanisms are used to ensure guaranteed QoS for multimedia applications as MPLS is independent technology from underlying access technologies. MPLS performs the traffic engineering by distributing SIP traffic load on all available links instead of using the same route and decrease the handoff disruption time and make effective resource utilization. For differential treatment of SIP Traffic of IP Multimedia Subsystems over Access Networks SPQM implement QoS routing protocols (extensions of OSPF) to support QoS Routing. Link State Advertisement (LSA) gives the available bandwidth on a link.

6.3 SPQM Architecture

SPQMs Architecture is composed of two Functions:

- SIP-based QoS Monitoring Function
- SIP-based QoS Control Function

6.3.1 SIP-based QoS Monitoring Function

Update link state information is used in monitoring function to examine the IMS SIP traffic over access networks. It also maintains identifiers for routing purposes in MPLS aware diffserv network. To monitor congestion in network these modules maintain information about available bandwidth, transmission delay and also use identifiers for routing mechanism. QMF monitors the network for troubleshooting and also issues commands if any degradation occurs in network.

6.3.2 SIP-based QoS Control Function

QCF controls the QoS over network via monitoring function and manages the QoS for applications. It is main agent that tells how to manage the network for the multimedia applications. QCF retrieves information from QMF and deduced the QoS parameters to manage the degradation in network.

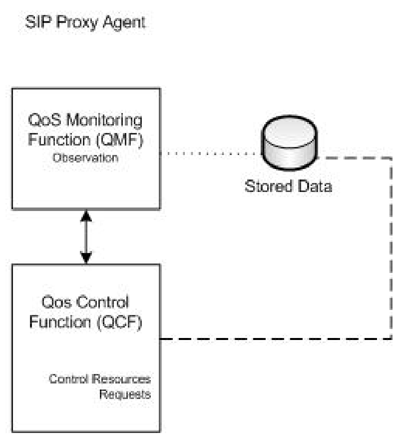

Figure 6.2: Conceptual Block Diagram of QoS in SIP Proxy Agent or Module

The routing mechanism used to control the QoS provisioning over access networks is shown in table. The two methodologies used for QoS provisioning in SIP modules are:

- Differentiated Services
- Multi-Protocol Label Switching

a) <u>Differentiated Services</u>

Heterogeneous multimedia applications require service differentiation to fulfill the user expectations. DiffServ provide preferential treatment to data packets in order to satisfy the performance requirements of users defined in SLAs. The edge routers classify the data packets into traffic classes and these classes have different service and priority levels. The DiffServ framework defined by IETF provides different levels of delivery services for differentiated traffic flows.

Numbers of functional elements are composed to provide per-hop forwarding behaviors, packet classification and traffic conditioning function. Edge routers receive the incoming packet in DiffServ domain and classify packets in order to find there associated SLA. The packet is reshaped or dropped in over-sending situation [22].if packet is not dropped, it is marked with DSCP to determine the Per-Hop-Behavior and then routers store and forward the packet using the appropriate scheduling and queuing mechanisms to core routers .Core routers identify PHB by DSCP. DiffServ uses three forwarding techniques i.e Best Effort, Assured and Expedited Forwarding.

b) <u>Multi-Protocol Label Switching</u>

MPLS is a label swapping technology in which fixed-length label is attached to every packet entering the network. Labels are locally significant identifiers to identify the streams of data. In conventional networks, all routers make an independent decision to route the packet and ordinarily follow the shortest path but in MPLS, first router decides the entire path of packet. MPLS provides the feasibility of implementing the traffic engineering and move the heavy processing to the edge routers. Edge Routers classify and label the packets and interior routers just perform label lookup and swapping. Label stack is used to place several labels.

In MPLS network, packets enters through the edge routers that contain FEC. Forward Equivalence Class specifies the packet forwarding sequence. After classification of FEC, label is assigned to the packet and forwarded to the next LSR LSR looks up the label and swap it with the label associated to that FEC. Label look up and swapping procedure continue until the last router remove the label (shim header) and forward packet to host or next router in adjacent domain by looking up the IP header of packet. The Path through which a labeled packets traverse is called Label switching path. MPLS even allows the path establishment in the architecture which operate in disconnected mode.

c) MPLS and Traffic Engineering

In conventional IP networks, congestion control is complicated as common gateways always choose shortest path for data transfer. Traffic engineering solves the congestion problem by selecting the less congested paths instead of the shortest paths. Traffic engineering in MPLS networks provides a predefined path for packet forwarding, being established between routers. Traffic Engineering can be stated as the mapping of traffic flows onto existing physical topologies. To avoid congestion and utilization of network resources, traffic engineering balance the traffic load on various links. The performance goals of traffic engineering are categorized as resource or traffic oriented.

6.4 SIP QoS Routing Mechanism

SIP QoS routing, incorporating policy based routing and traffic engineering functionalities into it. SIP QoS routing enables a demand driven, resource reservation aware, routing paradigm to co-exist with current topology driven hop by hop Internet interior gateway protocols.

To support DiffServ over MPLS packets with a variety of Differentiated Services Code Point (DSCP) values need to get the proper QoS at each LSR in the network. Because MPLS label standards were defined before DSCP standards, MPLS labels used 3 bits for Experimental bits that came from the old 3-bits Precedence in the IP header. The problem is that DiffServ can have 64 possible DSCPs whereas MPLS/Exp bits can only address up to 8 possible types of Per-Hob Behaviors (PHBs).

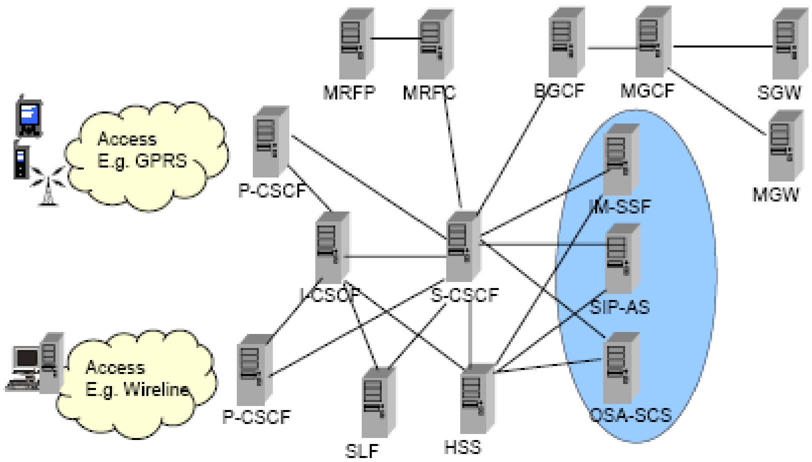

Figure 6.3: IMS over wire-line and wireless network [7]

6.4.1 Routing Algorithm for SPQM

% New SIP Session for IMS Core

while (QoS request == available) do

if PDF (Yes)

Process User Request for Multimedia Services

end while

% IMS traffic over Access Networks

a) IMS Packets arrives at the IP Access Network

While (packets== SIP)

Enable SIP-QoS Module

- Remove Packets header
- Insert SIP-QoS header

Else Process other packets

b) Compute the shortest path between two nodes

c) Process SIP Packets to next SIP module

d) Update the link state information.

end while

Else Process Other available data

SIP-Based QoS architecture manages the SIP multimedia traffic over IP networks to provide the best QoS. The implementation of SIP QoS routing modules are described below.

6.5 Implementation of Routing Mechanism for SPQM

6.5.1 Packet Forwarding and Routing

A node in simulation is represented by an instance of Class Node. Node itself is an aggregate object consisting of a set of classifiers. It also contains:

• An address or id, monotonically increasing by 1 as new nodes are created

• A list of neighbors

• A list of agents

• A list of routing modules

6.5.1.1 Class Classifier

The function of a node when it receives a packet is to examine the packet's fields, usually its destination address, and on occasion, its source address. A classifier provides a way to match a packet against some logical criteria and retrieve a reference to another simulation object based on the match results. Each classifier contains a table of simulation objects indexed by slot number. The job of a classifier is to determine the slot number associated with a received packet and forward that packet to the object referenced by that particular slot.

6.5.2 Routing Module

An ns node is essentially a collection of classifiers. The simplest node contains only one address classifier and one port classifier. When one extends the functionality of the node, more classifiers are added into the base node, and each of these blocks requires its own classifiers. A provide a uniform interface to organize these classifiers and to bridge these classifiers to the route computation blocks is provided through the concept of routing modules.

In general, every routing implementation in ns consists of three function blocks:

- Routing agent exchanges routing packet with neighbors.
- Route Logic uses the information gathered by routing agents to perform the actual route computation.
- Classifiers sit inside a Node. They use the computed routing table to perform packet forwarding.

A routing module manages all these function blocks and interfaces with node to organize its classifiers. In order to know which module to register during creation, the Node class keeps a list of modules as a class variable. The default value of this list contains only the base routing module. The Node class provides the following two procs to manipulate this module list:

- Node: enable-module [SIP-QoS] If module RtModule/ [SIP-QoS] exists, this proc puts [SIP-QoS] into the module list.
- Node::disable-module{[SIP-QoS]} If [SIP-QoS] is in the module list, remove it from the list.

When a node is created, it goes through the module list of the Node class, creates all modules included in the list, and register these modules at the node.

6.5.3 QoS Routing

This SIP-Based QoS architecture implements routing module for ns using MPLS and diffserv. It consists of following components:

- A routing logic that computes shortest path between two nodes with given available bandwidth.

- An agent — Agent/QOS that exchanges link state routing packets between various SIP QoS agents.
- Routing module RtModule/QOS which interfaces the above with Node

configure-qos : After all links have been created and queues initialized, configure-qos should be called to create SIP QoS routing agents on SIP nodes and initialize them. In the case of other routing modules eg. Mpls module have been enabled, they may install their own classifiers. Therefore it is important that configure-qos is called after other routing modules have been initialized.

6.5.4 Linkstate Updates

An important consideration in QoS routing implementation is when should routing updates be send. One option is to mandate periodic updates, where the period of updates is determined based on a tolerable corresponding load on the network and the routers. The main disadvantage of such an approach is that major changes in the bandwidth available on a link could remain unknown for a full period and, therefore, result in many incorrect routing decisions. Ideally, routers should have the most current view of the bandwidth available on all links in the network, so that they can make the most accurate decision of which path to select. Unfortunately, this then calls for very frequent updates, e.g., each time the available bandwidth of link changes, which is neither scalable nor practical. In general, there is a trade-off between the protocol overhead of frequent updates and the accuracy of the network state information that the path selection algorithm depends on.

Therefore, a good strategy could be to send link state updates when available bandwidth on a link changes by a certain minimum percent. It could be coupled along with regular periodic updates. Also a periodic timer constraint in the form of a hold down timer can be applied so that link state updates are not sent too frequently.

Chapter 7

Implementation & Experimental Evaluation

The aim of this simulation is to underline the QoS Architecture for IP Multimedia Subsystems. The environment consists of ns-2 network simulation software in Linux operating system. The ns-2 patches, the Diffserv, MPLS and SIP patch are applied to execute the simulations. Session Initiation Protocol is the Basic building block for the IP Multimedia Subsystems.

7.1 Simulation Setup and Details

We exercise the different topologies to simulate the SIP QoS Module using MPLS and Diffserv over Access Networks. The SIP LSRs, represented in scenario, shows the SIP Modules to handle the SIP traffic over Access networks. This is implemented using the SIP traffic agent with customizable packet sizes and inter-packet intervals. The link capacities are one Megabit per second (Mbps). The idea is that SIP and both the SIP and UDP traffic are allowed to mix on a single link between nodes. This fast rerouting scheme helps DiffServ networks by allowing continuous data flow even in the presence of link failures in the networks.

7.1.1 Simulation Experiments

In this section, we describe the simulation experiments that were used with MPLS and Diffserv QoS mechanisms and to compare the three simulation scenarios with simple network without SIP QoS Modules. Moreover, we present results that depict the improvement in network efficiency due to improved bandwidth sharing in SIP QoS Modules compared to general QoS Models in IMS. We conduct a set of experiments and compare the QoS results and the total bandwidth utilization under various scenarios. These statistics depict a realistic representation of an IP Multimedia Subsystem's goal of providing best QoS, keeping in view the applicable rules for multimedia traffic and bandwidth utilization. Moreover, it is important to consider the bandwidth associated with the accepted multimedia requests since it is possible for a less efficient scheme to accept a large number of small bandwidth requests compared to a more efficient scheme which accepts fewer requests which comprise a greater amount of bandwidth.

Therefore, in order to provide a better comparison, we present statistics for the amount of bandwidth placed on the network in addition to the number of SIP packets received on destination. In our simulations, we use the QoS scheme proposed in [56]; the algorithms proposed therein for the computation of bandwidth and utilization of paths are integrated with this QoS architecture. The simulation experiments are conducted on a SIP network topology for IMS. The computation of primary and QoS routes for a multimedia request depends on the simulation scenario. In case of SIP, for each request, if it is possible to route the requisite according to the QoS requirements, then the request is immediately accepted; otherwise another attempt is made to place the request by calculating the other routes, failure of which results in the rejection of QoS request.

The Table 7.1 shows the routing Mechanism for IMS SIP traffic over Access Networks. IMS network allows services to users when user QoS requests are accepted according to the available resources and When SIP data arrives at the Access network, the SIP Modules are enabled for SIP traffic, then these modules handle that traffic by providing the best path utilization and required resources to data. When both the SIP and other data arrives at these modules then its gives priority to the SIP traffic then process the other available data.

Algorithm: SIP-QoS-Routing for IMS
% New SIP Session for IMS Core
while (QoS request == available) do
if PDF (Yes)
Process User Request
for Multimedia Services
end while
% Access Networks
while (packets == SIP) **do**
while(SIP-QoS-Module==enabled) **do**
Process SIP Data
end while
end while
Else Process Other available data

Table 7.1: SIP QoS Routing Module

7.2 Scenario1

In this Scenario, the source node is represented by Node 0 and Destination Node is represented by Node1. The MPLS Access network contains the nine LSR Routers to process the network traffic and data. The Links between theses routers are expressed in three bandwidths i.e 1MB, 2MB and 3MB. When node0 sends SIP data to destination Node1 the route established between these routers are different in all the three times according to the available paths and resources. IMS SIP traffic is processed through the SIP LSPs. These SIP LSPs implements the QoS Module and process the SIP traffic for best QoS. The LSP routers established different paths for traffic but after applying the QoS on these LSPs efficient and effective utilization of resources is achieved. After SIP QoS Negotiation, Diffserv LSRs Performs the Marking, Policing and Shaping on the incoming SIP packets. QoS Modules calculate unreserved and available Bandwidth Between nodes in established path.

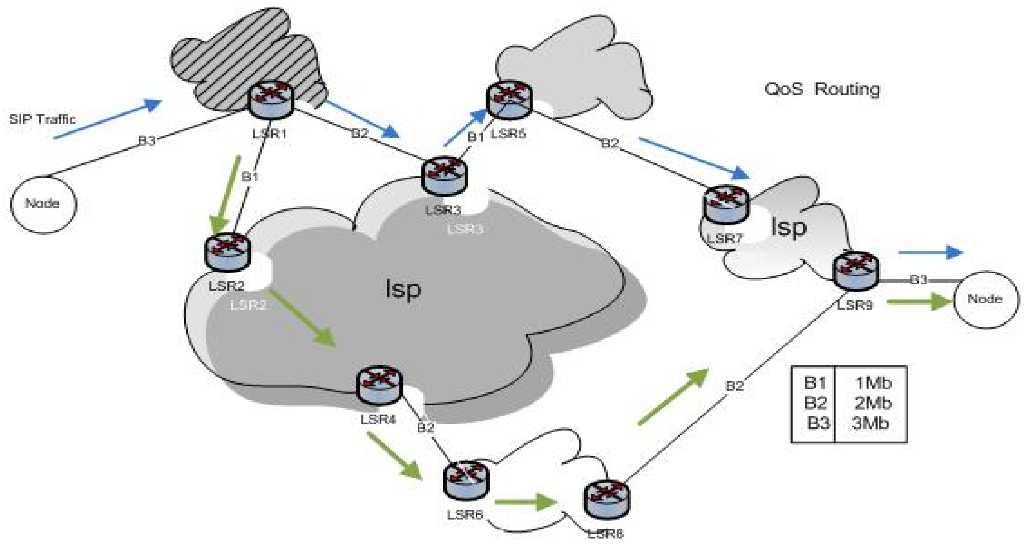

Figure 7.1: Simulation Scenario 1 for IMS QoS

7.2.1 Route establishment

Source starts sending data to destination at 1.09ms through the MPLS network [IP Backbone], First route established for SIP Traffic is 1_3_5_7_9 for traffic1. The link bandwidths are B2, B1, B2 and B3. The second path established for SIP traffic 2 is 1_2_4_6_8_9 for traffic 2 after calculating the available bandwidth and 2nd route is established. Similarly 3rd route is established 1_3_4_6_5_7_8_9 for traffic 3.

7.2.2 Observations

Simulation starts at 1.09 ms and stops at 4.10ms. Total Time to establish the 1st route for sip traffic through Diffserv LSR 1, 3,5,7,9 is 1.0833280000000001. 2nd route for 1, 2, 4, 6, 8, 9 is established at 1.3047039999999999ms . 3rd route for 1,3,4,6,5,7,8,9 is established at 1.5473230476190447. It is observed that when the links are available for processing traffic, the route is established in minimum time and it processed data efficiently. After resource reservations the other paths takes more time to establish routes. 1st route establish the shortest path for packet traversing but after that it also traverses the

SIP packets for unreserved path and available bandwidth. Average Throughput for utilizing bandwidth for all the paths established is

Average Throughput [kbps] = 13801.73kbps

Packets Output	No. of Packets
No. of Packets Sent	10286
No. of Packets Received	10249
No. of SIP Data Packets Received	9000
No. of Packets Lost	37
Average Throughput of Data	99.64%
Average throughput for SIP Data	87.50%

Table 7.2: SIP Evaluation for scenario 1

7.2.3 Evaluation

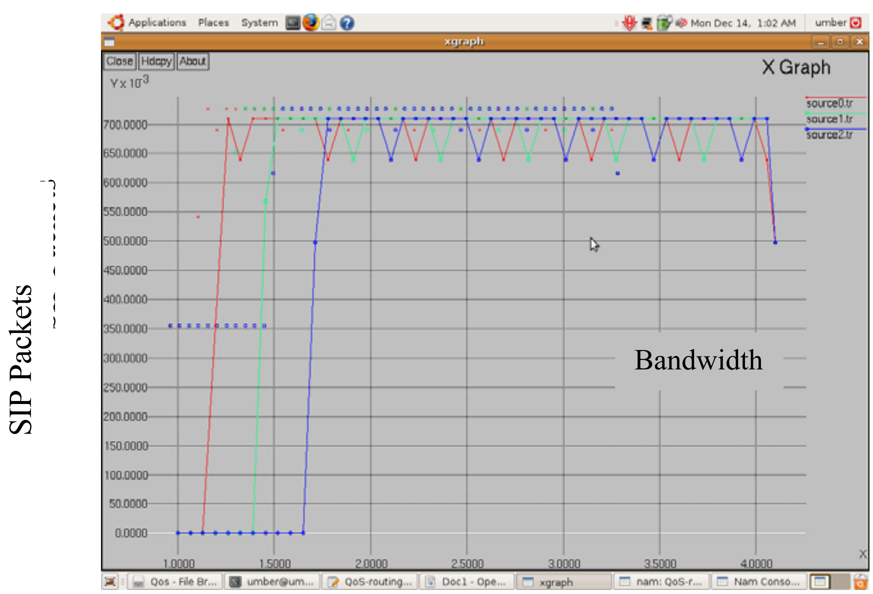

Figure 7.2: bandwidth utilization and QoS for Source0,1,2

7.2.4 Analysis of Scenario#1

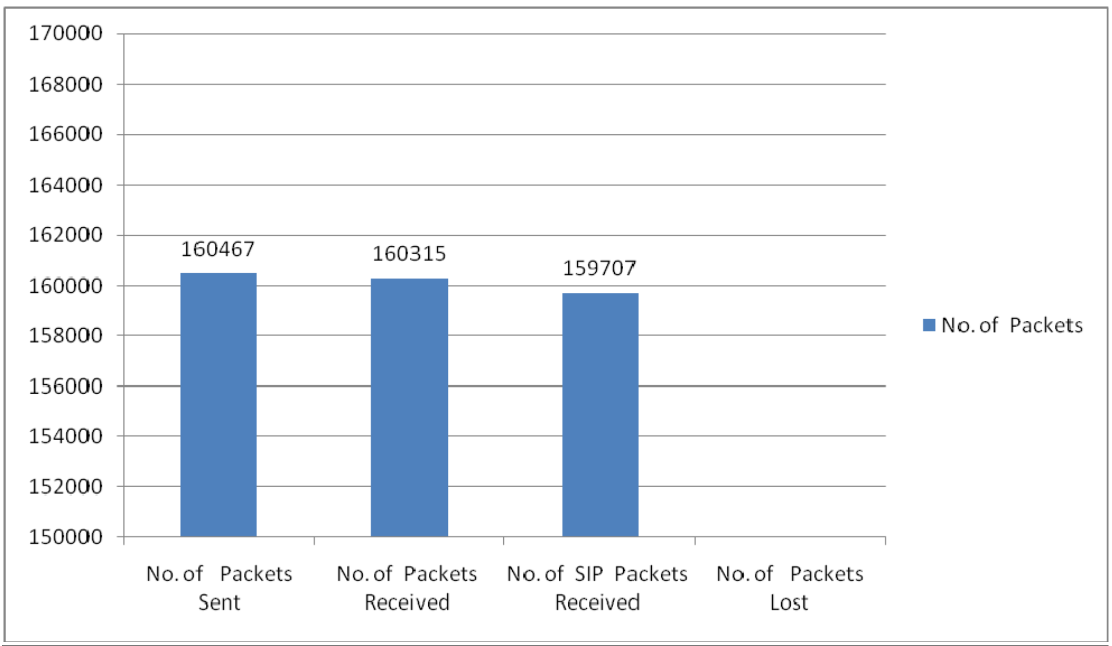

Figure 7.3: Analysis of scenario1

The Table 7.1 shows the simulation results for the scenario 1. The total Number of SIP packets send by the source and received by the destination include the session establishment packets and also data packets. The Average throughput for SIP data is 87% and link utilization is 99.64%. This shows that the number of packets lost in this case are minimum as resources are available for data.

Figure 7.5 shows the graphical representation of bandwidth utilization for SIP traffic 1, 2 and 3. The graph shows that these modules make the better bandwidth utilization for traffic. Figure 7.6 shows the total number of packets sent, received, lost and SIP data packets reached to destination. The results show that these modules can provide better QoS services to the IMS SIP traffic over MPLS and Diffserv network.

7.3 Scenario#2

Simulation Scenario 2 contains the two SIP Servers that receive the SIP data from IMS core Network. The aim of this simulation is to send SIP data from IMS core to Other IMS core network through SIP QoS Modules implemented on Diffserv LSRs. The network contains the four Diffserv LSRs to process data on network. On these Diffserv LRSs we enables the QoS Modules for SIP traffic. Both SIP Servers start sending traffic at time 1.06. This setup also contains the other sources for sending UDP data on same link.UDP Traffic Service starts at time 0. Voice and data is sent at same link for observing the performance of SIP QoS Modules for SIP traffic and UDP traffic. These SIP Servers send SIP traffic over SCTP instead of UDP.

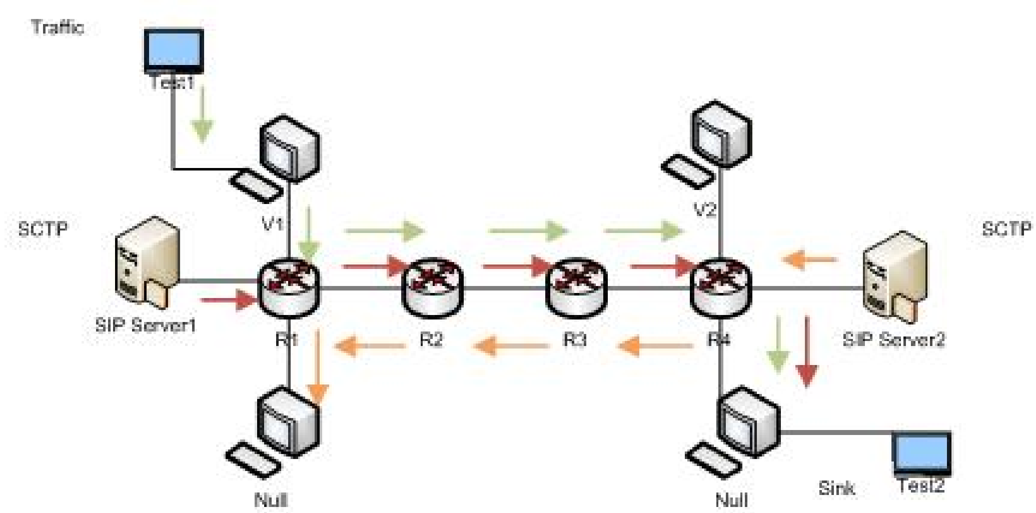

Figure 7.4: Simulation Scenario 2

7.3.1 Route establishment

Both SIP servers start sending data to sink at 1.09ms through the MPLS network [IP Backbone], and the UDP server starts its traffic at time0 ms. First route established for

SIP Traffic is shown by red arrows traffic flow. The second route established for SIP traffic 2 is shown by the orange arrows traffic flow. The link bandwidths are 255MB for each router. The traffic flow for UDP traffic path established is shown by the green arrows.

7.3.2 Observations

Simulation starts at 1.09 ms and stops at 3599.99ms. It is observed that when the links are available for processing traffic, the route is established in minimum time and it processed UDP data efficiently. But when SIP servers start sending data the QoS Modules prioritized the SIP traffic as compared to the UDP data. After the arrival of SIP traffic over network, as the SIP servers send SIP data on link the Average Throughput for Utilizing bandwidth for SIP voice traffic is

Average Throughput [kbps] = 56.80Kbps

As QoS Module also manages the bandwidth for SIP data traffic so the Average Throughput for Utilizing bandwidth for SIP Connections Establishment is

Average Throughput [kbps] = 35.59Kbps

The QoS module manages the links for SIP traffic and provides the Best QoS for IMS services. Average throughput for bandwidth utilizations for SIP traffic is

Average Throughput [kbps] = 92.39Kbps

SIP Traffic Analysis with QoS Established	
Packets Output	%
No. of Packets Sent	160467
No. of UDP Packets Received	608
No. of SIP Packets Received	159707
No. of Packets Lost	152
Average Throughput of Data	99.90%

Average throughput for SIP Data	99.52%

Table 7.3: Throughput of scenario2

7.3.3 Analysis of Scenario2

The Table 7.2 shows the simulation results for the scenario 2. The Total Number of SIP packets send by the source and received by the destination includes the session establishment packets and also data packets for IMS Multimedia voice services and also UDP data traffic. The Average throughput for SIP traffic is 99.52% and link utilization is 99.90%. This shows that the number of packets lost in this case are minimum as resources are available for data.

Figure 7.8 shows the total number of packets sent, received, lost and SIP data packets reached to destination. This graph also represents the Total number of UDP packets processed by the network. The results show that these modules can provide better QoS services to the IMS SIP traffic as compared to the UDP traffic over MPLS and Diffserv network. This graph shows that the utilization of resources for SIP traffic is made possible instead of UDP traffic. Multimedia services require more resources as compared to the simple data traffic. Numbers of SIP packets traversed are more than the UDP packets. So the results show that SIP-Based QoS Modules are more efficient for transmission of multimedia traffic of IMS over IP Access Networks.

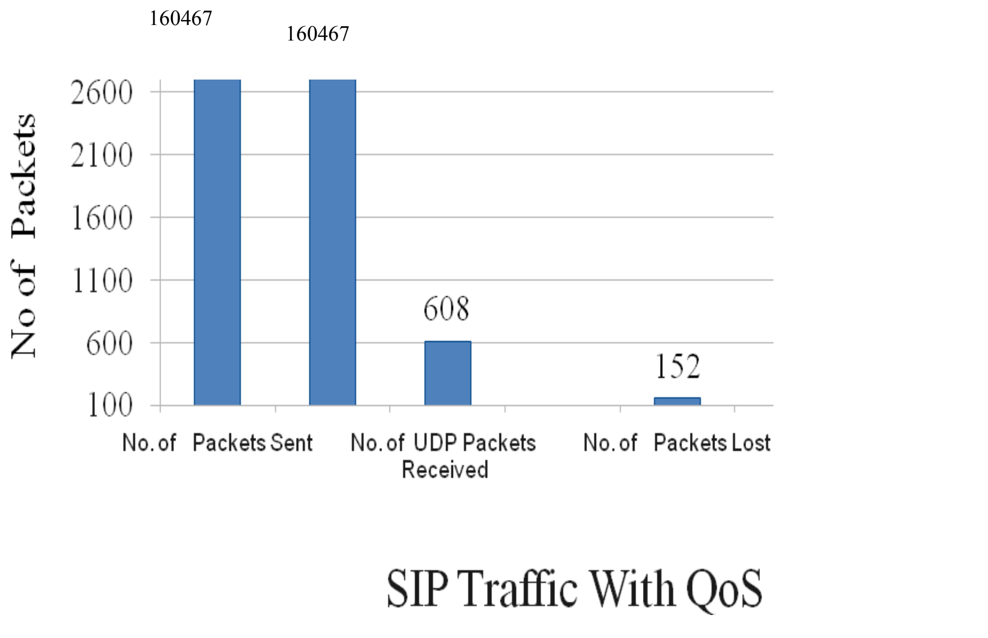

Figure 7.5: Analysis of scenario2

7.4 Fairness Results

In this section we will compare our approach with simple QoS mechanism over Access networks. Table 7.5 shows the results of SIP data processing over Access networks with SIP-Based Modules disabled.

SIP Traffic Analysis with QoS Established	
Packets Output	%
No. of Packets Sent	160467

No. of UDP Packets Received	608
No. of SIP Packets Received	159707
No. of Packets Lost	152
Average Throughput of Data	99.90%
Average throughput for SIP Data	99.52%

Table 7.4: Throughput for SIP Traffic with QoS SIP module

SIP traffic Analysis without QoS	
Packets Output	%
No. of Packets Sent	155495
No. of UDP Packets Received	1817
No. of SIP Packets Received	153478
No. of Packets Lost	200
Average Throughput of Data	98.70%
Average throughput for SIP Data	97.71%

Table 7.5: Throughput for SIP Traffic without SIP QoS Module

Simulation results of Table 7.5 show that the total number of UDP packets traversed by the network is 1817 and SIP multimedia traffic packets are 153478. The average throughput of link is 97.71%. The Total number of packet lost are 200 as compared to the SIP-Based Module enabled where packet lost are just 152.

7.4.1 Observations

Simulation starts at 0.66ms and stops at 3599.99ms. It is observed that when the links are available for processing traffic, and there is no SIP traffic to be processed. These modules process the other available traffic. But when SIP servers start sending data the QoS Modules prioritized the SIP traffic and ignore the other data traffic as the UDP data.

After the arrival of SIP traffic over network where SIP QoS Modules are disabled, Average Throughput for Utilizing bandwidth for SIP voice traffic is

Average Throughput [kbps] = 54.59Kbps

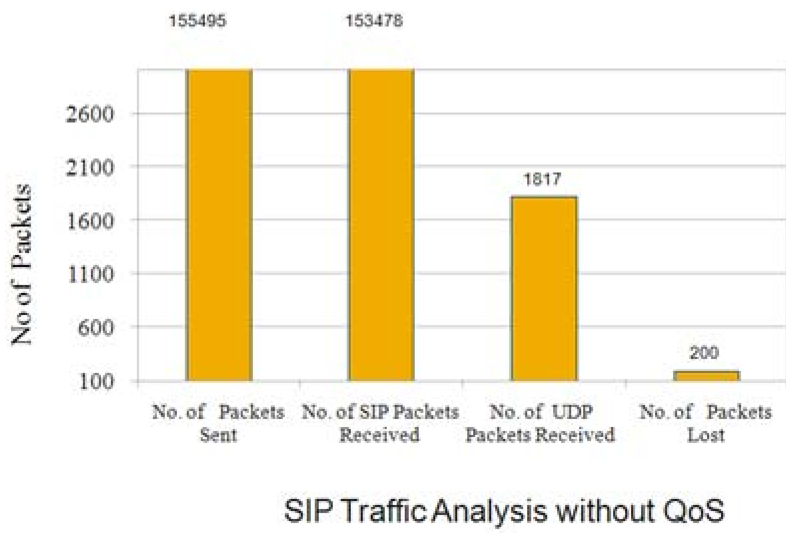

Figure 7.6: Analysis without QoS

As QoS Module also manages the bandwidth for SIP data traffic but when SIP QoS modules are disabled the network routers also process the UDP data and the Average Throughput for utilizing bandwidth for other UDP data is:

Average Throughput [kbps] = 34.56Kbps

Average throughput for bandwidth utilizations for SIP traffic and UDP traffic is:

Average Throughput [kbps] = 89.15Kbps

By comparing the both scenarios we conclude that the overall performance of SIP Data sent through QoS SIP Module provides better utilization of bandwidth and paths. These Modules not only use the available links but also utilize unreserved links. Traffic used to send is SIP-SCTP Instead of SIP-UDP. IMS traffic through SIP Module is now more valuable and well-organized.

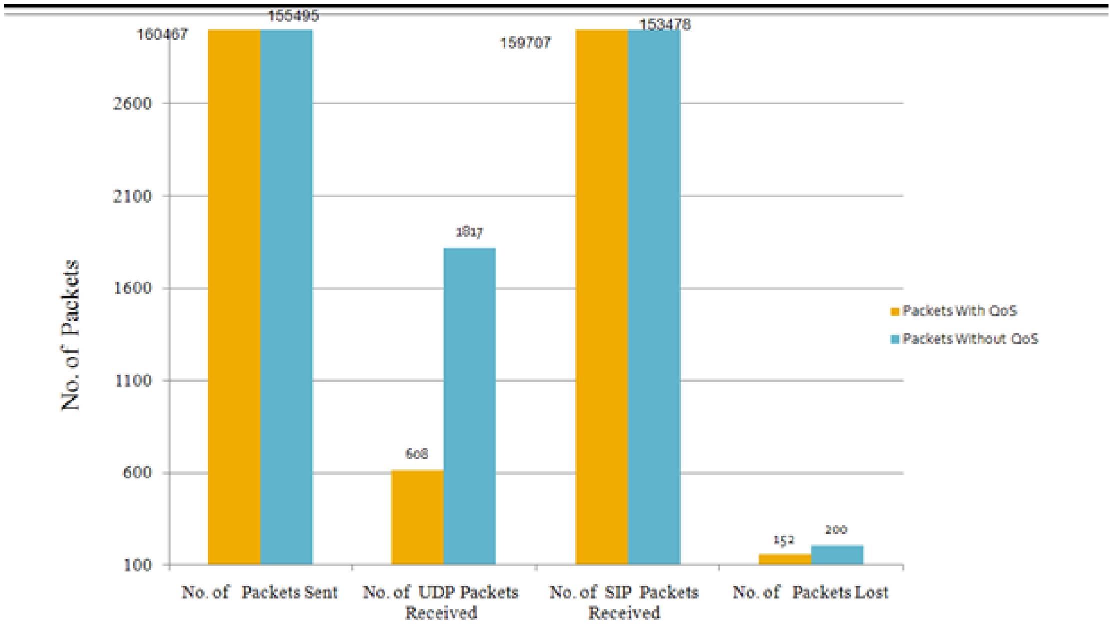

Figure 7.7: Comparative Analysis

7.5 Simulation Results

The simulation results are in the form of packets traversed with bandwidth in Mbps. Here, figure of scenario1 shows a case where only SIP traffic is present in the network from node 0 to node 1. Figure of sceanrio2 shows a case where UDP traffic and SIP traffic is present in the network from node 0 to node 1. The SIP bandwidth shown in figure is as measured from the source. Figure 7.7 shows the results for the simulation experiment with high rate of SIP traffic and UDP traffic being mixed in the same link between nodes 0 and 1 via MPLS nodes. Figure 7.10 shows that UDP is getting negatively affected on a significant increase in SIP traffic. Table 7.3 shows the results for the simulation experiment with QoS enabled in the network.

The simulation results show that performance of SIP modules is better in both network scenarios as shown in the figure 7.7.

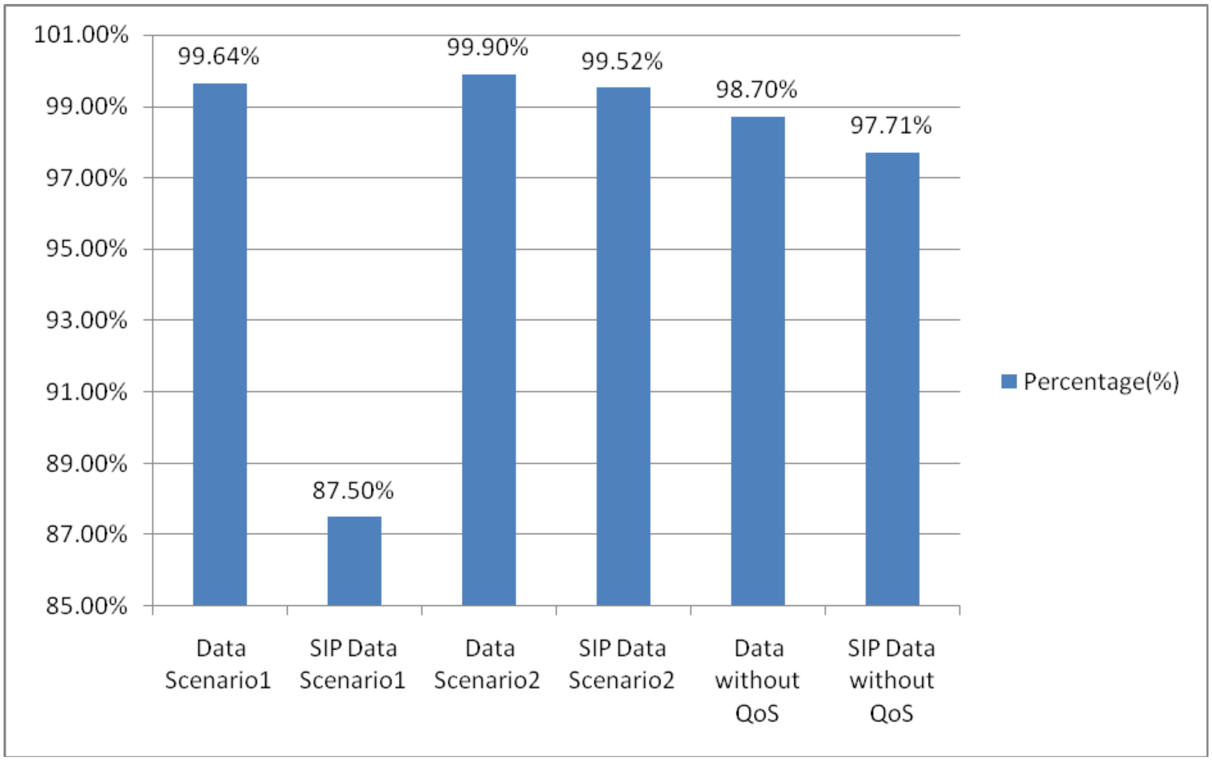

Figure 7.8: Performance of SIP Module

Chapter 8

Conclusion & Future Recommendations

In the previous sections, SIP-Based QoS framework has been proposed for IMS and NGN All-IP network and the architecture of QoS has been reviewed for IMS and IP Access Networks.

This Thesis proposed a SIP-Based QoS framework for IMS and its underlying access networks to distinguish the IMS services over IP access networks. To introduce the QoS provisioning this framework supports the SIP-Based QoS Modules based on the mobility management. In our approach we have used the combination of Diffserv and MPLS for QoS provisioning over access networks. To guarantee QoS, QoS SIP Modules reserve resources in advance from available resources to provide best QoS to IMS users over heterogeneous networks. To obtain a more efficient use of these SIP Modules, our approach allows the use of SCTP protocol instead of UDP due to its multi-streaming feature. We developed an analytic model for our proposed QoS management scheme. This analytic model is different from the existing approaches due to the combination of Diffserv and MPLS in IMS underlying access networks. The simulation results show that SIP-based QoS architecture provides the comparable results for transmission of IMS services.

- Following research contributions have been made by this work:
    - QoS Modules have been proposed for handling IMS data traffic with integration of Diffserv & MPLS
    - The performance of SIP modules has been checked using two different network Scenarios
    - The simulation results have shown that the performance of the proposed architecture is better in both scenarios due to implementation of SIP QoS modules

- The Comparison of different architectures shows that the proposed architecture has provided better results
- Based on Simulation experiments, it has been observed that:
  - SIP module provides the best path and reduces congestion
  - The proposed architecture efficiently utilizes the available bandwidth

8.1 Future Work

- The proposed framework was tested for MPLS and Diffserv techniques. It would be interesting to investigate the performance of SIP-Based QoS Modules using other QoS techniques such as Intserv and RSVP.
- We proposes an extended SIP-Based QoS management architecture based on SIP proxy modules. The main objective of introducing this architecture is to distinguish IMS networks from other available IP networks to manage the QoS more efficiently and to reduce the hand off disruption time over IMS and underlying access networks. We introduce the SIP-based proxy QoS modules to handle IMS traffic over heterogeneous networks.

  - SIP based QoS Module which is proposed for future work is by use of integration of following QoS mechanisms
    - Integrated Services
    - Multi-Protocol Label Switching
    - Resource Reservation Protocol
    - Differentiated Services

References

[1] Camden Associates San Bruno,White Paper Simulation, Testing and Monitoring In IP Multimedia Subsystem Networks CA, May 2006.

[2] Copyright © 2006 Princeton Research Solutions LLC. Terms of Use. http://www.princetonrs.com/

[3] 3GPP TS 23.228, “Service Requirements for the Internet Protocol (IP) Multimedia Core Network Subsystem (IMS), Stage 1.”

[4] 3GPP TS 23.228, “IP Multimedia Subsystem (IMS); Stage 2.”

[5] IP Multimedia Subsystems, Overview and applications, 3G Americas, july 2004.

[6] Umber Iqbal, Email:umber.iqbal@yahoo.com , Dr. Shaleeza Sohail, Email:shaleezas@hotmail.com, Dr. Younas Javed, Email:myjaved@ceme.com, SIP-Based QoS Management Architecture for IP Multimedia Subsystems over IP Access Networks , Department of Computer Engineering, College of Electrical & Mechanical Engineering, National University of Science & Technology, Islamabad, Pakistan

[7] Dan Leih, Strategic Marketing Manager Dave Halliday, Applications Architect Motorola, Inc.,IMS White Paper Series: Part 1 Introduction to IMS Standards, protocols, architecture and functions of the IP Multimedia Subsystem, Embedded Communications Computing.

[8] Edward Oguejiofor Philippe Bazot, Bruno Georges, Rebecca Huber, Callum Jackson, Jochen Kappel, Cameron Martin, Bala S. Subramanian, Developing SIP and IP Multimedia Subsystem (IMS) Applications.

[9] http://en.wikipedia.org/wiki/IP Multimedia Subsystems.

[10] http://portal.etsi.org/portal_common/home.asp?tbkey1+TISPAN

[11] www.itu.int/ITU-T/ functional architecture Release 1 document details, in accordance with the ITU-T (), the general reference model for next-generation networks (NGN).

[12] IP Multimedia Subsystems(IMS): A Standardized Approach to All-IP Converged Networks Issue Date: January 2006

[13] Qiu Qingxian & Ji to,BUILDING E2E IMS NETWORK SECURITY Nov, 2006.

[14] rest Roland Lenz , IP Multimedia Subsystems Version 1.0 Last change Copyright See Sources, Homepage http://www.2cool4u.ch History Version Datum Document created 1.0 19th of May 2008

[15] iikka Poikselka, Nokia, Finland, Georg Mayer, Nokia, Finland, Hisham Khartabil, Telio, Norway, Aki Niemi, Nokia, Finland, THE IMS IP Multimedia Concepts and Services, SECOND EDITION

[16] Zehan Zeb, IP Multimedia Subsystem (IMS): Concepts and Architecture, Contribution to: NANC FoN, October 4, 2006, Contact: zzeb@telcordia.com, (732) 699-6163

[17] Allied Telesis, QoS White Paper, White Paper

[18] Tekes Dnro 2625/31/03 Drawn up by: Jarmo Prokkola, Mikko Hanski, QoS Measurement Methods and Tools, Project's full name: Easy Wireless, Allow seamless roaming between wireless networks while maintaining Quality of Service, Project's short name: EW Project number: E4SU00169 Customer reference: 01.06.2005

[19] Ankur Chadda B.E., University of Bombay(Mumbai), Quality of Service Testing Methodology India (1999) THESIS Submitted to the University of New Hampshire in Partial Ful¯llment of the Requirements for the Degree of Master of Science in Computer Science December, 2004.

[20] SANS Institute InfoSec Reading Room,Latency and QoS for Voice over IP , URL:http://www.riverstonenet.com/pdf/qos.pdf

[21] Dr Ying Dar Lin, Internet QoS: Diffserv and Intserv, High Speed Network Lab, Department of computer information Science , National Chiao Tung University.

[22] Braden, R., Clark, D. Shenker, S., "Integrated Services in the Internet Architecture: an Overview ". Internet Engineering Task Force, Request for Comments (RFC) 1633, June1994.

[23] Broadband Radio Access for IP-based Networks, IST-1999-10050, http://www.istbrain.org

[24] Geo® Huston, \Internet Performance Survival Guide QoS strategies for Multi-service Networks", *Wiley Computer Publishings*, 2000.

[25] Braden, R., Zhang, L., Berson, S., Herzog, S., Jamin, S., .Resource ReSerVation protocol (RSVP) -- Version 1, Functional Specification.. Internet Engineering Task Force, Request for Comments 2205, September 1997.

[26] Wroclawski, J., “The use of RSVP with IETF integrated Services”, IETF RFC 2210, 1997.

[27] Blake, S., Black, D., Carlson, M., Davies, E., Wang,, Z., Weiss, W., “An Architecture for Differentiated Services”. Internet Engineering Task Force, Request for Comments (RFC) RFC, 2475, Dec. 1998.

[28] Nichols, K., Blake, S., Baker, F., Black, D., “ Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers”, RFC 2474 December 1998.

[29] Jacobson, V., Nichols, K., Poduri, K., “An Expedited Forwarding PHB”, IETF RFC 2598, 1999.

[30] Heinanen, J., Baker, F., Weiss, W., Wroclawski, J., “Assured Forwarding PHB group”, IETF RFC 2597, 1999.

[31] Jacobson, V., Nichols, K., Poduri, K., “An Expedited Forwarding PHB”, IETF RFC 2598, 1999.

[32] Broadband Radio Access for IP-based Networks, IST-1999-10050, http://www.istbrain.org

[33] Heinanen, J., Baker, F., Weiss, W., Wroclawski, J., “Assured Forwarding PHB group”, IETF RFC 2597, 1999.

[34] Y. Bernet, S. Blake, D. Grossman, A. Smith , “An Informal Management Model for Diffserv Routers”, Internet Draft , draft-ietf-diffserv-model-06.txt, February 2001.

[35] Giovanni Costa, Intserv over Diffserv for IP QoS in Radio Access Networks.

[36] Gilles Bertrand, The IP Multimedia Subsystem in Next Generation Networks, May 30, 2007

[37] A Set-top Box for End-to-end QoS Management and Home Network Gateway in IMS Muhammad Shoaib Siddiqui, Syed Obaid Amin, non members and Choong Seon Hong, Member, IEEE, 2009

[38] Alexandre Lung-Yut-Fong, 3G IP Multimedia Subsystem: QoS and Access Convergence, December 5, 2006.

[39] Tomislav BoroSa, Bruno Marsic , Sergej Pocuca, QoS Support in IP Multimedia Subsystem Using DiffServ,

[40] A new architecture for wireline access to the 3G IP multimedia subsystem. In Information, Communications and Signal Processing, 2005 Fifth International Conference on, pages 493–497, December 2005.

[41] F.C. de Gouveia and T. Magedanz. A framework to improve QoS and mobility management for multimedia applications in the IMS. In Multimedia, Seventh IEEE International Symposium on, December 2005.

[42] M. Mani and N. Crespi. New QoS control mechanism based on extension to SIP for access to UMTS core network via different kinds of access networks. In Wireless And Mobile Computing, Networking And Communications, 2005. (WiMob'2005), IEEE International Conference on, volume 2, pages 150–157, August 2005.

[43] Shun-Ren Yang, Member, IEEE, and Wen-Tsuen Chen, Fellow, IEEE, SIP Multicast-Based Mobile Quality-of-Service Support over Heterogeneous IP Multimedia Subsystems

[44] Raouf Boutaba and Issam Aib David R, Policy-Based Management: A Historical perspective Cheriton School of Computer Science, University of Waterloo, Canada

[45] Reiner Ludwig, Hannes Ekström, Per Willars and Niklas Lundin, An Evolved 3GPP QoS Concept Ericsson Research & Development

[46] Elwalid Sidahmed and Hesham H. Ali, Department of Computer Science, A QoS Based Model for Supporting Multimedia Applications over Cellular IP, College of Information Science and Technology, University of Nebraska at Omaha, Omaha, NE 68182-0116, {esidahmed | hali}@mail.unomaha.edu

[47] J. Postel, ”Internet Protocol”, RFC 791, September 1981

[48] Mohamed H. Ahmed "Scheduling of Multimedia Traffic in Interface-limited Broadband Wireless Access" Proc. of the 5th, International Symposium in Wireless, Personal, Multimedia Communications (WPMC 02)

[49] Prathima Agrawal, Auburn UniversityJui-Hung Yeh and Jyh-Cheng Chen, National Tsing Hua UniversityTao *Zhang* IP Multimedia Subsystems in 3GPP and 3GPP2: Overview and Scalability Issues, *Telcordia Technologies, Inc.*

[50] Fabricio Carvalho de Gouveia , A Framework to Improve QoS and Mobility Management for Multimedia Applications in the IMS *Technical University of Berlin Franklinstr. 28-29, D-10587 Faculty IV / Next Generation Networks Berlin, Germany* gouveia@cs.tu-berlin.de Thomas Magedanz *Technical University of Berlin Franklinstr. 28-29, D-10587 Faculty IV / Next Generation Networks Berlin, Germany tm@cs.tu-berlin.de*

[51] F.Weber1,2, W.Fuhrmann3, U.Trick2, U.Bleimann3, B.Ghita1, A framework for improved QoS evaluation and control in SIP-based NGN 1 Network Research Group, University of Plymouth, Plymouth, United Kingdom 2 Research Group for Telecommunication Networks, University of Applied Sciences Frankfurt/M., Frankfurt/M., Germany 3 University of Applied Sciences Darmstadt, Darmstadt, Germany e-mail: weber@e-technik.org

[52] Frank Siqueira and Vinny Cahill *Distributed Systems Group,*Quartz: Supporting QoS-Constrained Services in Heterogeneous Environments *Department of Computer Science, Trinity College Dublin, Ireland E-Mail: Frank.Siqueira@cs.tcd.ie, Vinny.Cahill@cs.tcd.ie*

[53] Guest Editorial: Quality of Service in Multimedia Networks

[54] Wei peng chin, Takeo Hamada, Jason Jeyhsin Yao, Hung-Yu wie,QoS management and peer-to-peer mobility in fixed mobile convergence.

[55] V. Sarangan and J.-C. Chen, "Comparative Study of Protocols for Dynamic Service Negotiation in Next Generation Internet," *IEEE Commun. Mag.*, vol. 44, Mar. 2006,pp. 151–56.

[56] W. Zhuang *et al.*, "Policy-based QoS Architecture in the IP Multimedia Subsystem of UMTS," *IEEE Network*, vol.17, May–June 2003, pp. 51–57.

[57] Amardeep Singh Gagan Mittal,QoS and Traffic Engineering: MPLS, DiffServ and Constraint Based Routing, A project report done in partial fulfillment of the requirements for the degree of Bachelor of Technology Submitted by May, 2004

[58] Fahad Rafique Dogar, Zartash Afzal Uzmi and Shahab Munir Baqai, CAIP: A Restoration Routing Architecture for DiffServ Aware MPLS Traffic Engineering Computer Science and Engineering Lahore University of Management Sciences, Pakistan {fahad,zartash,baqai}@lums.edu.pk

[59] Wei-Chu Lai, Kuo-ChingWu, and Ting-Chao Hou*, Design and Evaluation of Diffserv Functionalities in the MPLS Edge Router ArchitectureCenter for Telecommunication Research and Department of Electrical Engineering, National Chung Cheng University, Chia-Yi, Taiwan, 62102 * tch@ee.ccu.edu.tw

[60] Raymond Law and Srihari Raghavan, DiffServ and MPLS – Concepts and Simulation, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, {rlaw,srihari}@vt.edu

[61] Quality of Service and MPLS Methodologies , www.ipinfusion.com

[62] Thomas Telkamp Director, Data Architecture & Technology Global Crossing Telecommunications, Inc. *telkamp@gblx.net*

[63] Forschungszentrum Telekommunikation WienJohannes Stadler,"IP Multimedia Subsystems

[64] Raymond Law and Srihari Raghavan, "DiffServ and MPLS – Concepts and Simulation", Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, {rlaw,srihari}@vt.edu

[65] La Monte H.P. Yarroll, Karl Knutson, "SIP over SCTP "Linux Kernel SCTP : The Third Transport"http://lwn.net/2001/features/OLS/pdf/pdf/sctp.pdf