Route Selection for QoS over Mobile IP

Azimeh Sefidcon

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

June 2007

Your file Votre référence
ISBN: 978-0-494-31143-1
Our file Notre référence
ISBN: 978-0-494-31143-1

ABSTRACT

Route Selection for QoS over Mobile IP

Azimeh Sefidcon, PhD
Concordia University, 2007

Influenced by the availability of powerful portable computers, and the expansion of Internet-based services, consumers demand mobile Internet. This implies supporting user mobility while sustaining network connectivity which in turn requires more sophisticated routing methods than simple static Internet protocol. Routing optimization is critical for the efficiency of mobile Internet as it directly impacts resource utilization in the network. On the other hand, while connected to the Internet, users want to enjoy real-time services. Granting quality of service for real-time multimedia applications is one of the major concerns for Internet Service Providers.

In the mobile Internet context, as a result of a signalling conflict, providing quality of service declines in presence of routing optimisation. This conflict must be resolved to achieve efficient resource utilization in the networks and provide quality of service for real-time Internet-based services.

This thesis presents a solution for resolving the aforementioned conflict. Initially a set of requirements are established for a new solution. The fundamental idea is developed based on exercising routing optimization according to quality of service requirements and network conditions. A new protocol architecture for Mobile IP (MIP) is built based on a cross-layer design technique. This design

implies that a new data flow collects the necessary parameters in the network. This data is passed to a new entity in the network layer of the MIP protocol stack that inter-connects IP, routing, mobility and resource reservation protocols.

Based on network conditions and the quality of service requirements the optimal path is selected over which the resources are reserved for the duration of the connection. The new solution named MIP with Routing Optimization and QoS (MIP-ROQS) removes the signalling conflict and furthermore, results in increased network performance. Simulation of this solution demonstrates reduction in application end-to-end delay and delay variation, improved core network by reducing the amount of control traffic and dropped traffic and improved access network by reducing its delay.

Moreover, we integrate Multi Protocol Label Switching (MPLS) architecture and MIP-ROQS to take advantage of MPLS labelling mechanism to label the optimal path selected by MIP-ROQS for the duration of a session. This integration makes IP-in-IP tunnelling in data forwarding redundant. MPLS is used to switch packets and label the optimal path over which the resources are reserved. Using this approach, the transmission delay, and packet processing overhead are further reduced.

Acknowledgement

It has been a journey of challenge and discovery, a process of learning and self realization. In this journey I am grateful to the presence of many.

First and most of all I would like to express my sincere gratitude to my supervisor Dr. Ferhat Khendek for his technical and moral advice, inspiration and encouragement during the course of this study.

I would like to thank my examining committee members Dr. Agarwal, Dr. Butler, Dr. Seffah, and Dr. Hamouda for reviewing my work and giving valuable feedback at every milestone of this research with special thanks to Dr. Samuel Pierre for accepting to be my external examiner.

Research department of Ericsson Research Canada for providing me with the access to Opnet tool and supporting me through my research.

I would like also to extend my appreciation to Dr. Lata Narayanan for her time and valuable technical discussion and guidance during my research.

No word can express my deepest appreciation to my life companion Niklas Sandberg and to my beautiful children Ghazal and Sara who gave the time that belonged to them. To my sister Forough who set my standard for perseverance and strength. To my Father who was the first teacher I knew and to the spirit of my mother who has been with me and will be forever.

Dedicated to

my parents,

my life companion, Niklas

my lovely children, Ghazal and Sara

and
my spiritual master, Eckhart Tolle

There is so much more to a human being than thoughts activity, there is so much more intelligence beyond the world of thought, in the realm where intuition, creativity and sudden realisation come from.
Eckhart Tolle

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AA	Agent Advertisement
AS	Agent Solicitation
BGP	Border Gateway Protocol
BSC	Base Station Controller
CBR	Constraint Based Routing
CDMA	Code Division Multiple Access
CN	Correspondent Node
CoA	Care of Address
DiffServ	Differentiated Services
DRP	Directional Routing Protocol
ENCAP	Encapsulation
FA	Foreign Agent
FDA	Foreign Domain Agent
FEC	Forwarding Equivalent Class
GCoA	Global Care of Address
HA	Home Agent
HMIPv6	Hierarchical Mobile IP Version 6
HMRSVP	Hierarchical Mobile RSVP
IntServ	Integrated Services
ICMP	Internet Control Message Protocol
IGRP	Interior Gateway Routing Protocol

IRDP	Internet Router Discovery Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPv6	Internet Protocol Version 6
ISP	Internet Service Provider
LAN	Local Area Network
LCoA	Local Care of Address
LDP	Label Distribution Path
LER	Label Edge Router
LR	Location Register
LSR	Label Switch Router
LSP	Label Switch Path
MAC	Medium Access Control
MANET	Mobile Ad-Hoc Network
MAP	Mobility Anchor Point
MIP	Mobile Internet Protocol
MN	Mobile Node
MPLS	Multi Protocol Label Switch
MRSVP	Mobile RSVP
OSI	Open System Interconnection
OSPF	Open Shortest Path First
PARA	Penalty-based Adaptable Reservation Admission
PK	Packet

QoS	Quality of Service
QP	Quality Parameters
REG	Registration
RIP	Routing Information Protocol
ROQS	Routing Optimization with Quality of Service
RRP	Registration Reply
RRQ	Registration Request
RSVP	Resource Reservation Protocol
TCP	Transport Control Protocol
TE	Traffic Engineering
TIMIP	Terminal Independent MIP
UDP	User Data
VoIP	Voice over IP
WLAN	Wireless Local Area Network
RO	Route Optimisation
TOS	Type of Service

GLOSSARY

Cell: The smallest area of direct connectivity in a mobile network.

DiffServ: A mechanism for providing QoS in IP in which at each router in the network, QoS preferences are assigned to traffic aggregates, which are composed at the network edge. This is by marking of packets in a special field in the IP header.

Encapsulation: The act of transporting one protocol inside another protocol.

Handoff/handover: The act of moving from one cell to another. The term handoff usually implies a cell controller initiated action, as opposed to the term cell switching.

Host: A computer connected to the network, where services or services run.

IETF: An open organization, now under the Internet Society, whose mission is to develop protocols and other standards for the Internet.

IGRP: The Interior Gateway Routing Protocol is a routing protocol that was developed in the mid-1980s by Cisco Systems, Inc.

Internet: An IP-based inter-network, spanning the entire globe, and consisting of national and regional networks, which in turn interconnect organizational networks.

IntServ: A mechanism for providing QoS in IP by signalling per flow QoS requirement from the source to the destination. It uses Resource Reservation Protocol (RSVP) for end-to-end signalling protocol.

Mobile Node: The term may have the more restrictive meaning of a mobile host as specified by Mobile IP.

Mobile host: A host which can move in a network while some property concerning it, such as its address, its connections, etc., remains unchanged.

Mobility: The property mobile hosts have for being able to move in a network.

RFC: An Internet document specifying a protocol or other standards for use in the Internet. There also exist informational RFCs which give information as opposed to specifying a standard.

Router: A layer three device that forwards packets between networks based on their network address.

Routing: The act of forwarding packets between networks based on their network address.

Triangular routing: A routing where packets from host A to host B follow a direct route, but packets from B to A have to go through a third host

Tunnelling: Bypassing of normal routing procedures to get packets across a part of the network that cannot properly route them. Tunnelling is usually accomplished either by source routing or by encapsulation.

CHAPTER 1

INTRODUCTION

1.1 Research Context

The last two decades have brought tremendous success to the Internet and its technologies. The importance of wireless Internet access has grown as the continuation of this success. This has inspired the development of a variety of new services in business and consumer markets. In addition to the traditional data services currently provided over Internet, new voice and multimedia services are being developed and deployed.

The high traffic volume of services such as Voice over Internet Protocol (VoIP) and other multimedia traffic such as IPTV is imposing great challenges on the networks. These new services have driven the demand for increased and guaranteed bandwidth requirements in the network.

Providing network access for the mobile hosts, that change their location physically, needs a more sophisticated mechanism than traditional network connectivity and routing methods. “How to communicate with the mobile hosts” is the concern of mobility protocols.

At the same time the objective of Internet connectivity for a mobile user is to profit from the Internet-based services. Different services have different requirements to yield adequate performance. Between all the services, multimedia services are one of the main concerns as they are dependant greatly on the availability of the required resources in the network.

Mobile hosts want to use the same services as fixed hosts; however the complexity of providing QoS for multimedia services in the fixed network escalates greatly when adding the mobility effects to the picture.

This complexity arises because at the beginning Internet Protocol (IP) was neither designed for mobility nor for handling the real time multimedia with QoS. Special support is needed in order to maintain connectivity to the Internet for a Mobile Node (MN) as it has a fixed IP address at its home network and can obtain a new address in a foreign network. Mobile IP (MIP) protocol [61] developed by IETF provided this support at the IP layer. Regardless of vast research effort in this domain, wireless-IP mobile hosts are still facing important problems. Some of the main issues are security, QoS, resource management and wireless link adaptation.

MIP was chosen as the context of this research. In this context we look closely at the Routing Optimization (RO) for the sake of network performance and QoS for service performance and examine the issues when both QoS and RO are required.

1.2 Problem Statement and Research Issues

The performance of the host mobility depends on the efficiency of the routing mechanism used to route data from a source to a moving destination. MIP protocol on IPv4 [61] suffers from triangle routing, which is forwarding packets destined for MN via its Home Agent (HA). While on a foreign network MN receives a temporary address, and registers this temporary address with HA. HA can encapsulate the packets received for MN and send them to it. On the way back the packets are sent directly from MN to the Correspondent Node (CN).

The problem is tunnelling of packets via HA although supporting mobility creates traffic congestion at HA and hides information in the packet from the routers along the tunnel. Assuming a packet contains a control message (such as PATH message) of a resource reservation protocol (such as ReSource reserVation Protocol (RSVP) [2] in Integrated Services (IntServ) architecture [30]). This message will not be recognized by routers along the tunnel and the resource reservation mechanism will fail.

Research efforts on MIP and how to optimize the path via which the packets travel have resulted in many different mechanisms to provide RO for MIP. On the other hand providing QoS for Internet-based real-time services has been the focus of other researchers so far [69-70].

The early solutions addressed the need for efficient routing of packets as they traversed the network, but they did not address the QoS requirements of the information contained in the packets. Also, most of the routing protocols deployed today are based on algorithms designed to obtain the shortest path in the network for packet traversal and do not take into account additional metrics such as delay, and traffic congestion, which can further diminish network performance.

Available techniques for RO change the triangle path to a direct path, however they rely on the first packet (possibly a control message) to follow the triangle path to learn the location of the MN. Therefore the above problem is still in place. Moreover, a more crucial concern is that the direct path may not be the optimal path in terms of available resources for providing QoS.

The obvious way to convey mobility and QoS is to blend a mobility protocol with a QoS providing mechanism such as IntServ or DiffServ [3]. Unfortunately, this approach is not adequate without further modification. Reservation signalling interferes with the mobility support and as a result the global solution is not satisfying the primary goal. In addition, taking care of packet forwarding on the shortest path is not necessarily in tune with providing QoS in terms of available resources for the service in demand.

Therefore, the two enhancements, "QoS" and "RO" cannot simply be offered beside each other for MIP environment without blocking one another. Additional mechanisms are needed to select the optimized routes based on the quality of service requirements. The objective of this research is to resolve the conflict in providing simultaneous QoS and RO in MIP in order to improve both network performance and QoS guaranteed for multimedia services.

1.3 Contributions of This Thesis

Through investigation we found the key for resolving the above mentioned issue to be "selecting an optimal path based on the network conditions and QoS requirements prior to reserving the necessary resources". Considering network conditions while selecting a route is constraint-based routing that can be either static or dynamic.

Static constraint-based routing takes into account necessary network conditions before the actual communication session takes place. Dynamic constraint-based routing evaluates the path conditions right at the time of choosing the route. The trade off is the accuracy versus processing speed. For developing the new solution we adapt a dynamic constraint-based routing in MIP.

To choose the optimal path, we need to gather necessary information from the network. The most important information is network conditions in terms of traffic, link utilization, delay, etc. Collecting this information can not be achieved in current standard layered network architecture. We need to adapt a cross-layer design [106] in order to break the gap to certain extent between different layers of the MIP protocol stack and to collect hidden information from inside these layers.

To avoid introducing new signalling, we use the existing signalling for routing, mobility and resource reservation and enrich them by adding necessary parameters. These parameters are input to an optimal path selection mechanism that will analyse the network conditions against QoS requirements. The outcome of this analysis is the optimal route which satisfies the QoS requirements and harmonizes traffic load in the network. The analysis is performed using proposed position analysis and routing analysis algorithms.

The contributions of this thesis include establishing a set of ***"requirements",*** and arguing the need for a ***"cross-layer design"***, based on which a new ***"architecture"*** is built. This new architecture is enhanced with a ***"new entity"*** at network layer in the MIP protocol stack that inter-connects relevant protocol entities.

A new ***"information flow"***, carries necessary data to the new entity to serve as input to an ***"optimal path selection"*** mechanism. This mechanism is performed using ***"two new algorithms"***, for position analysis and routing analysis. As a result, the message sequence is modified from the original MIP message sequence.

The performance evaluation of new solution is done through ***"simulation with Opnet"*** and the result of this evaluation shows improved service response time and increased network performance.

The new solution which embraces all of the above mentioned modifications is referred to as MIP_ROQS (short for Mobile IP with Routing Optimisation and QoS) hereafter.

As an extension, the ***"integration of MPLS and MIP-ROQS"*** is proposed to facilitate labelling of the optimal path after it is chosen by the optimal path selection mechanism. This also overcomes the scalability concerns of the resource reservation protocol.

1.4 Organization of This Thesis

This thesis is comprised of two parts. The first part focuses on the background and the description of the problem as well as introduction and evaluation of existing solutions. The second part presents new solution, its evaluation, and its extension.

<u>Chapter Two</u>

This chapter delimitates the context of this research. At an abstract level, it starts with the basics of routing and QoS in the Internet, and emphasizes on the necessity of mobility. It provides a detailed description of MIP in order to setup the context. It describes in detail the issues of RO and QoS for MIP environment and further defines the conflict in providing both at the same time. A classification and evaluation of existing approaches and mechanisms is followed by a summary of the advantages and drawbacks of each solution.

Chapter Three

This chapter presents the design approach and architectural design of the proposed solution. Primarily a set of requirements for any new solutions for providing QoS over optimal paths are defined. Based on these requirements, a cross-layer design technique is devised. The description of cross-layer design and some of the existing solutions based on this technique are presented. The architectural design of new solution and modified protocol stack are described in this chapter.

Chapter Four

This chapter presents the detailed design of the new solution for providing QoS over optimal routes in MIP. The information flow and QoS parameters are explained. The protocols engaged in the communication session enhanced by new parameters are detailed. The ROQS entity is presented in terms of its different components. Finite state machines and state variables as well as QoS constraint-based algorithms are described in detail.

Chapter Five

This chapter presents the performance evaluation of the proposed solution using simulation. The components of new model implemented in Opnet environment are presented. To evaluate the effectiveness of the solution at different layers (service, core network and access network), relevant statistics are collected. Simulating different network topologies and various scenarios show the impact of the proposed solution in different contexts. By analyzing the measured statistics, we show the improvements made in application response time, as well as access and core network performance.

Chapter Six

In this chapter as an enhancement to the main proposal, the triangle routing in MIP is viewed as a "traffic engineering" problem and the integration of MIP-ROQS and Multi Protocol Label Switching (MPLS) is proposed for labeling the optimal path chosen by MIP-ROQS. The purpose of this integration is to address the scalability concern of MIP and RSVP. Using this integration MIP takes care of supporting mobility, ROQS mechanism selects the optimal route, and MPLS labeling mechanism labels the optimal path for resource reservation during the session.

Chapter Seven

This chapter recapitulates this research and with a conclusion examines the goals and contributions of this thesis. Furthermore, possible improvements and other issues that are still present in this context are enumerated.

CHAPTER 2

BACKGROUND AND RELATED WORKS

2.1 Introduction

Rapid growth of "wireless equipment" as well as "IP services" makes the convergence of the two very attractive. The increasing demand for mobility support has led to considerable development of wireless communications. Any solution for dealing with issues of IP environment should take into account mobility concern, in order to be applicable to MIP environment.

Mobility support deals with many different issues such as mobility management, resource allocation, call admission control, security and billing. A user roaming across the two domains is usually based on the mobility management protocols of the cellular networks. All these issues have been the subject of significant research efforts between which we enumerate few here.

Study in [8] emphasizes on flexibility while providing mobility support and [10] is based on the well known cellular system for mobility and works toward providing IP access.

Discussion of [16] is about a domain-based approach for mobility support called Hawaii. Study in [19] is a comparison between MIPv4 and MIPv6. Proposal in [20] adds a hierarchical infrastructure for MIPv6 to reduce traffic congestion for MIPv6. Authors of [27] are proposing integration of MIP and cellular IP.

Study in [46] is a comparison of micro-mobility protocols. Proposals in [47, 48] are focusing on IPv6 and MIPv6. Proposals in [55, 59] are trying to reduce the amount of binding update for MIP. Study in [67] is a mailbox-based scheme for improving MIP performance. Method introduced in [68] is addressing simultaneous host mobility for MIP.

Scheme in [72] is a dynamic hierarchical temporary home agent scheme in Mobile IP. Authors of [73] are proposing location management and multimedia communication service based on Mobile IP and Cellular IP network. Proposal in [85] is a handoff mechanism in micro-domain. Proposal in [90] is mobility support in IPv6 and [92] is an adaptive hierarchical Mobile IPv6 using mobility profile.

As we can see mobility support has been studied from many aspects and there are many approaches available for each corner of the problem. With that consideration as the base of the research we assume the mobility support is already in place. To support mobility we chose the well known Mobile IP protocol developed by IETF [61] and explore it below.

2.2 Background

2.2.1 Mobile Internet Protocol

The Internet protocol requires all nodes connected to one network to share the same network prefix portion of their IP addresses. If a node moves from one network to another, the network prefix of its IP address would no longer be equal to the network prefix assigned to its current network.

Changing a node's IP address as it moves makes it impossible for the node to maintain any ongoing communications. Mobile IP [61] developed by IETF is a network layer routing protocol that specifies how hosts can move across IP subnets and retain the same IP address. It does this primarily by arranging for packets to be forwarded via HA, (the MN's routing proxy) to the MN when it has moved to a foreign network.

The key feature of the MIP design is that all required functionalities for managing mobility information are embedded in well defined entities, the HA, Foreign Agent (FA) and MN. The MIP protocol allows the MNs to retain their IP addresses regardless of their point of attachment to the network.

Allowing the MN to use two IP addresses is solving the problem of mobility. The first one is the home address, which is static and is used to identify higher layer connections, for example TCP and the second one is the Care of Address (CoA), which has to identify the MN's new point of attachment with respect to the network topology. FA manages the CoA by including it in agent advertisement message before sending it out.

MIP functionality is realised using three main mechanisms:

1) Discovering the CoA
2) Registering the CoA
3) Tunnelling to the CoA

MIP Agent Discovery: A MN discovers the FA and HA during agent discovery. HA and FA advertise their services using Internet Control Message Protocol (ICMP) messages and extensions to Internet Router Discovery Protocols (IRDP). MNs listen to the advertisements and determine if they are in their home or in a foreign network.

Advertisement includes:

- Status as HA or FA
- CoA
- Supported MIP features
- Registration lifetime

MN may send a solicitation message to cause advertisement. When the MN detects FA, it performs registration. Agent Advertisement (AA) messages consist of an ICMP router advertisement with an AA extension that identifies the node as a mobility agent. The mobility AA extension follows the ICMP router advertisement fields. It is used to indicate that an ICMP router advertisement message is also an AA being sent by a mobility agent. The Mobility Agent Advertisement Extension is defined as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |        Sequence Number        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Registration Lifetime      |R|B|H|F|M|G|r|T|   reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 zero or more Care-of Addresses                |
|                              ...                              |
```

MIP Registration: The MN registers its current location with the FA and HA during registration. MN constructs a registration message using data from Agent Discovery and previously provisioned configuration:

- Provisioned configuration
  - MN's home address
  - Address and security association of HA
- Learned from Agent Discovery
  - Care-of Address
  - Encapsulation mechanism
  - Lifetime
- MN sends registration to HA either directly (if using CoA) or via the FA
- Registration is always for a limited lifetime and always authenticated

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |S|B|D|M|G|r|T|x|          Lifetime             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Home Address                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Home Agent                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Care-of Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                         Identification                        +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Extensions ...
+-+-+-+-+-+-+-+-
```

In the registration reply the UDP header is followed by the Mobile IP fields shown below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Code      |           Lifetime            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Home Address                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Home Agent                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+                         Identification                        +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Extensions ...
+-+-+-+-+-+-+-+-
```

Considering Figure 1, MIP works in the following sequence of events:

1) HAs and FAs advertise their presence on any attached networks by periodically broadcasting AA.

2) MNs listen to these AA and examine their contents to determine whether they are connected to their home network or a foreign network. While connected to their home network, a MN acts just like a stationary node and

uses no other MIP functionality. The rest of the steps which follow therefore assume that a MN has discovered that it is connected to a foreign network.

3) A MN connected to a foreign network acquires a temporary address called CoA. A FA CoA can be read from one of the fields within the FA's AA.

4) The MN registers its CoA acquired in step three with the HA, using a message exchange defined by MIP. In this registration procedure, the MN asks for service from a FA. The registration messages are required to be authenticated in order to prevent the access by irrelevant users.

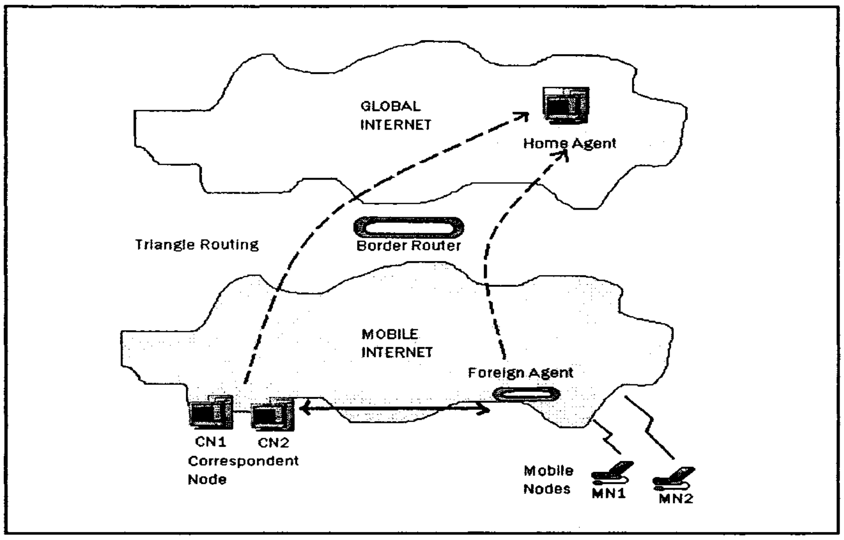

Figure 1: MIP Events Sequence

5) The HA or some other routers on the home network advertise their reachability to the network prefix of the MN's home address. When a CN wants to send packets to a MN, it sends them to the HA. The HA intercepts

the packets and tunnels them to the CoA that the MN has registered in step four.

6) At the CoA (at either the FA or one of the interfaces of the MN) the original packet is extracted from the tunnel and is delivered to the MN.
7) In the reverse direction, packets sent by the MN are routed directly to their destination, without any need for tunnelling. The FA serves as a router for all packets generated by a visiting MN.

The steps above described the movement of the MNs in the Internet based on the MIP protocol. However, the efficiency and quality of the connections depends on the routing, amount of signalling on the HA and the mechanism which guarantees the QoS.

MIPv6 [90] enables IPv6 nodes to cache the binding of a MN's home address with its temporary address, and to send any packets destined for the MN directly to it at its temporary address. Hierarchical Mobile IPv6 (HMIPv6) [20] is based on MIPv6 with the purpose of reducing registration control signalling. It separates micro and macro mobility with the help of an intermediate mobility agent, called the Mobility Anchor Point (MAP).

However, HMIPv6 still suffers from congestion in the MAP. The method in [92] enhances HMIPv6 by temporal locality. This is done using a new profile for management of the locality information. Based on the information in the profile, some packets are directly delivered to MN, if the MN seems to be residing for a long time in its current subnet.

To support the mobility, MIPv6 [90] and consequently HMIPv6 [20] and HMIPv6 with profile management [92] place some special requirements on the IPv6 nodes such as ability to maintain a binding update list and performing encapsulation and de-capsulation.

2.2.2 MIP and Routing

To send a packet to another mobile host, the sender constructs a source route in the packet's header giving the address of each host in the network through which the packet should be forwarded in order to reach the destination host. The sender then transmits the packet over its wireless network interface to the first hop identified in the source route.

When a host receives the packet, if it is not the final destination of the packet, it simply transmits it to the next hop identified in the source route in the packet's header. Once the packet reaches its final destination, the packet is delivered to the network layer software on the host. Each mobility host participating in the network maintains a route cache in which it caches source routes that is has learned.

When one host sends a packet to another host, the sender first checks its route cache for a source route to the destination. If a route is found, the sender uses this route to transmit the packet. If no route is found, the sender may attempt to discover one using the route discovery protocol. While waiting for the route discovery to complete, the host may continue normal function and may send and receive packets with other hosts.

The host may buffer the original packet in order to transmit it once the route is learned from route discovery, or it may discard the packet, relying on higher-layer protocol software to retransmit the packet if needed. Each entry in the route cache has associated with it an expiration period, after which the entry is deleted from the cache.

While a host is using any source route, it monitors the continued correct operation of that route. For example, if the sender, the destination or any of the other hosts named as hops along the route move out of wireless transmission range of the next or previous hop along the route, the route can no longer be used to reach the destination.

This monitoring of the correct operation of a route in use is called route maintenance. When route maintenance detects a problem with a route in use, route discovery may be used again to discover a new, correct route to the destination.

Using MIP, when a mobile is in a different subnet all data to the mobile is intercepted by the HA. The packets are then encapsulated and sent to the FA residing in the subnet the MN is currently located. The advantage of this method is that all senders need to know only the home address of MN. The HA keeps track of the current location of the MN. The path the packets take from CN to reach the MN is via the HA while the MN can send the packets directly to CN.

This is known as “triangle routing” and is shown in Figure 2. It means packets from the CN to the MN must in general travel via three sub networks: the CN’s subnet, the HA’s subnet, and the subnet where the MN is currently located.

Triangle routing adds potentially significant overhead, delay and network resources consumed for communication with MN and therefore is considered a main routing issue in MIP.

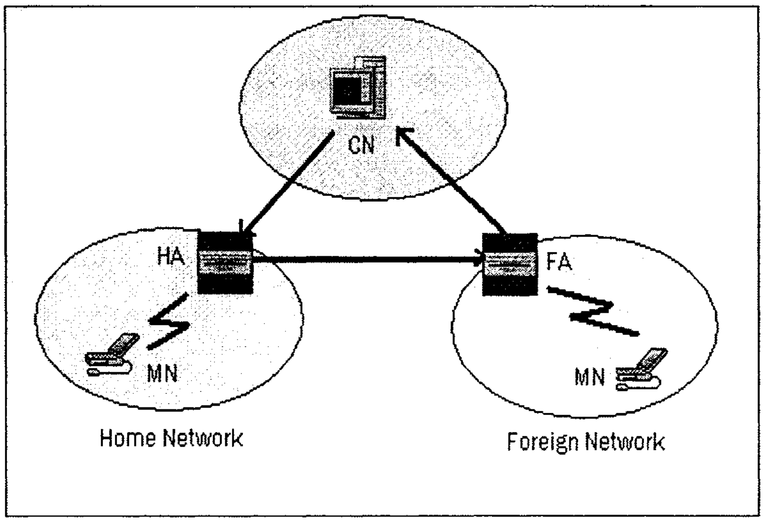

Figure 2: Triangle Routing

More importantly, even if the distance between the MN and CN is very small, still the packets have to go to HA and come back to FA to reach the MN. This unnecessary long travel for packets is a waste of network resources. There has been research specifically on resolving the triangle routing problem and some solutions already exist. Some of the main solutions are described below.

2.2.3 MIP and RO

To reduce the routing overhead for MIP, as shown in Figure 3, a RO approach was proposed by [22] to directly tunnel packets from a sending node to a FA where the receiving MN is visiting. Following this approach, a CN may learn the binding of MN's home address and corresponding CoA from the binding update message sent by the HA. CN can store a binding of home address and CoA of the MNs to optimize its own communication with the MNs.

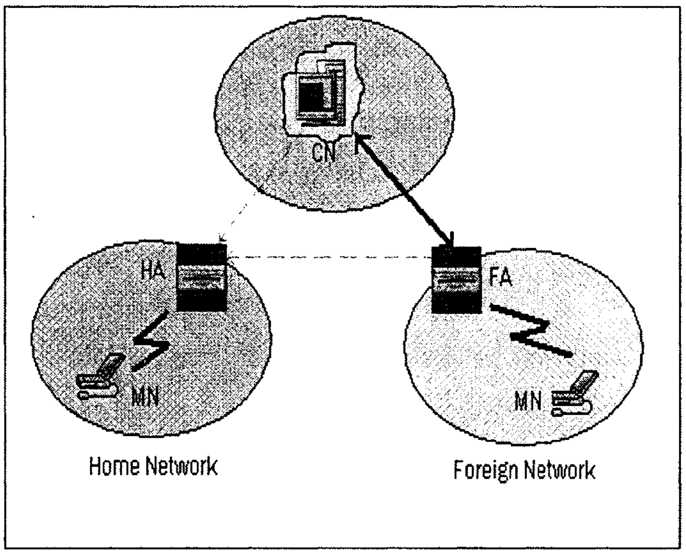

Figure 3: Optimized Routing

Although this would improve the performance of packet routing for MIP, it implies that the CN be enhanced to support RO.

To obtain the transparency for the CN, an agent based RO approach was proposed by [29] to move the tasks of maintaining binding caches and encapsulating messages away from individual CNs. In addition, as shown in Figure 4, a tunnel set up by one correspondent agent could serve multiple CNs simultaneously.

This would reduce the total number of control messages to set up individual tunnels for all CNs nearby the correspondent agent. Although the above RO techniques were improving the routing efficiency for MIP to a certain extent, they focus on the routing direction from the CN to the MN and they were not working well in most networks that follow strict and secure routing policies.

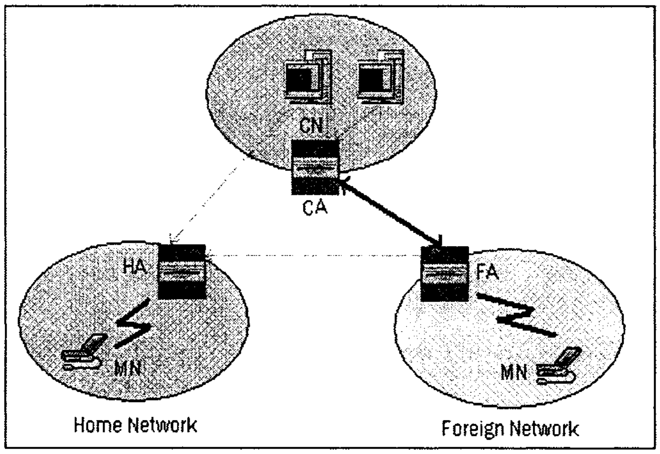

Figure 4: Agent-based Optimized Routing

For example, to prevent the possible attackers that use IP source address, many routers are filtering on the source address and will drop a packet whose address is not topologically correct. This is also called network ingress filtering in which the routers disallow the MN to send packets directly from the visited foreign network to CNs using its home address as the source address.

To tackle this ingress filtering constraint, the reverse tunnelling scheme was proposed by [43] to establish a topologically correct reverse tunnel from the CoA of the MN to the HA. The packets sent from the MN are reverse tunnelled to HA and then forwarded to the destined CN after de-tunnelling. This scheme results in the triangle routing problem in the reverse direction as shown in Figure 5 and makes the current RO techniques to become unidirectional or asymmetric.

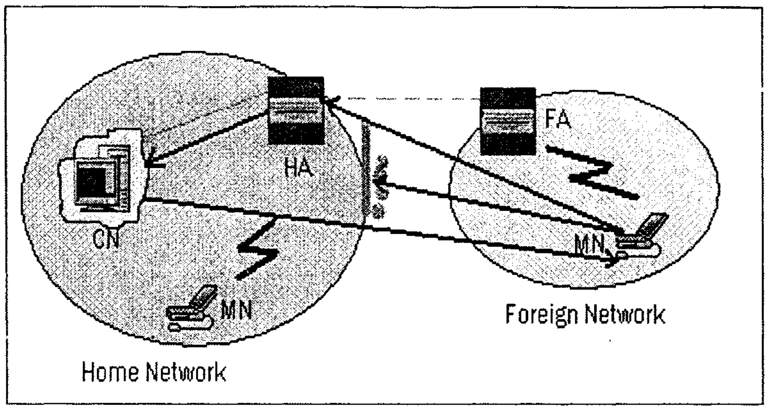

Figure 5: Triangle Routing in the Reverse Direction

To resolve the reverse triangle routing, [50] proposed supporting direct routing in CNs and direct tunnelling in MNs, in order to provide bi-directional RO. In this proposal MN uses its CoA as the source address in the IP header of packets it sends.

It may also directly send new binding update messages to certain CNs to reduce handoff latency. To achieve this, the techniques for MIPv6 draft was adapted beside a new subnet based tunnelling and RO techniques. This method is visualized in Figure 6.

Study in [48] is a survey on multicast routing protocols and [49] is focusing on multicast routing for mobile hosts. RFC in [4] is an IETF effort on Tunnelling for MIP. Approaches in [58], [71] and [77] are specifically addressing the RO for MIP. On the above proposals, the question still remains valid, how to provide QoS while optimizing the routing.

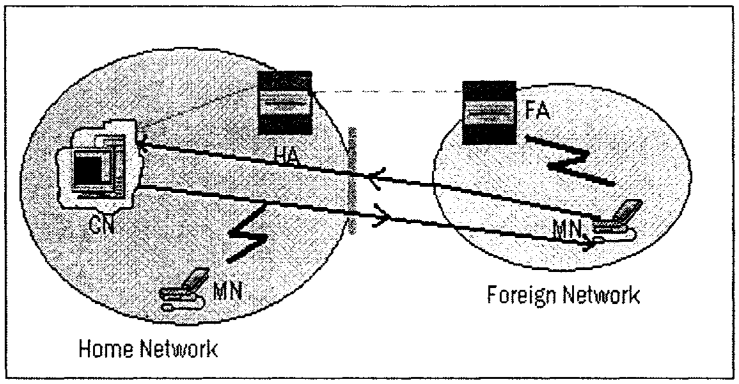

Figure 6: Bi-directional Routing

2.2.4 MIP and QoS Concern

Internet QoS paradigms (IntServ and DiffServ) were supposed to compensate for IP's lack of service differentiation and performance guarantees. MIP in addition to inheriting this problem further imposed new challenges associated with mobility.

When a MN moves from one location to another with an open connection, the data flow path changes. As a result, the packet propagation time and the congestion delay along the new path may be different or the new cell may be so congested that the minimum QoS requirements of all services cannot be satisfied. Due to handoff, additional delays and packet loss may occur.

Since the route from CN to MN is via HA and the route from MN to CN is a direct path, is a major issue for reserving the resources for the session between CN and MN.

Several problems occur when DiffServ is used in conjunction with MIP. These problems can be classified as follows:

- Network provisioning in mobile environment
- Lack of dynamic configuration

- Definition and selection of service level agreements

Furthermore, DiffServ method needs a sophisticated signalling protocol as one can not expect that each network a MN visits has already established service level agreements that are sufficient to support the MN's requirements in general. Such a signalling protocol must support service negotiation among ISPs and among the sub networks and their ISPs. In addition some kind of signalling protocol among the MN and routers must be supported.

On the other hand when IntServ is used in conjunction with MIP, some other issues arise:

- If a CN sends a PATH message to a MN, the MN would make reservations on the direct route from the MN to the CN, even though this route will not be taken when the CN sends further packets to the MN.
- PATH message gets hidden in the tunnel from HA to MN and as a result can not serve its original purpose.

2.2.5 MIP and QoS Methods

One way to provide QoS guaranteed for mobile users is to make advanced resource reservations in all cells where the MN may visit during a session. However, this is a waste of the limited wireless resources.

The proposal in [69] introduces the integration of IP mobility and QoS for wireless access. This is done by keeping track of the active flow. In [70] MIP is designed based on a Terminal Independent MIP (TIMIP) to make the QoS regardless of IntServ or DiffServ mechanism used in the network.

Solution in [96] is based on the domain resource manager concept and introduces mobility-aware QoS signalling in IP network. To assure QoS while MN is moving, it uses advance resource reservation mechanism as well. All the approaches for providing QoS in MIP environment ignore RO impacts.

2.3 Conflict in Providing Simultaneous QoS and RO

A problem of RO is that packets sent by the CN still use the triangle route until the CN receives the binding update message. An important implication arises when using protocols such as RSVP for providing QoS guarantees for communication between CN and MN.

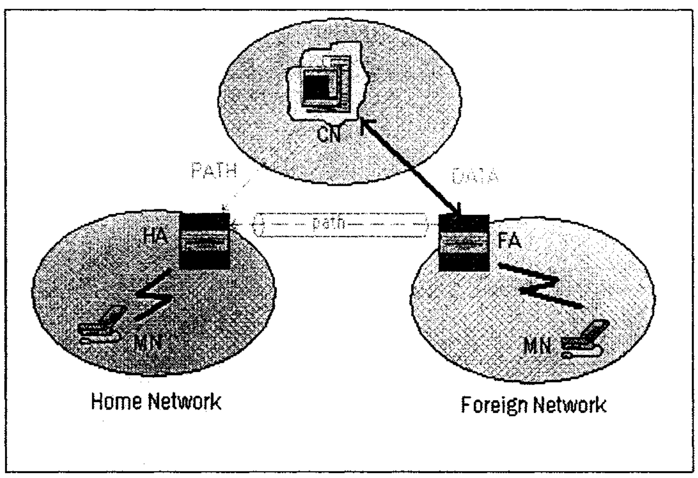

Figure 7: RSVP and RO

As mentioned previously the RSVP's PATH and RESV messages contain a description of the flow for which resource reservations are being requested. This flow description contains a list of packet header fields that a router can use to distinguish packets for which the requested QoS must be provided from all other packets that might pass through the router.

The following problems arise when RSVP is used in conjunction with MIP or MIP with RO:

1) Routers will not be able to recognize a PATH message encapsulated while tunnelled from HA to the MN as shown in Figure 7, and thus will not record the information required for reservations to be involved in resource reservation.

2) Even if the first issue is resolved, the resources will only be reserved along the triangle route from the CN to the MN. Since RSVP protocol issues PATH messages periodically (in order to overcome the effect of routing changes that may take place in a fixed network), eventually resources will be reserved along the direct route to the MN, but unnecessary delay and resource consumption will still result, and the desired QoS guarantees may not be achieved since packets sent along the triangle route receive different treatment than those sent directly.

3) Another challenge to real time traffic introduced by MIP is the fact that MNs change their locations. Because RSVP reserves resources only along a specific path, the fact that the path from sender to MN can change relatively frequently implies that new reservations will be required every time a MN changes network.

If RO is used, the MN must reserve resources from its CoA back to the original sender. Whether this can be considered an improvement or not, depends on the relative locations of the sender, the HA, and the MN's current network.

In some cases there might be rather large opportunities for reducing the number of nodes from which resources must be reserved as the MN moves. A complication introduced by MIP is the presence of the tunnel itself. Some of the methods developed to overcome the above mentioned problem are presented here.

2.4 Survey and Evaluation of Related Work

The main QoS approaches can be classified based on their service in the MIP network. Mainly they can be categorized in two groups. The first group is based on the combination of MIP and IntServ architecture and the second group is based on MIP and DiffServ architecture.

Authors of [14], [17], [24], [25], [31], [63] and [65], are focusing on providing QoS for MIP environment using IntServ mechanism while [15] and [34] specifically trying to provide QoS for MIP using DiffServ mechanism.

All the above studies have addressed QoS issue for MIP and have left out the impacts of routing and RO while providing QoS.

Similar to this research [7], [18], and [87] are trying to provide QoS and optimized routes at the same time.

2.4.1 Classification of Different Approaches

The approaches used in existing solution for aforementioned problem can be classified as follows:

- Modifying RSVP to support Mobility
- Modifying MIP to avoid tunnelling
- Creating a New Mobility Framework to Facilitate QoS Mechanism

For each group some solutions are explained below in more details.

2.4.1.1 Modifying RSVP to Support Mobility

"History based RSVP setup" [78] (HBRSVP-78) uses moving history of the MN to provide QoS support. It integrates RSVP with MIP regional registration, and makes advance resource reservations. When MN is about to move from one cell to another the registration request message is transmitted to a gateway. On receiving the registration message, the gateway activates the RSVP path between the gateway and MN in the new cell by using the passively reserved RSVP path, if the registration message is transmitted by this cell.

"Mobile RSVP" [30] (MRSVP-30) assumes a hierarchical architecture where each subnet in a domain has a proxy agent. When a MN reaches a foreign domain, it uses the proxy discovery protocol to find the proxy agents of the subnets in this domain. Then, the MN will initiate a reservation protocol to the domain's subnets in advance, so when it is in a certain subnet, the reservation is already made.

"Hierarchical Mobile RSVP" [66] (HMRSVP-66) integrates RSVP with a hierarchical MIP using regional registration. To reduce advance reservation in (MRSVP-27) the reservation is made only for long duration handoffs. When a MN visits a foreign network, during the initial reservation signaling with the CN, the

Gateway Foreign Agent (GFA) initiates a tunnel to the MN's local FA. When the MN moves to an area served by a different FA under the same GFA, only a tunnel between the GFA and the new FA needs to be established.

The "Pointer Forwarding Chain" [81] (POFOCH-81) analyzes the effect of applying different advance resource reservation mechanisms within HMRSVP in terms of localization of passive reservation messages. Three options are considered: making passive reservations in all branches; making passive reservations in the two neighboring FAs; and, leaving a forwarding pointer at the old FA prior to handoff, eventually creating a pointer forwarding chain with that. It concludes the third option is the best.

The "RSVP Mobility Proxy" [56] (RSVPMP-56) based on hierarchical MIP uses two CoAs, the global CoA (GCoA) applicable in a domain and the local CoA (LCoA). An RSVP enhanced edge router with mobility support keeps track of the LCoA/GCoA bindings in addition to their reservations. Exchange of PATH/RESV messages between the RSVP-MP and the CN is done using the GCoA, while internal exchange is done using the LCoA.

"Aggregation of RSVP for IPv4 and IPv6 reservations" [39] (AGRSVP-39) modifies RSVP at the correspondent and foreign networks such that they acknowledge the MIPv6 addressing scheme. To avoid re-reservations whenever the MN moves to a different foreign network, it uses a new RSVP object. By including the MN's home address, this object informs the intermediate routers that this is the same MN with a different CoA, and thus maintains the previously reserved states.

2.4.1.2 Modifying MIP to Avoid Tunnelling

With the goal of avoiding tunnelling "MIP with location register" [5] (MIPLR-5) as shown in Figure 8 uses a database called location registers (LR) to maintain the location of the MN. When a MN moves, it registers its location with LR. A CN must query the LR to obtain the address of the MN, and use this address to send packets directly to the MN; the CN caches the MN's address to avoid querying the LR for every subsequent packet.

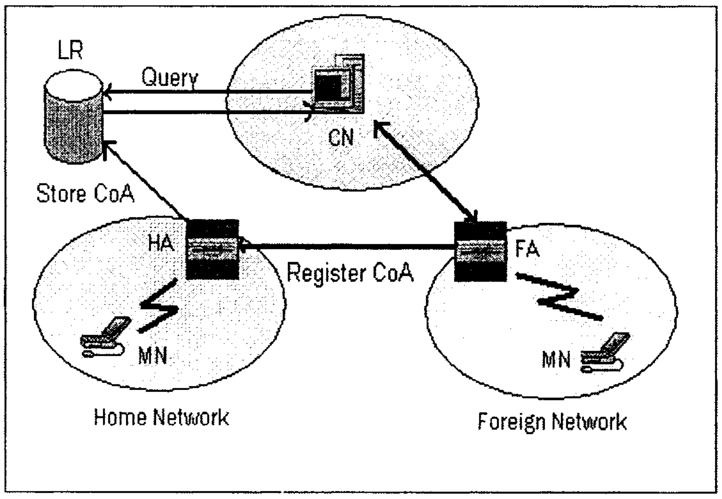

Figure 8: MIP with Location Register

"A simple QoS signalling protocol" [11] (QOSSP-11) as shown in Figure 9 proposes a separate RSVP session between the tunnel end points. The tunnel session can exist independently from end to end sessions that are created via a management interface. When an end to end RSVP session crosses an RSVP capable tunnel it is mapped to a tunnel RSVP session. Then a reservation is made from tunnel exit point to tunnel entry point for the data crossing the tunnel.

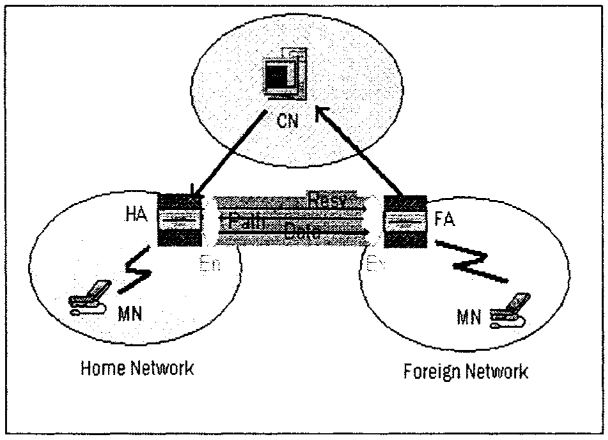

Figure 9: Resource Reservation inside the Tunnel

"Hierarchical Mobile IPv6 Using Mobility Profile" [6] (HMIPV6-6) introduces a profile for management of locality information. According to the information in the profile, some packets are directly delivered to the MN if the profile indicates the MN is standing at one place for a long time.

2.4.1.3 Creating a New Mobility Framework to Facilitate QoS Mechanism

"Domain based QoS" [9] (DBQOS-9) adapts hierarchical architecture of node mobility for organizing the QoS. Each gateway router in one domain loads HA/FA functionality. At least one RSVP session is created between the gateway routers, which play the role of HA/FA. Each RSVP session between domains reserves more bandwidth than required to be used when MN moves at the macro level. At micro level QoS, this method adapts the DiffServ mechanism for which the FA performs the functionality of bandwidth broker. It also can use RSVP at the micro level.

"Hierarchical QoS guarantee and routing" [28] (HQOSG-28) proposes a hierarchical approach for both QoS and routing. It includes a hierarchical architecture for guaranteeing QoS based on the concept of QoS region and a hierarchical architecture for fast routing based on routing region using advantages of route table changes.

"Penalty-based adaptable reservation admission" [82] (PARA-82) is an approach for QoS differentiation at network layer and MAC layer. It enables resource reservation, allocation, and admission control to be adaptive to traffic conditions and to the stress conditions at individual nodes. In PARA, differentiated service provisioning is under the control of network operators or individual nodes.

2.4.2 Classification of Different Mechanisms

Another classification can be done based on the mechanism used in each solution regardless of the main approach. The mechanisms are as follows:

- Using storage for future reference
- Using hierarchical network architecture
- Using advance resource reservation

We re-group the aforementioned existing solutions in these categories.

2.4.2.1 Using Storage for Future Reference

[MIPLR-5], [HBRSVP-78], [HMIPV6-6] rely on the storage of certain MN information in order to proceed with routing or resource reservation. However maintaining any storage/database and keeping it up to date, which contains address, moving history or certain profile implies processing overhead.

2.4.2.2 Using Hierarchical Network Architecture

Approaches in [DBQOS-9], [HQOSG-28], [MRSVP-30], [HMRSVP-66], [POFOCH-81], and [RSVPMP-56] need hierarchical architecture to facilitate the routing and QoS mechanisms. However, one can not expect the existence of such structure in the actual networks.

2.4.2.3 Using Advance Resource Reservation

[DBQOS-9], [HBRSVP-78], [HMIPV6-6], [MRSVP-30], [HMRSVP-66], and [POFOCH-81] use different degree of advance resource reservation. Advance resource reservation however, is considered to be waste of network resources and implies extra signalling in some part of the network while the MN may not even move to these places during an active session.

2.4.3 Evaluation of Existing Techniques

To compare the existing methods applicability or coverage in the network, routing method, QoS method, assumption, overhead and advantages of each method are considered and a summary is presented in Table 1. The existing methods have different applicability and sometimes use specific routing or QoS methods. However, the approaches and mechanisms are quite limited. They are categorized mainly in three different approaches:

1) Modifying RSVP to support mobility
2) Modifying MIP to avoid tunnelling
3) Creating a new framework for QoS and mobility

Regardless of these approaches, there are certain mechanisms that are often used:

1) Using storage to keep a profile of MN to be used later for routing or resource reservation.
2) Introducing a hierarchical network architecture which facilitates development and analysis of any new approach.
3) Advance resource reservation to some degree which although seems to be inevitable for supporting smooth handoff, it should be limited as it has negative impact on the usage of network resources and signalling overhead.

Beside the advantages that each method has, each has some drawbacks and some assumptions. Considering the multi operator context of MIP, assuming the availability of specific conditions in which the method is functional is far from reality.

Most of the solutions have added signalling overhead instead of reducing it. Modifying RSVP signalling to support mobility and implement MIP as originally described by IETF, may be efficient in terms of isolating the amount of necessary changes. Implementing a new framework for QoS support and mobility although may seem more sophisticated can encounter implementation issues as it is not feasible to change the existing network infrastructure.

Modifying MIP to avoid tunnelling and support QoS may be the best approach as it can enhance the mobility protocol to a higher degree of performance. It can become more adequate for usage of multimedia services with QoS dependency and more efficient in terms of routing. Therefore, future research in this area must focus on enhancing MIP regardless of IP version.

Except (MIPLR-3) and (AGRSVP-36) all the others have assumed that the tunnelling of packets is inevitable and tried to come up with a solution that works with tunnelling.

The proposals that have integrated MIP and RSVP can be categorised in two groups. The first group relies on advance resource reservation and the second group is performing the resource reservation at the time it is requested. The existing proposals also can be categorised based on their applicability, some of the proposals specifically focus on the access network and some specifically deal with the core network.

Whether the mechanism can take care of itself or operator intervention is required also defines some aspects of existing proposals.

METHOD	Applicability	Routing Method	QoS Method	Assumption	Overhead/ Drawback	Advantage
HBRSVP-78	Core Network	Tunnelling	IntServ	Availability of Moving history	Storing moving history	It is done at setup time
MIPLR-5	Macro Mobility	Bi-Directional	IntServ	CN is Mobility aware No Fire Wall	Extra signalling for querying the HLR	Removes Tunnelling and encapsulatio n
HMIPV6-6	End-to-End	Tunnelling	DiffServ / IntServ		Modification to all the Nodes	Reduces waste of resources
QOSSP-11	Macro Mobility	Tunnelling	Intserv	CN is Mobility aware No Fire Wall	Extra RSVP session in the tunnel	Simplicity, Not much change in existing IP
HQOSG-28	End to End	Tunnelling	IntServ		Modification to architecture	Maintain QoS and provide fast handoff
DBQOS-9	End to End	Tunnelling	IntServ DiffServ	HA and FA should perform as BB	Reserves more bandwidth than necessary	Reduces the size and the update frequency of the route
MRSVP-30	Access Network	Tunnelling	IntServ	MN knows address of all sub-networks	Additional signalling for proxy discovery	Enhanced RSVP
HMRSVP-66	Access Network	Tunnelling	IntServ	MN knows address of all sub-networks	Extra processing time	Hierarchical architecture
POFOCH-81	Access Network	Tunnelling	IntServ	MN knows address of all sub-networks	Advance resource reservation	Reduces handoff delay
RSVPMP-56	Access Network	Tunnelling	IntServ	Specific architecture in place	Special Edge Router functionality	Hierarchical architecture
AGRSVP-39	Access Network	Bi-Directional	IntServ	Presence of RSVP objects and triggers	Advance resource reservation	Prevents multiple reservations
PARA-82	Core Network	----	DiffServ	Stable traffic conditions	Operator supervision	Considers the actual status

Table 1: Evaluation of Existing Methods

To reduce processing time, building a hierarchical approach is important; however, the proposal should be still valid where such architecture is not available. Another observation is that none of the hierarchical solutions implements a passive reservation mechanism. This could be the result of distributed resource management.

Finally, it should be noted that considering the MIP environment, the proposal for providing QoS and RO should consider both access and core network. Based on the observations of the existing methods we derived the requirements for the new solution. Next chapter starts by elaborating on the requirements for the new solution.

CHAPTER 3

PROVIDING QOS OVER OPTIMAL ROUTES IN MIP: ARCHITECTURE

3.1 Introduction

Based on the observations from existing solutions we establish a set of requirements prior to developing a new solution. Obviously, the most significant requirement is to resolve the issue described in the previous chapter as the conflict in simultaneously providing QoS and RO. Nonetheless, we need to consider the specific MIP environment as it has a hybrid wired/wireless IP-based architecture. Besides resolving this issue, it is required to achieve higher network performance with lower cost in terms of signaling and processing overhead.

Through the investigations we believe that the key for resolving the conflict is to select an optimal path at the MIP level before attempting to reserve resources on it. This could be seen as QoS constraint-based routing in MIP. QoS routing in IP environment addressed by [7], [87], [91] and [114-120] has a long history in the research community but has not been specifically considered for MIP environment.

To find the optimal path we need to analyze the network conditions against the QoS requirements. This is not possible unless we have access to the information from other layers than the one MIP is acting in.

We argue that to get the necessary information it is required to adopt a cross-layer design technique. Based on this design a cross-layer interface and a new information flow are necessary to collect the QoS parameters from different layers of the MIP protocol stack.

The cross-layer interface should be integrated in the network layer where it has the minimum impact on the signaling and processing overhead. We show an abstract view of cross-layer interface and develop a new version. The cross-layer interface in the new design is more than simple exchange of information but rather a sophisticated entity which enhances the MIP network layer for selecting the optimal path based on the collected QoS parameters.

In this chapter, we present the derived requirements, the new design approach, architectural design of the new solution and its impacts and preliminary evaluation.

3.2 Requirements for QoS over Optimal Routes in MIP

According to the characteristics of the hybrid wired-wireless IP-based environment and the evaluation of existing methods the following requirements are established for any solution attempting to solve the problem of QoS and RO in MIP environment:

1) The solution should resolve the conflict between providing QoS and RO. This is the most important requirement as this conflict prevents resource reservation for services.

2) The solution should provide the optimal route based on the network conditions while guaranteeing the required QoS. The reason is that changing triangle routing to direct routing regardless of the network condition may result in diverting the traffic. However the direct path may not have the necessary resources available.

3) The solution should be an end-to-end solution that includes all the components of MIP environment consisting of service, core and access network. The reason is that ignoring one part of the network in the hybrid wired-wireless MIP environment will be based on the assumption that this

part of the network will function perfectly while in reality this is not always possible.

4) The solution should improve the application end-to-end delay as this is directly impacting the service performance.

5) The solution should improve the performance of the core network by reducing amount of traffic dropped as this is impacting the service performance.

6) The solution should at least not degrade access network performance. The reason is that the available resources in access network are limited.

7) The solution should not increase the signaling overhead. This is because resolving the conflict between QoS and RO should not be achieved by adding a burden on the network performance.

8) The solution should take into account capability of different entities involved in the session in order to be applicable with a wide range of equipment. For example, an adaptive MN can use several modes depending on the available QoS support of each mode and based on the routing method available at the time of communication.

9) The solution should improve the performance of MIP protocol. The reason is that, high processing time and signaling in the MIP protocol directly impacts the resource utilization in the network.

10) The solution should not interfere with the base functionality and only be triggered when the necessary data is available. The reason is that the

solution should at least be performing like the original MIP without the proposed solution.

We are going to assess these requirements after the presentation of the new solution and its performance evaluation in Chapter 5. In the next section we explain that these requirements cannot be satisfied unless a cross-layer design is adopted.

3.3 A Cross-Layer Architecture for MIP-ROQS

On the standard layering architecture, QoS problems are caused by lack of information from transport layer congestion and link utilization. The mobility problems are related to the effects of handover on transport layer connection and QoS signaling. Wireless link problems can be caused by packet corruption and losses that are perceived by TCP as congestion indications, resulting in poor performance.

Based on a cross-layer design, proper information exchange between different layers of the MIP protocol stack is required. This exchange includes the actual status of wireless links from the access network and the characteristic established by Internet Service Provider (ISP) in the core network. The congestion in the network may create extra delay on the routers, which are performing the RO or QoS establishment.

Providing QoS over optimal path becomes very challenging when there is no sufficient information for choosing the optimal path at the network layer where

MIP is acting in. As explained in the previous chapter, this causes degradation of resource utilization in the network. Even when the MIP tunneling is changed to direct routing (for example in MIPv6), its impact on the network performance and service requirements are not considered.

Hybrid wired-wireless MIP networks although relatively new, are a merge of two older technologies. These two technologies being wireless network and Internet are both implemented based on a standard layering architecture. In this traditional layering standard, based on the predefined and strict interface between the layers, the internal information in each layer is hidden from the neighboring layers.

In MIP, a path from the source to the destination consists of wired links in the Internet (core network) and wireless links in the access network. MAC layer in the protocol stack contains the information regarding the links, but hides this information from the network layer. In order for MIP, to choose the optimal path based on the links' (wired and wireless) condition, it has to have access to this information.

Considering IP protocol stack, providing QoS and RO in MIP at network layer, requires information from service layer about QoS requirements and from data link layer (specifically MAC layer) regarding the link conditions and available resources.

To be able to take network condition into account for selecting the optimal path, there is an absolute need for collecting information regarding these conditions. This data can not be collected without a cross-layer design. Below we expand the concept of cross-layer design and introduce its different classifications and present some solutions based on cross-layer design technique.

3.3.1 Cross-Layer Design

So far the standard layering design has provided a simple and fast way of developing new protocols and network architectures. In TCP/IP protocol stack model there are pre-defined interactions among neighboring layers. MIP protocol stack is built on layering architecture of TCP/IP model with a difference in data link and physical layers. In MIP architecture, both wired and wireless data link and physical layers are present.

Using the standard layering architecture, although it simplifies the development of network and protocols, limits how much one can improve the performance of the wired-wireless IP-based services. The reason is lack of information in one layer where it is needed while this information is available in other layers.

In the framework of reference layered architecture, the designer has two choices at the time of protocol design. Protocols can be designed by respecting the rules of the reference architecture or violating the reference architecture for example by allowing direct communication between protocols at nonadjacent layers or sharing variables between layers [106].

The whole idea behind cross-layer design is to combine the knowledge available in different parts of the network and create an environment, which can be highly efficient. This means sharing state information between different modules in the system. Breaking the boundary to some extent between the different layers of standard model leads to a cross-Layer design [106]. This will make information sharing possible. This information can serve decision-making when it is necessary for achieving a higher network performance.

3.3.1.1 Classification of Cross-layer Design Techniques

Layered architecture can be violated in the following ways [106]:

- Creation of a new interface
- Merging of adjacent layers
- Design coupling without any new interface
- Vertical calibration across layers

When creating a new interface in order to share information between layers, a new information flow needs to be defined. Possible information flows with a new interface are as follows:

- Upward: From lower layers to a higher layer
- Downward: From higher layer to a lower layer
- Back and forth: Iterative between two layers

To fit the new design in the above classification of cross-layer designs, we follow "creation of a new interface" in protocol architecture and create upward and downward information flow in order to collect necessary data for analysis and selection of optimal path.

However, the new design is a variation of the existing model for "creation of a new interface" described in [106]. In the new design, the interface is put inside the network layer and is more than simple collection of information. This interface is rather an enhancement to network layer as it includes optimal path selection mechanism integrated in the same entity.

3.3.1.2 Some Solutions Based on Cross-layer Design

The design in [95] is an analytical framework for optimal soft-handoff in cellular CDMA system, involving the link layer and network layer. Using the network layer parameter call-blocking-probability, the admission region in Link Layer is calculated.

Proposal in [103] has used cross-layer design in Directional Routing Protocol (DRP) in Mobile Ad-hoc Networks (MANET). In this proposal, cross-layer is between routing and MAC layer. This allows a route discovery mechanism in routing layer to use the information of MAC layer in order to establish and maintain a directional routing and directional neighbour tables.

The proposal in [104] uses information sharing between Service, Physical and Link layer to improve resource allocation in CDMA channels.

However, as emphasized by [109], a cross-layer design should be exercised with appropriate caution. It should not prohibit further development of the system and it should not need complete re-design of the system following any new update.

Authors of [93], [95], [98], [103-104], and [107-110], have developed different solutions based on cross-layer designs. They all show that going beyond certain level of performance in the network without breaking the walls between different layers of standard layering architecture is not possible.

3.3.2 MIP Architecture and Protocol Stack for ROQS

In this section, the current MIP architecture and the way we modify it based on a cross-layer design are presented. MIP architecture in practical terms consists of three major components called services, core and access networks. Each component has different characteristics and provides different functionalities to support end-user services.

The importance of considering practical components of MIP network is that each component is built by a different service provider. These providers as it is today do not share the detailed characteristics and capabilities of their networks with each other. At the same time from user point of view all these components are seen as one entity for providing QoS-sensitive applications.

Figure 10 shows the MIP network components. Brief description of these components and their interactions are as follows:

1) **Access Network**: This is wireless telecommunication system which provides the access from and to the mobile users. An example of this layer is the well-known cellular network.
2) **Core Network:** Wide area network consisting of many networks that each runs TCP/IP protocols. The networks are connected via routers and the transmission of packets through the networks is guided by routing

algorithms. To provide mobility in this level MIP and ICMP protocols are used.

3) **Service** refers to any service that can run on MIP and is of interest for the end user. Many of these services are time, security, or QoS sensitive. The example of such a service is video conferencing and on-line banking.

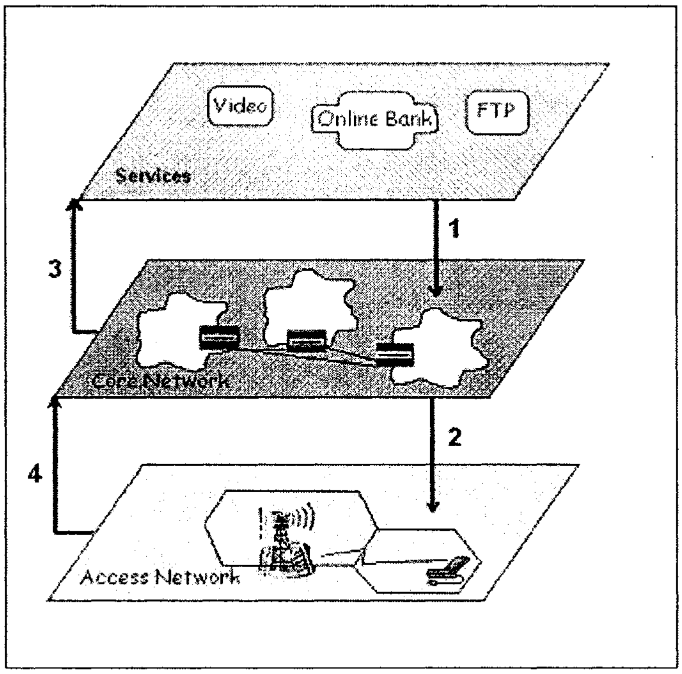

Figure 10: MIP Network Components

Each component interacts with the neighboring components respecting strict boundaries. The simple interactions between different components of MIP architecture are as follows:

1) The services provide a set of parameters that show their concern and sensitivity to certain aspects of quality.

2) The core network based on the implemented protocols provides the transmission for a packet data call. So it provides the means of routing and forwarding the packets from one sub-network to another.

3) The access network provides the transmission of control and data packets to the receiver, which in this case is the mobile node. This part of the network has different nature than core network and therefore has its own limitations.

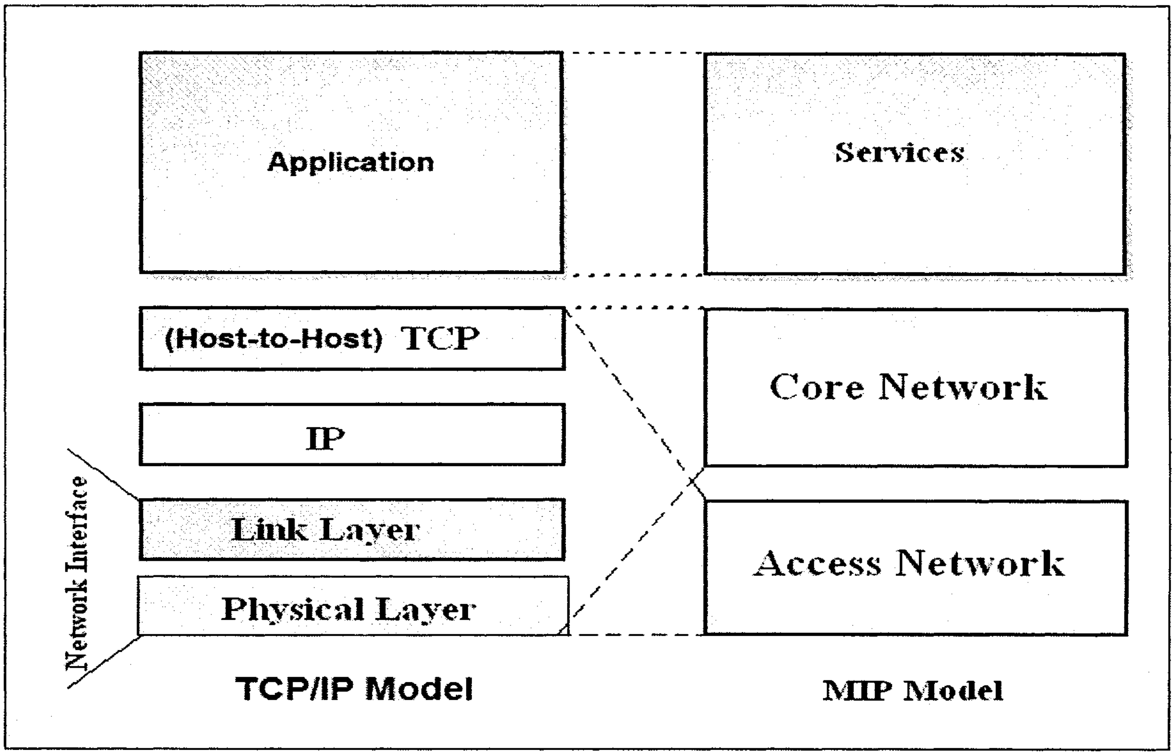

Figure 11: Mapping Network Architecture to Protocol Stack

Core network and access network in the above architecture are built based on the TCP/IP protocol stack model. The core network referred to as Internet in MIP context is actually a TCP/IP protocol stack running on wired media. MAC layer in the core network could be of type IEEE 802.3. The access network on the other hand is a wireless network, either WLAN, WCDMA, which has a TCP/IP protocol stack running on a wireless MAC layer for example of type IEEE 802.11

Figure 11 shows the mapping of MIP network components to the TCP/IP protocol stack model which indeed can be mapped to the standard OSI 7-layer protocol stack.

The MIP network always has two different MAC devices providing wired and wireless link setup. This implies not only collecting information from MAC devices in wired network but also MAC devices in wireless part of the MIP network.

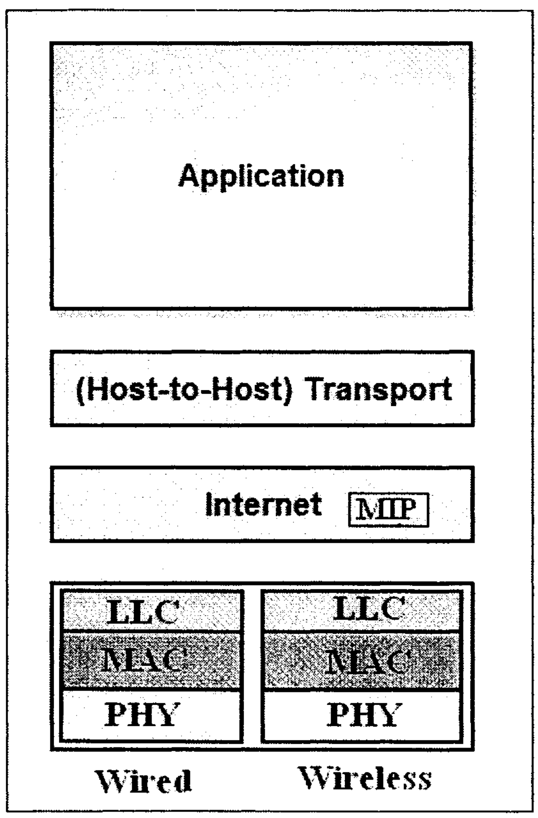

Figure 12: Expanded MIP Protocol Stack

In fact the MIP protocol stack can be expanded as shown in Figure 12. In each stack the information flows both ways between adjacent layers in a strict way. The internal information about link conditions is not shared with the network layer.

Furthermore, information between two MAC protocols is not shared with each other, as one represents the wired link and the other is responsible for accessing wireless media. Information from different MAC layers is to be shared with network layer.

The details of QoS requirements need to be transferred from application layer to network layer. The modification of the protocol stack is divided into two parts:

- Collecting data using a new information flow
- Analysing the collected data using a new entity called ROQS in the network layer

These two parts are described below.

3.3.2.1 Cross-layer Information Flow

In an abstract level collection of the QoS parameters from all the layers of the TCP/IP protocol stack needs a new interface as shown in Figure 13. Information from application, transport, network and different link layers are collected using a cross-layer interface. Any entity in the network layer can access this data for further analysis.

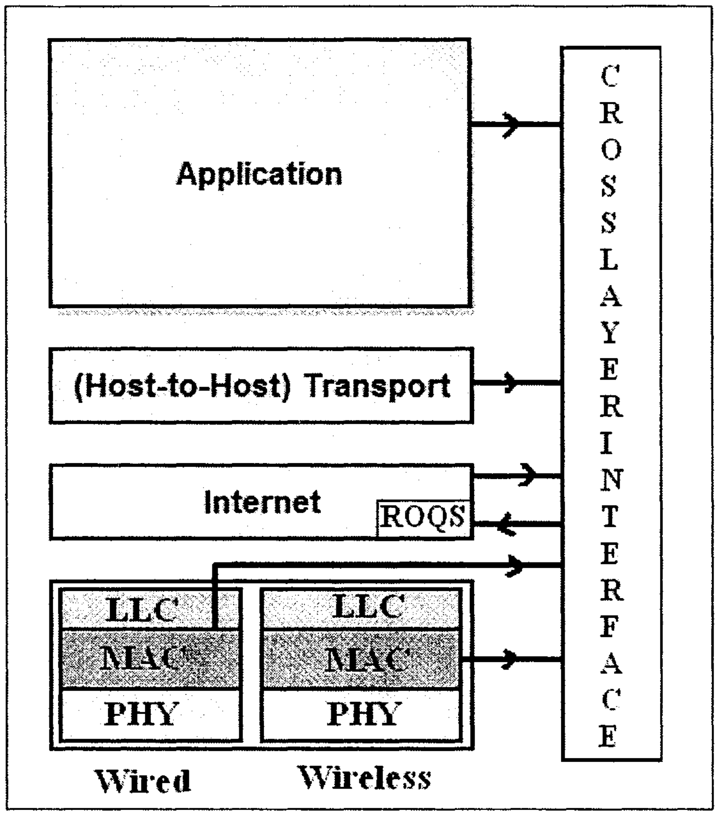

Figure 13: Cross-Layer Interface

To realize this concept, one may consider an actual storage where all the layers can store the detailed information and any layer can access this data for further analysis. This approach, although simple introduces extra signaling and function calls. While respecting the same concept, in order to avoid introducing new signaling (to satisfy requirement numbered 7) in the new realization of the above concept, the information flow uses the modified version of the existing signaling of the protocols involved in mobility, routing and resource reservation. These signals bring the QoS parameters. Figure 14 shows the QoS information flow in the new realization of cross-layer design technique.

In the new architecture, two types of signaling interfaces can be distinguished:

- Signaling within a domain
- Signaling between domains

The signaling within a domain is used for the new purpose. To implement the MIP-ROQS solution new data members were added to the messages of MIP, RSVP and Routing protocols.

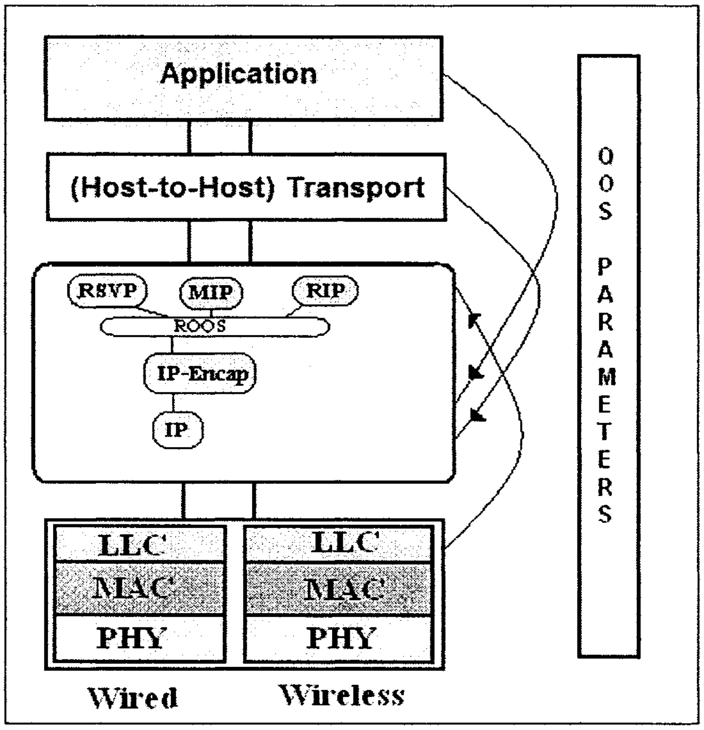

Figure 14: Cross-Layer Information Flow

In brief, the mobility (MIP) and routing signaling will be more QoS aware and resource reservation (RSVP) signaling will be more routing concerned. The messages for the protocols of concern are as follows:

RSVP: RESV and PATH

MIP: Registration Request (RRQ), Registration Response (RRP),

Routing: Agent Advertisement (AA), Agent Solicitation (AS),

The details of the QoS parameters and the signaling of protocols involved are presented in the next chapter.

3.3.2.2 ROQS Entity

The ROQS entity shown in Figure 14, consists of several components as shown in Figure 15.

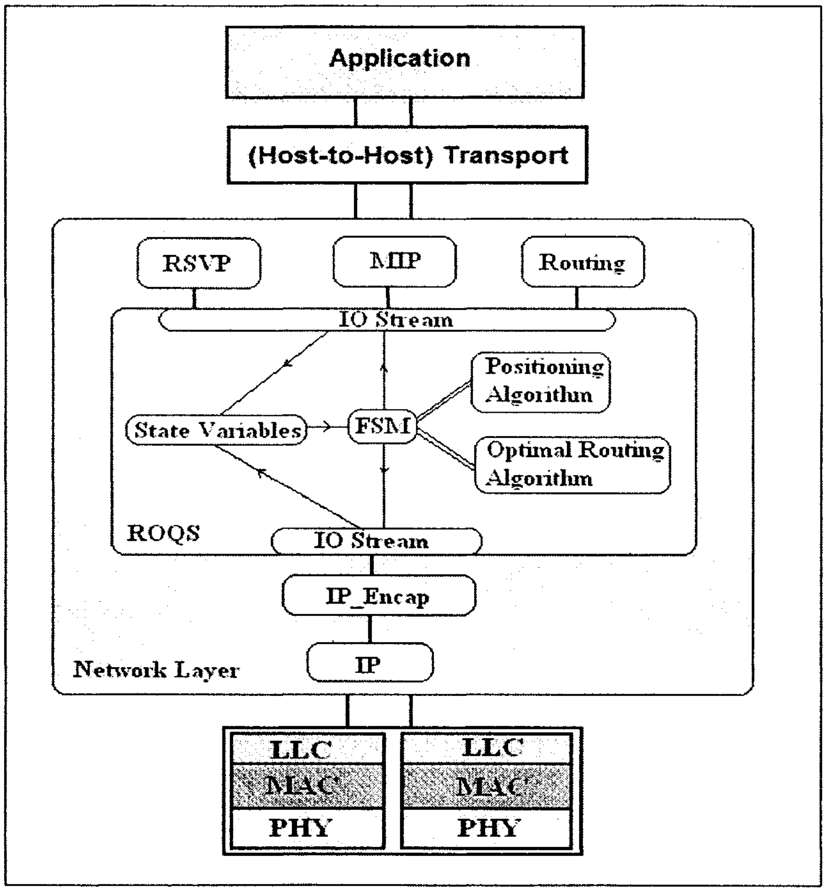

Figure 15: Modified Protocol Stack

It includes a Finite State Machine (FSM), a set of state variables and necessary algorithms. This entity also collects the QoS parameters. ROQS interacts with IP encapsulation module from one side and RSVP, MIP and Routing protocol from the other side. The FSM is responsible for managing the state transition based on different events in the network and invoking algorithms for selecting the optimal path in proper state.

Furthermore, ROQS is responsible for retrieving and temporarily storing the QoS requirements and network conditions, and analysing them to find the optimal path. This ROQS is detailed in the following components:

- Input and output stream to interact with MIP, Routing, Resource Reservation and IP encapsulation modules
- Temporary storage for QoS requirements and network conditions
- Finite state machine
- Position analysis algorithm
- Routing analysis algorithm

To describe in which physical entities ROQS will reside, we consider the simplified MIP network architecture shown in Figure 16. In this network, physical entities that communicate with each other are MN, BSC, HA, FA, and CN. BSC is the representative of wireless cellular network that relays the packets to MN. FA and HA are two mobility aware routers and can communicate with each other in the same manner as normal routers via Internet.

Both HA and FA are routers enhanced with mobility support and have the same infrastructure. CN and MN are hosts of the same nature, whether connected via wireless link or wired link to the Internet. Internet on the other hand can be seen as many routers connected to each other.

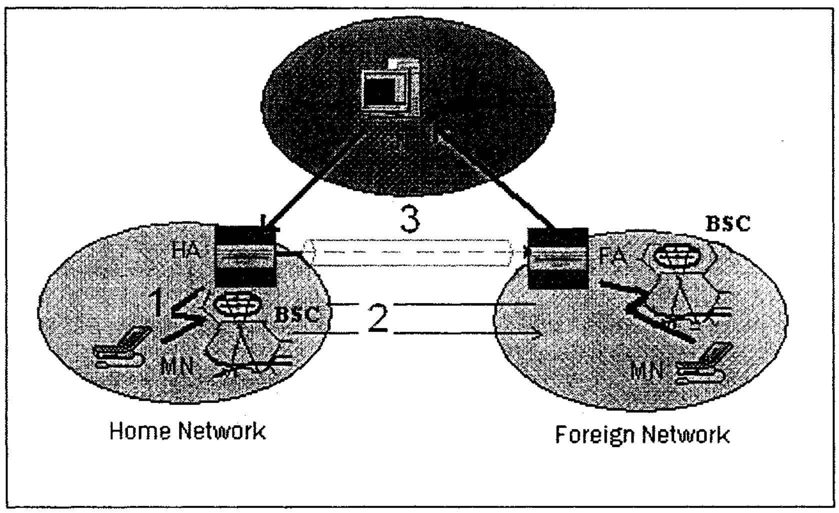

Figure 16: MIP Entities

Each of these entities (Agent, Host and Internet) functions based on interworking of many protocols including TCP, UDP, RSVP, MIP, IP, IP_ENCAP, Routing Protocol, OSPF. All these entities have been developed respecting layering architecture described earlier. Figure 17 shows the protocol stack of a common host.

The service running on the MN sends its control or data message toward physical layer to the wireless LAN ports through necessary protocol layers.

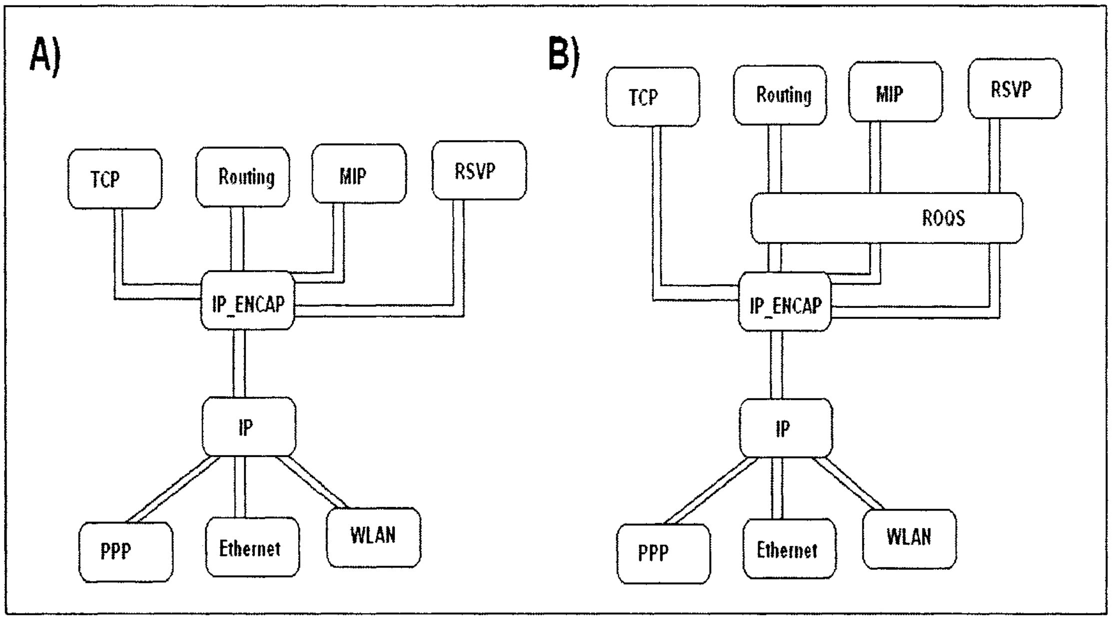

Figure 17: Host Internal Architecture

To include the ROQS entity the internal architecture of the host is modified as shown in right side of Figure 17 by adding the ROQS entity interconnecting MIP, RSVP, RIP and IP_Encap modules.

Figure 18 shows the protocol inter-relation for a typical agent (HA or FA). An agent is mainly a router with different routing protocols such as OSPF, RIP, IGRP, etc, for normal routing and MIP for mobile routing. It is connected to a wired and wireless port to accommodate different kind of connections (wired and wireless).

After the messages from MN pass through an agent, they pass through Internet to reach the destination. The protocol stack of agent shown in part A of Figure 18 is enhanced by ROQS entity in part B.

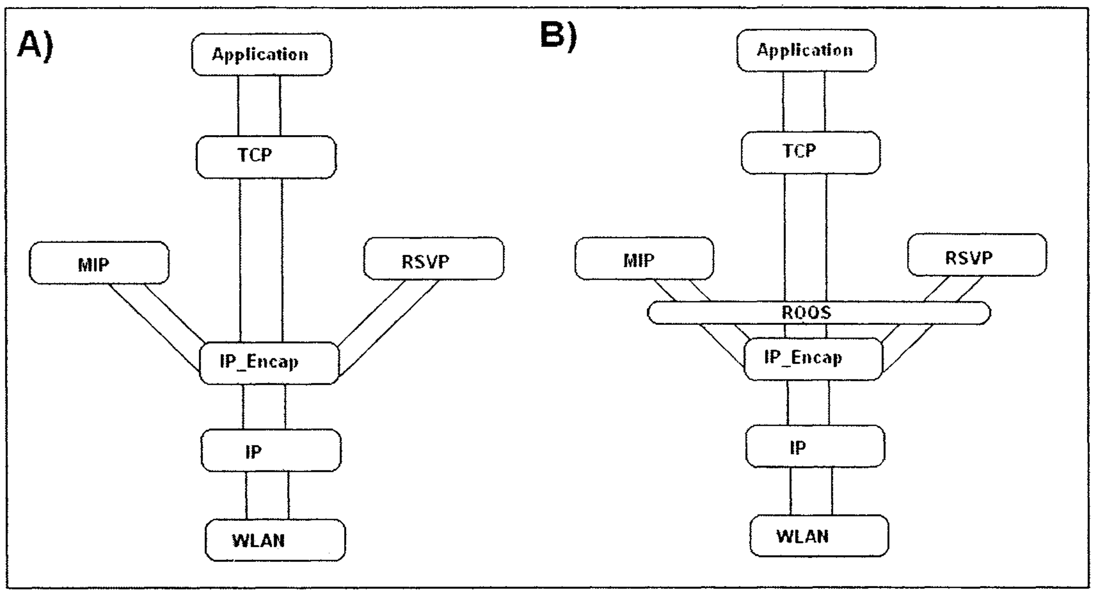

Figure 18: Agent Internal Architecture

The common protocols in different entities deal with mobility, routing, and resource reservation. They are MIP, RSVP and Routing Protocol. The relevant messages of these three protocols carry the necessary information to the ROQS entity. The details of ROQS components (FSM, position analysis algorithm and routing analysis algorithm) are presented in the "Optimal Path Selection" chapter.

3.4 Impacts and Preliminary Evaluation

Based on the design described in the previous sections, the main events and message sequences are unchanged. The changes to the event and message sequence only include the intervention of ROQS entity, and if it finds direct path as optimal path, it will divert traffic on the direct path. Below, the modified event sequence and message sequence are provided.

3.4.1 Processing Stages

Using MIP-ROQS solution the selection of an optimal path over which resources can be reserved is shown in Figure 19.

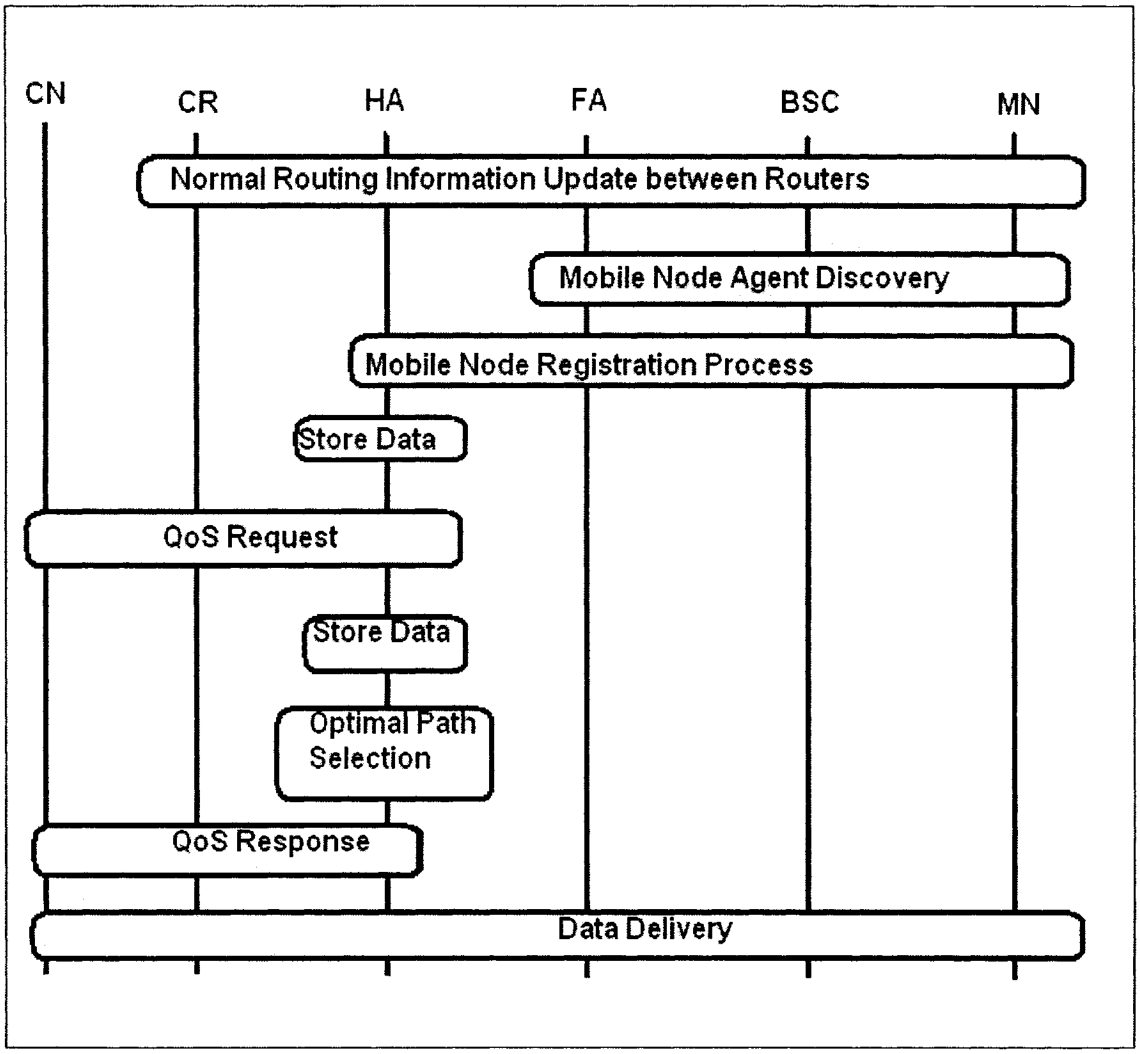

Figure 19: Processing Stages in a Session

1) Routing Update

   At the normal routing updates, the characteristics of the links in the core network are collected.

2) Agent Discovery

During Agent discovery stage the data regarding the wireless link in the access network is collected.

3) Registration

During registration stage, the data describing the condition of wireless links and capabilities of MN, Agents and the wired link in the core network are collected.

4) ROQS QoS Parameters retrieved and stored

ROQS entity scans the messages and retrieves the QoS parameters

5) Resource Reservation Request

When actually a session is about to be established and a QoS request is received by HA.

6) ROQS QoS Parameters retrieved and stored

The necessary data is retrieved and stored in state variables and the optimal path selection starts.

7) ROQS Optimal path selection

Position analysis and routing analysis are performed and the optimal path is identified.

8) Resource Reservation Response

Based on the results of the previous step, adequate response to resource reservation request is generated and requested resources are reserved on the optimal path.

9) Data Delivery

   Actual data delivery starts on the optimal path.

3.4.2 Message Sequence

For ease of comparison message sequence is shown without ROQS and with ROQS solution.

3.4.2.1 Message Sequence without ROQS

Based on MIP and RSVP protocols the interaction between entities involved in the session without ROQS solution is depicted in Figure 20. FA and HA are routers enhanced with MIP protocol and wireless LAN ports. Internet is not directly connected to access/wireless network but through these agents.

1) Agent Discovery: MN sends out Agent Solicitation.
2) Agent Discovery: FA sends Agent advertisement which goes from FA to the MN.
3) Registration: MN sends the RRQ to register its temporary address with HA.

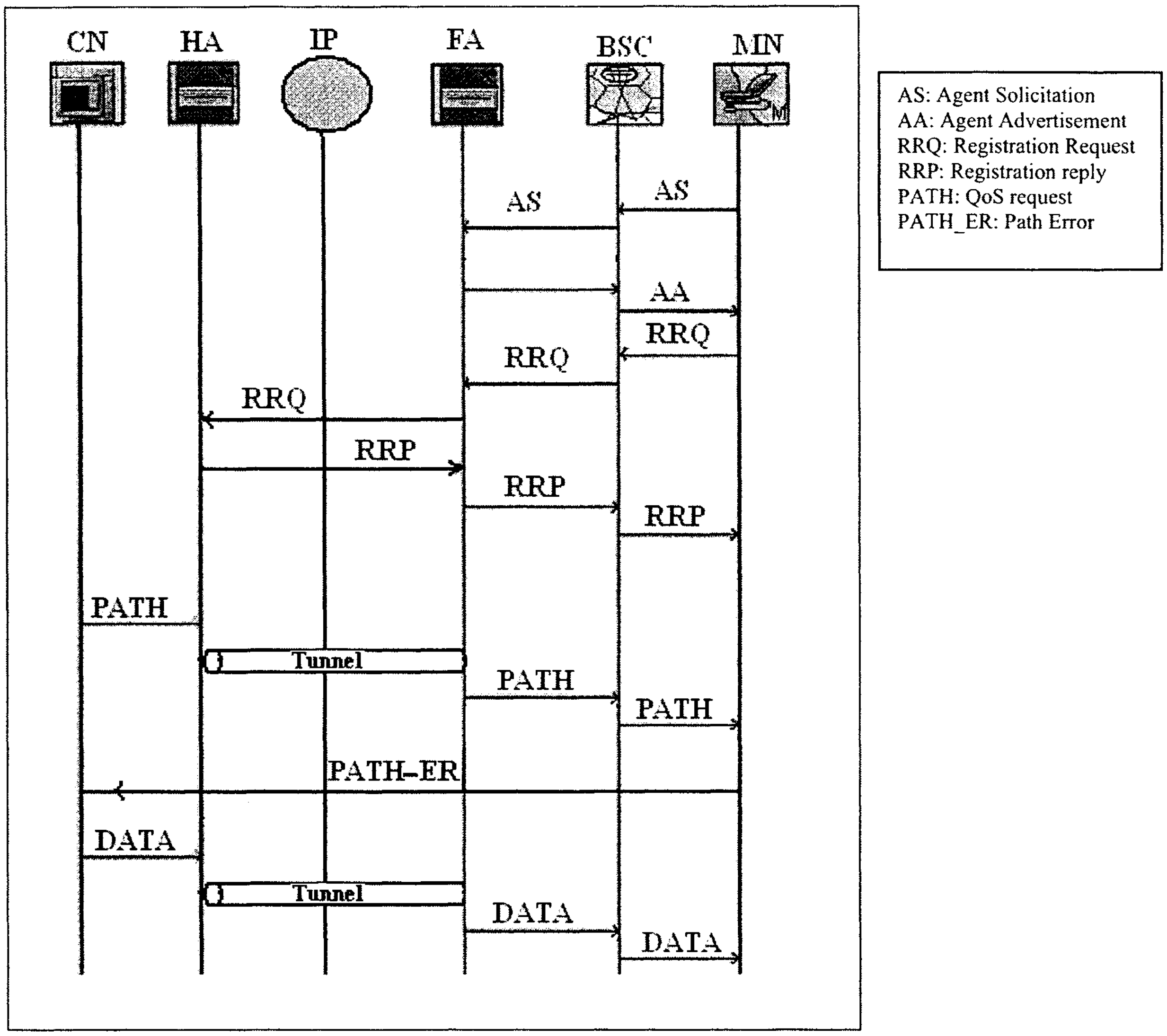

Figure 20: MIP-RSVP Message Sequence

4) Registration: HA respond back to MN with RRP.

5) Resource Reservation: When needed, a CN will send out Resource reservation request to MN home address, which will be encapsulated and sent to FA and MN.

6) Resource Reservation: MN will answer reservation request via FA or directly to CN.

7) Data Delivery: After completion of resource reservation, data delivery happens.

3.4.2.2 Message Sequence with ROQS

Based on MIP-ROQS and RSVP protocol the interaction between entities involved in the communication session is shown in Figure 21.

In summary the sequence of events will be as follows:

1) Agent Discovery: MN sends out Agent Solicitation with additional QoS capability parameters.

2) Agent Discovery: FA sends Agent advertisement which goes from FA to the MN. It takes with it the QoS characteristics of the foreign network and wireless link.

3) Registration: MN sends the RRQ to register its temporary address with HA. It has to add its QoS capability to it, as this message goes from node to node to HA it adds and takes QoS characteristics of each node with it.

4) Registration: HA respond back to MN with RRP.

5) Resource Reservation: When needed, a CN will send out Resource reservation request to MN home address, from which the necessary QoS parameters are retrieved. The optimal path selection starts.

6) Resource Reservation: MN will answer reservation request via FA or directly to CN based on the result of optimal path selection.

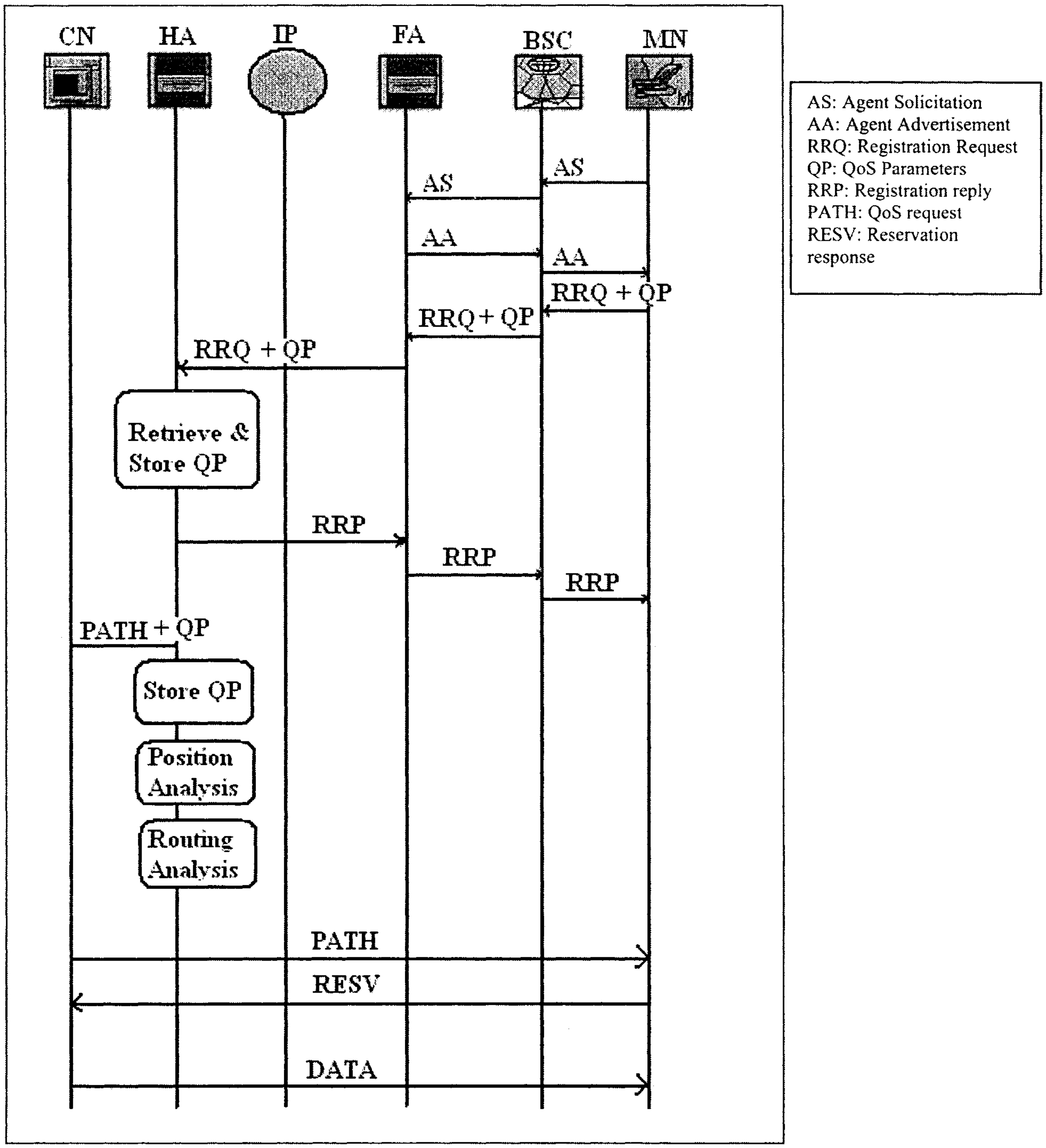

Figure 21: MIP-ROQS-RSVP Message Sequence

7) Data Delivery: After completion of resource reservation, data delivery happens.

3.5 Conclusion

Fast pace growing wireless IP-based networks and their services require solutions for providing QoS that can not necessarily be achieved in the standard layered architecture. Information sharing between different layers seems to be the adequate design although it should be done with caution. In this chapter we described MIP-ROQS solution in detail as a cross-layer designed architecture for addressing the QoS guarantee in MIP environment and resolved the problem of providing simultaneous RO with QoS.

Based on [109] one of the concerns about cross-layer designs is that the amount of new interactions can be so large that it makes the introduction of new extensions very complex. One of the main highlights of MIP-ROQS design is that the amount of interactions is not deviating from the original system and therefore is not a hindrance for any new development.

The global MIP-ROQS system is not preventing any new solution or changing the architecture of the system in a way that the future design needs to be dependent on it.

At the conceptual level the new perception of addressing the issues of mobile environment with a cross-layer design is applicable to other issues such as security and link adaptation. At practical level, resolving the conflict between MIP protocol and IntServ/RSVP can be generalized. Any mobility protocol which relies on learning the actual address of the MN using the first interaction and any resource reservation mechanism that needs to signal the requirements can use the new solution.

MIP-ROQS solution is easily configurable to be turned off when not necessary by configuring a dynamically settable parameter or statically in the system. We elaborate on this in the Chapter 5 while describing the implementation details. Therefore, in the case the ROQS optimal routing is disabled by setting the configurable parameter, the system will act as original MIP is dictating. The detail design of the architecture presented here is described in next chapter and the evaluation of the proposed solution with respect to the derived requirements described in Section 3.2 will be discussed in Chapter 5.

CHAPTER 4

ROQS OPTIMAL PATH SELECTION

4.1 Introduction

Based on the architecture presented in the previous chapter a new information flow is required to collect the QoS parameters from different layers of the MIP protocol stack and a sophisticated entity that analyzes the collected data and select the optimal path. In this chapter the details of the information flow and the optimal path selection mechanism are presented.

QoS parameters at the service level are certain conditions that should be in place for a service to perform adequately. A service may need the guarantee that all the packets for a specific session are delivered within 200 ms or guarantee that no more than x number of packets will be lost. In the core network, packet loss, network congestion and delay are the QoS restrictions. In the access network the limitation of radio transmission capacity, bandwidth and number of simultaneous calls are the boundaries.

To have the new information flow without adding new signaling, we bridge the gap to certain extent between the protocols involved in mobility, routing, and resource reservation. In other words, the mobility and routing signaling will be more QoS aware and resource reservation signaling will be more aware of routing. The new information flow will collect the QoS parameters, which describe the network conditions and QoS requirements using existing signaling for routing, mobility, and resource reservation.

The optimal path selection mechanism is responsible for selecting the optimal path. This selection is managed using a Finite State Machine (FSM), a set of state variables to temporarily hold the network conditions and necessary algorithms to select the optimal path. The FSM invokes retrieving and storing state variables at proper stages of a session. It also invokes the required algorithms to analyze the data and select the optimal path in the relevant stages.

The optimal path is used for data delivery and therefore, the resources can be reserved on it.

4.2 Collecting QoS Parameters

QoS parameters are the data describing the service requirements and network conditions. Network performance is measured in general based on the amount of traffic it can handle. This can be detailed further as the packet loss, bit error rate, queuing delay, and throughput in the wired network or the amount of data dropped, delay and throughput in the wireless network.

In this Section the QoS parameters and the information flow using the existing signaling that collects these parameters are presented.

4.2.1 QoS Parameters

As explained earlier, in the context of MIP, network is a general word that refers to three components of MIP architecture, services, core and access network. QoS requirements or restrictions have different characteristics in each part of MIP network. These requirements can be described as follows:

The type of QoS requirement specified in service layer could include:

- Guarantee that all packets for this session will be delivered within 200ms, provided no more than 2 Mbit/s is sent.

- Guarantee that only 1% of packets will be lost, when measured over a specific period.

These requirements for each service are passed from application layer to the network layer. It means that in each packet there should be enough parameters that specify which quality aspects are needed and what is the priority that should be adapted without affecting the outcome for the end-user.

Core network has its own restriction that should be highlighted as data elements to facilitate the routing decision. The nature of core network is impacted by packet loss, router congestion, and delay. Since there is no fix number of the constraints of these types, these data elements can be also taken into account as probabilities of occurrence.

Access network also has its own limitation that is varied from one network to another. Since mobility is the subject, to travel from one access network to another can change the resulting QoS for the end user. The capacity of the network, bandwidth, and the number of simultaneous sessions are considered as another category of input data for decision algorithm.

We have chosen a set of parameters to describe the conditions of different parts of the network (core network and access network), different entities in these networks (host and agents) and the QoS requirements of application. These parameters are presented in Table 2.

	Service	Core Network	Access Network	Host	Agent
QoS Parameters	-Minimum guaranteed Bandwidth -Maximum tolerable delay -Maximum tolerable Delay variation -Maximum tolerable packet loss -Weight of each element above	-Number of hops -Packet loss -Bit error rate -Queuing delay -Throughput	-Data dropped -Delay -Throughput	-Media Access Delay -Throughput	-Control Traffic -Processing delay -Binding Delay -Tunnelled traffic delay -Tunnelled traffic overhead -RO Traffic delay -RO Overhead

Table 2: QoS Parameters in Different Network Layers and Nodes

However, there can be other parameters that one may consider important. Additional parameters do not alter the fundamental idea of this solution. Any parameter can be added or removed from this list as they are fundamentally constraints that different parts of the network and the equipment involved in the communication session put on a scenario. How to carry these parameters using the modified signalling to the optimal path selection entity is described in next section.

4.2.2 Modified Signalling

In order to satisfy the requirements and avoid introducing a new set of signals for collecting QoS parameters from the network, we use the existing signaling of routing, mobility and resource reservation and modify them to collect the necessary data. These signals are described below.

4.2.2.1 Mobility Signalling

As the nature of mobility implies; the conditions and QoS specification of the system elements (core network and access network) change pretty often. Each time a mobile user moves from one network to another throughout a single session, the new network possibly will have a new set of constraints. Consequently, the pertinent data should be refreshed for optimal path selection.

q1: wireless link condition

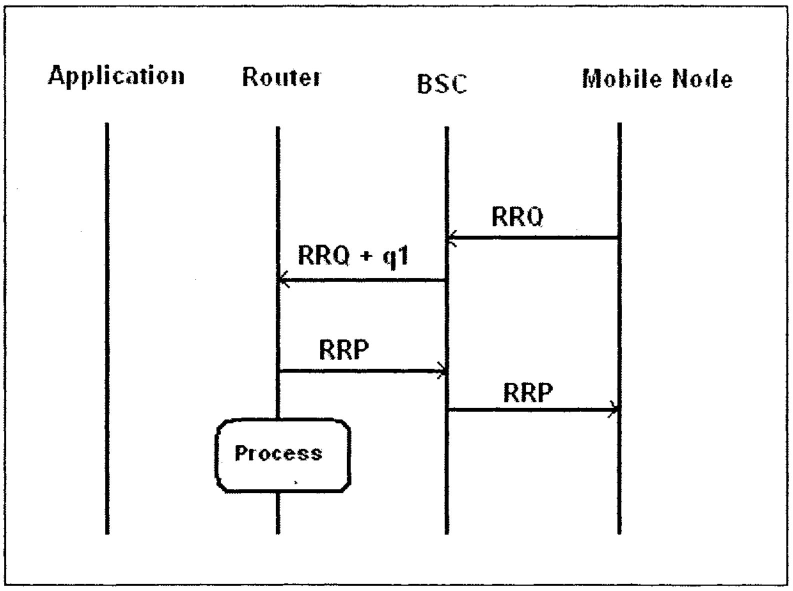

Figure 22: Modified Mobility Signalling

Figure 22 shows updating the QoS specification in registration procedure via mobility signaling. When MN sends registration request (RRQ), it passes via the access network which will incorporate its own QoS restriction to the message prior to passing it to the core network. The router evaluates the QoS restrictions and requirements again and chooses the best routing path according to the new specification.

4.2.2.2 Routing Signalling

Figure 23 shows the routing update that happens periodically in the network. Correspondent Router (CR) in this Figure is the router on the edge of the network where CN resides.

While Routing Information Protocol (RIP) [121] acts based on single hop count metric to evaluate the links for packet forwarding, Interior Gateway Routing Protocol (IGRP) [121] is a more sophisticated routing protocol that uses a composite metric calculated by factoring weighted mathematical values for delay, bandwidth, reliability, and load. Network administrators can set the weighting factors for each of these metrics. We use the information available by IGRP routing protocol to choose the optimal path in ROQS entity. Furthermore, Agent Advertisement (AA) and Agent Solicitation (AS) as router discovery messages piggyback MN capabilities and conditions of wireless link.

q1: MN Capability

q2: Wireless link condition

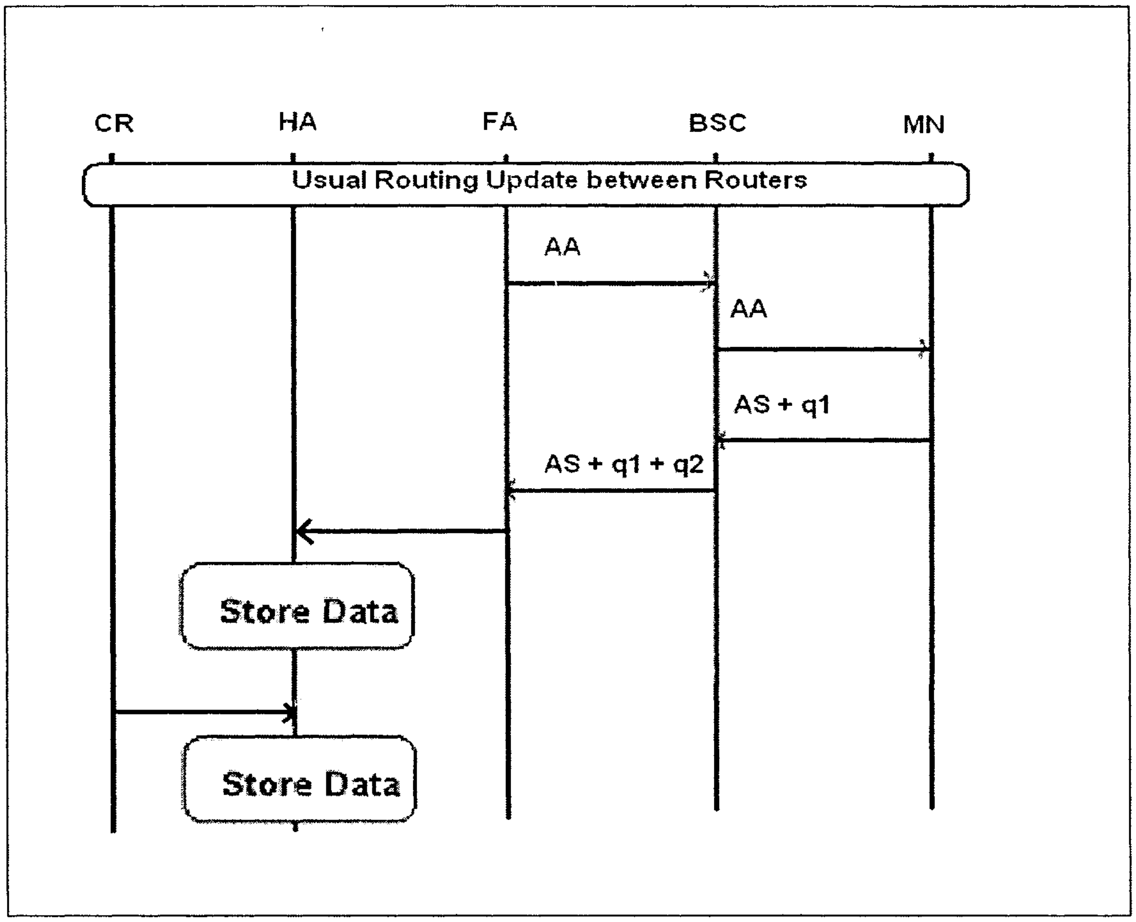

Figure 23: Modified Routing Signalling

4.2.2.3 Resource Reservation Signalling

In order to request and reserve resources along the path of autonomous systems from the CN to the MN, signaling messages must be exchanged between adjacent CN, HA, FA and MN. Figure 24 shows the message sequence for resource reservation signaling in general. This could be considered simply a request for certain level of quality and a response on whether that quality might be granted.

QoS request and response or PATH and RESV messages will convey not only a collection of what the application requires but also distance from CN to HA based on the hub count and application performance limits. When the HA receives these data, it has enough information to analyze the situation.

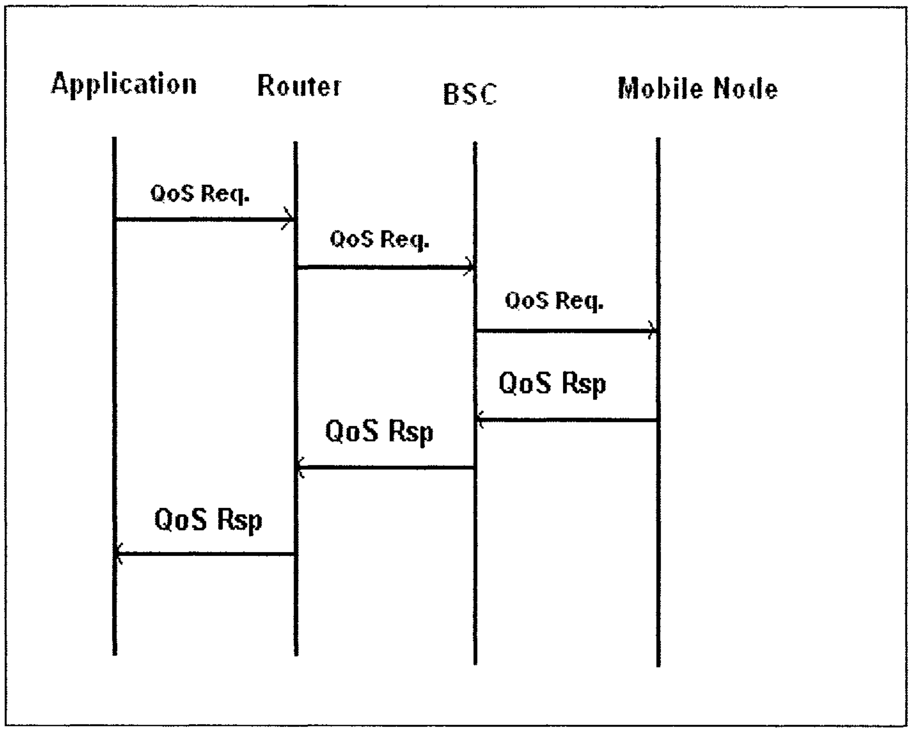

Figure 24: Modified Resource Reservation Signalling

4.3 Optimal Path Selection Mechanism

In the frame of the high level architectural design presented in previous chapter, we collect the necessary QoS parameters and carry them over the signaling of MIP, RSVP and routing. These parameters are the input for the optimal path selection mechanism. As described previously MIP network layer is enhanced by optimal path selection mechanism. This mechanism is handled by an entity which receives the necessary parameters and analyse them for optimal path selection. The result of the optimal path selection is a route to the MN. This entity is comprised of a FSM and a set of variables and necessary algorithms.

As long as the service requirements or the characteristics of the network have not changed, the optimal path remains the same. Any change in the service layer, core network, or access network that affects the total QoS will result in a route recalculation. The new parameters are again transferred to the ROQS entity via the Routing update and MIP binding update in the core network and refresh signals in the access network.

The new route to the destination is re-calculated as a background task. When the QoS specification of the network drops down in a way that the best route does not satisfy the QoS requirement any more, a message is sent to the application while the data transfer continues with available quality.

The normal routing algorithms may need to provide routing information for all the nodes in the network. When there are too many nodes, this approach is not efficient. Treating each node as an equal participant in a large network generates too much information to share and send throughout the network.

An alternative is to have some nodes perform the routing task for the others. Therefore, the optimal path selection is performed in certain routers in the core network instead of all routers.

4.3.1 FSM and State Variables

Considering the distributed nature of MIP environment and several stages that ROQS entity is involved in, state transition is defined and managed using a FSM. As shown in Figure 25, ROQS FSM model, based on its functionality, has necessary states to wait for the reception of required data and perform the optimal route selection.

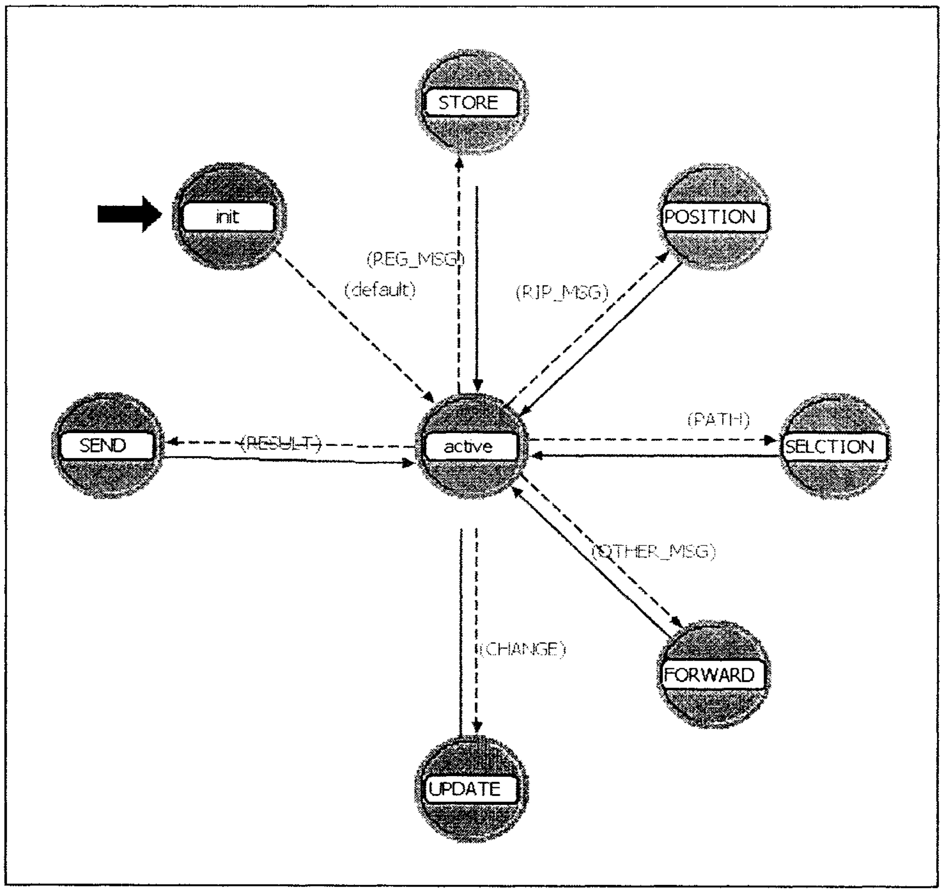

Figure 25: Optimal Path Selection State Machine

This state machine is responsible for handling the following events:

- Reception of MIP Registration messages (REG_MSG) which are registration request and Registration response
- Reception of RSVP PATH message which is the request for certain level of quality
- Reception of Routing Protocol message which are routing update and agent advertisement and agent solicitation
- Reception of any other message in order to cover all the messages that
- Handling the result of optimal path selection
- Handling change in the network which happens every time one of the above messages arrives and indicates a change in the network condition.

Handling of these events is described below in more detail. However, handling some of the events requires temporary storage to save the QoS parameters. State variables are the temporary storage needed to store the retrieved information regarding QoS requirements and network condition before the optimal path selection can start analysing these data.

Table 3 presents the state variables used to store the collected QoS parameters for a session before analysing them to find the optimal path. Each variable is described briefly in the Table below.

	Name	Description		Name	Description
1	MGB	Minimum guaranteed Bandwidth	16	QDD	Queuing delay on Direct Route
2	WMGB	Weight of MGB	17	THT	Throughput on Triangle Route
3	MTD	Maximum tolerable delay	18	THD	Throughput on Direct Route
4	WMTD	Weight of MTD	19	MAD	Media Access Delay in MN
5	MTDV	Maximum tolerable Delay variation	20	THMN	Throughput in MN
6	WMTDV	Weight of MTDV	21	DDWLL	Data dropped in Wireless Link
7	MTPL	Maximum tolerable packet loss	22	DWLL	Delay in Wireless Link
8	WMTPL	Weight of MTPL	23	THWLL	Throughput Wireless Link
9	NHT	Number of hops on Triangle Route	24	CTPD	Control Traffic Processing delay in HA
10	NHD	Number of hops on Direct Route	25	BD	Binding Delay in HA
11	PLT	Packet loss on Triangle Route	26	TTD	Tunnelled traffic delay in HA
12	PLD	Packet loss on Direct Route	27	TTO	Tunnelled traffic overhead in HA
13	BERT	Bit error rate on Triangle Route	28	ROTD	RO Traffic delay in HA
14	BERD	Bit error rate on Direct Route	29	ROOV	RO Overhead in HA
15	QDT	Queuing delay on Triangle Route			

Table 3: List of State Variables and their Description

Handling of the different events in ROQS FSM is as follows:

- Reception of "REG_MSG":

At the reception of registration request and registration reply messages of MIP protocol, initially it verifies whether the optimal routing feature is active. If this is the case, it checks which node it is acting in. If it is in FA, it includes the QoS parameters in the message before sending it out. If it is in the HA, it retrieve the QoS parameters and store them in the temporary variables.

- Reception of "RIP_MSG":

At the reception of routing update messages, if optimal routing feature is active, the QoS parameters are retrieved and stored in temporary variables.

- Reception of "PATH" (Resource reservation request) message:

At the reception of PATH or resource reservation request, if optimal routing feature is active, the optimal path selection algorithms are invoked. It starts by calling the position analysis algorithm (will be elaborated in the next section). If the result of this algorithm is the optimal path, it will invoke handling the result (explained below). If further analysis is required, routing analysis algorithm (elaborated in the next section) is invoked.

- Reception of any other message "OTHER_MSG":

Any other message that is passed through the FSM will be simply forwarded to the destination without any further action.

- “Change” in the network:

Since routing update and mobility registration signalling are repeated periodically, they refresh the QoS parameters as well. If the QoS parameters have been modified compare to the previously stored information, the optimal path selection mechanism is invoked again (similar to reception of PATH message).

- “Result” of optimal route selection:

When the optimal path selection mechanism is performed, its result is handled in this state. If the optimal path is the tunneling, PATH message is forwarded to CoA. If the optimal path is the direct path, a PATH error response is sent to CN which includes the CoA of the MN. By receiving PATH error response, CN will issue another PATH message toward the CoA directly.

4.3.2 Algorithms

The task of optimal path selection is divided into two phases: Position Analysis and Routing Analysis. Position Analysis discovers the relative position of MN, CN and HA, and the Routing Analysis estimates the available resources on each route and compares it to the service requirement and service constraints and decides which path is the optimized path for the communication.

These two phases are executed using position analysis algorithm and routing analysis algorithm described below.

4.3.2.1 Position Analysis

Position analysis is about making the primary decision on routing path based on the relative distance of CN, MN and HA. The 3 elements can have relatively 6 positions based on their distance from each other. The distance between nodes is based on the hub count extracted from routing table and routing update messages.

As shown in Figure 26, analyzing the addresses of CN, HA and FA as 3 corners of the triangle, and the hub counts as the edges of the triangle, one of the following could be the case.

- CN and HA are in the same network or domain as cases 1 and 2 (Triangle path is the optimal path)
- CN and MN are in the same network or domain as cases 3 and 4 (Direct path is the optimal path)
- HA and MN are in the same network or domain. (no need for further analysis as the MN is at home)

If none of the above relation is satisfied then Routing Analysis stage is necessary.

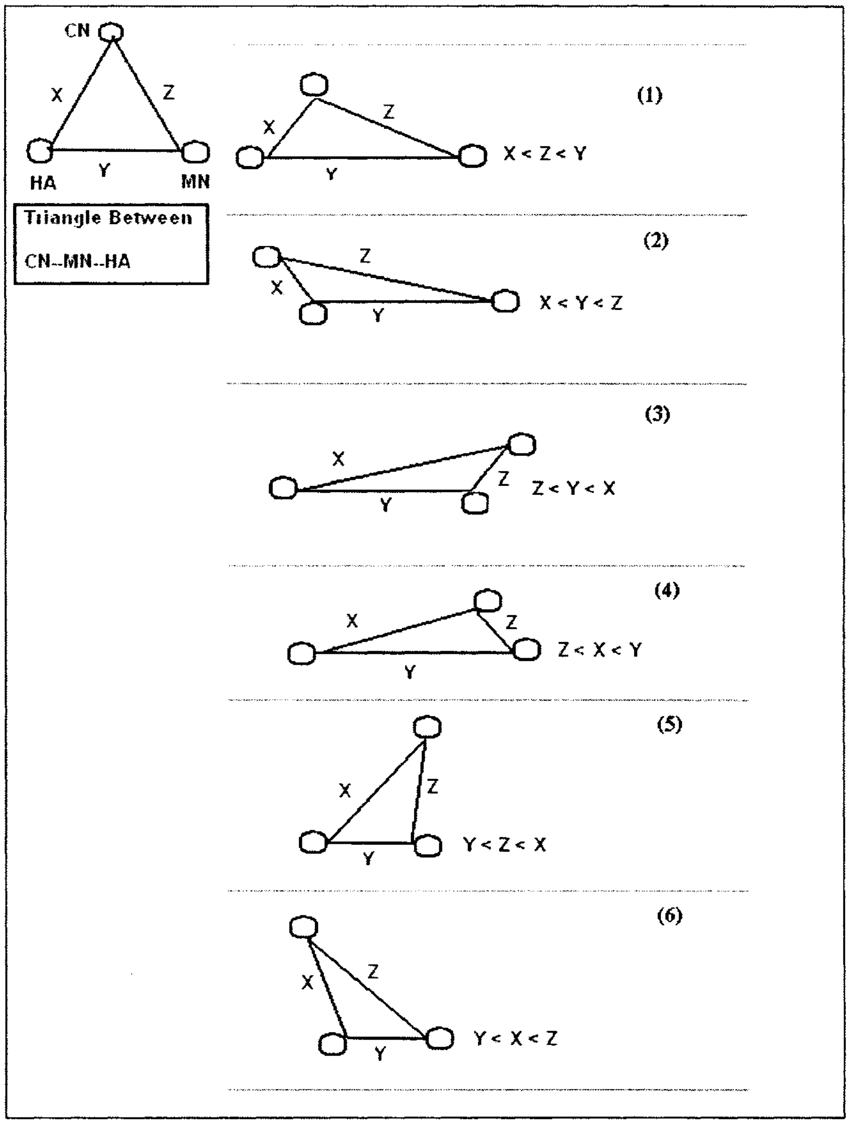

Figure 26: Possible Positions

Figure 27 shows the pseudo code of position analysis algorithm. Initially it retrieves the network and subnet prefix of the IP address of the three nodes (CN, HA and FA). The analysis includes verifying whether the nodes are in the same network or sub-network. Further analysis is based on relative distance of the three nodes.

```
Get the distance of CN and HA in X
Get the distance of HA and FA in Y
Get the distance of CN and FA in Z

Fetch Network Prefix of CN IP address in A1
Fetch Network Prefix of HA IP address in A2
Fetch Network Prefix of FA IP address in A3
Fetch SubNet Prefix of CN IP address in SA1
Fetch SubNet Prefix of HA IP address in SA2
Fetch SubNet Prefix of FA IP address in SA3

IF SA1 = SA2 or A1 = A2
        Triangle Route is Optimal
ELSE
    IF SA1 = SA3 or A1 = A3
        Direct Route is Optimal
    ELSE
        IF SA2 = SA3 or A2 = A3
            Triangle Route is Optimal
        ELSE
            IF  (X< Z & Z <Y)   or (X<Y & Y<Z)
                Triangle Route is Optimal
            ELSE
                IF (Z<Y &Y<X) or (Z<X & X<Y)
                    Direct Route is Optimal
                ELSE
                    IF (Y<Z & Z<X) or (Y<X & X<Z)
                        Triangle Route is Optimal
                    ELSE
                        Perform Routing Analysis
                    ENDIF
                ENDIF
            ENDIF
        ENDIF
    ENDIF
ENDIF
```

Figure 27: Position Analysis Algorithm

The distance between these nodes is retrieved from routing table, based on which the optimal path can be chosen without further need for analysis by routing analysis algorithm.

4.3.2.2 Routing Analysis

Routing analysis is about the evaluation of two paths against a service requirement to find out which path satisfies the service requirements while increasing the network performance. On each path total condition of wired link, wireless link and the delay and processing time in different devices are calculated and compared against service requirements. The result of this comparison yields one of the following five results:

1. Direct path satisfies all the service requirements and Triangle path does not satisfy any (Optimal path is Direct path).
2. Triangle path satisfies all the service requirements Direct path doesn't satisfy any (Optimal path is Triangle path).
3. Triangle path satisfies some of the service requirements and direct path satisfies the other requirements: In this case the weight of different requirements is verified to see which requirement is the most crucial and choose the path accordingly.
4. Both Triangle path and direct path satisfy all service requirements: In this case the selection of the path must aim at minimizing the blocking

probability of future requests. This means to calculate the remaining resources after assigning the requested resources on each path and compare the result to find out which path after allocation of the resources for this session will remains with more resources. This will ensure unified resource utilization and minimum blockage for future sessions.

5. None of the paths satisfy any of the service requirements: In this case an error message is sent back to CN as a response to PATH message informing that the QoS can not be granted.

Figure 28, shows the more detailed pseudo-code of the routing analysis algorithm. The inputs to the algorithm are 3 arrays, holding the conditions of the direct path, conditions of the triangle path and the QoS requirements respectively. The QoS requirements are stored in the array based on the order of their importance, where the requirement in index (i) has a lower priority than the requirement in index (i + 1).

The index is the representative of the weight of the requirement. When comparing the conditions of the link versus the requirements, the weight of each requirement (index) is added as a factor. This will ensure that the requirements with higher priorities are granted with higher priority.

For example if one path can satisfy two of the requirements and the other path can satisfy two other requirements, the sum of priorities of the requirements is the basis for selecting the optimal path.

```
Let:
   DPC[0-n] be Direct Path characteristics
   TPC[0-n] be Triangle Path characteristics
   APR[0-n] be Application requirement in order of priorities from 0 to n, when n is the highest
priority
   Td be Traffic on Direct path
   Tt be Traffic on Triangle path

Initialize flag[n][2] to zero
for (i=0, i<n, i++)
         if (DPC[i,0] == APR[i])
                  Set flag[i,0] to 1
         else
                  Set flag[i,0] to 0

        endif
         if (TPC[i,0] == APR[i])
                  Set flag[i,1] to 1
         else
                  Set flag[i,1] to 0
         endif
endfor

Dop, Top = 0;
for (i=n, i>0, i--)
         if (flag[i,0]< flag[i,1])
                  APR[i] can be granted by Triangle path
                  Top = Top + i
         elseif (flag[i,1]< flag[i,0])
                  APR[i] can be granted by Direct path
                  Dop = Dop + i
         elseif (flag[i,0] == 1)
                  Both paths can satisfy APR[i]
                  Dop = Dop + i
                  Top = Top + i
         else
                  QoS cannot be granted
         endif
endor

if ( Dop> Top )
         Direct path is the optimal path
elseif ( Dop <Top )
         Triangle path is the optimal path
elseif ( Dop == 0 )
         No QoS characteristics of APR[0-n] can not be granted
elseif ( Td < Tt )
         Direct path is the optimal path
elseif ( Tt < Td )
         Triangle path is the optimal path
else
         Triangle path is the optimal path
endif
```

Figure 28: Optimal Routing Algorithm

After the above selection, one of the following procedures is performed accordingly:

- In the case direct path is the optimal path: A PATH error message is sent to CN. Included in this error message will be the CoA (current location of the MN). CN can retrieve the CoA and send another PATH message directly to the MN using this CoA.
- In the case triangle path is the optimal path: Reserve the resources along the path and follow the same path for the consequent packets.

4.4 Conclusion

In this chapter the details of collecting and analyzing the QoS parameters from the different components of the MIP network were presented. Mobility agents are armed with an optimal path selection mechanism. This mechanism is managed using a FSM, which relies on getting the information about resource availability in access and core networks. This data is piggybacked over existing signaling of mobility, routing, and resource reservation protocols. It invokes position and routing analysis algorithms to find the optimal path from CN to MN.

Position analysis makes the initial route selection based on the IP address and the distances of the CN, MN and HA. Routing analysis is about evaluating conditions of two paths against service requirements.

CHAPTER 5

PERFORMANCE EVALUATION THROUGH SIMULATION

5.1 Introduction

Evaluating MIP-ROQS solution for providing QoS on optimal routes in MIP requires a suitable environment and proper metrics to measure its impacts on different components of MIP architecture. Performance evaluation can be either based on simulation and measuring pre-defined parameters in the system's model or based on analytical and theoretical estimation.

We chose simulation for performance evaluation of MIP-ROQS system. To evaluate the efficiency of MIP-ROQS, a model of the system was built with and without the ROQS mechanism. In the actual environment, the network topology, type of service and the characteristics of communication scenarios affect system performance. Therefore, to have proper conclusion we built different network topologies, run different services and followed different communication scenarios.

5.2 Evaluation Environment

We used Opnet [107] as the simulation environment as it provides the model for most of existing networks. Using consistency checking of Opnet, the implemented model was verified to ensure the validity of ROQS software entity. Network configuration was performed by setting available attributes. New behaviour was introduced by extending the existing attributes and acting upon their values when necessary. The modular nature of Opnet accommodated implementing ROQS CBR algorithm in host and agent.

To measure the performance we created a model of the system, by configuring existing MIP model and modifying some of its components. The next step was to identify measurable metrics of the system, and configure them accordingly.

Running the simulation performed the model checking and collected the statistics on the configured metrics for the duration of the simulation run. Finally we interpreted the result to describe the impact of MIP-ROQS.

The assumptions made in the simulation are as follows:

- The IP network is considered one node with settable values for delay and packet lost.
- Simple network configuration is evaluated and hierarchical network architecture with domain definition assumed to show similar or improved results.
- Video-Conferencing service is considered a QoS-sensitive service.
- Simple traffic demand configuration is similar to a non-QoS-Sensitive service such as email, etc.
- Only 75% of resources on each link are available to reflect the actual network condition.

We compare the performance of MIP-ROQS against original MIP and MIP enhanced with RO. The comparison metrics are detailed in sub-section 5.4.3.

5.3 Creating MIP Network Model

To build a network model for performance evaluation, in an incremental approach, first the wired-MIP network was configured. Based on the wired-MIP, a wireless MIP model with roaming trajectories was configured. On each network topology, tunnelling and RO was configured. After configuring a video-conferencing application, then the resource reservation was configured.

5.3.1 Project Level

At project level, the nodes and their connections are depicted. Type of nodes and connections, and their specific characteristics and functionality are explained below.

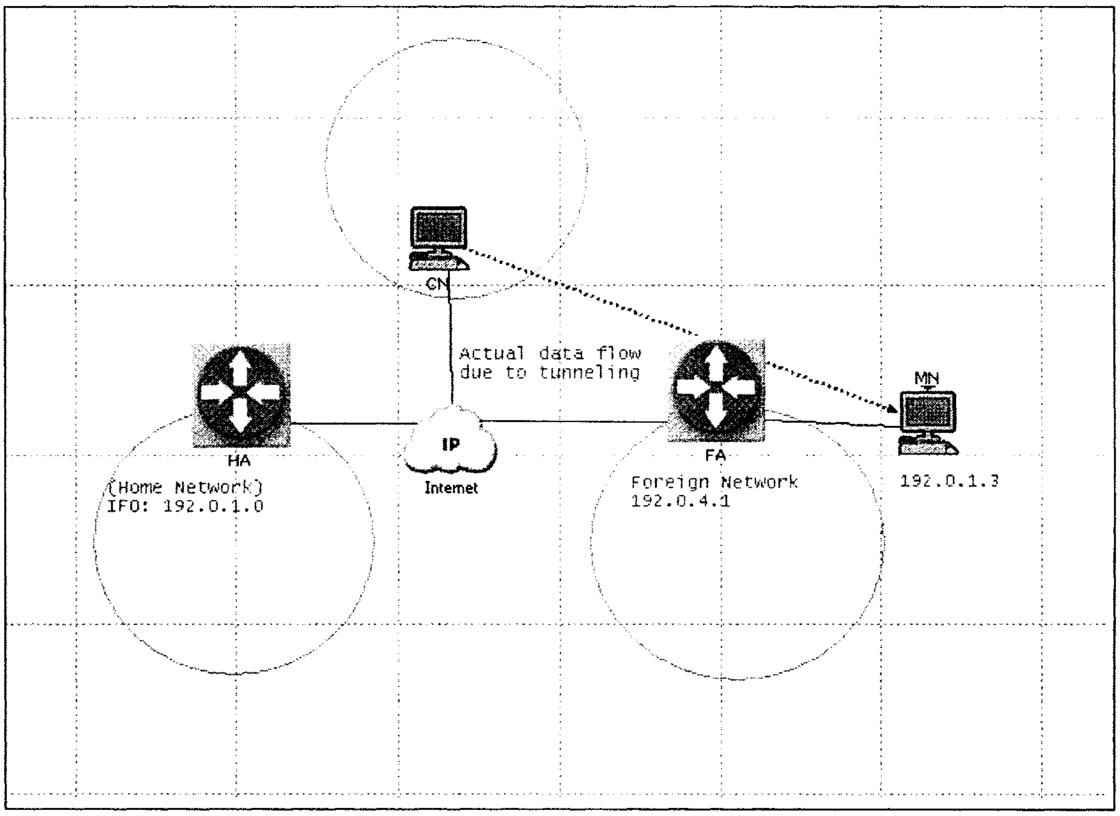

Figure 29: Wired MIP Topology

5.3.1.1 Wired MIP Topology

As shown in Figure 29, the wired network topology consists of IP cloud, HA, FA, CN, and MN. MN is moving from its home network with IP address prefix of 192.0.1 and has home IP address of 192.0.1.3. Later the MN is connected to foreign network with IP address prefix of 192.0.4.

CN, FA, and HA are connected via T1 links configured with distance-based propagation delay. CN and MN are workstation of type mip_ppp_wkst_adv model and HA and FA are router of type mip_atm2_ethernet32_frelay4_slip32 model. MN is connected to FA using a T1 link. IP node representing the Internet is an ip8_cloud model for which we set the packet latency as a constant 0.1-second, and packet discard ratio to 1% before running the simulation.

While connected to the foreign network MN acquires a temporary address called CoA. The MN registers its CoA with the HA, using the registration message exchange. The HA still advertise its reachability to the network prefix of the MN's home address.

For sending packets to the MN, CN send them to the HA. The HA intercepts the packets and tunnels them to the CoA of MN. The tunnel is created between HA and FA for the duration of data delivery.

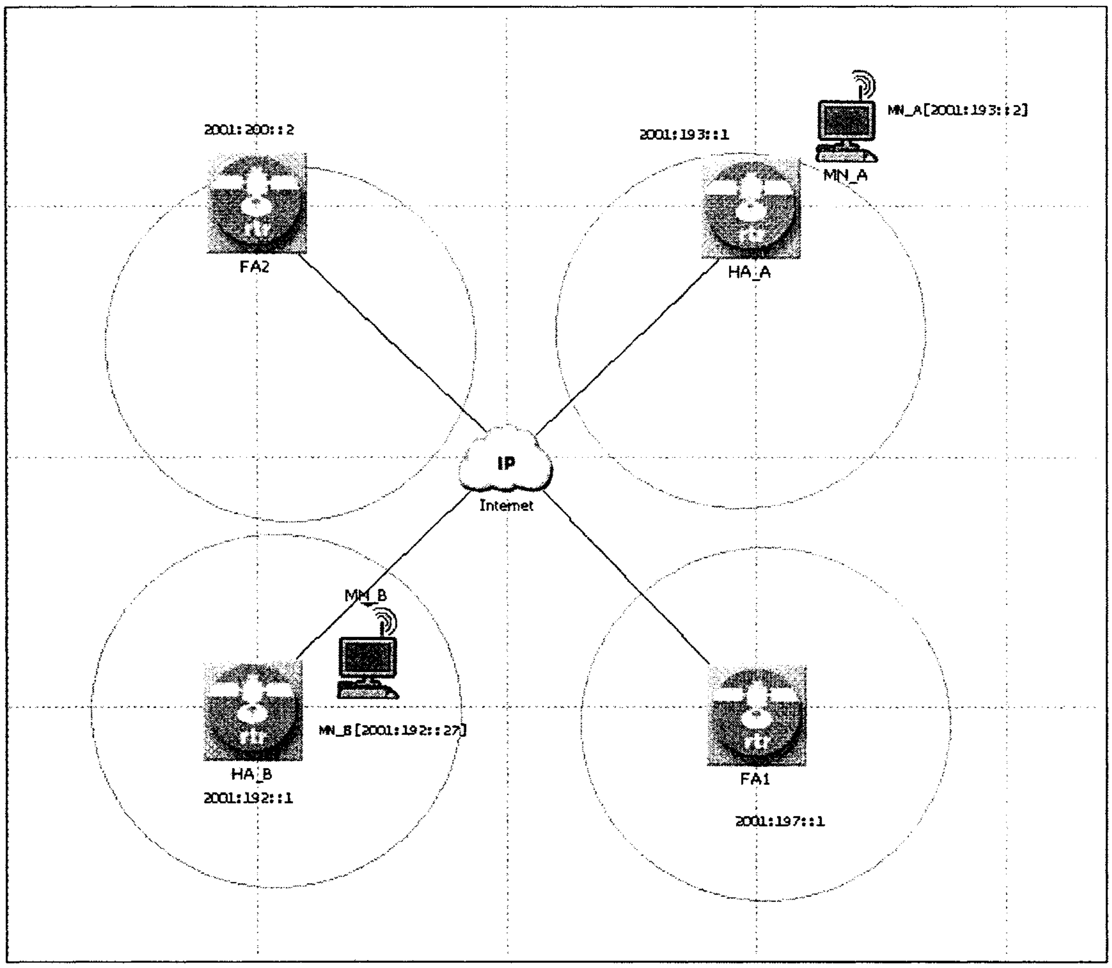

Figure 30: Roaming Topology

5.3.1.2 Roaming Topology

In a MIP topology shown in Figure 30, the IP network (ip8_cloud model) is connected to four Wireless Local Area Networks (WLANs). For each WLAN, there is one router acting as mobility agent (FA/HA). The core network is the IP cloud, with a constant latency of 0.1 seconds.

To show the roaming effects, the direction MN_A and MN_B will travel during the simulation run is defined in specific trajectories.

MN_A and MN_B are communicating with each other while moving on their pre-defined trajectories. They use MIP protocol to register their new location each time they move from one network to another.

The routers in this model are of type wlan_ethernet_slip4_adv node model, which are configurable as mobility-agent while acting as a wireless router. MN_A and MN_B are of type wlan_wkstn_adv that can roam across the network. Links that are connecting the nodes to each other are T1 links.

Node and process level description of MIP-ROQS model is provided in the Appendix.

5.3.2 MIP-ROQS Setup

The incremental approach for configuration and implementation of MIP-ROQS model is as follows:

- Configuration of Basic MIP model
- Configuration of video-conferencing service
- Configuration of RSVP on basic model
- Implementation of RO
- Configuration of RSVP with RO
- Implementation of ROQS module
- Configuration of RSVP with ROQS module

The detail of the implementation of MIP-ROQS model is provided in the Appendix.

5.4 Statistic Collection

Statistic collection is the act of measuring the values of certain characteristics/parameters in the network in pre-defined intervals and storing them in an output file. Statistic collection happens during simulation period from the start to the end of the simulation time. Collecting statistics at regular intervals shows the way the value of a parameter change during the course of the simulation. The statistics can be categorised in different groups. Statistics that are collected for the whole system are called global statistics. The one collected for specific objects is called node statistics. Furthermore, statistics can be collected for specific links, which are called link statistics.

5.4.1 Model Verification

The model of system evolved in an incremental approach and new features were added. As supported by Opnet, at the beginning, the purpose of running a simulation was to verify the system specification. Verification involved checking the following criteria:

- Progress: Observing simulation progress and processing events and reaching completion without encountering serious errors.

- Flow of Data: Observing basic flow of packets in the system model to ensure that packet generation, connections, and links are properly configured.
- Key Events: Observing particular events that are expected to occur at certain stage.

All the above were performed in each iterations of simulation run of different scenarios.

5.4.2 Available Statistics

The statistics that were collected during a simulation run are divided into four groups, Global, Node, Link and Demand statistics:

Global statistics

Global statistics are the end-to-end behaviour of different protocols or services running on the model. For example RSVP protocol statistic in the whole network is the total number of PATH messages sent in the network.

Node Statistics

For each node in the network, statistics are collected for different modules which indeed are different parts of protocols. For example, if chosen, on each node (HA, MN, FA, CN), RSVP related statistics are collected that are for example number of PATH messages sent by that specific node.

Link Statistics

Different attributes of a link, as a connection between two nodes can be collected. For example delay, throughput, etc, when chosen are collected for all the links in the network.

5.4.3 Collected Statistics

To show the impact of this contribution statistics were collected for different components of MIP architecture (service, core network and access network) in different available statistics explained above. Table 3 shows these parameters/metrics.

	Global Statistics	**Node Statistics**	**Link Statistics**
Service Video Conferencing	Packet delay variation Packet ETE delay Traffic received Traffic sent	Packet delay variation Packet ETE delay Traffic received Traffic sent	
Core Network	Background traffic delay Traffic dropped		Throughput → Throughput ← Utilization → Utilization ←
Access Network	Delay Load Throughput		
Protocols	**MIP:** Control traffic received Control traffic sent Tunnelled traffic overhead Tunnelled traffic received Tunnelled traffic sent **RSVP:** RSVP control traffic received RSVP control traffic sent	**MIP:** Control traffic received Control traffic sent Tunnelled traffic overhead Tunnelled traffic received Tunnelled traffic sent **RSVP:** RSVP control traffic received RSVP control traffic sent	

Table 4: Statistics Parameters

5.5 Simulation Run

Simulation run initiated discrete event simulation of the model configured with specific scenarios in order to collect a pre-established set of statistics and store them for further analysis. Duration of the simulation run and the seed were set. Duration of simulation was set based on the expected behaviour of the system. In the case of wired MIP topology, main controlling actions (Registration, resource reservation) are happening at the beginning of the simulation run and the rest of the simulation is the data transfer.

Therefore, within 15 minutes of simulation run time, the behaviour of the system can be observed. In the case of roaming topology, duration of simulation run was long enough so MNs can traverse their pre-defined trajectories, as this was the behaviour that needed to be observed. The scenarios that show the behaviour of the network in all possible combinations are detailed below.

5.5.1 Scenario 1: MIP-Tunnelling

MIP_tunneling is the most basic scenario. As described in chapter two (background), MIP creates a tunnel from HA to FA in order to encapsulate and send the packets received for MN. In both Wired and Roaming topologies, MIP_tunneling is configured by setting the necessary attributes of MN, HA, and FA. MN resides in foreign network but has a permanent address with the network prefix of home network. Furthermore, the service profile and definition is also configured.

5.5.2 Scenario 2: MIP-Tunnelling-RSVP

Duplicated from the MIP_tunneling scenario, and by setting RSVP attributes as explained in the "configuration of RSVP on basic model" Section in the Appendix, on both wired and roaming topologies, MIP_tunneling_RSVP scenario is built.

5.5.3 Scenario 3: MIP-RO

Duplicated from MIP_tunneling scenario, to build MIP_RO, it is sufficient to set the ROQS_RO attribute to enable. That triggered the routing optimization mechanism implemented in the model.

5.5.4 Scenario 4: MIP-RO-RSVP

Duplicated from the MIP_RO scenario, and by setting RSVP attributes as explained in the "configuration of RSVP on basic model" Section in the Appendix, on both wired and roaming topologies, MIP_RO_RSVP scenario is built.

5.5.5 Scenario 5: MIP-ROQS

Duplicated from MIP_tunnelling scenario, to build MIP_ROQS, it is sufficient to set the ROQS_Optimal Routing attribute to enable. This attribute has precedence over ROQS_RO attribute. This means if both RO and Optimal Routing attributes are set, the Optimal Routing will be performed. Enabling ROQS-Optimal Routing triggers the functionality of ROQS entity implemented in the Model.

5.5.6 Scenario 6: MIP-ROQS-RSVP

Duplicated from the MIP_ROQS scenario, and by setting RSVP attributes as explained in the "configuration of RSVP on basic model" Section in the Appendix, on both wired and roaming topologies, MIP_ROQS_RSVP scenario is built.

5.6 Simulation Analysis

By configuring individual statistics and running the simulation in MIP with tunnelling, MIP with RO and MIP with ROQS the necessary data for comparison were collected. The impacts of the MIP-ROQS solution is categorized and explained below based on different layers of network architecture, Application, Core network and Access network.

5.6.1 Impacts on the Application Layer

Figure 31 shows the comparison of the application ETE delay in MIP_Tunnel, MIP_RO and MIP_ROQS. The application ETE delay is calculated as the current simulation time minus the time when the packet was generated. As expected, we observe that the ETE delay is the lowest (less than 75% of MIP_RO) in MIP_ROQS and highest in MIP_Tunnel scenario.

This is due to the fact that optimal path selection has chosen the best path in terms of available resources for data delivery.

Table 5 summarizes different values of performance parameters measured for a video conferencing application in different scenarios. As it can be observed, using MIP_ROQS, packet delay variation has dropped in average from 0.0059 in MIP_Tunnel scenario and 0.034 in MIP_RO scenario to 0.0022.

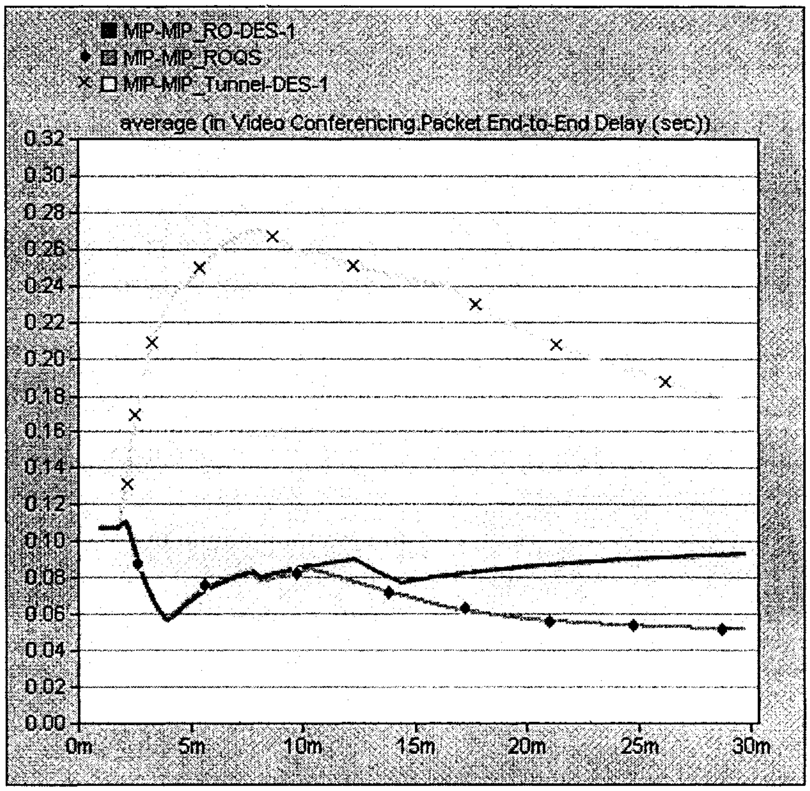

Figure 31: Application ETE Delay

Maximum ETE delay in ROQS scenario drops by 75% compared to the same metric in RO scenario and by 61% compared to Tunnelling mechanism.

This is clearly a major improvement in application response time which directly impacts service performance. While the traffic sent in three scenarios is the same, the maximum amount of traffic received is slightly higher in MIP_ROQS but in average slightly lower compared to pure tunnelling or RO scenarios.

	MIP_Tunnel			MIP_RO			MIP_ROQS		
Statistic	**Av.**	**Max**	**Min**	**Av.**	**Max**	**Min**	**Av.**	**Max**	**Min**
Packet Delay Variation	0.0059473	0.0099657	0.0000001	0.034043	0.091009	0.000000	0.0022110	0.0038489	0.0000001
Packet ETE Delay (sec)	0.17754	0.31238	0.10676	0.09602	0.48318	0.00537	0.09276	0.12313	0.00537
Traffic Received (bytes/sec)	622.45	688.00	0.00	629.71	688.00	0.00	618.72	697.56	0.00
Traffic Received (packets/sec)	3.6189	4.0000	0.0000	3.6611	4.0000	0.0000	3.5972	4.0556	0.0000
Traffic Sent (bytes/sec)	665.10	688.00	0.00	665.10	688.00	0.00	665.10	688.00	0.00
Traffic Sent (packets/sec)	3.8672	4.0000	0.0000	3.8672	4.0000	0.0000	3.8672	4.0000	0.0000

Table 5: Video-Conferencing Performance Measurement

5.6.2 Impacts on the Core Network

By comparing Tunnelling and RO it was observed that RO re-directs traffic from HA-IP link to CN-FA link by adding control traffic overhead. Furthermore when the packet travels through the tunnel not only has to go through a longer path, but also endures encapsulation and de-capsulation processing time at HA and FA.

On the core network, the amount of traffic dropped is one of the important aspects of network performance which as shown in Figure 32, it is the highest in RO scenario and lowest in ROQS scenario.

As shown in Table 6, the maximum traffic dropped in the case of ROQS mechanism is decreased by 30% compared to RO and 47.5% compared to Tunnelling mechanism.

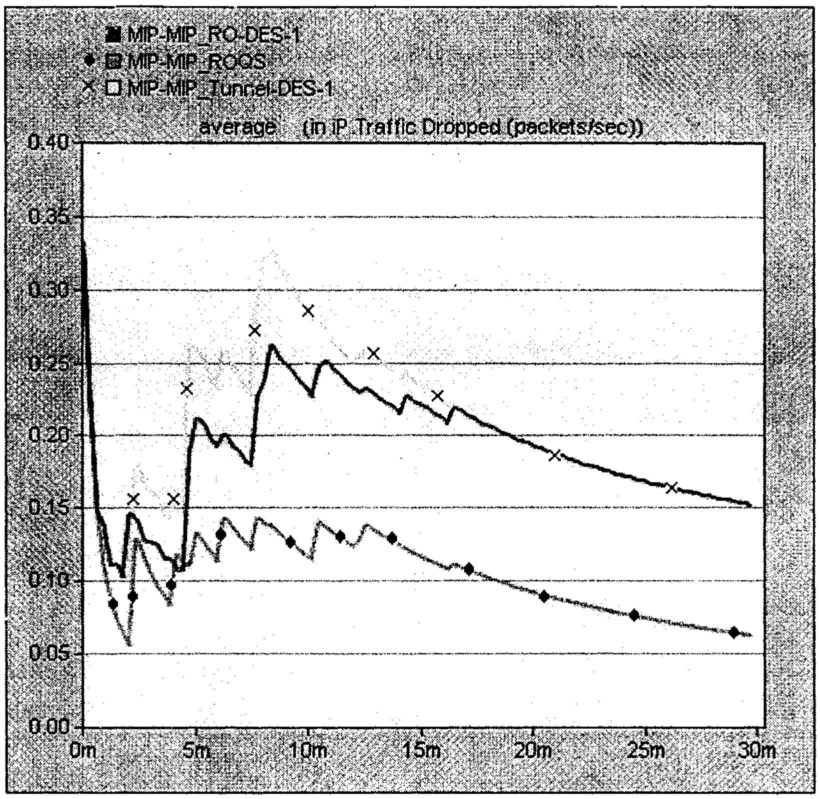

Figure 32: Core Network Traffic Dropped

	MIP_Tunnel			MIP_RO			MIP_ROQS		
Statistic	**Av.**	**Max**	**Min**	**Av.**	**Max**	**Min**	**Av.**	**Max**	**Min**
IPv6 Traffic Dropped (packets/sec)	0.1517	2.0000	0.0000	0.1517	1.5000	0.0000	0.0622	1.0556	0.0000

Table 6: Internet Performance Measurement

5.6.3 Impacts on MIP Performance

The average amount of control traffic as shown in Figure 33 is close for the RO and ROQS scenarios, while still slightly lower for ROQS method.

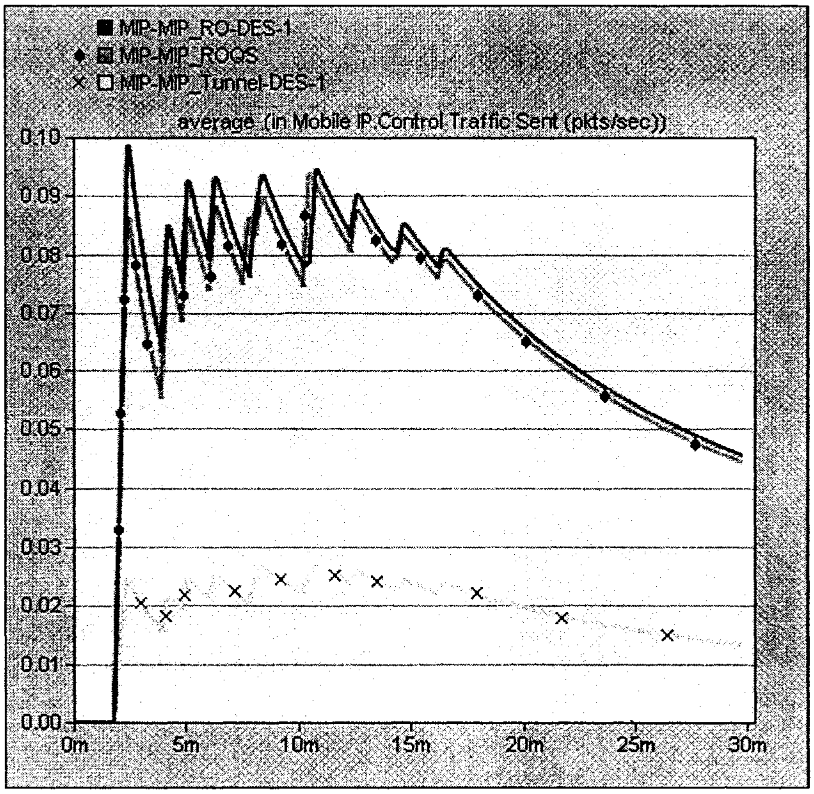

Figure 33: MIP Control Traffic Sent

As shown in Table 7, ROQS scenario has higher maximum amount of control traffic sent and received compare to RO and tunnelling scenarios while it improves the signalling overhead in tunnelled traffic overhead and tunnelled traffic sent and received. The maximum amount of control traffic sent in the case of ROQS mechanism is increased by 7% compared to RO case. The maximum amount of control traffic sent to the network is increased by 14% compared to Tunnelling scenario.

	MIP_Tunnel			MIP_RO			MIP_ROQS		
Statistic	**Av.**	**Max**	**Min**	**Av.**	**Max**	**Min**	**Av.**	**Max**	**Min**
Control Traffic Received (pkts/sec)	0.01333	0.11111	0.00000	0.04333	0.61111	0.00000	0.04333	0.66667	0.00000
Control Traffic Sent (pkts/sec)	0.01333	0.11111	0.00000	0.04556	0.66667	0.00000	0.04444	0.77778	0.00000
Home Agent Binding Delay (sec)	1.3854	1.6894	1.1339	1.4122	1.6145	1.2331	1.4455	1.6298	1.1882
RO Overhead (bits/sec)	0	0	0	0.13326	0.21818	0.10909	0.13258	0.21818	0.10909
RO Traffic Received (pkts/sec)	0	0	0	3,351.3	8,576.0	0.0	3,503.1	8,576.0	0.0
RO Traffic Sent (pkts/sec)	0	0	0	3,869.3	8,576.0	0.0	3,913.7	8,576.0	0.0
Tunnelled Traffic Delay (sec)	0.03280	0.34720	0.00234	1.9022	4.0000	0.0000	1.9267	4.0000	0.0000
Tunnelled Traffic Overhead (%)	0.090909	0.090909	0.090909	0.00896	0.10495	0.00102	0.00150	0.10570	0.00102
Tunnelled Traffic Overhead (bits/sec)	429.6	1,280.0	0.0	0.29289	0.35714	0.08197	0.30764	0.35714	0.08197
Tunnelled Traffic Received (pkts/sec)	2.4711	8.0000	0.0000	58.6	2,328.9	0.0	44.4	1,137.8	0.0
Tunnelled Traffic Sent (pkts/sec)	2.6850	8.0000	0.0000	59.2	2,222.2	0.0	45.1	1,137.8	0.0

Table 7: MIP Performance Measurement

5.6.4 Impacts on the RSVP Signalling

As shown in Figure 34, the amount of RSVP signalling in the case of MIP_Tunnelling_RSVP is higher and stays high as the PATH message is repeated.

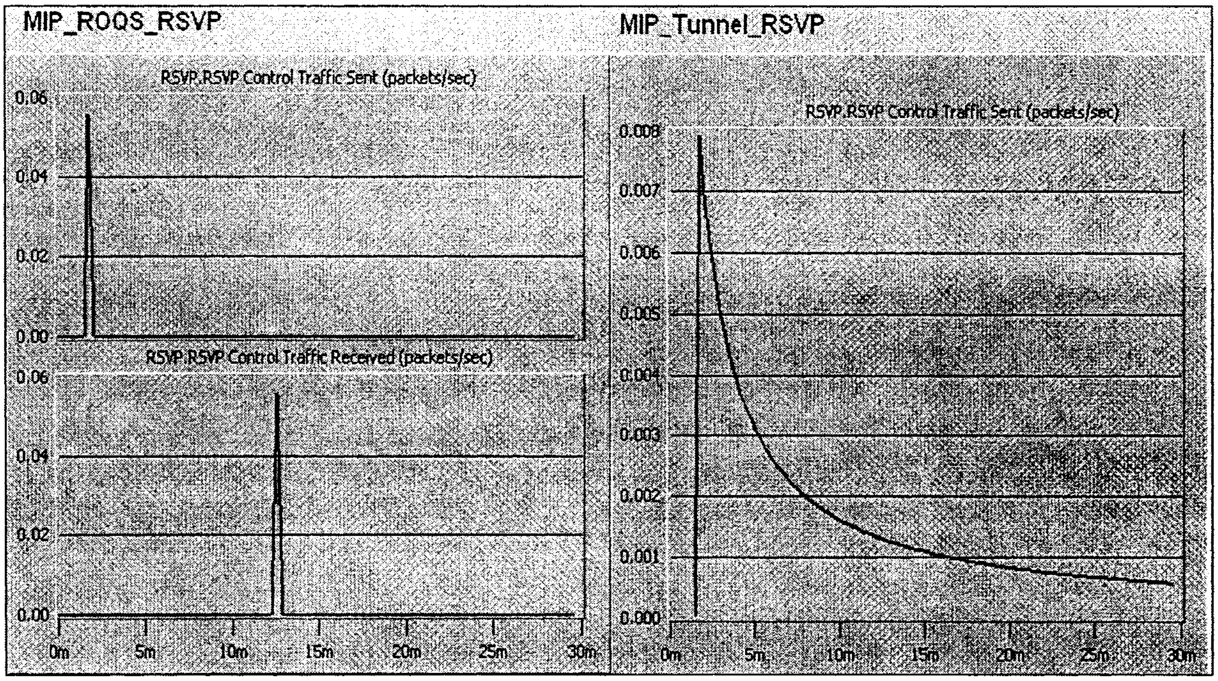

Figure 34: RSVP Control Traffic Received ROQS_RSVP/Tunnel_RSVP

The amount of RSVP control traffic that is sent for the MIP_Tunnel_RSVP scenario is very high at the beginning and eventually is updated in regular intervals.

As shown in Figure 35, similar to MIP_Tunnelling_RSVP, the amount of RSVP signalling in the case of MIP_RO_RSVP is higher and stays high as the PATH message is repeated. This is because the identity of the PATH message gets lost in the tunnel and RSVP protocol re-tries sending the PATH message.

As shown in Figure 36, the RSVP control traffic is high just at the beginning of the communication session and does not stay high compared to MIP_Tunnel_RSVP or MIP_RO_RSVP.

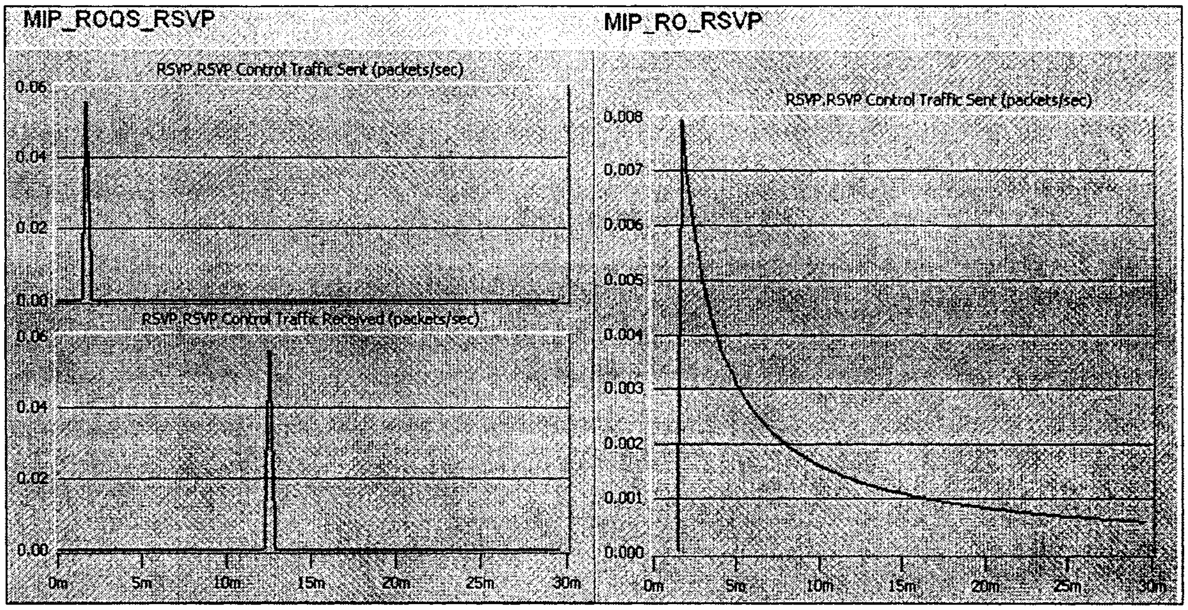

Figure 35: RSVP control Traffic Sent ROQS_RSVP/RO_RSVP

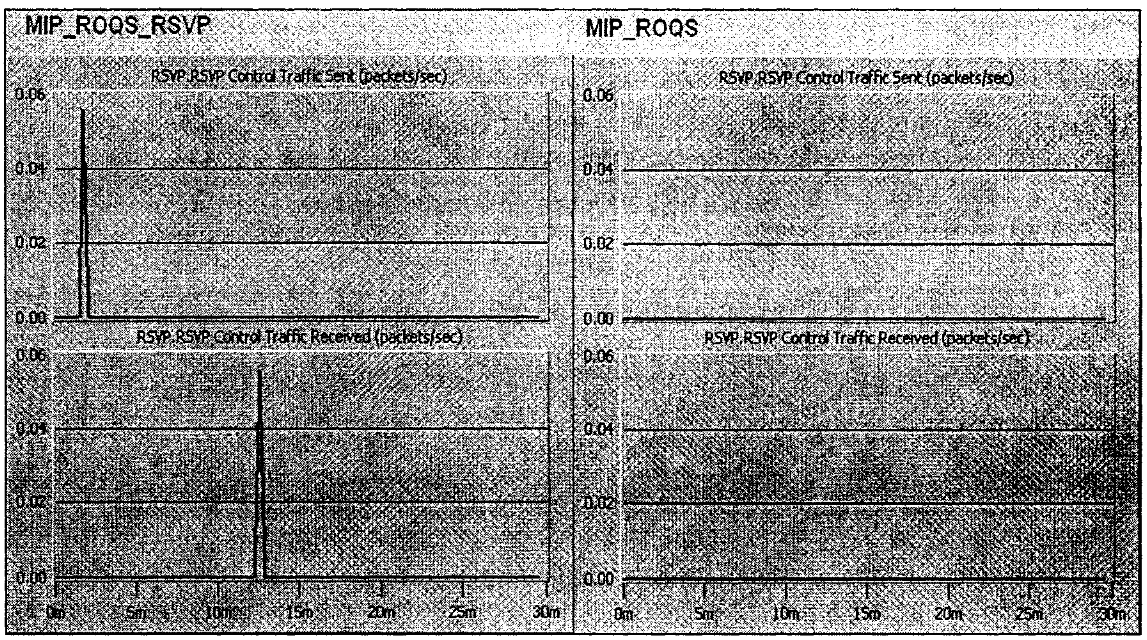

Figure 36: RSVP Control Traffic ROQS_RSVP/ROQS

5.6.5 Impacts on the Access Network

On the access network (WLAN), the delay is measured based on the time it takes to reach MN. As shown in Figure 37, the delay in access network is the highest in Tunnelling scenario and the lowest in RO scenario. In the case of ROQS scenario, the delay in WLAN is increased around 5 minutes after the start of simulation because the MIP control signalling starts at that time.

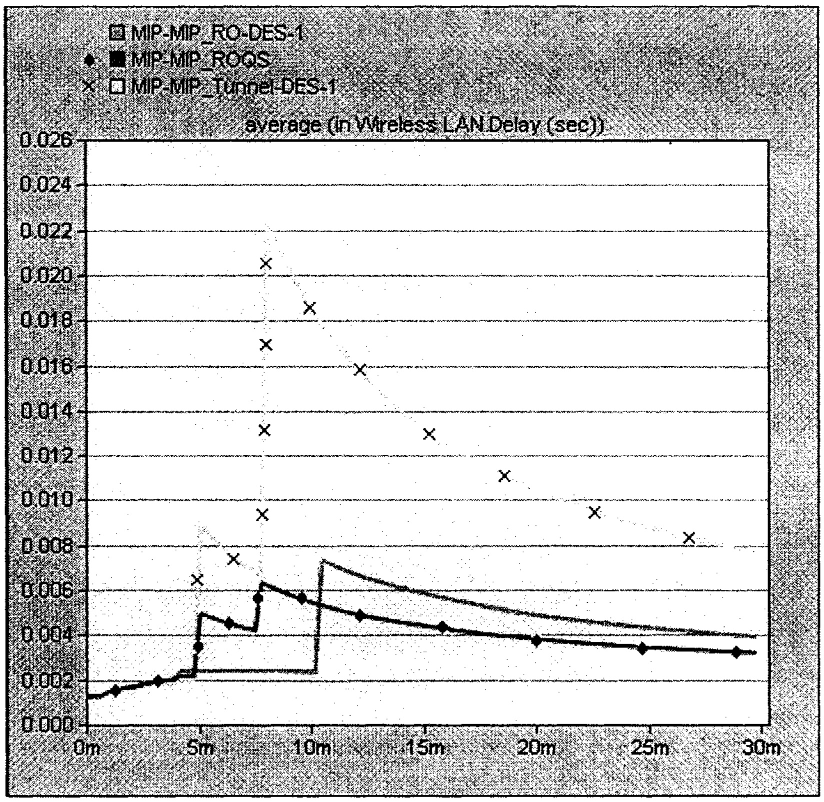

Figure 37: Access Network Delay

Considering the performance in access network in terms of the amount of control traffic sent to support mobility by routing the packets toward the destination, we achieved further interesting result with ROQS method.

Table 8 shows the access network measured delay, load, and throughput. Delay in the case of RQOS mechanism is dropped by 66% compared to RO scenario and by 84% compare to Tunneling scenario.

Throughput and delay in wireless LAN are calculated as:

Wireless LAN Throughput = Total number of bits sent to higher layer

Wireless LAN Delay = current simulation time – time the packet was generated

	MIP_Tunnel			MIP_RO			MIP_ROQS		
Statistic	**Av.**	**Max**	**Min**	**Av**	**Max**	**Min**	**Av**	**Max**	**Min**
Delay (sec)	0.00770	0.44578	0.00122	0.00394	0.18002	0.00124	0.003204	0.061265	0.001220
Load (bits/sec)	18,456	20,412	4,514	18,873	21,728	4,732	19,044	22,649	4,814
Throughput (bits/sec)	15,702	17,913	2,272	16,289	19,556	2,375	16,229	20,096	2,407

Table 8: Access Network Performance Measurement

The number of packets transferring control information (registration messages, and binding update to CN) is higher in the case of RO, simply because there is extra interaction with CN to inform it about CoA of MN. As it is expected, the number of control messages will go higher when resource reservation is used beside RO.

The amount of control traffic, which consists of registration, RO and resource reservation, is much higher in the case of MIP_RO_RSVP compared to MIP_Tunnel_RSVP scenario. The amount of control traffic is higher in the case of ROQS optimal routing.

When using ROQS, traffic is distributed between links HA-IP and CN-FA. In the case of MIP_RO the traffic is always blindly re-directed to CN-FA link and in the case of MIP_ROQS it is less in CN-FA and is more on IP_HA link

The amount of oontrol traffic sent in MIP_RO is higher than MIP_ROQS. This is because optimal route selection only in the case it selects direct route as optimal route, will send control traffic to inform CN about CoA of MN.

However, it is still noticeable that in the case of ROQS_RSVP the amount of control traffic is higher compared to the case when resource reservation is not needed.

Tunnelled Traffic Overhead = packet size – (inner packet size), which is zero in the case of RO and in the case of tunnelling the packet size is increased by the length of new addresses. Obviously, the amount of tunnelled traffic is much higher in the MIP_Tunnelling scenario. Tunnelled Traffic received is the number of packets per second that are encapsulated in HA and passed through the tunnel.

In the case of MIP_Tunnel_RSVP, the number of tunnelled packets is higher, and in MIP_RO_RSVP, close to zero, as only the first packet toward MN is encapsulated. In the case of MIP_Tunnel_RSVP, the amount of tunnelled packets is higher as the PATH message is re-sent several times.

In the case of RO, the first packet from CN activates RO and the CN is informed about actual location of MN, so it continues with direct communication to MN. While including RSVP with MIP_Tunnel or MIP_RO, it is expected that PATH message looses its identity in the tunnel and as the result no resource reservation is made. In the case of ROQS, the first resource reservation message activates the optimal path selection and another resource reservation will be performed on the direct route from CN to the MN.

As it was expected the tunnelling traffic and overhead dropped. In the case of MIP_ROQS_RSVP, the resource reservation is successful and the ETE delay for video conferencing is reduced.

As explained previously, MIP enhanced with ROQS optimal routing resolves the conflict in providing simultaneous resource reservation and RO, therefore, as it is expected resource reservation happens smoothly and the RSVP control messages sent and received do not lose their identity and actually reserve necessary resources on the path. This also allows the service to have less ETE delay.

5.7 Conclusion

In this chapter evaluation of MIP-ROQS solution was presented. Evaluation consisted of creating a model of the system, selecting appropriate metrics, building different scenarios, and comparing the result of simulation runs against each other. MIP-ROQS model was compared to MIP-Tunnelling and MIP-RO. In an incremental approach, the impacts of RO, resource reservation and ROQS optimal routing were examined.

5.7.1 Impact of RO

By comparing different aspects of network behaviour in the cases of MIP_Tunnelling and MIP_RO it was observed that RO re-direct traffic from HA-IP link to CN-FA link while adding control traffic overhead. RO also did not allow resource reservation and as result, the maximum ETE delay of application is actually increased while the amount of control traffic is increased.

5.7.2 Impact of Resource Reservation

By comparing MIP_Tunnelling versus MIP_Tunneling_RSVP, MIP_RO versus MIP_RO_RSVP, and MIP_Tunnelling_RSVP versus MIP_RO_RSVP, we observed the impact of attempting to reserve resources in original MIP that uses tunnelling and MIP enhanced with RO. The amount of control traffic was increased while the ETED for service remained the same.

5.7.3 Impact of ROQS

To show the impacts of ROQS optimal routing several comparisons were conducted. MIP_ROQS versus MIP_Tunnelling and MIP_RO to show the link utilization and ETE service delay even without resource reservation.

Comparison of MIP_ROQS_RSVP versus MIP_Tunnelling_RSVP and MIP_RO_RSVP shows the impacts on the control traffic overhead and the ETED of service.

Comparison of the proposed MIP-ROQS solution with the existing solutions presented in Chapter 2 based on the same metrics presented in Table 1 is as follows:

Applicability: MIP-ROQS applies to the total network including access network and core network.

Routing Method: Tunnelling or bi-directional only when it is optimal.

QoS Method: The solution is developed based on the restriction of IntServ/RSVP conflict with RO; however it applies to any similar resource reservation signalling

Assumption: There is no specific assumption about the network status in this solution

Drawback/overhead: As shown in this chapter, control traffic sent and received is much less than the original MIP but still not zero. RSVP control signalling is the minimum required for reserving resources.

Advantages: It removes the conflict in simultaneously providing QoS and RO, therefore provides QoS for application, while reducing signalling overhead. It increases network performance in terms of resource utilization while preventing congestion in some part of the network.

5.7.4 Evaluation of MIP-ROQS Solution vs. Derived Requirements

In Chapter 3, a set of requirements were derived based on evaluation of existing methods. Now, after presenting the new solution in Chapter 3 and 4 and discussing its performance evaluation in this chapter, the same requirements are assessed as follows:

1) The solution resolves the conflict between providing QoS and RO. As we can see in Figure 35, the reservation signalling is sent and received at the start of the session and there is no need to repeat the PATH message.
2) The solution provides the optimal route based on the network condition while guaranteeing the required QoS. This is based on processing sequence of MIP-ROQS, and the fundamental nature of this solution.
3) The solution is an end-to-end solution as it includes all the layers of MIP environment consisting of application, core network and access network. This is based on the fact that the characteristics of all the elements are collected to feed the optimal path selection mechanism.
4) The solution improves the application end-to-end delay. This was observed in Figure 31 and Table 5.
5) The solution improves core network by reducing control traffic. This was demonstrated in Figure 32 and Table 6, as the reduction in the amount of traffic dropped in the IP network.
6) The solution improves access network performance. This was shown in Figure 36 and Table 8.
7) The solution does not increase the signalling overhead. As shown in Table 6, 7, and 8 the average amount of signalling is reduced in different part of the network. This is also further confirmed in Figure 31, 32, and 33 by reducing the amount of RSVP signalling.

8) The solution takes into account the capability of different entities involved in the session as part of the data collected from the network.

9) The solution improves the performance of the MIP protocol. As shown in Table 7, the average amount of signalling is reduced.

10) The solution does not interfere with the base functionality and only is triggered when the necessary data is available.

CHAPTER 6

INTEGRATION OF MIP-ROQS AND MULTI PROTOCOL LABEL SWITCHING

6.1 Introduction

In this chapter, as an extension of ROQS solution, we introduce the cooperation of MIP enhanced with ROQS QoS constraint based routing mechanism and

MPLS architecture. In MIP-ROQS after selection of the optimal path, we rely on the original RSVP mechanism in order to perform the resource reservation and implicitly label the path by marking all the routers on the way from source to the destination by sending the PATH message and receiving back the RESV message. RSVP though is claimed to lack scalability for acting in IP-based networks.

Proposing the integration of MPLS and MIP enhanced with ROQS mechanism for selecting the optimal path, MPLS labelling mechanism can be used for marking the optimal path. MPLS labelling mechanism is faster than routing table lookup. Using MPLS labelling mechanism and RSVP-TE [84] for resource reservation on the optimal path, we overcome the scalability issue of IntServ mechanism.

This integration makes IP-in-IP tunnelling in data forwarding redundant. MPLS is used to switch packets and label the optimal path over which the resources are reserved using RSVP-TE [84]. Using this approach, the transmission delay, and packet processing overhead is reduced.

MPLS has the mechanism to label the path but it does not decide which path is the optimal path. MPLS replaces IP forwarding by a simple label lookup mechanism. MPLS combines layer three routing and layer two switching, which enhances network performance in terms of scalability, complexity, latency, and control message overhead. Its usage seems to become more and more widespread. ROQS Constraint Based Routing (CBR) fits easily into MPLS network, which can be considered, as portability of ROQS solution.

By integrating MIP-ROQS and MPLS, we use the benefits of both protocols. MIP is taking care of the mobility support, ROQS mechanism takes care of identifying the optimal path, and MPLS deals with traffic engineering and labelling the optimal path and resource reservation on the optimal path. This improves the network performance in terms of resource handling and guaranteeing QoS.

6.2 Overview of MPLS

Traffic distribution in the network can be unified using Traffic Engineering (TE). Performing TE using MPLS [122] moves traffic to where the resources are available. Constraint based routing (CBR) algorithms beside MPLS are responsible for finding a path in the network for each traffic flow that is able to guarantee certain quality parameters such as bandwidth, delay, delay variation, bit-error rate or packet loss.

A CBR algorithm can be dynamic or static. When a new connection request arrives at the edge of the network, a dynamic CBR algorithm can select a route to satisfy the QoS for that incoming request. Using static CBR algorithms, the routes can be established prior to data transmission.

In MPLS, data transmission occurs on Label-Switched Paths (LSPs). The LSPs are the paths selected by CBR algorithm. Label creation can be topology-based (OSPF [121], BGP [121]), request-based (RSVP) or traffic-based (upon reception of a packet). Shown in Figure 38, the sequence of events in MPLS network is as follows:

- Label creation and Label distribution
- Table creation
- Label switched path creation
- Label insertion/table look-up
- Packet forwarding

The devices that participate in MPLS protocol mechanisms can be classified into Label Edge Routers (LERs) and Label Switching Routers (LSRs). An LSR is a high speed router device in the core of an MPLS network that participates in the establishment of LSPs using the appropriate label signalling protocol.

An LER is a device that operates at the edge of the access network and MPLS network. LERs support multiple ports connected to different networks and forward this traffic into the MPLS network after establishing LSPs. LER mainly assigns and removes the labels.

In explicit routing which is similar to source routing, LSR specifies the list of nodes through which the flow must traverse. This path can be non-optimal as well. However, if the path is chosen particularly based on certain requirements, it eases traffic engineering throughout the network.

Another main concept in MPLS is the Forward Equivalent Class (FEC). FEC is a representation of a group of packets that share the same requirements for their transport. All packets in such a group are provided the same treatment on route to the destination. A label in its simplest form identifies the path a packet should traverse.

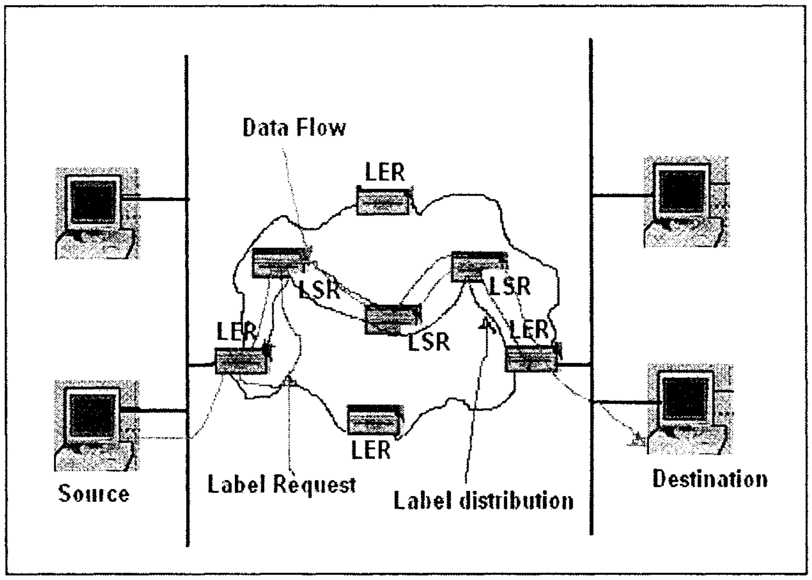

Figure 38: MPLS Network Components

A label is carried or encapsulated in a layer-2 header along with the packet. The receiving router examines the packet for its label content to determine the next hop. Once a packet has been labelled, the rest of the journey of the packet through the network is based on label switching.

The label should be bound to an FEC as a result of some event or policy that indicates need for such a binding. These events can be either data-driven bindings or control-driven bindings.

6.3 MPLS Usage for Mobility and QoS

Authors of [35-41], [91], [94], and [97] are proposing enhancement of MPLS network with mobility to address Traffic Engineering (TE) and provide QoS in mobile environment.

The proposal in [35] is an architecture for MPLS and MIP at the same time. In this architecture FA and HA are edge Label Switch Routers (LSRs) and reside in the same MPLS domain.

When the MN moves to a foreign network and gets a CoA from FA, the FA configures its entries, then forwards the registration to the HA as in original MIP. HA adjusts its MN's binding and issues the Label Distribution Protocol (LDP) request to the FA and the LSP is eventually set up. When the MN visits a new FA, the same procedure is repeated.

Proposal in [37] is an extension of [35] where a hierarchical architecture is set in the network and the whole network is MPLS-aware. There is a two level hierarchy with a new mobility agent, called the foreign domain agent (FDA), at every MPLS domain.

Although [35] and [37] are valuable proposals in terms of architecture, they leave out the necessary CBR algorithm specifically for MIP environment. MIP is a protocol for hybrid wired-wireless IP network that is impacted by both access and core network conditions and can not be treated as simple IP network.

Therefore, it requires specific optimal path selection algorithm which is built based on its specific nature. Here is where MIP-ROQS solution can play an important role. Figure 39 shows the message sequence based on the MIP/MPLS integration proposed in [37].

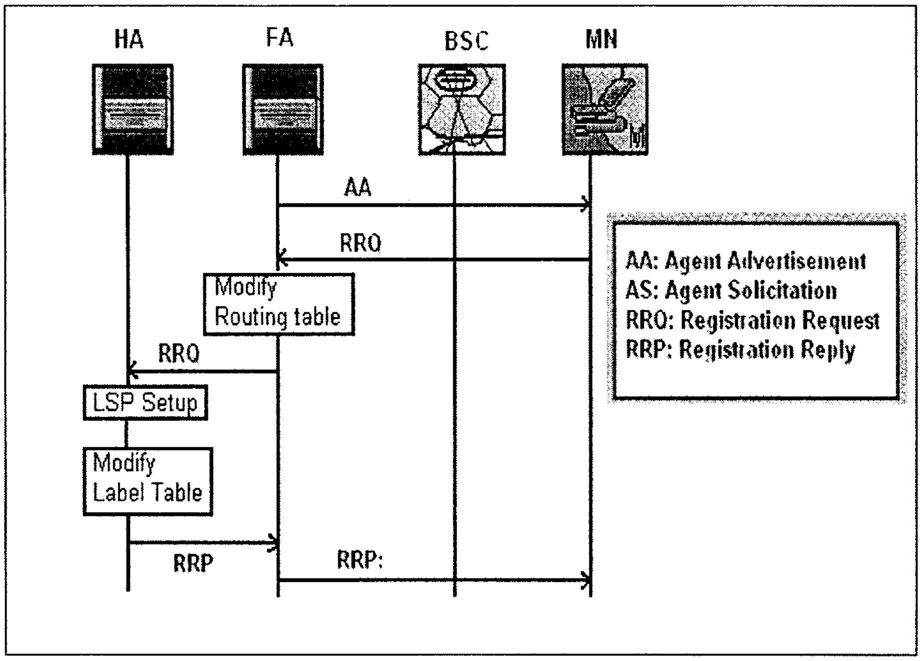

Figure 39: MIP/MPLS Integration

6.4 Cooperation of MIP-ROQS with MPLS

The cooperation of MIP and MPLS protocols needs a management interface for invoking the proper module at proper time. This interface also interacts with normal IP routing module. As described earlier in MIP-ROQS, HA is enhanced with a dynamic CBR algorithm. The place of this algorithm in the MIP/MPLS network is shown in Figure 40. Inputs to CBR process include traffic attributes, wired and wireless links' conditions, resource specification, and service requirements.

This process is incorporated in layer three of the protocol stack (network layer). As explained before, the CBR process can be implemented in each router acting as a mobility agent and co-exist with current routing protocol. ROQS CBR has interconnection with MIP, MPLS, and RSVP-TE.

The MPLS architecture does not assume a single label distribution protocol. RSVP-TE [84] developed by IETF is an extension to RSVP for establishing label switched paths (LSPs) in MPLS networks. In the integration of MIP-ROQS/MPLS, RSVP-TE is used to reserve necessary resources on the optimal path.

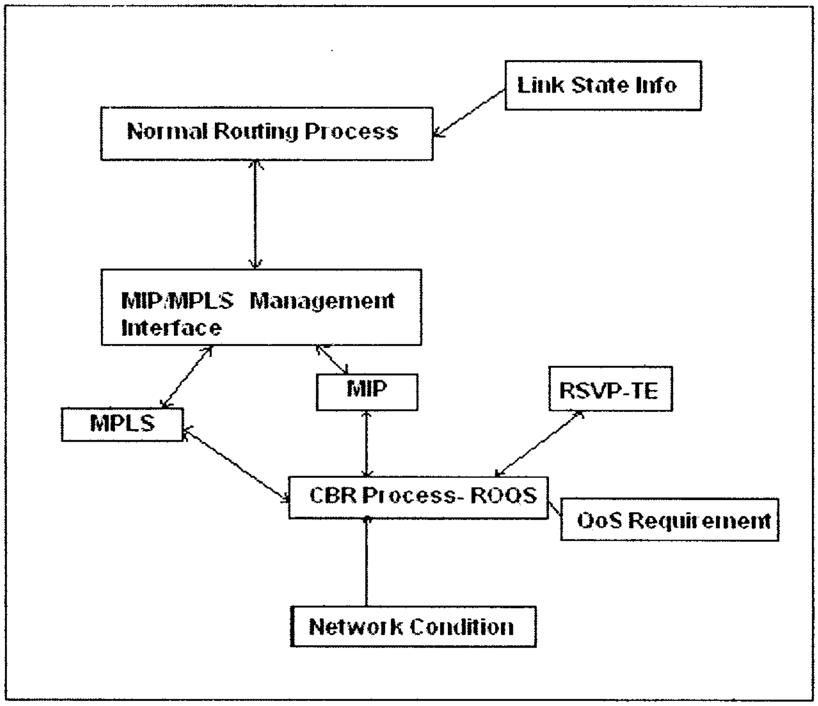

Figure 40: MIP-ROQS/MPLS information flow

6.5 MIP-ROQS/MPLS Event Sequence

We assume a network topology as shown in Figure 41. Home network and foreign network have their edge router acting as a mobility agent (for MIP) and edge LSR (for MPLS). Furthermore, the edge router on the correspondent network's edge acts as an edge LSR (for MPLS). The optimal path that is found in MIP-ROQS needs to be labelled for the incoming data flow.

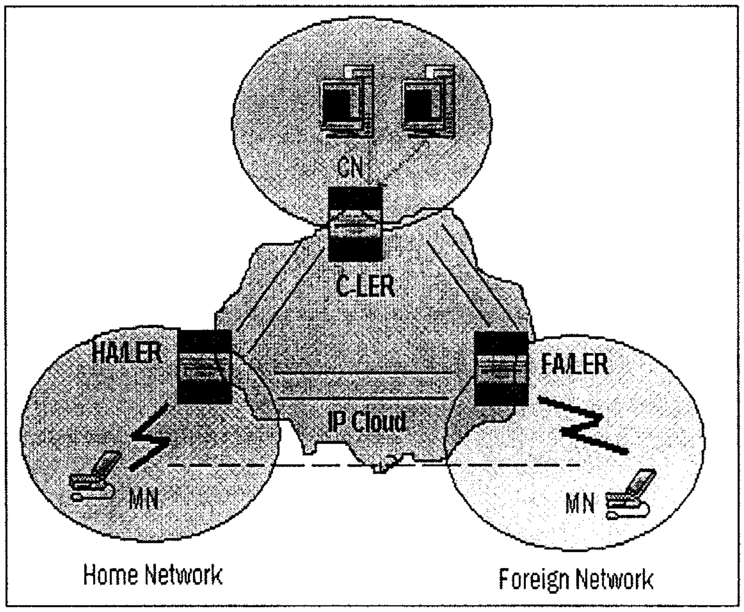

Figure 41: MIP/MPLS Network Topology

HA and FA acting as Label Edge Routers (LER) receive an IP datagram (an unlabeled packet) with a destination address and map it to Forwarding Equivalent Class (FEC) and assign a label to the packet and forward it to the next hop on the LSP. The rest of the routers on the path acting as LSR will not perform the normal layer 3 routing mechanism but simply forward the packets by label swapping.

The sequence of events in the integration of MIP-ROQS and MPLS technologies can be divided into two stages: Registration, and Data delivery.

6.5.1 Registration

Registration process shown in Figure 42 is as follows:

1) (MIP) HA and FA advertise their presence by sending agent advertisement (AA) messages
2) (MIP) MN receives AA and detects that it is in a foreign network
3) (MIP) MN receives a CoA from FA
4) (ROQS) MN adds its QoS capability to the RRQ and sends it to FA via the BS
5) (ROQS) BS controller as the representative of the wireless network is aware of the conditions of assigned radio link to this MN, it adds the capability of radio link to RRQ before delivering it to FA
6) (MPLS) FA/LER receives RRQ and configures its entries in the label distribution table
7) (MIP) FA forwards the registration request to the HA using regular IP routing
8) (MIP) HA receives RRQ and updates its LER table
9) (MIP) HA updates its binding table with CoA
10) (ROQS) HA retrieves all the QoS parameters from RRQ and send RRP to FA

6.5.2 Data Delivery

Data delivery process shown in Figure 43 is as follows:

1) (MIP) When a QoS-sensitive service on the CN wants to send data to the MN, it sends a PATH message (RSVP protocol) to HA via the LER on the CN's home network.

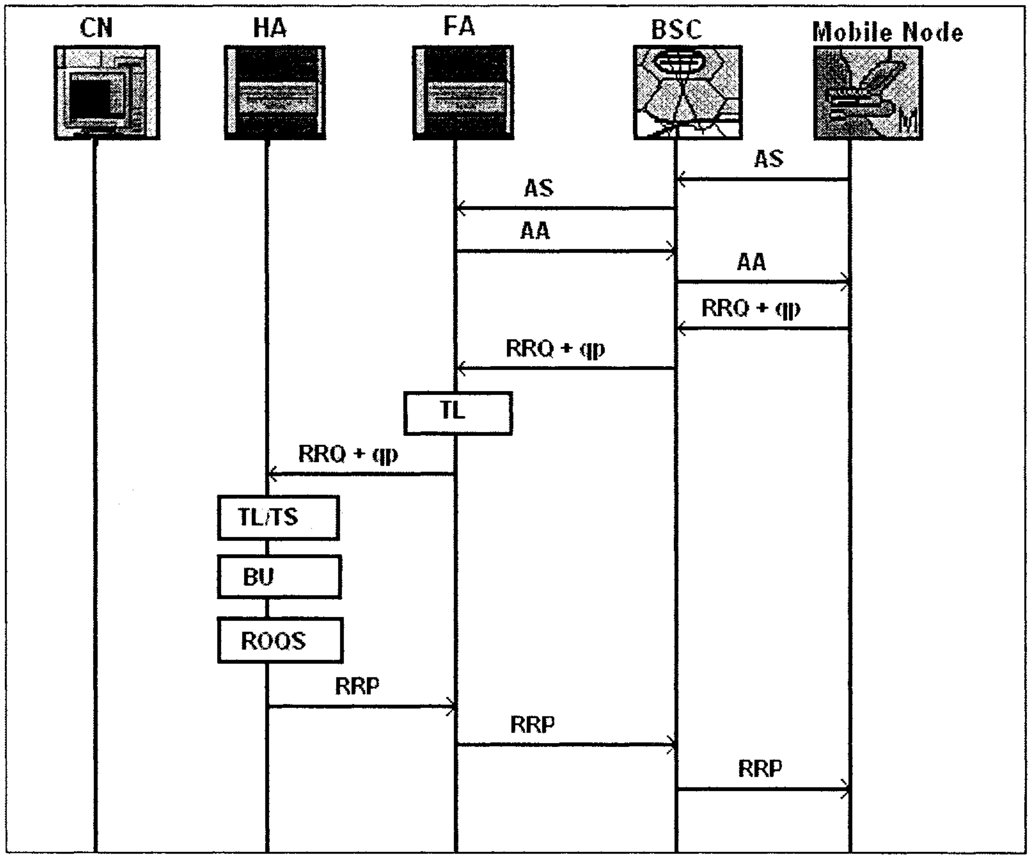

Figure 42: MIP-ROQS/MPLS Registration

TL: Table Look up

TS: Table Setup

2) (ROQS) HA receives the PATH message along with the specification of the link from CN to HA and routing table of Correspondent LER

3) (ROQS) HA retrieves the data and calculates the optimal path

There are two possible cases:

Path via HA is the Optimal Path:

4) (MPLS) HA issues a label creation request to FA
5) (MPLS) Label Switch Path (LSP) is set up between HA and FA.
6) (MPLS) Data flow happens on LSP

Path from CN to MN is the Optimal Path:

4) (MIP) HA sends back a binding update to CN
5) (MIP) CN will update its binding table
6) (MPLS) LER on the edge of the Correspondent network will update its label lookup table
7) (MPLS) Correspondent-LER will issue a label creation request to FA
8) (MPLS) LSP is setup between CN and FA
9) (MIP) When the MN visits a new FA, registration event sequence is repeated

In the new proposal, label creation and LDP are performed dynamically after reception of PATH message for initiating a new packet flow as opposed to after registration stage [2] shown in Figure 39. At this stage, the quality parameters (QP) of both sides of the communication have been collected and a more accurate decision on optimal path can be made.

Furthermore, a registration stage is not necessarily followed by a data flow, and setting LSP after registration stage could become a wasted signaling overhead in the network. MN may move to many foreign networks and update binding table at HA before an actual data flow happens.

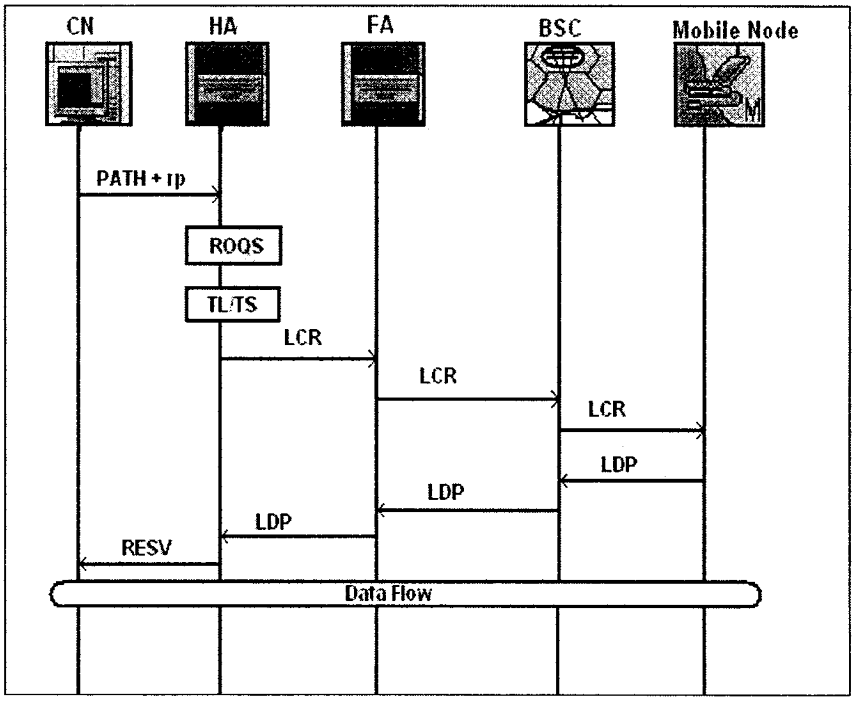

Figure 43: MIP-ROQS/MPLS Data Delivery

LCR: Label Creation Request

LDP: Label Distribution Process

6.6 Conclusion

In this chapter we showed the integration of MIP-ROQS and MPLS networks and presented the details of event sequence in registration and data delivery stages. In this enhancement to the main proposal research, we deal with traffic congestion in MIP based on the cooperation of MIP-ROQS and MPLS network architectures.

The way it works is that in MIP-ROQS/MPLS, the optimal QoS-constraint-based path is obtained for each data flow and MPLS explicit routing is applied to label the path. The ROQS CBR not only takes into account the characteristics of the core network but also considers the radio link conditions and mobile device capability before choosing the optimal path.

The way it works is that the forwarding process is done at the MPLS layer and HA does not need to use IP layer for IP tunnelling. This also increases the scalability of MIP. Since there is less overhead per packet handling, in total traffic overhead is reduced.

The benefit of this integration is that by optimizing the forwarding path, the end-to-end delay of packet delivery is reduced which increases the network performance in terms of resource handling and signalling overhead. Using this approach, packet processing overhead is reduced.

In previous chapter we evaluated MIP-ROQS solution through simulation and showed that it improves application, core and access network performance. On the other hand as stated by [35], HA processing delays in MIP increase with the increase of routing table size. But in MIP/MPLS integration, the HA processing delay is almost constant and lower than the processing delay in MIP. The lower value is the result of having the entire HA data forwarding process executed in the MPLS layer. No IP routing table search is executed. Since label table search is much faster than the routing table search, HA performance is much improved.

The same argue applies for the throughput and round-trip delay which are constant in the integration of MIP/MPLS as opposed to increasing with the table size in MIP. Routing table search is simply more time consuming for lager routing table, which impacts the processing delay at HA. However in MPLS, the time needed for label look up is constant.

Besides the improvements in processing delay at HA, TCP performance and round trip delay in MIP and MPLS, by enhancing MIP with ROQS mechanism, other aspects of network performance can be optimized. This is because, MIP-ROQS improved application, core network and access network performance by choosing the optimal path and diverting the traffic accordingly.

Therefore, the integration of MIP-ROQS and MPLS enhances the application and network performance further as the combined improvements based on MIP-ROQS (presented in Chapter 4) and MIP/MPLS integration.

CHAPTER 7

SUMMARY AND FUTURE WORK

In this chapter we summarize the research issue and contributions, then we discuss potential future work.

7.1 Contributions

Considering the triangle routing issue in MIP, if IntServ mechanism and RSVP signalling are used to reserve resources from the source to the destination some problems will arise:

1) The PATH message after reaching the HA will be encapsulated and go via a tunnel toward the FA to reach MN. Inside the tunnel the characteristics of the messages are hidden and therefore no reservation can be made.
2) The RESV message will go from different direction to the CN
3) The resources are reserved on a path different from the path actual data will be delivered on. Therefore, packets sent along the triangle path will receive different treatment than the ones sent via direct path.

When changing triangle route to an optimized/direct route in order to avoid the above issues, there is no guarantee the direct route is optimal in terms of necessary resources to grant the QoS. To resolve this problem, we combined the objectives of providing QoS and RO in MIP to one sole objective. The solution was a cross-layer design that bridged the gap to certain extent between different layers of MIP protocol architecture.

MIP network follows Internet layered architecture, which consists of application, Transport layer, IP layer and physical layer. The new solution collects QoS related parameters from all layers of MIP network. This data is carried over existing signalling of mobility, routing, and resource reservation protocols and passed to an entity for providing RO and QoS (ROQS).

This entity performs position and routing analysis and selects the optimal path. The chosen path will be the path over which the resources are reserved for the duration of the communication session. We designed, built, and evaluated MIP-

ROQS architecture and showed that it improves application performance and resource utilization in the access and core networks.

We enhanced the main proposal to integrate MIP-ROQS and MPLS network. This integration addressed scalability concerns of RSVP and MIP. Using MPLS labelling mechanism optimal path selected by MIP-ROQS is labelled.

The contribution emerges primarily as a solution for guaranteeing QoS over optimal routes in MIP. This contribution can be detailed as follows;

- Establishing a set of requirements for providing QoS over optimal routes in MIP environment
- The notion of QoS Constraint-based routing in MIP
- A Cross-layer MIP architecture with :
    - A new information flow
    - A new entity that embraces:
        - Constraints based optimal path selection finite state machine
        - Position analysis algorithm
        - Routing analysis algorithm
- Performance evaluation of MIP-ROQS through simulation
- MIP-ROQS/MPLS integration

With these contributions we achieved:

- Resolving the conflict between providing QoS and RO, and
- Improving application and network performance

The content of Chapter 3 was presented in WAMICON2006 and published in proceeding of this conference under the title of “A Cross-Layer Design Technique for QoS over Optimized Routes in MIP” [124]. More detailed version of the new solution along with the result of performance evaluation has been submitted under the title of “Guaranteed QoS for MIP-based Multimedia Services” [125] for publication to Elsevier journal on Computer Networks. The content of Chapter 6 has been submitted under the title of “QOS Constraint-based Routing in MIP/MPLS Network” [126] for publications to ISCIT2007.

7.2 Future Work

The course of this research delivers a different perspective on designing network protocols. There are six main evolutionary paths stemming from this work:

- **Security consideration with cross-layer design:** Security in mobile environment is a significant issue that is interesting to study without and with ROQS solution. The observations so far indicates that security issues can be addressed well with a cross-layer design as it is impacted by all layers of network architecture.
- **Impact of different handoffs:** Handoff mechanisms such as inter-domain, intra-domain, hard handoff, and soft handoff in mobile environment are always of interest to the wireless community. How different types of handoff can be handled or whether they affect the

network performance in presence of ROQS solution can be further investigated.

- **Extending the constraint-based routing algorithm to take into account the future traffic predictions in the network:** One of the ideas during development was whether we can have the optimal path selection not only based on the current status of the network but also based on the predictions of future traffic. This possibly can improve further the unified resource utilization in the network. It also can reduce the need for re-selection of the optimal path.

- **Static ROQS CBR:** ROQS optimal path selection is currently a dynamic CBR, it will be worthwhile to examine an alternate version of this algorithm as a static CBR, where based on the prediction of network status and link conditions the optimal path can be selected in advance.

- **Examining MIP-ROQS solution in an hierarchical network architecture:** To develop a new model we chose a simple network topology, if we choose a hierarchical network topology where the network is divided in domains and sub-domains, and domain gateways are acting as the major players in supporting mobility and QoS, theoretically we can further reduce the signalling overhead in the network. This can be examined by modelling and network performance measurement.

REFERENCES

[1] Klerer, S.M., The OSI management architecture, an overview, IEEE Network Magazine, March 1988, Vol. 2, No 2.

[2] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification. September 1997, RFC2205.

[3] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, An Architecture for Differentiated Services, IETF RFC 2475, December. 1998.

[4] W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zom, B. Palter, Layer Two Tunnelling protocol, L2TP, RFC 2661, August 1998.

[5] R. Jain, T. Raleigh, C. Graff and M. Bereschinsky, Mobile Internet Access and QoS Guarantees Using MIP and RSVP with Location Registers, Proceedings of IEEE International Conference on Communications 1690-5, Vol.3 June 1998.

[6] I. Mahadevan and K. M. Sivalingam, An Experimental architecture for providing QoS guarantees in Mobile Networks using RSVP, Proceedings IEEE PIMRC 98, (Boston, MA), 50-4, Vol.1, September 1998.

[7] S. Chen and K. Nahrstedt, An overview of QoS Routing for Next-Generation High speed Networks, Problem and solutions, IEEE Networks, December 1998.

[8] X. Zhao, C. Castelluccia, and M. Baker, Flexible Network Support for Mobility, Proceedings of 4th Int. Conference on Mobile Computing and Networking, 145-56, IEEE MobiCom 1998.

[9] I. Mahadevan and K.M. Sivalingam, An architecture for QoS guarantees and Routing in Wireless/Mobile Networks, 1998.

[10] A.T. Campbell, J. Gormez, An overview of Cellular IP, September IEEE Wireless Communications and Networking Conference, 606-10 Vol.2, 1999.

[11] A. Terzis, M. Srivastava, L. Zhang, A Simple QoS Signalling Protocol for Mobile Hosts in the Integrated Services Internet, IEEE INFOCOM 1999. Proceedings of Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. 1999, 1011-18 Vol.3.

[12] G. Fodor, P. Barta, Á. Marquetant, A. Rácz, T. Henk, Modeling and analysis of routing and resource allocation techniques in multi-service networks,

International Journal of Communication Systems, Vol. 12, No. 2, March/April 1999.

[13] X. Xiao and L.M. Ni, Internet QoS, A Big Picture, University of Michigan State, IEEE Networks, April 1999.

[14] G Fankhauser, S Hadjiefthymiades, N Nikaein, L. Stacey, Interworking RSVP and Mobile IP Version 6 in Wireless Networks, Proceedings 4th ACTS Mobile Communications Summit, June 1999.

[15] I. Mahadevan and K. M. Sivalingam, QoS in Wireless Networks Based on Differentiated Service Architecture, Proceedings 8th International Conference Computer Communications and Networks, October. 1999.

[16] R. Ramjee, A Domain-based Approach for supporting Mobility in Wide-area Wireless Networks, Proceedings International. Conference Network Protocols, November 1999.

[17] A. Terzis, J. Krawczyk, J. Wroclawski, L. Zhang, RSVP Operation Over IP Tunnels. RFC 2746, January 2000.

[18] I. Mahadevan and K.M. Sivalingam, A hierarchical architecture for QoS guarantees and Routing in Wireless/Mobile Networks, Journal of Parallel and Distributed Computing, April 2000.

[19] S. Kivisaari, A comparison Between Mobile IPv4 and Mobile IPv6, Internet Report, May 2000.

[20] C. Castelluccia, HMIPv6, A Hierarchical mobile IPv6 proposal, ACM Mobile Computing and Communications Review 2000.

[21] Routing Protocols Uyless Black, Prentice Hall Series 2000.

[22] H. Petander, Mobile IP RO, Internet Report 2000.

[23] S. Ala-Luukko, Mobility Management in IETF and GPRS Specifications, Helsinki University Internet Report, May 2000.

[24] W.T. Chen and L.C. Huang, RSVP Mobility, A signalling Protocol for Integrated Services Internet with Mobile Hosts, Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, 2000, 1283-92 Vol.3

[25] P. Pan, E. Hahne, and H. Schulzrinne, BGRP, a framework for scalable resource reservation, Internet Draft, January 2000.

[26] B. Vandalore, W. Feng, R. Jain, and S. Fahmy, A survey of Service layer techniques for adaptive streaming of multimedia, Journal of Real Time Systems, January 2000.

[27] M. Carli, A. Neri, A.R. Picci, Mobile IP and cellular IP integration for inter access networks handoff, ICC-2001. IEEE International Conference on Communications 2001, 2467-71 Vol.8.

[28] K. Kim, and S. Kim, Domain Based Approach for QoS Provisioning in Mobile IP, GLOBECOM'01, IEEE Global Telecommunications, 2001, 2230-4 Vol.4.

[29] R. Vadali, J. Li, Y. Wu, and G. Gao, Agent Based RO for Mobile IP, IEEE 54th Vehicular Technology Conference. VTC Fall 2001, Proceedings 2001, 2731-5, Vol.4.

[30] A. Talukdar, B. Badrinathm, and A. Acharya, MRSVP, A Resource Reservation Protocol for an Integrated Services Network with Mobile Hosts, Wireless Networks, Vol. 7, January 2001.

[31] G.C. Lee, T.P. Wang and C.-C. Tseng, Resource Reservation with Pointer Forwarding Schemes for the Mobile RSVP, IEEE Commun. Letters, Vol. 5, no.7, July 2001.

[32] S. Paskalis, A. Kaloxylos, E. Zevas, An Efficient QoS Scheme for Mobile Hosts, Proceedings LCN 2001, 26th Annual IEEE Conference on Local Computer Networks. 2001, 630-7.

[33] G. Stattenberger, T. Braun, QoS Provisioning for Mobile IP Users, 1st International. Wksp. Svcs. and Apps. in the Wireless Public Infrastructure, July 2001.

[34] T. Braun and G. Stattenberger, Providing Differentiated Services to Mobile IP Users, Proceedings 26th Annual IEEE Conference Local Comp. Network, Nov. 2001.

[35] Z Ren, CK Tham, CC Foo, CC Ko, Integration of Mobile IP and Multi-Protocol Label Switching, Proceedings IEEE ICC, Vol. 7, June 2001.

[36] H Kim, KSD Wong, W Chen, CL Lau, Mobility-Aware MPLS in IP-Based Wireless Access Networks, IEEE GLOBECOM, Vol. 6, November 2001.

[37] T. Yang and D. Makrakis, Hierarchical Mobile MPLS, Supporting Delay Sensitive Services Over Wireless Internet, Proceedings of International Conference Info-Tech and Info-Net, Vol. 2, November 2001.

[38] D Awduche, L Berger, D Gan, T Li, V Srinivasan, G. Swallow, RSVP-TE, extensions to RSVP for LSP tunnels, Internet Draft, August 2001.

[39] F. Baker, C. Iturralde, F. Le Faucheur, B. Davie, Aggregation of RSVP for IPv4 and IPv6 reservations, Internet Draft ietf issll rsvpaggr 03, February 2001.

[40] C.H. Yeh, Scalable QoS supports for multimedia services in the next-generation Internet, Proceedings IEEE Real Time Technologies and Services Symp., May/June 2001.

[41] C.H. Yeh, Medium access control with differentiated adaptation for QoS management in wireless networks, Proceedings IEEE International Conference Mobile and Wireless Communication Networks, August 2001.

[42] C.H. Yeh, Scalable, adaptive, and reliable resource management in high-speed and mobile networks, Proceedings IEEE International Conference Computer Communications and Networks (IC3N'01), 2001.

[43] Reverse Tunnelling for Mobile IP, revised, RFC 2344, 2001.

[44] E. Rosen, A. Viswanathan, R. Callon. Multiprotocol Label Switching Architecture, January 2001, RFC3031.

[45] K.I. Kim, D.H. Kim, Hierarchical Admission Control Scheme for supporting mobility for MIP, 2002 Military Communications Conference. Proceedings 2002, 431-5 Vol. 1.

[46] J Gomez, S Kim, CY Wan, Comparison of IP Micromobility Protocols, IEEE Wireless communications Feb 2002 9(1) 72-82.

[47] S. Hagen, O'Rirlly, IPV6 Essentials, 2002.

[48] M. Ramalho, Intra and inter-domain multicast routing protocols, A survey and Taxonomy, Alcatel corporate research centre, 2002.

[49] C. Jelger, T. Noel, Multicast for Mobile Hosts in IP Networks, Progress and Challenges, IEEE wireless communication October 2002.

[50] C.H. Wu, A.T. Cheng, S.T. Lee, J.M. Ho, D.T. Lee, Bi-directional RO in mobile IP over wireless LAN 2002 IEEE 56th Vehicular Technology Conference Proceedings 2002, 1168-72 Vol. 2.

[51] M. Ergen, S. Coleri, B. Dundar, R. Jain, A. Puri, P. Varaiya, Service of GPS to mobile IP and routing in wireless networks, 2002 IEEE 56th Vehicular Technology Conference Proceedings 2002, 1115-19 Vol. 2.

[52] S. Yasukawa, J. Nishikido, H. Komura, Scalable QoS support mobile resource reservation protocol for real-time wireless Internet traffic GLOBECOM'02 IEEE Global Telecommunications Conference. 2002, 1475-9 Vol. 2.

[53] C.H. Chu, C.M. Weng, Pointer forwarding MIPv6 mobility management, GLOBECOM'02 IEEE Global Telecommunications Conference. 2002, 2133-7 Vol. 3.

[54] K.I. Kim; S.H Kim; J.M. Moon; Y.J. Kim, Hierarchical admission control scheme for supporting mobility in mobile IP 2002 Military Communications Conference. Proceedings. 2002, 431 5 Vol. 1.

[55] H. Kim, J. Song, A time-based binding update in mobile IP networks with very high mobility, 8th International Conference on Communication Systems. ICCS 2002 Nov. 2002.

[56] S. Paskalis, A. Kaloxylos, E. Zervas, L. Merakos, Evaluating the RSVP mobility proxy concept, Proceedings of 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC 20020), Sept. 2002.

[57] P.M. Ruiz, E. Garcia, Improving user-perceived QoS in mobile and wireless IP networks using real-time adaptive multimedia services, Proceedings of 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Sept. 2002.

[58] H. Chen, L. Trajkovic, Simulation of RO in Mobile IP, Proceedings LCN 2002. 27th Annual IEEE Conference on Local Computer Networks, Nov. 2002.

[59] S.Y. Sung, Enhancement of binding update for mobile IP, Proceedings LCN 2002. 27th Annual IEEE Conference on Local Computer, Nov. 2002.

[60] C. Perkins, Recent Developments with Mobile IPv6, RFC 2002.

[61] C. Perkins, Mobile IP, IEEE Communications Magazine, May 2002.

[62] C.H. Yeh, H.T. Mouftah, and H. Hassanein, Signaling and QoS guarantees in mobile ad hoc networks, Proceedings IEEE International Conference Communications (ICC'02), Apr. 2002.

[63] B. Moon and H. Aghvami, Reliable RSVP Path Reservation for Multimedia Communications Under an IP Micro-mobility Scenario, IEEE Wireless Communications, Oct. 2002.

[64] J. Manner, A. Lopéz Toledo, A. Mihailovic, H.L. Velayos Muñoz, E. Hepworth, Y. Khouaja, Evaluation of Mobility and QoS Interaction, Computer Networks, Vol. 38, Feb. 2002.

[65] S. Paskalis, A. Kaloxylos, E. Zervas, L. Merakos, An efficient RSVP-Mobile IP interworking scheme, Mobile Networks and Services Vol. 8, no.3 2003; 8(3), 197-207.

[66] CC Tseng, GC Lee, RS Liu, TP Wang, HMRSVP, A Hierarchical Mobile RSVP Protocol, Wireless networks, 2003 9(2).

[67] L. Zhang; J. Cao, S.K. Das, A mailbox-based scheme for improving mobile IP performance , Proceedings 23rd International Conference on Distributed Computing Systems Workshops p. IEEE , May 2003.

[68] K.D. Wong, D. Ashutosh, K. Young, H. Schulzrinne, Managing simultaneous mobility of IP hosts, IEEE Military Communications Conference, Oct 2003.

[69] J.W. Jung, R. Mudumbai, D. Montgomery, H.K. Kahng, Performance evaluation of two layered mobility management using mobile IP and session initiation protocol, GLOBECOM '03. IEEE Global Telecommunications Conference, Dec. 2003.

[70] A. Grilo, M. Nunes, G. Sergio, N. Ciulli, Integration of IP mobility and QoS for heterogeneous wireless access in MOICANE, 14th IEEE 2003 International Symposium on Personal, Indoor and Mobile Radio Communications. Sept. 2003.

[71] M.A. Vergo, H.R. Sharif, A bi-directional tunnelling extension to improve mobility scalability in mobile multicast, International Conference on Software, Telecommunications and Computer Networks. Oct. 2003.

[72] Y.B. Wen, X.W. Chen, A dynamic hierarchic temporary HA scheme in Mobile IP, Mini Micro Systems. April 2003.

[73] J.S. Chiang, Y.H. Wang. C.H. Tsa, C.P. Hsu, Location management and multimedia communication service based on Mobile IP and Cellular IP network, Proceedings 17th International Conference on Advanced Information Networking and Services. March 2003.

[74] Y.W. Chen, Z.J. Yan, Effect of the label management in Mobile IP over MPLS networks, Proceedings 17th International Conference on Advanced Information Networking and Services. March 2003.

[75] S. Ohzahata, S. Kimura, Y. Ebihara, Adaptive handoff algorithm for efficient network resource utilization, Proceedings 17th International Conference on Advanced Information Networking, March 2003.

[76] Y.J. Lee, I.F. Akyildiz, A new scheme for reducing link and signalling costs in mobile IP, IEEE Transactions on Computers Vol. 52, no.6, June 2003.

[77] Y. Takagi, T. Ihara, H. Ohnishi, Mobile IP RO method for next –Generation Mobile Networks, Electronic and communications in Japan, Nov 2003.

[78] S.G. Choi, J.K. Choi, S.H. Yang, S.G. Jong, History Based RSVP Setup Mechanism for Supporting QoS Guarantee on the Mobile Environment, Mobile Communications. Revised Papers Lecture Notes in Computer Science Vol.2524. 2003, 413-19.

[79] R. Bless, Quality-of-Service Signaling in Wireless IP-Based Mobile Networks, Proceedings VTC Fall 2003, Oct. 2003.

[80] J. Hillebrand, Design of Quality-of-Service Signaling for IP-Based Mobile Networks, Proceedings WPMC '03, Oct. 2003.

[81] Y Min-Hua, Z Ying, Z Hui-Min, L Yu, A Modified HMRSVP Scheme, Proceedings 57th Semiannual VTC, Vol. 4, Apr. 2003.

[82] C.H. Yeh, Performance evaluation for penalty-based resource management, An adaptive QoS scheme for service differentiation and adaptation in heterogeneous wireless networks, Proceedings IEEE International Performance Computing and Communications Conference, Apr. 2003.

[83] C.H. Yeh, QoS Differentiation Mechanisms for Heterogeneous Wireless Networks and the Next-Generation Internet, Proceedings IEEE International Symp. Computer Communications, June/July 2003.

[84] L. Berger, Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions. January 2003, RFC3473.

[85] E.D. Kim, H.S. Kim; S.H. Lee; Y.J. Kim, An efficient handoff mechanism in micro-domain, 6th International Conference on Advanced Communication Technology, IEEE, Feb. 2004.

[86] L. Wei, T. Kato, S. Ueno, S. Itoh, A proposal of scalable QoS guaranteed mobile IP communication using MPLS path with pre-allocated bandwidth, IEICE Transactions on Communications, March 2004.

[87] B. Zhang, H.T. Mouftah, QoS routing through alternate paths in wireless ad hoc networks, International Journal of Communication Systems, Vol. 17, Issue 3, April 2004.

[88] P. Conforto, C. Tocci, V. Schena, L. Secondiani, N. Blefari-Melazzi, P. M. L. Chan, F. Delli Priscoli , End-to-end QoS and global mobility management in an integrated satellite/terrestrial network, International Journal of Satellite Communications and Networking, Vol. 22, Issue 1, January/February 2004.

[89] C.H. Yeh, The Heterogeneous Hidden/Exposed Terminal Problem for Power-controlled Ad Hoc MAC Protocols and Its Solutions, Proceedings IEEE Vehicular Technology Conference IEEE Communications Society May 2004.

[90] D.B. Johnson. C. Perkins, J. Arkko, Mobility support in IPv6. IETF Internet draft, June 2004.

[91] A. Capone, F. Martignon, Analysis of Dynamic QoS Routing Algorithms for MPLS Networks, 2004 IEEE International Conference on Communications IEEE. 2004, 1192-6 Vol.2.

[92] S.H. Hwang, B.K. LeeYoun, H. Han, C.S. Hwang, An Adaptive Hierarchical Mobile IPv6 Using Mobility Profile, Wireless Communications and Mobile Computing magazine, Vol. 4, Issue 2, March 2004.

[93] G. Carneiro, J. Rulea and Manual Ricardo, Cross-Layer Design in 4G Wireless Terminals, IEEE Wireless Communications, April. 2004.

[94] A. Chung-Yu and M. Naraghi-Pour, Path Restoration with QoS and Label Constraints in MPLS Networks, 2004 IEEE International Conference on Communications IEEE 2004, 1278-82 Vol.2.

[95] J. Yao, T.C. Wong, Y.H. Chew, Cross-Layer Design on the Reverse and Forward Links Capacities Balancing in Cellular CDMA Systems, 2004 IEEE Wireless Communications and Networking Conference IEEE 2004, 2004-9 Vol.4.

[96] J. Hillebrand, C. Prehofer, R. Bless and M. Zitterbart, Quality-of Service Signalling for Next-Generation IP-Based Mobile Networks, IEEE Communication Magazine, June 2004.

[97] G. Pavlou, Traffic Engineering and QoS Management for IP-Based Next Generation Networks, 2004 IEEE/IFIP Network Operations and Management Symposium IEEE 2004, 923 Vol.1.

[98] J. Chen, T. Lv, H. Zheng, Joint Cross-Layer Design for Wireless QoS Content Delivery EURASIP Journal on Applied Signal Processing. 1 Feb. 2005; 2005(2), 167-82.

[99] H. Fathi, S. Chakraborty, R. Prasad, Mobility management for VOIP, Evaluation of Mobile IP-Based Protocols, ICC 2005 - IEEE International Conference on Communications, no. 1, May 2005.

[100] Abd-Elhamid M. Taha, Hossa S. Hassanien, and Hussein F. Mouftah, Extensions for Internet QoS Paradigms to Mobile IP, A Survey, IEEE Communication Magazine, May 2005

[101] A. Lera, A. Molinaro, S. Polito, G. Ruggeri, End-to-End QoS Provisioning in 4G with Mobile Hotspots, IEEE Communication Magazine, Sept./Oct. 2005.

[102] W. Song, H. Jiang, W. Zhuang, X. Shen, Resource Management for QoS Support in Cellular /WLAN Interworking, IEEE Communication Magazine Sept./Oct. 2005.

[103] H. Gossain, T. Joshi, C. Cordeiro, D.P. Agrawal, A Cross-Layer Approach for designing Directional Routing Protocol in Manets, 2005 IEEE Wireless Communications and Networking Conference 2005, 1976-81 Vol. 4.

[104] H. Jiang, W. Zhuang and X. Shen, Cross-Layer Design for Resource Allocation in 3G Wireless Networks and Beyond, IEEE Communication Magazine, Dec. 2005.

[105] L. Fang, J.L. Le Roux, Inter-provider IP-MPLS Services, Requirements, Implementations and Challenges, IEEE Communication magazine, June 2005.

[106] V. Srivastava, M. Motani, Cross-Layer Design, A Survey and the Road Ahead, IEEE Communications Magazine, Dec. 2005.

[107] C.C. Tseng, L.H. yen, H.H. Chang and K.C. Hsu, Topology-Aided Cross-Layer Fast Handoff Design for IEEE 802.11/Mobile IP Environment, IEEE Communications Magazine, Dec. 2005.

[108] E. Setton, T. Yoo, X. Xhu, A. Goldsmith and B. Girod, Cross-Layer Design of Ad-Hoc Networks for Real-time Video Streaming, IEEE Wireless Communications, August. 2005.

[109] V. Kawadia, P.R. Kumar, A Cautionary Perspective on Cross-Layer Design, IEEE Wireless Communications, February 2005.

[110] H. Jiang, W. Zhuang, Cross-Layer Resource Allocation for Integrated Voice/Data Traffic in Wireless Cellular Networks, IEEE Transaction on Wireless Communications, Vol 2, No2, Feb. 2006.

[111] R.L. Aguiar, S. Sargento, A. Banchs, C.J. Bernardos, M. Calderon, I. Soto, M. Liebsch, T. Melia, P. Pacyna, Scalable QoS-Aware Mobility for Future Mobile Operators, IEEE Communications Magazine, Vol. 44, no. 6, June 2006.

[112] X. Hsiao, H. Song, M. Zhang, Cross-Layer-Based Modeling for Quality of Service Guarantees in Mobile Wireless Networks, IEEE Communications Magazine, Vol. 44, no. 1, January 2006.

[113] H. Koga, S. Kashihara, Y. Fukuda, K. Iida, Y. Oie, A Quality-Aware VoWLAN Architecture and Its Quantitative Evaluations, IEEE Wireless Communications, Vol. 13, no. 1, February 2006.

[114] X. Lin, N. Shroff, An Optimization-Based Approach for QoS Routing in High-Bandwidth Networks, IEEE/ACM Transactions on Networking, Vol. 14, no. 6, December 2006.

[115] UC. Kozat, I. Koutsopoulos, L. Tassiulas, Cross-Layer Design for Power Efficiency and QoS Provisioning in Multi-Hop Wireless Networks, IEEE Transactions on Wireless Communications, Vol. 5, no. 11, November 2006.

[116] K. Yang, Y. Wu, H.H. Chen, QoS-Aware Routing in Emerging Heterogeneous Wireless Networks, IEEE Communications Magazine, Vol. 45, no. 2, February 2007.

[117] J. Spencer, J. Griem, M. Boucadair, P. Morand, M. Howarth, N. Wang, G. Pavlou, A. Asgari, P. Georgatsos, Interdomain Routing through QoS-Class Planes, IEEE Communications Magazine, Vol. 45, no. 2, February 2007.

[118] X. Masip-Bruin, M. Yannuzzi, R. Serral-Gracià, J. Domingo-Pascual, J. Enríquez-Gabeiras, M.Á. Callejo, M. Diaz, F. Racaru, G. Stea, E. Mingozzi, A. Beben, W. Burakowski, E. Monteiro, L. Cordeiro, The EuQoS System, A Solution for QoS Routing in Heterogeneous Networks, IEEE Communications Magazine, Vol. 45, no. 2, February 2007.

[119] G. Xue, A. Sen, W. Zhang, J. Tang, K. Thulasiraman, Finding a Path Subject to Many Additive QoS Constraints, IEEE/ACM Transactions on Networking, Vol. 15, no. 1, February 2007.

[120] K. Yang, Y. Wu, and H.H. Chen, QoS-Aware Routing in Emerging Heterogeneous Wireless Networks, IEEE Communications Magazine, Vol. 45, no. 2, February 2007.

[121] http://www.ietf.org/iesg/1rfc_index.txt

[122] MPLS Overview, Web ProForoum Tutorials http://www.iec.org

[123] Opnet, www.opnet.com

[124] A. Sefidcon, F. Khendek, A Cross-Layer Design Technique for QoS over Optimized Routes in MIP, Proceeding of WAMICON 2006, Dec 2006.

[125] A. Sefidcon, F. Khendek, QOS Constraint-based Routing in MIP/MPLS Network, Submitted to ISCIT2007.

[126] A. Sefidcon, F. Khendek, Guaranteed QoS for MIP-based Multimedia Services, Submitted to Elsevier Journal on Computer Networks.

Appendix: MIP_ROQS Model

In this appendix, the details of MIP-ROQS model in node and process level are described.

1. Node level

The difference between the wlan_wkstn_adv used in roaming topology and the mip_ppp_wkst_adv used in wired topology is the port type connecting them to the rest of the network. The port used in mip_ppp_wkst_adv is a ppp port and the port used in wlan_wkstn_adv is a wlan port.

The difference between wlan_ethernet_slip4_adv used in roaming topology and mip_atm2_ethernet32_frelay4_slip32 used in wired topology is also the type of connection they have with the rest of the network. The available connection types in mip_atm2_ethernet32_frelay4_slip32 are ppp, atm and ethernet connection and the connection types used in wlan_ethernet_slip4_adv are ppp, ethernet and wlan connection.

In this Section the node level architecture of the common nodes used in two above topologies are described. The node level architecture is composed of hierarchy of necessary modules. Each module is a process implemented as a state machine.

a. FA/HA

As shown in Figure 44, regardless of type of connection that the mobility router/agent supports, IP packets are received from lower layers (physical and data link, whether wired or wireless) and are entered to IP module.

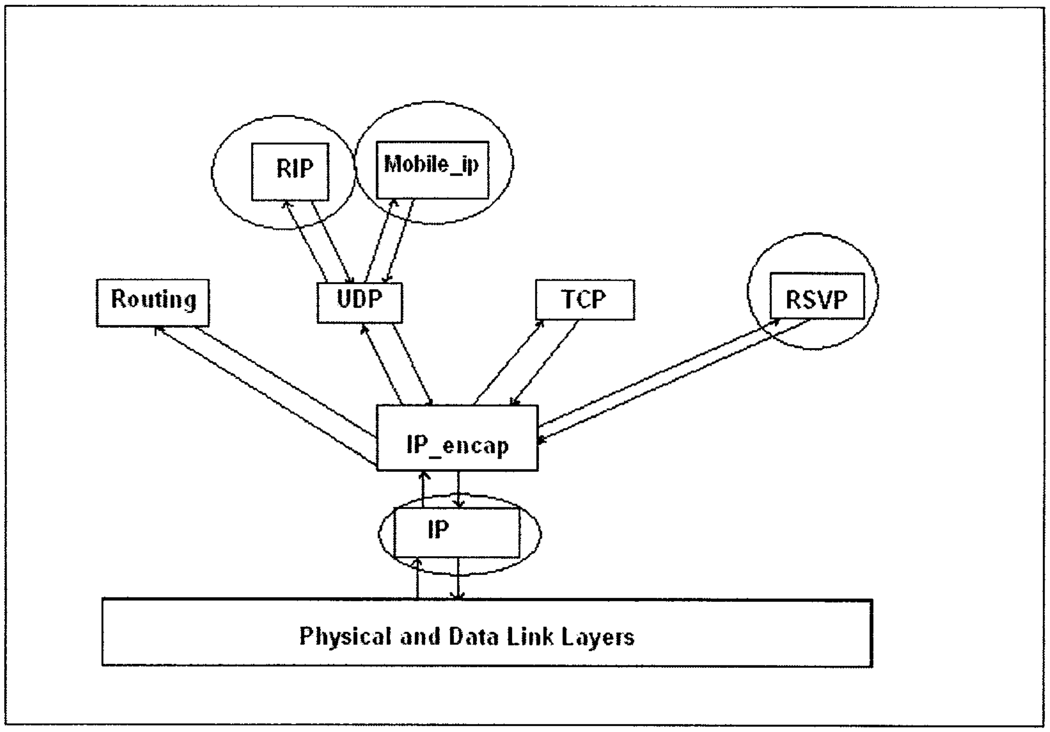

Figure 44: Agent Internal Structure

IP module in turn passes the packets to ip_encap module to remove the encapsulation and discover which module they are intended for.

The received IP packets could be one of the following:

- Routing
- RSVP
- UDP (special routing messages either related to MIP or Routing Protocol)
- TCP

The modules that we will describe further in process model Section are IP_encap, RSVP, MIP, and Routing Protocol.

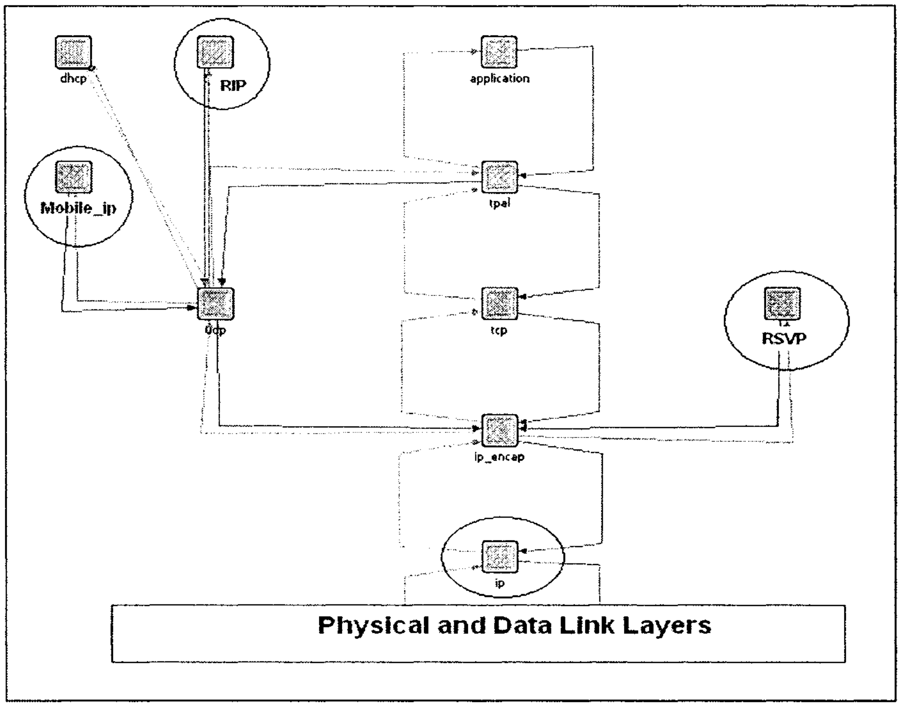

Figure 45: Host Internal Structure

b. CN / MN

As shown in Figure 45, a mobility aware workstation whether wired or wireless (configured through physical and data link layer) needs an IP module to receive IP packets and ip_encap module to encapsulate or de-capsulate the packets.

This module then will send the packets to TCP, UDP, or RSVP modules where it applies. TCP packets are directed toward service module and UDP packets are either MIP, or routing packets, which are sent to relevant modules.

In the next Section the node level interaction between MN, CN, HA and FA is described.

2. Node level Interaction

Putting all the nodes on the same page, actual communication between Client, HA / IP / FA, MN as shown in Figure 46 is via physical and data link layers.

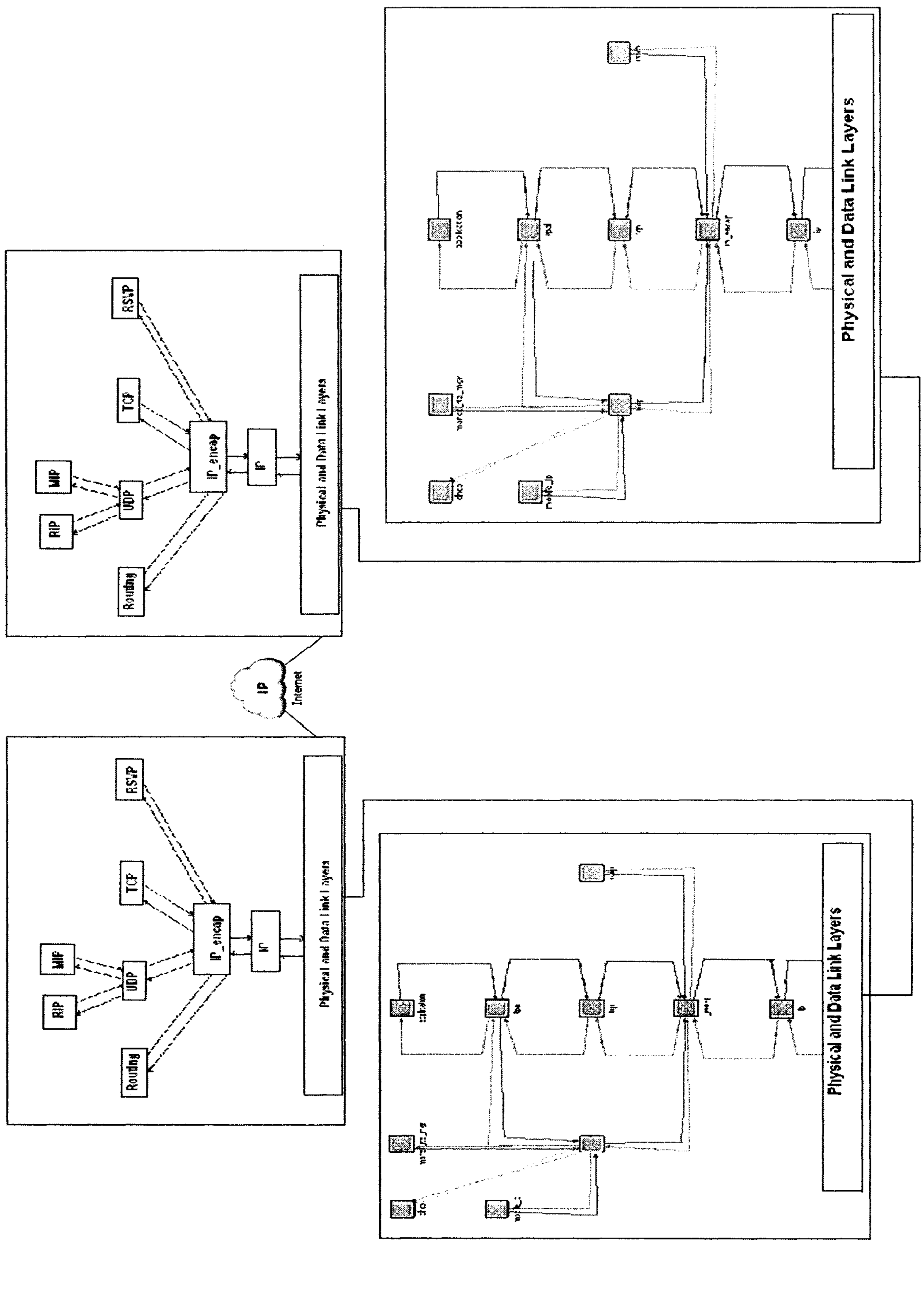

Figure 46: Node level interaction

3.Processes

In all nodes, packets are passed via ip module. To keep an easy to follow description of model under study, first the sequence in which ip module is invoking other processes is described followed by the description of process models.

a. Process Invocation

Shown in Figure 47, ip_dispatch is the process of IP module in the node model, and mobile_ip_reg_mgr is the mobile_ip process in node model. These two processes are created statically as shown before.

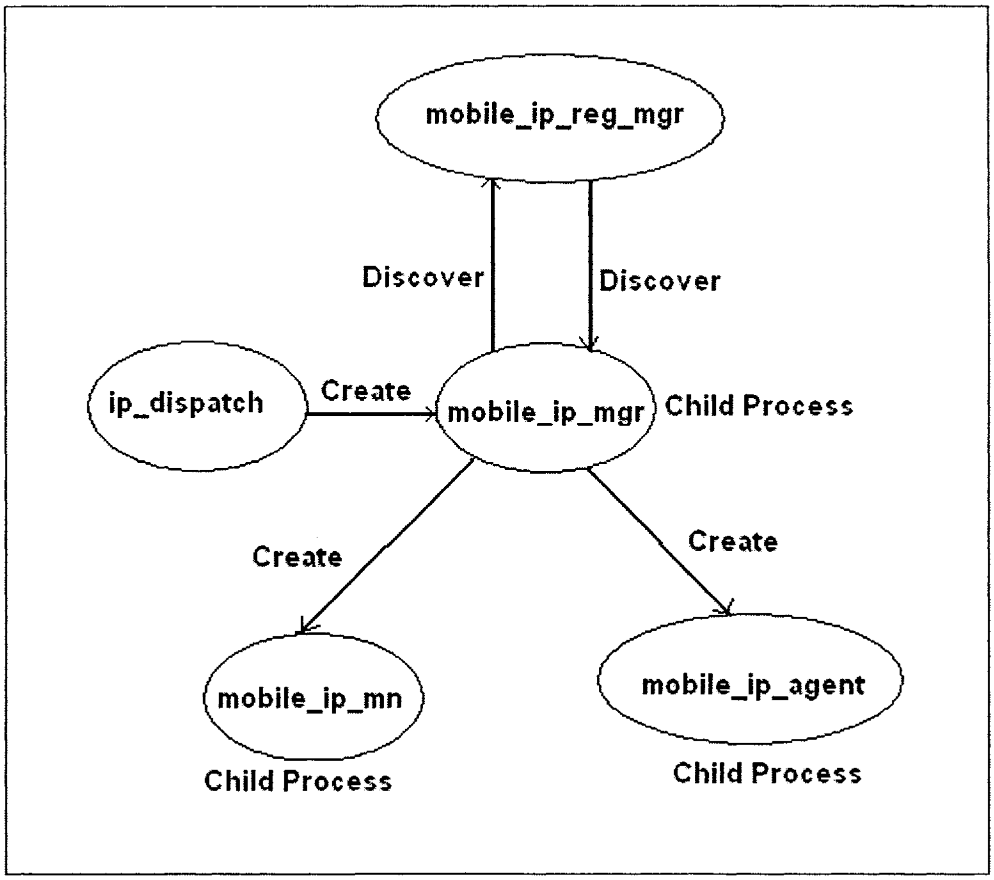

Figure 47: IP Dispatch Process Invocation

Ip_dispatch process will create dynamically mobile_ip_mgr as a child process when it receives a message destined for MIP protocol. mobile_ip_mgr based on the type of the received message will create one of the child processes for MN or agent (FA/HA). After a process is created dynamically, relevant messages will be passed to it for further processing. mobile_ip_mgr will find out the process handle of mobile_ip_reg_mgr using discover function in order to activate this process. MIP processes are explained below in more detail.

b. mobile ip reg mgr

As shown in Figure 48, Mobile_ip_reg_mgr process is responsible for management of MIP processes. The other MIP processes are mobile_ip_mgr, mobile_ip_agent, and mobile_ip_mn that are created dynamically.

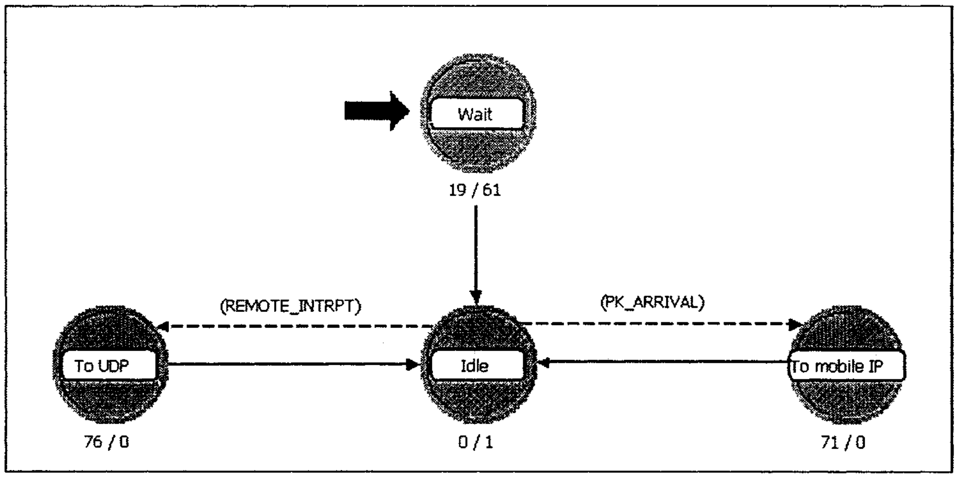

Figure 48: MIP Registration Manager Process

This process receives packets from UDP with PK_ARRIVAL interrupt, discovers mobile_ip_mgr process, and sends them for further processing to it. It also creates MIP messages (Registration request and Registration reply) and sends them to UDP port to be delivered to the other agent.

c. mobile_ip_mgr

Mobile_ip_mgr process is responsible for processing the registration messages (request and reply) and perform encapsulation of the ip packets (data sent by CN destined to MN) before sending them to the IP. As shown in Figure 49, a received packet is either a registration packet (REG_PK event), that changes the state to "Reg PK" for processing, or is an IP packet that changes the state to "IP PK" for encapsulating and sending it to MN's CoA.

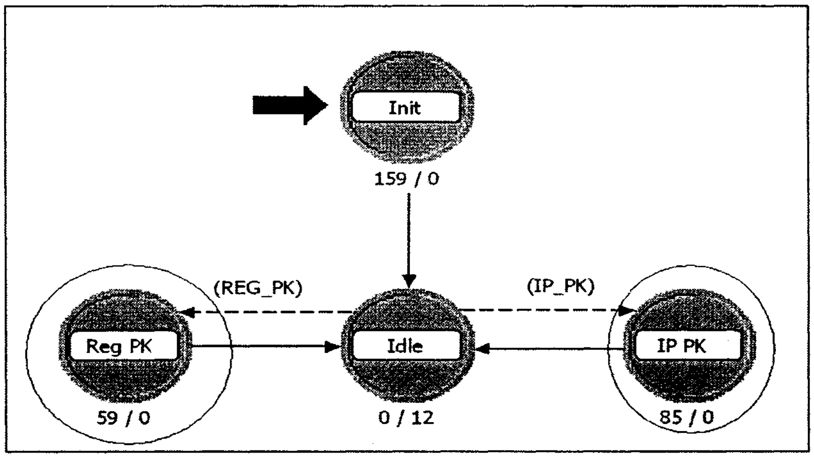

Figure 49: MIP Manager Process

While processing the registration messages, based on type of node, the mobile_IP_host or mobile_ip_agent attributes are set, the necessary child process is created for processing messages.

d. mobile ip agent

As shown in Figure 50, the events mobile_ip_mgr process is responsible for processing are:

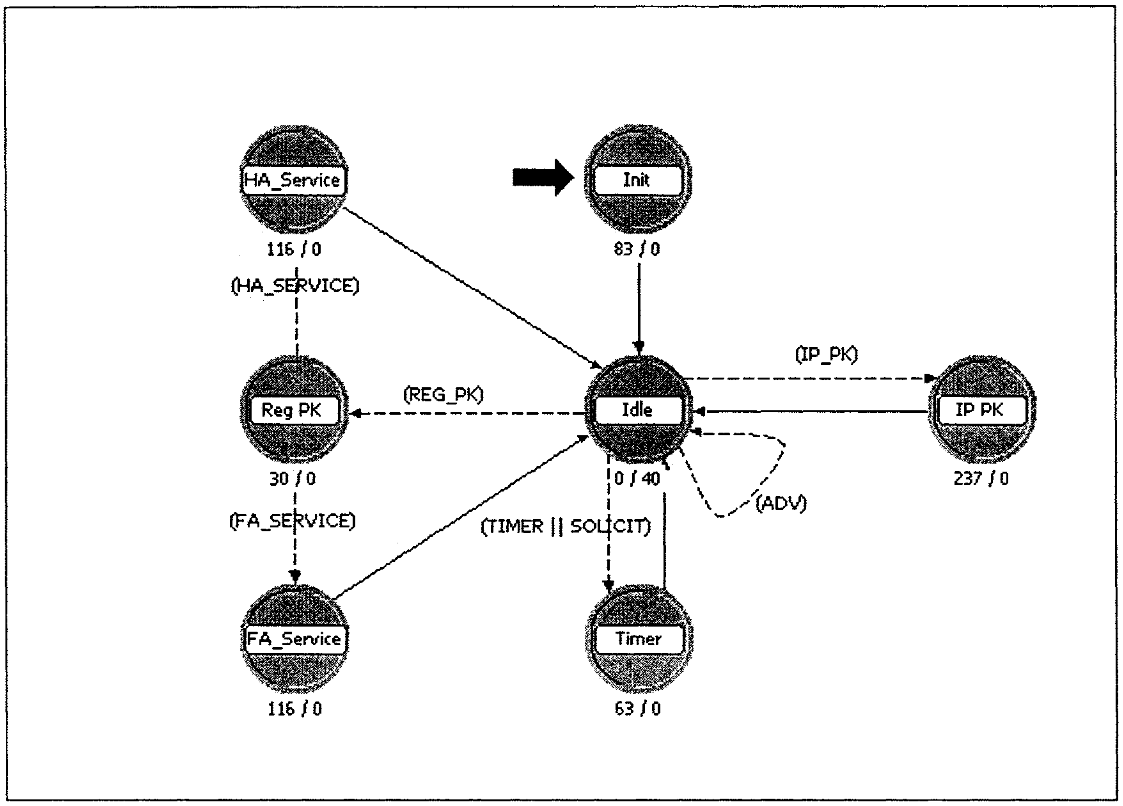

Figure 50: MIP Agent Process

- ADV: sending out the agent-advertisement message periodically
- SOLICIT: Receiving agent-solicitation message from MN and sending out agent- advertisement back to MN.
- REG_PK: Receiving registration messages and responding to them.
- IP_PK: Receiving IP packets destined for the MN that based on original mobile ip protocol must be encapsulated and sent toward CoA

e. mobile_ip_mn

As shown in Figure 51, the events mobile_ip_mn process is responsible for are:

- Receiving agent-advertisement
- Creating agent-solicitation
- Creating registration request and receiving registration reply
- Handling IP packets (service data)

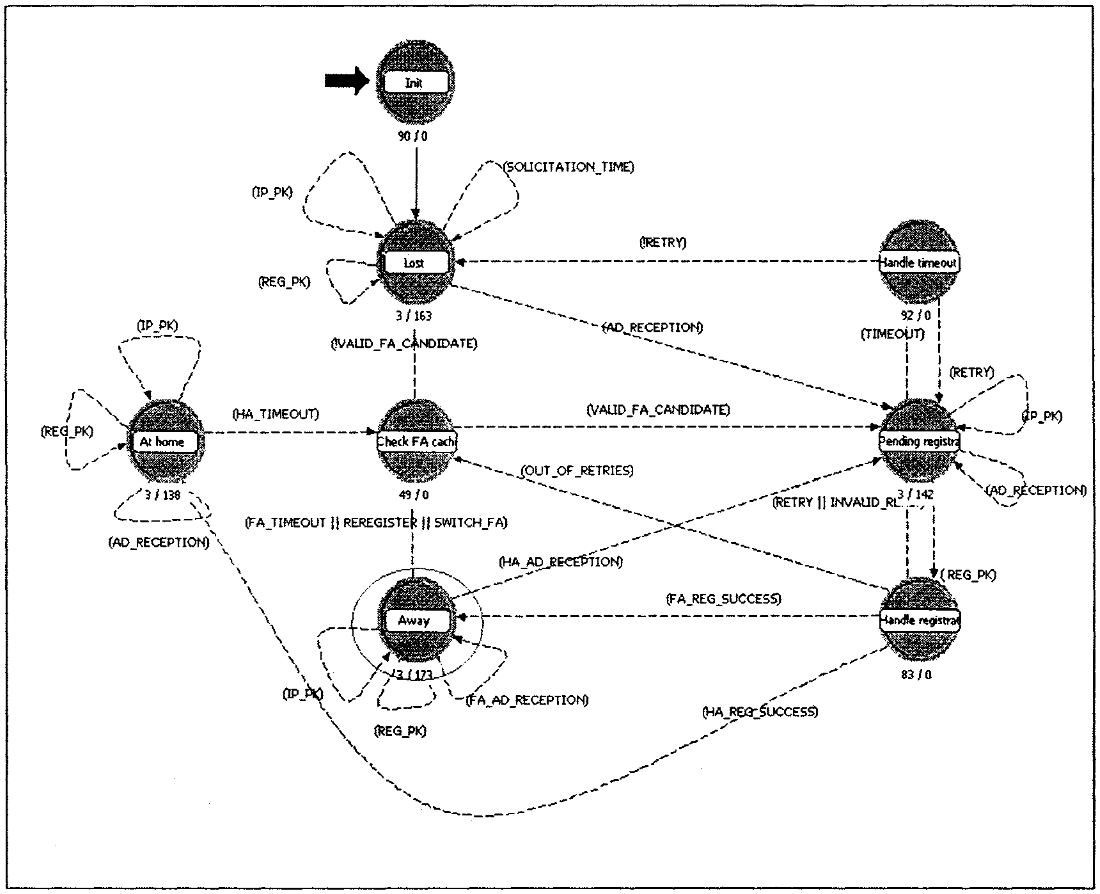

Figure 51: MIP MN Process

The above events are handled according to the location of MN. MN could be at home or away from home or in between (lost).

__f. RSVP__

RSVP process as shown in Figure 52, is responsible for handling the following events:

- Initiation by SENDER module
- Processing PATH message

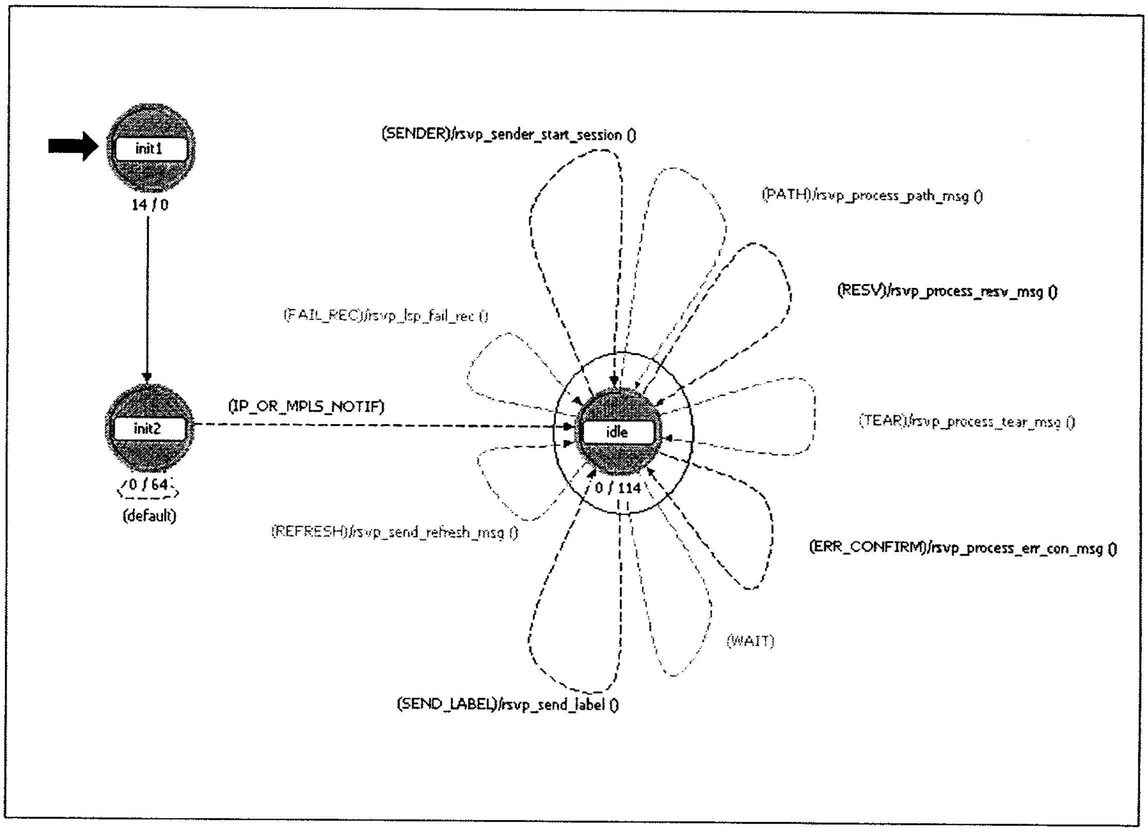

Figure 52: RSVP Process

- Processing RESV message
- Processing TEAR message
- Processing ERR_CONFIRM message
- Processing REFRESH message
- Processing FAIL_REC message

➤ Processing SEND_LABEL message

g. Routing Protocol

As shown in Figure 53, Routing Protocol process is responsible for initializing routing table at reception of IP_Notification and handling routing request and response, routing table update and link failure. Routing update happens periodically. In the wait state, the process waits for four possible events to occur:

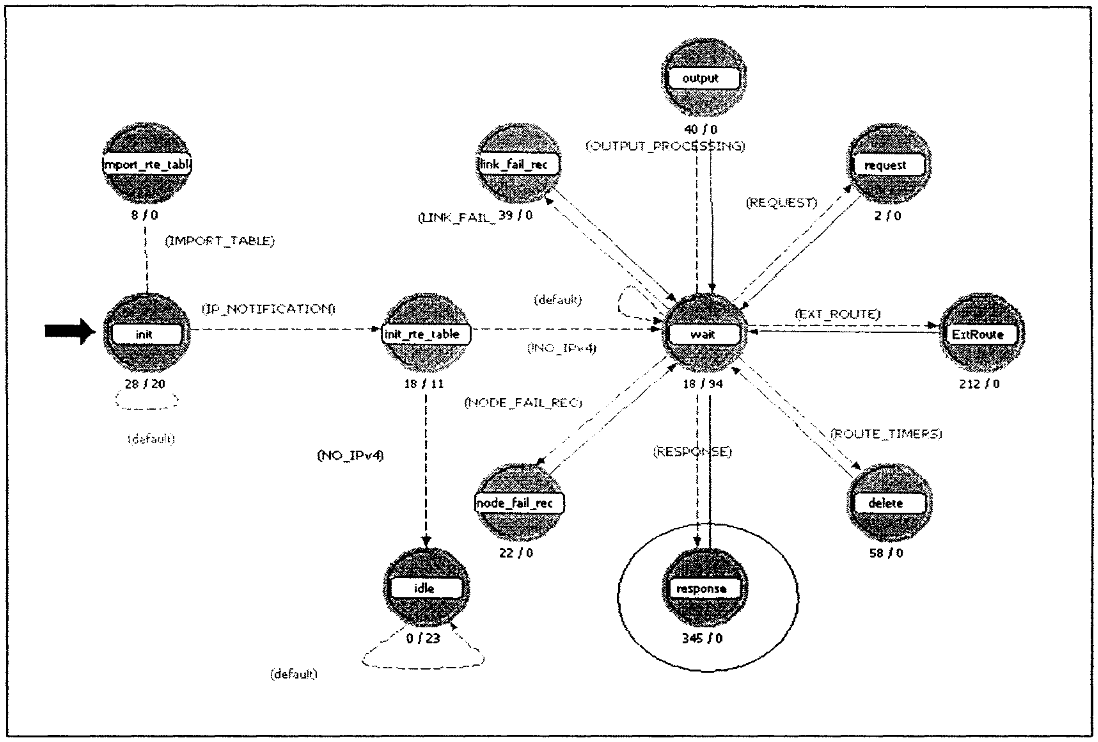

Figure 53: Routing Protocol Process

1) Arrival of a response
2) Arrival of a request
3) Expiration of update timer

4) Expiration of a route timer which can be either the timeout or garbage collection timer for a particular route.

4. Important messages

Important messages are the messages of MIP, RSVP and Routing Protocol that are performing mobility, resource reservation and routing. These messages are used to carry necessary data for ROQS processing.

a. mip

Mobile IP messages are:

Agent Advertisement sent by mobility agent (shown in Figure 54)

Agent Solicitation sent by MN (shown in Figure 55)

Registration Request sent by MN, handled by mobility agent (shown in Figure 56)

Registration Reply, sent by mobility agent and handled by MN (Figure 57)

We use extension in registration messages to carry the necessary data.

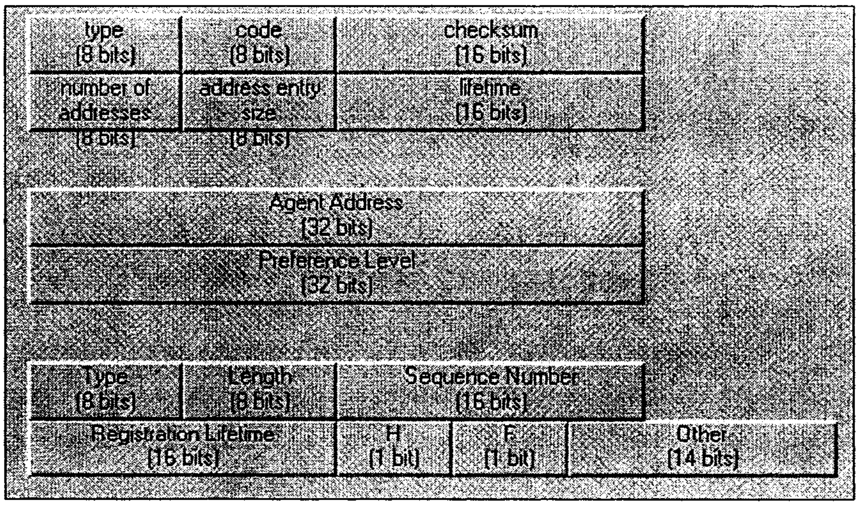

Figure 54: Agent Advertisement

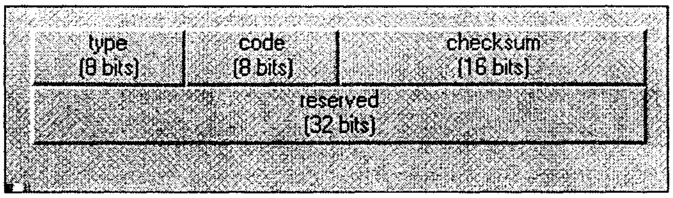

Figure 55: Agent Solicitation

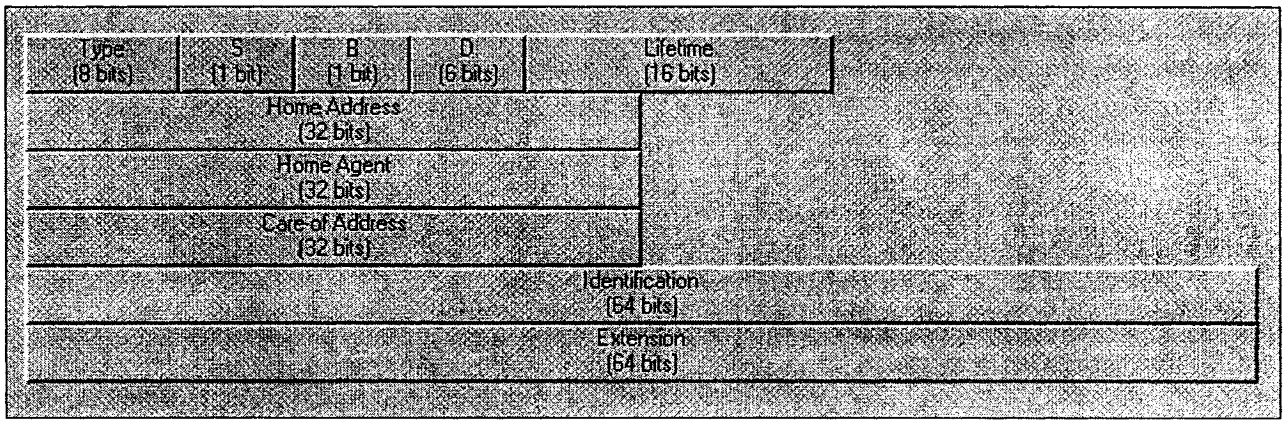

Figure 56: Registration Request

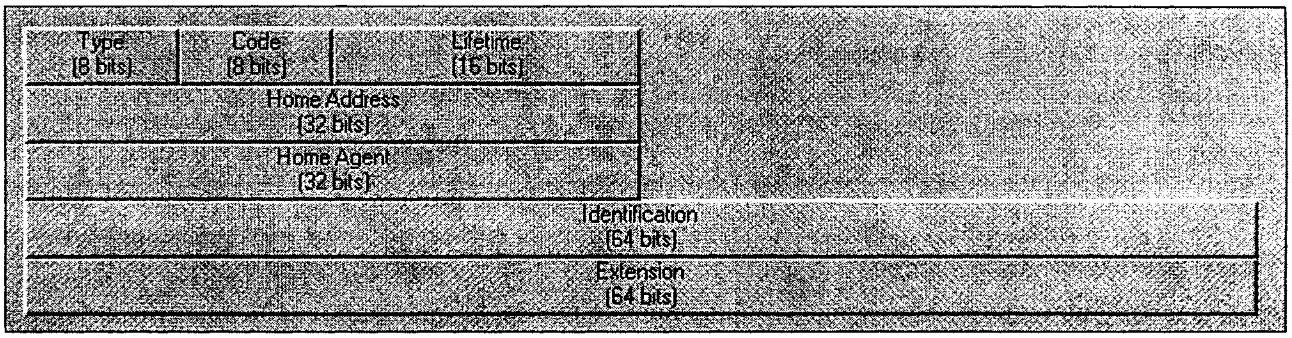

Figure 57: Registration Reply

b. rsvp

As shown in Figure 58, RSVP process has a unified message that is differentiated based on the header type. Header type can be one of the following:

RsvpC_Path

RsvpC_Resv,

RsvpC_Path_Error,

RsvpC_Resv_Error,

RsvpC_Path_Tear,

RsvpC_Resv_Tear,

RsvpC_Resv_Conf,

Rsvp_C_Path is modified to carry the ROQS specific data and RsvpC_Path_Error is modified to bring CoA back to CN in the case direct path is the optimal path.

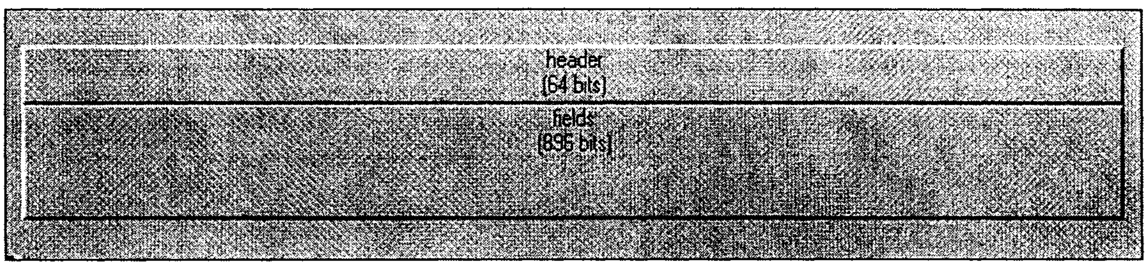

Figure 58: RSVP messages

c. rip

As shown in Figure 59, Routing Protocol messages follow a unified packet format. Differentiation of certain message is done by setting the command bits. Routing commands can be a request or response.

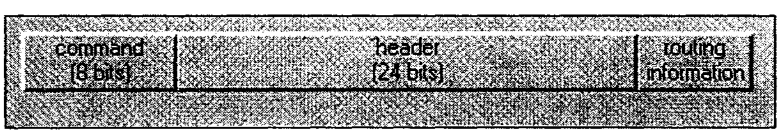

Figure 59: Routing Protocol Messages

a. Implementation of RO

Implementation of RO was done in stages (a) to (f) as described below. RO although incorporated in IPv6 model is not provided by Opnet on IPv4 model.

c.1 ROQS_RO Configurable attribute

Mobility agent attributes were extended to define a new class of attributes called ROQS. Under this category, a new attribute named "RO" was added as a configurable attribute to identify whether RO is enabled on this node or not. The attribute was promoted to be set before starting the simulation run.

c.2 Catch the first IP packet

When the first IP packet sent from CN toward the MN reaches HA, HA verifies if the ROQS_RO attribute is set for this scenario, it will avoid usual encapsulation process toward MN's CoA.

To do this part, as shown in Figure 60, Mobility_agent process model was modified.

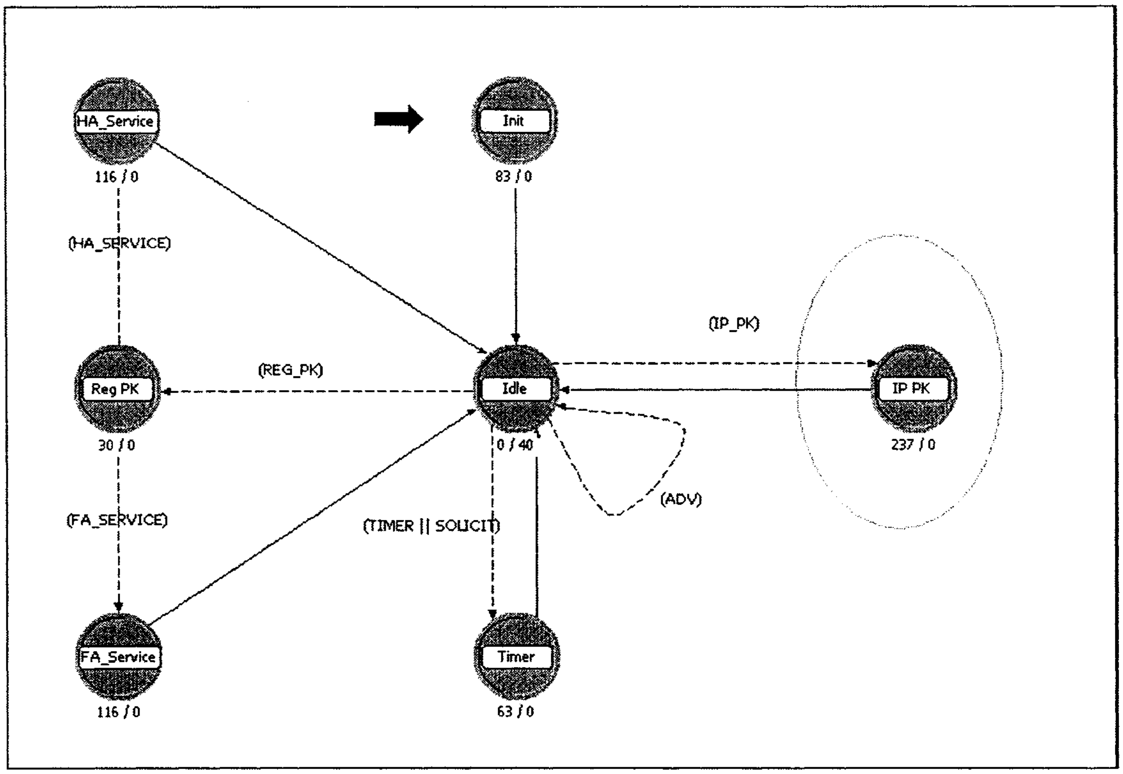

Figure 60: Impact on mobile_ip_agent Process

In this process, the algorithm for handling IP_PK event was changed to capture the first packet and act according to the routing mechanism.

c.3 Modified Packet Handling

The code in IP_PK state of mobile_ip_agent was modified according to the algorithm in Figure 61.

```
:
:
If (ROQS_RO) then
        If (this is the first packet)
                If (this is PATH message) then
                        - Search for MN in binding table
                        If (found) then
                                - Retrieve CoA
                                - Send a RsvpC_Path_Error message to CN inform it
about CoA
                        Else
                                Drop the packet
                        Endif
                Else
                        Continue
                Endif
        Else
                Continue
        Endif
Else
        Continue
Endif
:
```

Figure 61: RO Algorithm

CN-rsvp gets the Path_Error

When a QoS_sensitive service has sent a PATH message from CN to MN, the message has reached HA and based on the algorithm presented in previous Section, a RsvpC_Path_Error message has been sent back to CN, RSVP module in CN process. RSVP process model is modified to handle the Path_Error message and re-send PATH message directly to MN using CoA.

CN modified RSVP handling

The code in wait state of rsvp process was modified according to the algorithm in Figure 62.

```
:
If (Path_Error) then
        If (Retrieve CoA)
                If (Found) then
                        - Send PATH to CoA
                Else
                        Continue
                Endif
        Else
                Continue
        Endif
Else
        Continue
Endif
:
```

Figure 62: RSVP handling Path_Error

b. Configuration of RSVP with RO

With ROQS_RO attribute and supporting process model and algorithm in place, configuring RO is simply setting the ROQS_RO attribute to enable. To show the impact of RSVP besides RO, the rsvp attributes on different nodes and interfaces as well as the service are configured as explained in stage configuration of RSVP on basic model

c. Implementation of ROQS process

Implementation of "ROQS_Optimal Routing" was done in stages (a) to (e) as described below.

e.1 ROQS mechanism configuration attribute

Mobility agent attributes were extended to define a new attribute named "ROQS_Optimal Routing" as a flag to identify whether optimal RO is enabled on this node or not. The attribute was promoted to be set before starting the simulation run.

e.2 Node model

ROQS process interfaces between IP_Dispatch module and MIP, Routing Protocol and RSVP modules as it is responsible to collect QoS parameters and select the optimal path. Messages intended for MIP, Routing Protocol and RSVP that carry QoS parameters pass through RQOS Interface, so this process can retrieve necessary information before forwarding the messages to relevant modules. Figure 63 shows the node level model of ROQS. The ROQS process is incorporated in mobility agent node (HA/FA) model.

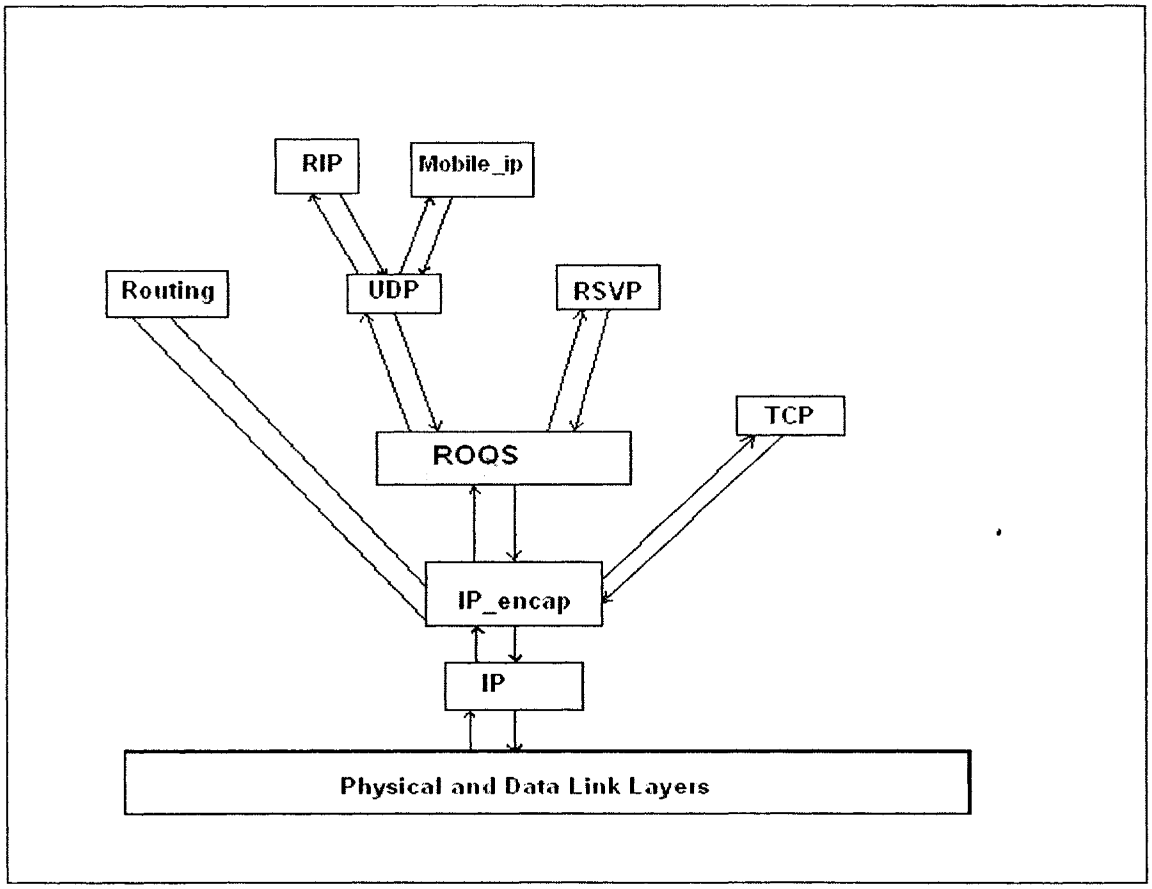

Figure 63: Agent Enhanced by ROQS