END-TO-END INTERNET QUALITY OF SERVICE WITH INTSERV/DIFFSERV, MOBILE IPV6 AND IEEE802.11E

WU WEI

(M.ENG OF NORTHWESTERN POLYTECHNICAL UNIVERSITY, PRC)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2003

Acknowledgment

I would like to express my deep gratitude to my supervisor, Dr. Winston K.G. Seah for all the guidance and support he has given me during the course of my graduate studies in NUS.

I would like to thank the Institute for Infocomm Research, a member of A*STAR for providing me with all the necessary facilities and congenial environment to carry out my research.

Finally, I would like to thank all my colleagues in NUS and I2R for providing hearty help for my research work. Of these, the following individuals deserve mention. I need to thank Mr. He DaJiang for providing IEEE802.11e model for NS2 implementation and support on the simulation work. I also want to thank Mr. Li Feng and Mr. Shen Qi for their discussion and advice on my research work.

Table of Contents

Acknowledgment i

Table of Contents ii

List of Figures v

List of Tables viii

Abbreviations ix

Summary xi

Chapter 1 Introduction 1

- 1.1 Overview 1
- 1.2 Thesis Contributions 4
- 1.3 Thesis Organization 5

Chapter 2 Internet QoS Architectures and Mobile IPv6 7

- 2.1 Introduction 7
- 2.2 Integrated Services and RSVP 7
- 2.3 Differentiated Services 10
- 2.4 Mobile IPv6 12
  - 2.4.1 Router Discovery 14
  - 2.4.2 Address Notification 15
  - 2.4.3 Packet Routing 16
- 2.5 IEEE802.11e Wireless LAN Standard 16
- 2.6 Summary 19

Chapter 3 An End-to-End QoS Architecture for Mobile Hosts: IntServ/DiffServ with Mobile IPv6 and IEEE802.11e 20

- 3.1 Introduction 20

3.2 Integration of IntServ and DiffServ 21

3.3 Combination of IntServ/DiffServ, Mobile IPv6 and IEEE802.11e 23

3.4 Services Mapping.. 26

3.4.1 Services Provided by IntServ.. 27

3.4.2 Services Provided by DiffServ.. 28

3.4.3 Services Mapping.. 29

3.5 Summary 31

Chapter 4 Improvement on Handoff QoS of the Architecture that Combines IntServ/DiffServ, MIPv6 and IEEE802.11e 32

4.1 Introduction.. 32

4.2 Handoff Procedure 33

4.3 Improving Handoff QoS Performance by Assigning *RADs* Higher Priority…………………………………………………………...35

4.4 Summary 37

Chapter 5 Simulation Studies 38

5.1 Introduction.. 38

5.2 Simulation Overview 39

5.2.1 QoS Models in Simulation System 39

5.2.2 Performance Criteria.. 41

5.2.3 Simulation Scenario 42

5.2.4 Simulation Configuration and Parameters 43

5.3 Simulation Study – Part 1: Simulation study on the handoff performance of the QoS architecture 44

5.3.1 Scenario 1: Improvement on Handoff Performance by Sending *BUs* to the Previous Base Station.. 45

5.3.2 Scenario 2: Improvement on Handoff Performance by Assigning *RADs* Higher Priority 49

5.4 Simulation Study – Part 2: Simulation study on the end-to-end QoS provided by the QoS architecture 54

5.4.1 Scenario 1: Compare the End-to-End QoS Provided by the Architecture of IntServ/DiffServ/MIPv6/IEEE802.11 with that Provided by the Architecture of IntServ/DiffServ/MIPv6/IEEE802.11e........ 55

5.4.2 Scenario 2: End-to-End QoS after Intra-Domain and Inter-Domain Handoff........58

5.5 Simulation Study – Part 3: Simulation Study on a More Practical Situation........64

5.5.1 Scenario 1: Average End-to End QoS Obtained by the Mobile Nodes before Movement 65

5.5.2 Scenario 2: Average End-to-End QoS Obtained by the Mobile Nodes after Intra-Domain and Inter-Domain Handoff........ 68

5.6 Summary 73

Chapter 6 Conclusion 74

6.1 Summary of Existing Work 74

6.2 Future Work 75

Publication 77

Reference 78

List of Figures

Figure 2.1 Basic RSVP Operation 9

Figure 2.2 Integrated Services Model 10

Figure 2.3 Differentiated Services Code Points (DSCP) 11

Figure 2.4 DiffServ Routers 11

Figure 2.5 Mobile IPv6 Architecture 14

Figure 2.6 Internal Contention of Different Traffic Categories 18

Figure 3.1 The reference network for the IntServ/DiffServ Framework 22

Figure 3.2 An End-to-End QoS Architecture for Mobile Hosts 24

Figure 3.3 Mobile Host's Intra-Domain Mobility 25

Figure 3.4 Mobile Host's Inter-Domain Mobility 26

Figure 4.1 The probability of dropping *RADs* in IFQ varying with network load 36

Figure 5.1 Basic simulation configuration 43

Figure 5.2 Simulation Configuration1 45

Figure 5.3 Packets tunneled from the previous base station that the mobile node just visited to the current base station 46

Figure 5.4 Intra-domain handoff latency varying with network load 47

Figure 5.5 Inter-domain handoff latency varying with network load 47

Figure 5.6 Improvement on intra-domain and inter-domain handoff latency with sending *BUs* to the previous base station 48

Figure 5.7 Handoff latency varying with network load 50

Figure 5.8 Handoff latency varying with last wired link delay 51

Figure 5.9 Handoff latency varying with overlap between base stations BS1 and BS2 52

Figure 5.10 Handoff latency varying with the speed of the mobile node 53

Figure 5.11 Simulation configuration 2 54

Figure 5.12 Goodput of GS, CL, BE flow vs. network load with IEEE 802.11 and with IEEE 802.11e 56

Figure 5.13 Average flow delay of GS, CL, BE flow vs. network load with IEEE802.11 and with IEEE802.11e 56

Figure 5.14 Packet drop rate of GS, CL, BE flow vs. network load with IEEE802.11 and with IEEE802.11e 57

Figure 5.15 Delay jitter of GS, CL, BE flow vs. network load with IEEE802.11 and with IEEE802.11e 57

Figure 5.16 Goodput of GS, CL and BE flow after handoff under light network load 59

Figure 5.17 Average flow delay of GS, CL and BE flow after handoff under light network load 60

Figure 5.18 Delay jitter of GS, CL and BE flow after handoff under light network load 60

Figure 5.19 Goodput of GS, CL and BE flow after handoff under heavy network Load 61

Figure 5.20 Average flow delay of GS, CL and BE flow after handoff under heavy network load 62

Figure 5.21 Delay jitter of GS, CL and BE flow after handoff under heavy network load 62

Figure 5.22 Packet drop rate of GS, CL and BE flow after handoff under heavy network load 63

Figure 5.23 Simulation configuration 3 – more practical situation 64

Figure 5.24 Average goodput of GS, CL and BE flows varying with the number of mobile nodes 66

Figure 5.25 Average flow delay of GS, CL and BE flows varying with the number of mobile nodes 67

Figure 5.26 Average delay jitter of GS, CL and BE flows varying with the number of mobile nodes 67

Figure 5.27 Average goodput of GS, CL and BE flow after handoff under light network load........ 69

Figure 5.28 Average flow delay of GS, CL and BE flows after handoff under light network load........ 69

Figure 5.29 Average delay jitter of GS, CL and BE flows after handoff under light network load........ 70

Figure 5.30 Average goodput of GS, CL and BE flows after handoff under heavy network load........ 71

Figure 5.31 Average flow delay of GS, CL and BE flows after handoff under heavy network load........ 72

Figure 5.32 Average delay jitter of GS, CL and BE flows after handoff under heavy network load........ 72

List of Tables

Table 3-1 Services mapping... 31

Table 5-1 The common parameters 44

Table 5-2 Traffic management... 44

Table 5-3 Different traffic type... 65

Table 5-4 Source rate with different number of the mobile nodes 66

Abbreviations

AF	Assured Forwarding
BB	Bandwidth Broker
BE	Best Effort
BR	Border Router
BS	Base Station
BU	Binding Update
CA	Collision Avoidance
CL	Controlled Load
CN	Correspondent node
CoA	Care of Address
CSMA	Carrier Sense Multiple Access
CW	Contention Window
DAD	Duplication Address Detection
DCF	Distributed Coordination Function
DiffServ	Differentiated Services
DIFS	DCF Inter-Frame Space
DSCP	Differentiated Services Code Point
EDCF	Enhanced DCF
EF	Expedited Forwarding
ER	Edge Router
GS	Guaranteed Service
HA	Home Agent
IEEE	Institute of Electrical and Electronics Engineers

IETF	Internet Engineering Task Force
IntServ	Integrated Services
IP	Internet Protocol
LAN	Local Area Network
MAC	Medium Access Vontrol
MIPv6	Mobile IPv6
MN	Mobile Node
MSDU	MAC Service Data Unit
NAR	New Access Router
NS	Network Simulator
PAR	Previous Access Router
PHB	Per Hop Behaviour
QoS	Quality of Service
RAD	Router Advertisement
RED	Random Early Detection
RSpec	Request Specification
RSVP	resource ReSerVation Protocol
TC	Traffic Category
TCP	Transmission Control Protocol
ToS	Type of Service
TSpec	Traffic Specification
UDP	User Datagram Protocol
VoIP	Voice over IP
WFQ	Weight Fair Queueing

Summary

Various mechanisms have been developed to address the need to provide quality of service (QoS) in the Internet where mobile devices are becoming the norm. Integrated Services (IntServ) and Differentiated Services (DiffServ) are two approaches to provide QoS guarantees in the Internet. Designed for wired non-mobile networks, IntServ provides per-flow QoS and is deemed to be only suitable for edge networks while DiffServ provides aggregate QoS and is suitable for use in core networks. For mobile hosts, IPv6 has built-in capability to support IP level mobility, which we refer to as Mobile IPv6 (MIPv6), and IEEE802.11e provides QoS support at the MAC layer of IEEE802.11-based WLANs. In this thesis, we use the QoS support in the MAC layer provided by IEEE802.11e to guarantee QoS on the wireless last hop, and propose an end-to-end QoS architecture for mobile hosts that combines IntServ/DiffServ, MIPv6 and IEEE802.11e.

In order to enhance the handoff performance of our QoS architecture, we also propose an approach to improve the handoff performance. It is found that the loss of *Router Advertisement* (*RAD*) messages of MIPv6 affects adversely the handoff performance. Therefore, the approach to improve handoff performance is assigning *RAD* messages higher priority to minimize the loss of *RADs* or totally eliminate the loss of *RADs*.

Handoff performance is essential to a good mobile QoS architecture. In this thesis, we combine the approach of assigning *RADs* higher priority with another approach that the MN sends *binding update (BU)* messages of MIPv6 to the previous base station that it just visited to improve the handoff performance of our end-to-end QoS architecture for mobile hosts. The latter approach is based on the idea of signaling and tunnel between the previous access router (PAR) and new access router (NAR) in Fast Handovers for Mobile IPv6 protocol.

Simulation is conducted to evaluate the performance of the QoS architecture that combines IntServ/DiffServ, MIPv6 and IEEE802.11e and to verify the efficiency of these two approaches to improve the handoff QoS performance. Simulation results have shown the achievable end-to-end QoS and improved handoff QoS.

Chapter 1

Introduction

1.1 Overview

Although current Internet is a widely deployed network with hundreds of millions of hosts, an Internet Protocol (IP)-based network still operates as a best effort network. It processes traffic as quickly as possible, but does not provide a reliable data delivery. It also cannot ensure timely delivery and cannot provide any guarantees on data throughput. These limitations have not been a problem for traditional Internet applications such as email, file transfer and Web applications. However, for some new applications such as real-time and multimedia applications, they demand high data throughput as well as low delay and jitter. Therefore there is a strong interest in introducing quality of service (QoS) to current IP networks.

The Internet Engineering Task Force (IETF) has proposed many service models and mechanisms to meet the demand for QoS [1][2]. Notable among them are the Integrated Services (IntServ)/RSVP model [3] and the Differentiated Services (DiffServ) model [4]. IntServ/RSVP model is characterized by resource reservation. The applications must set up paths and reserve resources before data are transmitted. RSVP [5][6] is such a signaling protocol for setting up paths and reserving resources. DiffServ model divides the traffic into a small number of classes, and packets in different classes receive differentiated services. IntServ provides per flow QoS while DiffServ provides QoS based on per aggregate. They are not mutually exclusive of one another, on the contrary, they complement each other in a perfect way. Some work has been done on integrating IntServ/DiffServ [7][32][34], and simulation results have reported that integrating IntServ and DiffServ can guarantee end-to-end QoS in wired networks[8].

In recent years, the number of portable computing devices has increased tremendously, such as laptop computers, palmtop computers, Personal Digital Assistants (PDAs). In the coming future, a major part of personal communication, no matter whether it is voice, data, images or video, will be wireless. With the increasing deployment of the wireless networks, proliferation of mobile computing devices, and emergence of new multimedia applications, there is an increasing need to provide QoS to mobile devices.

However, for a wireless network where hosts are likely to be mobile, the requirements of QoS are more difficult to achieve due to host mobility and the features of the wireless medium such as low bandwidth and high loss rate. The mobility management in the Internet world is dealt with Mobile IP. The next generation IPv6 protocol has built-in capability to support IP level mobility, which we refer to as Mobile IPv6 (MIPv6) [9]. Mobile IPv6 is based on the best effort delivery model and has no consideration of QoS. Moreover, both IntServ/RSVP and DiffServ are designed for wired non-mobile networks and become invalid under host mobility. There has been some research on extending RSVP or DiffServ to wireless mobile networks. The protocols on extending RSVP to wireless mobile network include Mobile RSVP [10], multicast-based model [11], RSVP tunnel model with Mobile IP [12], Mobile IP with location registers [13], Mobile IPv6 and RSVP integration model [14], and flow transparency-based model [15][16]. Some studies on extending DiffServ to wireless mobile network are also based on providing signaling support in DiffServ networks [17][18][19].

All these studies have focused on providing handoff QoS when mobile hosts move. Even though these studies set up RSVP resource reservation paths efficiently, most of these solutions have no QoS mechanism enough to prevent service disruption at a new cell during handoff [16]. Moreover, these studies have been carried out on the network layer or higher layer. Inherited from the basic idea of layered network that considering the case in each separated independent layer, these studies have not considered the lower layer, whereby the perfect QoS on lower layer is assumed. However this is not true for the existing IEEE802.11-based wireless access networks. The current IEEE802.11 wireless LAN standard [20] does not support QoS in medium access control (MAC) layer. Even though the mobile hosts do not move, the QoS for mobile hosts cannot be guaranteed. This led to the development of the IEEE802.11e wireless LAN standard [21] to provide QoS support in MAC layer. The basic 802.11 MAC protocol is the Distributed Coordination Function (DCF). In order to support QoS, Enhanced DCF (EDCF) is defined in the 802.11e MAC protocol. The QoS support in EDCF is realized with the introduction of Traffic Categories (TCs)[30].

Provision of QoS has been studied at various levels in the protocol hierarchy. The present thesis work will discuss the QoS provision to mobile hosts in a different perspective: combining QoS mechanisms of different levels, i.e. combining IntServ/DiffServ, MIPv6, and IEEE802.11e MAC protocol to provide the end-to-end QoS to mobile hosts.

In this thesis, we will evaluate and compare the end-to-end QoS achieved by mobile nodes using combination of IntServ/DiffServ/MIPv6/IEEE802.11 with the case of IntServ/DiffServ/MIPv6/IEEE802.11e. We will study how the IntServ/DiffServ, MIPv6 and IEEE802.11/IEEE802.11e interoperate in detail and how the requested services in different domains are mapped to each other. We will also investigate the end-to-end QoS achieved by mobile nodes after intra-domain and inter-domain movement.

Furthermore, handoff performance is essential to a good mobile QoS model and there have been much work done on this. As we have mentioned above, these QoS solutions do not have enough QoS mechanism to guarantee handoff performance[16]. Moreover, some of them are difficult to be implemented and some are costly. In this thesis, we will propose an approach to improve the handoff performance — assigning *Router Advertisement (RAD)* messages of MIPv6 higher transmission priority supported by IEEE802.11e MAC protocol.

In Fast Handover for Mobile IPv6 protocol, one key protocol operation is setting up a routing path between the previous access router (PAR) and new access router (NAR) to enable the mobile node (MN) to send and receive IP packets [23]. Based on this idea, we will let the MN send *binding update (BU)* messages of MIPv6 to the previous base station that it just visited to improve the handoff performance. In this thesis, we will combine these two approaches to improve the handoff performance of our QoS architecture. Simulation results will demonstrate that these two approaches are efficient to improve handoff performance.

1.2 Thesis Contributions

The major contributions of this thesis are as follows:

- Proposing an end-to-end QoS architecture that combines IntServ/DiffServ, MIPv6, and IEEE802.11e[22]. Analyzing how the IntServ, DiffServ, MIPv6 and IEEE802.11e in this architecture interoperate, how the QoS services in IntServ and DiffServ are mapped and how to realize the QoS services of IntServ/DiffServ in IEEE802.11e.

- Implementing the QoS architecture that combines IntServ/DiffServ, MIPv6, and IEEE802.11/IEEE802.11e with the Network Simulator (NS2).
- Comparing the QoS achieved by the architecture of combining IntServ/DiffServ/MIPv6 and IEEE802.11 with the case of combining IntServ/DiffServ/MIPv6 and IEEE802.11e. Subsequently, demonstrating that the QoS can be guaranteed in the architecture of combining IntServ/DiffServ /MIPv6 and IEEE802.11e in the quantitative point of view[22].
- Proposing an approach to improve handoff performance — Assigning *Router Advertisement (RAD)* messages of MIPv6 higher priority to transmit[22]. It is found that the loss of *RADs* affects adversely handoff performance. Therefore, minimizing or totally eliminating the loss of *RADs* will improve the handoff performance significantly. We achieve this by assigning *RADs* a higher priority. Simulation results show that the handoff performance of MIPv6 with higher priority for *RADs* supported by the IEEE802.11e MAC outperforms the handoff performance with no priority for *RADs*.
- Combining the approach of assigning *RADs* higher priority with the approach of sending *BUs* to the previous base station that proposed in Fast Handovers for MIPv6 to improve the handoff performance of our end-to-end QoS architecture. Conducting simulation to show how much improvement can be achieved by the combination of these two approaches.

1.3 Thesis Organization

The rest of the thesis is organized in a manner as follow:

Chapter 2 is an overview of existing QoS models/protocols: IntServ/RSVP, DiffServ, and QoS supported MAC protocol – IEEE802.11e wireless LAN standard. We will also introduce the mobility support protocol – Mobile IPV6 in this chapter.

Chapter 3 gives a general description of the end-to-end QoS architecture that combines IntServ/DiffServ, MIPv6, and IEEE802.11e. We will describe how IntServ, DiffServ, IEEE802.11e, and MIPv6 interoperate, and how the services that they provide respectively are mapped.

In chapter 4 we propose an approach to improve the handoff performance — assigning *RAD* messages of MIPv6 higher priority to transmit. We also introduce the method that the MN sends *BU* messages of MIPv6 to the previous base station that it just visited.

In chapter 5 we design a simulation to evaluate and compare the end-to-end QoS achieved by the architecture of combining IntServ/DiffServ/MIPv6/IEEE802.11 with the case of combining IntServ/DiffServ/MIPv6/IEEE802.11e. Improvement on handoff QoS performance through two approaches described in chapter 4 is also evaluated by simulation.

Finally the conclusion is given in chapter 6.

Chapter 2

Internet QoS Architectures and Mobile IPv6

2.1 Introduction

An overview of the existing Internet QoS architectures is presented in this chapter, i.e. the Integrated Services architecture, the Differentiated Services architecture and QoS supported MAC protocol — IEEE802.11e wireless LAN standard, as well as Mobile IPv6. We focus on their functionality, characteristics, advantages and disadvantages.

The section 2.2 describes the Integrated Services (IntServ) architecture. Resource ReSerVation Protocol (RSVP) is also given in this section. The section 2.3 presents the Differentiated Services (DiffServ) architecture. Mobile IPv6 is expressed in section 2.4, followed by an introduction of IEEE802.11e wireless LAN standard in section 2.5.

2.2 Integrated Services and RSVP

The Integrated Services (IntServ) architecture introduces a set of extensions to the current Internet architecture in order to enable services that go beyond the traditional best effort service. QoS in terms of IntServ is associated with the time-of-delivery of packets and is characterized by parameters such as bandwidth, packet delay and packet loss rate [3].

IntServ architecture is based on per-flow resource reservation. In order to obtain resource assurance, an application must set up the resource reservation along its path before it can transmit traffic into the network. IntServ model proposes two service classes besides Best Effort (BE) Service: Guaranteed Service (GS) for applications requiring strict delay bound and Controlled load (CL) Service for applications requiring reliable and enhanced best effort service. We will discuss these two kinds of services in detail in the next chapter.

RSVP (the ReSerVation Protocol) is such a resource reservation signaling protocol developed for an application to set up a reservation before transmitting traffic. RSVP is the most complex among all the QoS technologies and provides a high level of QoS guarantee. The basic RSVP operation is illustrated in Fig 2.1 [24]:

- The senders send PATH messages that contain the *traffic specification* (TSpec) information to the receivers to specify the characteristics of the traffic. TSpec includes the following parameters: token bucket rate (r), token bucket size (b), peak rate (p), maximum datagram size (M), minimum policed unit (m). Each RSVP-enabled router along the downstream route toward the receivers establishes a “path-state” that includes the previous source address of the PATH messages.
- After receiving the PATH messages, the receivers send a RESV message toward the senders along the reverse direction of the PATH messages to reserve resource. Besides the TSpec that describes the traffic flow, the RESV message includes a *request specification* (RSpec) that defines the desired QoS and a *filter specification* (filter spec) that characterizes the packets for which the reservation is being made (e.g. the transport protocol and port number). RSpec includes parameters as follows: service rate (R) and slack term (S) Together, the RSpec and filter spec represent a flow-descriptor that is used by routers to identify each reservation (a “flow” or a “session”). Guaranteed

service requires both the RSpec and the TSpec parameters; Controlled load service needs only TSpec.

- When each RSVP router along upstream the path toward the senders receives the RESV message, it uses the admission control process to authenticate the request and allocate the necessary resources. If the request cannot be satisfied (due to lack of resources or authorization failure), the router returns an error back to the receivers. If accepted, the router sends the RESV upstream to the next router. When the last router receives the RESV and accepts the request, it sends a confirmation message back to the receiver.
- When the senders or the receivers end an RSVP session, there is an explicit teardown process for a reservation.

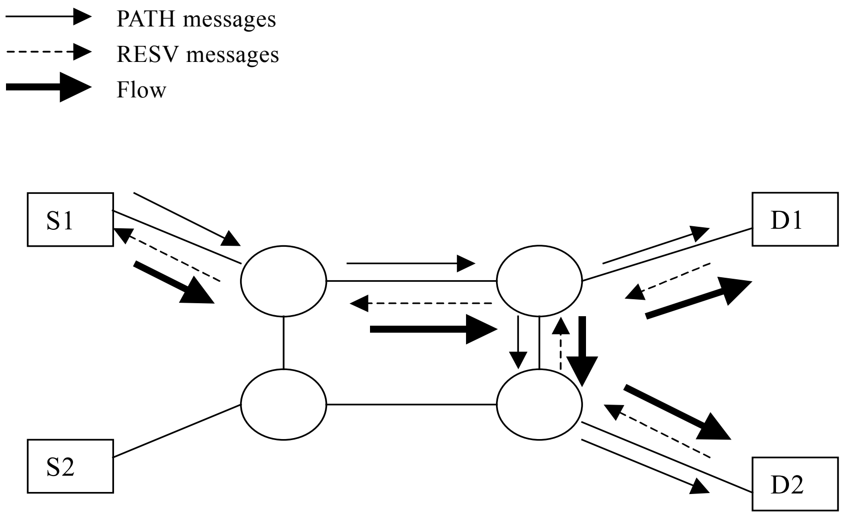


Figure 2.1 Basic RSVP Operation

IntServ is implemented by four components: the signaling protocol (RSVP), the admission control routine, the classifier and the packet scheduler. IntServ model can be logically divided into two parts (Figure 2.2): the control plane and the data plane [24]. The control plane sets up resource reservation and the data plane forwards data packets based on the reservation state.

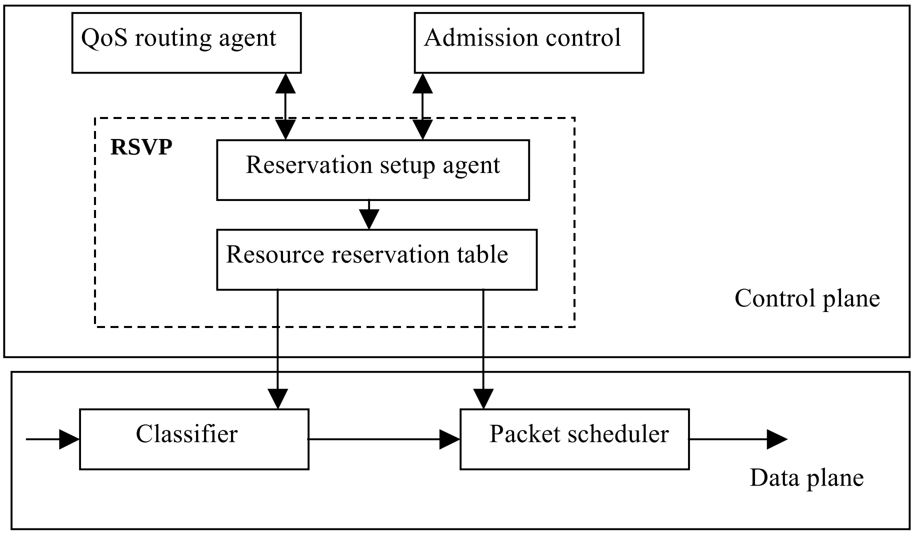


Figure 2.2 Integrated Services Model

Even though IntServ provides the means for end-to-end QoS, it is not widely deployed. The problems with the IntServ architecture are: 1) maintenance and control of per-flow states introduces severe scalability problems at the core networks, where there are millions of flows. Consequently, it is suitable for access network where the number of flows using reservation is modest; 2) the requirement on routers is high. All routers must implement RSVP, admission control, Multi-field (MF) classification and packet scheduling.

2.3 Differentiated Services

To solve the scalability and complexity problem of IntServ and RSVP, Differentiated Services (DiffServ) [4][25] is introduced. Scalability is achieved by offering services on aggregate basis rather than on per-flow basis. With DiffServ, traffic is divided into a small number of groups called *forwarding classes*. The forwarding class that a packet belongs to is encoded into a field in the IP packet header (TOS (Type-of-Service) octet in IPv4 header or Traffic Class octet in IPv6 header).

The TOS byte is divided into 6 bits Differentiated Service Code Point (DSCP) field and a 2-bit unused field [2].

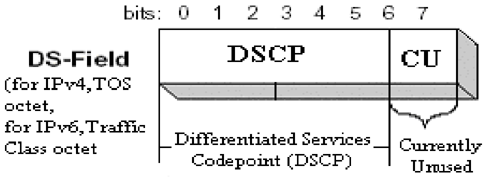


Figure 2.3 Differentiated Services Code Points (DSCP)

DiffServ is constructed by a combination of (i) marking packets with a DiffServ code point (DSCP) at boundary nodes, (ii) using the DSCP to determine how the nodes inside the domain forward packets, and (iii) conditioning the marked packets at boundary nodes.

In DiffServ, externally observable forwarding treatments at a single node are described by the term per-hop behavior (PHB). Each PHB is represented by a 6-bit value DSCP. In DiffServ architecture, PHBs are used as the basic building blocks for resource allocation to different behavior aggregates. There are currently two standard PHBs that represent two service levels (traffic classes): Expedited Forwarding (EF) and Assured Forwarding (AF).

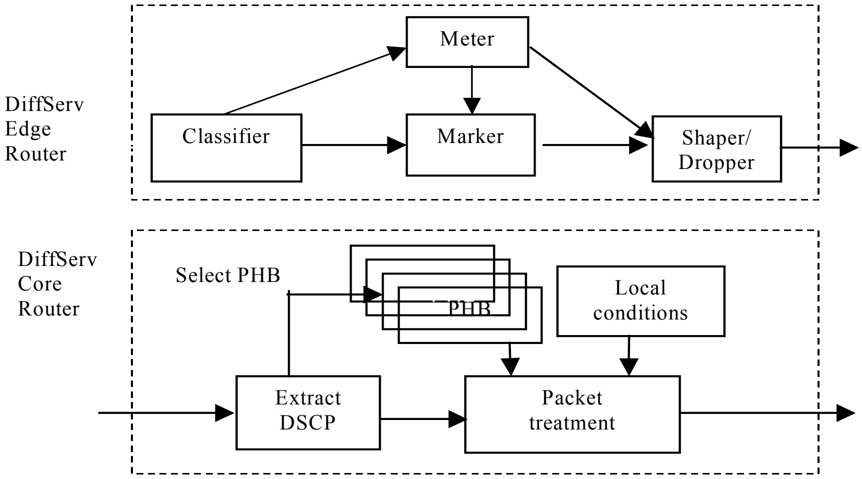


Figure 2.4 DiffServ Routers

In a DiffServ network, the routers at the boundary of the network (boundary routers or edge routers) and routers inside the network (interior routers or core routers) have different responsibilities (Figure 2.4). The edge routers perform the complex classification and traffic conditioning functions while the core routers only forward the packets based on the forwarding classes in the packet header.

DiffServ is quite different from IntServ. Firstly, resources are allocated to aggregated traffic rather than individual flows. In DiffServ, resources are allocated to individual classes that represent aggregated traffic. The performance assurance to individual flows in a forwarding class is provided through prioritization and provisioning rather than per-flow reservation. IntServ approach allocates resources to individual flows, which can run into tens of thousands in a large network. Secondly, there are only a limited number of service classes indicated by the DS field. Since service is allocated in the granularity of a class, the amount of state information is proportional to the number of classes rather than the number of flows. Therefore DiffServ is more scalable and suitable for the core network. Thirdly sophisticated classification, marking, policing and shaping operations are only needed at boundary of the networks. Core routers need only to implement Behavior Aggregate classification. Therefore, it is easier to implement and deploy DiffServ.

2.4 Mobile IPv6

The next generation Internet Protocol – IPv6 [27] has evolved from current IPv4 protocol and is an improvement over IPv4. The size of IPv6 address is 128-bit versus 32-bit of IPv4, which increases the address space by a factor of 2^{96}. Larger address space allows more levels of addressing hierarchy, which lead to more efficient network operations and network scaling. Another difference of IPv6 over IPv4 is that many of the optional fields in IPv4 have been moved into extension headers of IPv6, which allows for more efficient forwarding, less stringent limits on the length of options, and greater flexibility for introducing new options in the future. Moreover IPv6 includes many features for mobility support that are missing in current IPv4 such as Stateless Address Autoconfiguration [28] and Neighbor Discovery [29]. Improvement on Internet security is also one of the biggest differences between IPv6 and IPv4. All IPv6 nodes are expected to implement strong authentication and encryption features to enhance Internet security.

Mobile IPv6 (MIPv6) is designed to manage the mobility of mobile nodes between different IPv6 networks, possibly wireless [9]. It enables mobile nodes to maintain ongoing communications even when changing their points of attachment to the network. The three fundamental functional units within the protocol are the correspondent node (CN), the home agent (HA), and the mobile node (MN). Each mobile node is identified by two IP addresses: home address and care of address (CoA). Home address is the mobile node’s constant IP address that does not change as it moves from one network to another. Care of address is an IP address temporarily assigned to a mobile node when it is visiting a foreign network. The mobile node can configure its care of address by using Stateless Address Autoconfiguration and Neighbor Discovery.

Figure 2.5[26] shows the Mobile IPv6 architecture. In mobile IPv6 route optimization is a mandatory part of the protocol. When a mobile node remains in its home network, it communicates with correspondent nodes through its home address. When a mobile node moves to a new subnet, its home address is not valid anymore and it needs to acquire a new care of address in the visiting subnet. The mobile node registers this care of address with its home agent and correspondent nodes through *binding update (BU)* messages. Once the correspondent nodes know the mobile node's care of address they will send packets directly to mobile node's care of address

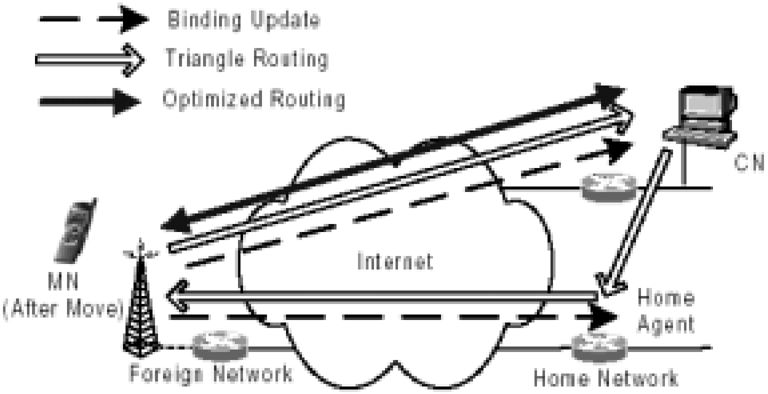


Figure 2.5 Mobile IPv6 Architecture

The above brief overview of Mobile IPv6 operation contains three key components: router discovery, address notification and packet routing.

2.4.1 Router Discovery

Router discovery determines whether the mobile node is currently connected to its home network or a foreign network and whether the mobile node has moved from one network to another. The process is specified in IPv6 Neighbor Discovery document[29]. The base stations (BSs) periodically broadcast *Router Advertisement (RAD)* messages. After receiving *RADs*, The mobile node may perform location and movement detection by examining the network-prefix contained in a received advertisement. If any of these prefixes match the network-prefix of the mobile node's current IP address, then the mobile node is still connected to the current network; otherwise, if none of the prefixes matches the network-prefix of the mobile node's current IP address, then the mobile node has moved to other network. If the mobile node has not received the periodic transmission of a *RAD* for some time, it may sent *Router Solicitation* to base stations to ask for a *RAD*.

Router discovery also helps the mobile node obtain a care of address in the foreign network through *Stateful Address Autoconfiguration* or *Stateless Address Autoconfiguration*.

2.4.2 Address Notification

Address notification is the process by which a mobile node informs both its home agent and various correspondent nodes of its current care of address. The messages used for notification include binding update, binding acknowledgment and binding request:

- A *binding update* message is used by a mobile node to announce that it has changed its point of attachment to the Internet or to renew an existing binding which is about to expire.
- A *binding acknowledgement* message is sent as a reply to a binding update if the “acknowledgement required”-flag in the binding update was set.
- If the correspondent node’s binding to a mobile node is about to expire, the correspondent node may ask the mobile node to renew that binding by sending *binding request* message.

Above three types of Mobile IPv6 address notification messages are encoded as options to be carried within an IPv6 Destination Options Header. Therefore, these messages are only examined by the last destination and not by any intermediate routers along the path.

2.4.3 Packet Routing

When a mobile node is connected to its home network, it sends and receives packets just as stationary node. When a mobile node is connected to a foreign network, packet routing is divided into two cases:

1) Packets are routed to the mobile node

Correspondent nodes have binding caches that contain the currently valid bindings. Each time when a correspondent node is about to send a datagram, it first checks if it has a binding for the destination. If the binding exists, the correspondent node sets the IPv6 destination address to the mobile node’s home address. The routing header is initialized to contain a single route segment, containing the mobile node’s care of address. If the binding does not exist, the packet will be intercepted by the mobile node’s home agent and tunneled to the mobile node’s current care of address.

2) Packets are routed from the mobile node

The mobile node must be able to determine a router that can forward packets generated by itself and then uses standard IP routing to deliver each packet to its destination. For all IPv6 routers that are required to implement router discovery, a mobile node can select any router from which it has received *RADs* and configures its routing table to send all packets to that router.

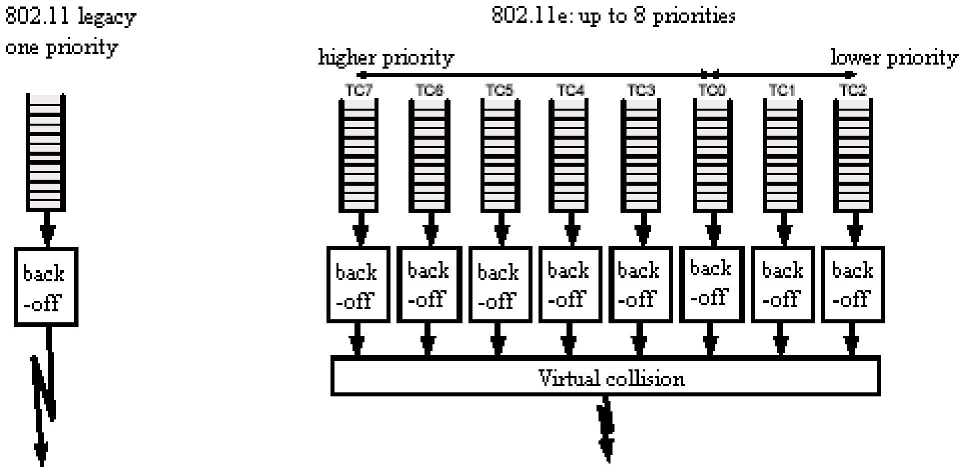
2.5 IEEE802.11e Wireless LAN Standard

IEEE802.11 wireless LAN standard is being accepted widely and rapidly for many different environments today. It can be considered as a wireless version of Ethernet, which supports best effort service [30]. However, the interest in wireless networks supporting QoS has recently grown. Accordingly, a new protocol was developed to enhance the current 802.11 MAC protocol to support applications with QoS requirements, which is named 802.11e.

The basic 802.11 MAC protocol is the Distributed Coordination Function (DCF) that works as listen-before-talk scheme, based on a CSMA/CA (the Carrier Sense Multiple Access /Collision Avoidance) mechanism. The base stations deliver MAC Service Data Units (MSDUs) after detecting that there is no other transmission in progress on the wireless medium. However, if two base stations detect the channel as free at the same time, a collision occurs. In order to reduce the probability of such collisions, a base station performs a backoff procedure before starting a transmission. It has to keep sensing the channel for an additional random time after detecting the channel as being idle for a minimum duration called DCF Inter-frame Space (DIFS). Only if the channel remains idle for this additional random time period, the base station is allowed to initiate the transmission. The duration of this random time is determined as a multiple of a slot time. Each base station maintains a Contention Window (CW), which is used to determine the number of slot times a base station has to wait before transmission. The CW size increases when a transmission fails. After any unsuccessful transmission attempt, another backoff is performed with a doubled size of the CW. The CW value shall be increased exponentially from a CW_{min} value until up to a CW_{max} value during each retransmission. This reduces the collision probability in case there are multiple base stations attempting to access the channel.

It can be seen from the basic DCF mechanism above, that at least two parameters can be used to provide channel access differentiation: the defer time DIFS and CW, based on which the random backoff timer is generated. Lower DIFS and CW values give higher priority for channel access [31]. Instead of treating all traffic with a single DIFS value and a single (CW_{min}, CW_{max}) set, EDCF defines that the channel access has up to eight Traffic Categories (TCs)[30], each with its own Defer Time called Arbitrary Distributed Inter-Frame Space (AIFS) and CW_{min}/CW_{max} values. MSDUs are now delivered through multiple backoff timers within one base station, each backoff timer parameterized with TC-specific parameters (Figure 2.6). One or more user priorities can be assigned to one TC and normally packets belonging to the same priority share one buffering queue.

After introducing the multiple backoff within one base station in EDCF, there exist two levels of channel access contention: internal contention among traffic of different priorities inside the same base station and external contention among traffic from different base stations. Different values of defer timer and backoff timer are used to enable prioritized channel access for different traffic. Higher priority traffic will most probably obtain the channel first and lower priority traffic will have to backoff [31].



One Backoff in DCF

Multiple Backoff in EDCF

Figure 2.6 Internal Contention of Different Traffic Categories

2.6 Summary

In this chapter we described IntServ/RSVP, DiffServ, IEEE802.11e MAC protocol and mobile IPv6. These models/protocols are the important components of the QoS architecture that we will study in the following chapters. In the next chapter, we will introduce an end-to-end QoS architecture for mobile host that combines IntServ/DiffServ, MIPv6, and IEEE802.11e.

Chapter 3

An End-to-End QoS Architecture for Mobile Hosts: IntServ/DiffServ with Mobile IPv6 and IEEE802.11e

3.1 Introduction

In the previous chapter, we have discussed the QoS architectures – IntServ and DiffServ, QoS supported Mac protocol ─ IEEE802.11e wireless LAN standard, and mobility management protocol ─ Mobile IPv6. In this chapter, we will propose an end-to-end QoS architecture for mobile hosts. Some work has been done on providing QoS for mobile hosts based on integrated network, which includes core networks and access networks. For example, Sangheon Pack and Yanghee Choi proposed an end-to-end QoS provisioning architecture in mobile network [40], which consists IntServ model and DiffServ model. This work has been focused on providing handoff QoS in diverse mobility situations. V. Rexhepi, G. Karagiannis and G. heijenk proposed a framework for QoS & mobility that integrated various QoS architectures and mobility protocols [41]. There are some other studies on providing QoS in wired backbone cum wireless network, we will not mention them one by one. These studies have all been carried out on the network layer or higher layer. They have not considered the lower layer, or just assumed the perfect QoS on lower layer. However this is not true for the existing IEEE802.11-based wireless access networks. The current 802.11 MAC protocol does not support QoS. Even though the mobile hosts do not move, QoS for mobile hosts cannot be guaranteed at the last wireless hop. QoS assurances are only as good as their weakest link, which means every segment of the route must have QoS support in order to guarantee end-to-end QoS. In order to compensate the limitation of IEEE802.11 wireless LAN standard, the 802.11 Working Group developed the IEEE802.11e wireless LAN standard which can provide QoS support in MAC layer. In this chapter, an end-to-end QoS provisioning architecture that combines IntServ/DiffServ, MIPv6, and IEEE802.11e is proposed[22].

This chapter is organized such that the integration of IntServ and DiffServ are described first in section 3.2, followed by the description about combination of IntServ/DiffServ, MIPv6 and IEEE802.11e in section 3.3, and ending with the discussion about services mapping among IntServ, DiffServ and IEEE802.11e in section 3.4.

3.2 Integration of IntServ and DiffServ

As be mentioned above, both IntServ and DiffServ architectures are designed to deploy QoS on the current best effort Internet. IntServ and DiffServ are fundamentally different QoS mechanisms. IntServ provides QoS to individual connections while DiffServ provides QoS to aggregates. However they are not mutually exclusive of one another, they complement each other nicely. The framework for Integrated Services operation over Differentiated Services [32] views the two architectures as complementary technologies for deploying end-to-end QoS, e.g. using IntServ at the access networks will enable the hosts to request and reserve resources for per flow by means of RSVP, and using DiffServ in the core networks will avoid the RSVP scalability and complexity problems.

The reference network for the proposed IntServ/RSVP over DiffServ framework described in RFC2998 is shown in Figure 3.1 [34].

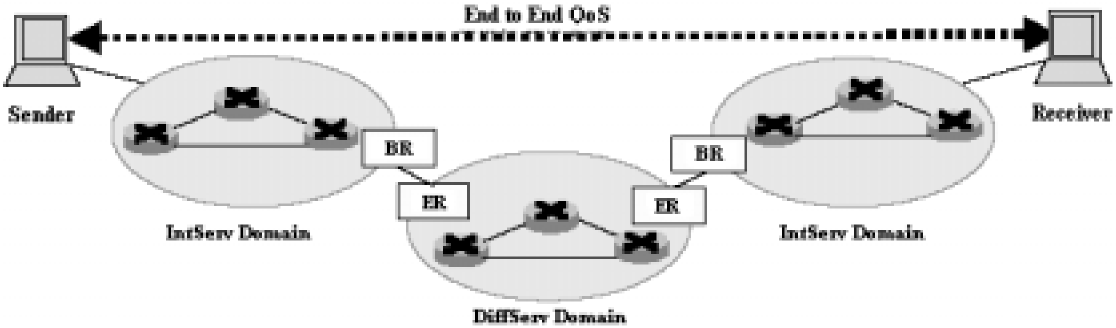


Figure 3.1 The reference network for the IntServ/DiffServ Framework

The above IntServ/DiffServ framework includes a DiffServ domain in the middle and two RSVP/IntServ domains in the edge of large network. In this framework, both sending and receiving hosts use RSVP to communicate the quantitative QoS requirements of QoS-aware applications running on the hosts [32]. RSVP signaling messages travel end-to-end between sending hosts and receiving hosts to support RSVP/IntServ outside the DiffServ domain. DiffServ domain may be RSVP-aware or RSVP-unaware. When DiffServ domain is RSVP-unaware, the routers in DiffServ domain pass RSVP messages transparently. When DiffServ domain is RSVP-aware, RSVP-aware routers in DiffServ domain may perform per flow signaling and admission control. In this thesis, we will mainly study the case when DiffServ domain is RSVP-unaware.

In the above IntServ/DiffServ architectural framework, the routers in IntServ networks are standard RSVP/IntServ routers, performing per-flow RSVP signaling, admission control, policing and scheduling. Border routers (BR) residing at the border of the RSVP/IntServ network are also standard RSVP/IntServ routers. In addition to basic function of RSVP/IntServ routers, BRs might perform some additional functions related to IntServ/DiffServ interoperability such as communicating with the Bandwidth Broker (BB) in DiffServ domain. The DiffServ Core routers are standard DiffServ routers that should apply appropriate PHB to packets based on their DS code point. They may do some limited traffic conditioning. The DiffServ Edge Router (ER) interconnects the DiffServ domain either to RSVP/IntServ domains or to other DiffServ domains. The functions of ER include traffic conditioning between peering domains, interoperating with RSVP/IntServ domains and tunneling RSVP messages.

Integration of IntServ and DiffServ is an efficient solution to provide QoS in traditional wired network. Simulation results have demonstrated that end-to-end QoS requirement of end applications can be met when IntServ is run over a well-provisioned DiffServ [8]. It becomes more difficult to introduce QoS in an environment of mobile hosts and wireless networks due to scarce resources. Therefore, IntServ/DiffServ interoperation would be more valuable if it is to be applied in a wireless environment.

3.3 Combination of IntServ/DiffServ, Mobile IPv6 and IEEE802.11e

In the previous subsection, we have discussed the QoS provision by the integration of IntServ and DiffServ in wired network. Now we discuss the QoS provided to mobile hosts that are linked by means of a wired backbone. As discussed above, Integration of IntServ and DiffServ can provide end-to-end QoS in wired network, 802.11e MAC protocol can support QoS in MAC layer, and Mobile IPv6 is the most perfect mobility management protocol at present. Therefore we propose an end-to-end QoS provisioning architecture that combines IntServ/DiffServ, mobile IPv6, and IEEE802.11e in wired backbone cum wireless network.

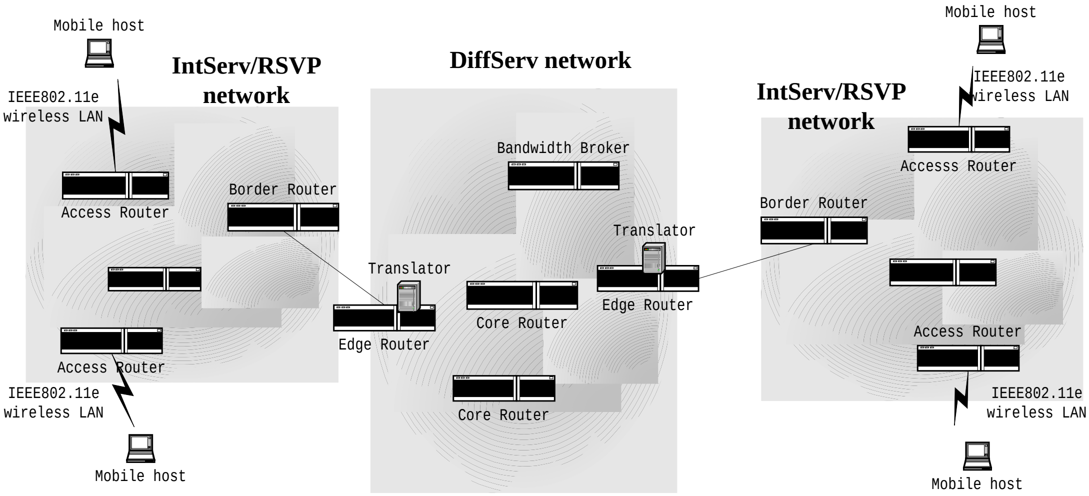


Figure 3.2 An End-to-End QoS Architecture for Mobile Hosts

Figure 3.2 shows an end-to-end QoS provisioning architecture that composes three parts: a wired core network (DiffServ network), two wired access networks (IntServ/RSVP networks) and several wireless access networks (IEEE802.11e-based access networks).

The DiffServ network is composed of bandwidth broker, edge routers, core routers and translators. Bandwidth broker is an agent responsible for managing resources for QoS services in DiffServ domain. It can be a router or a software entity. Resource management includes intra-domain resource management that deals with allocation of resources within a network or a domain, and inter-domain resource management that is concerned with provisioning and allocating resources at network boundaries between two domains. Edge router classifies, meters packets and marks DSCP field in the packets. Core router forwards the received packets according to PHBs. Translator plays an interface role between IntServ region and DiffServ region. It can be a router or a small program. Translator maps services and passes information between IntServ network and DiffServ network.

The access networks use IntServ/RSVP model. The routers in IntServ network are general routers capable of RSVP messages. Access routers that reside the border of IntServ networks are routers that act as base stations for mobile hosts.

In the wireless access networks, IEEE802.11e MAC protocol is used to support QoS for mobile hosts.

The mobility of mobile hosts includes intra-domain mobility and inter-domain mobility. Intra-domain mobility means that a mobile host moves within one access network and from one cell to another adjacent cell. Inter-domain mobility means that a mobile host moves from an access network to another access network.

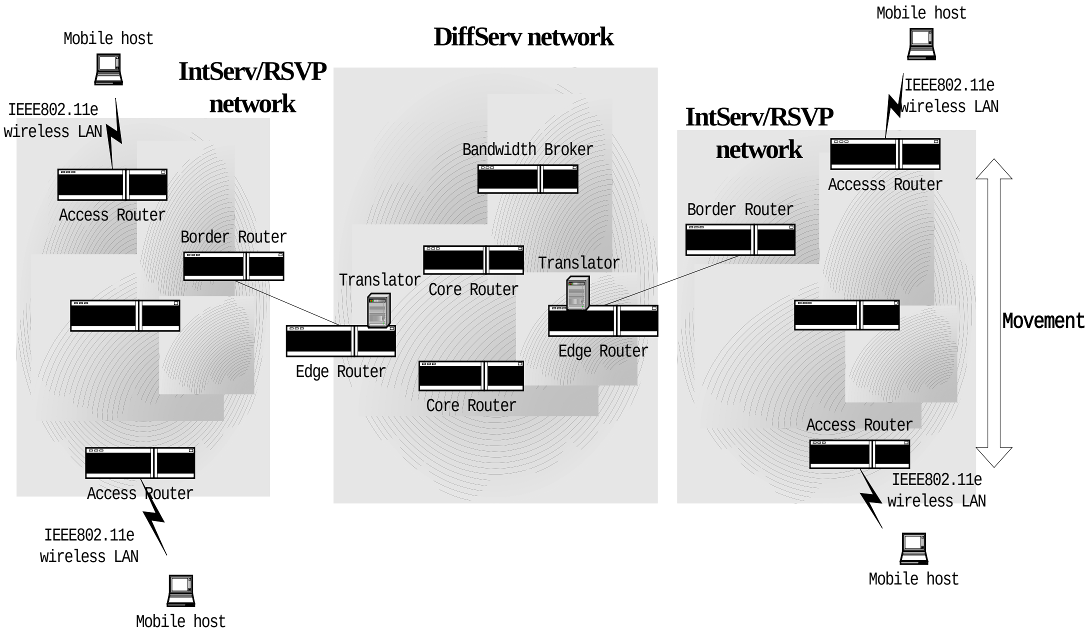


Figure 3.3 Mobile Host's Intra-Domain Mobility

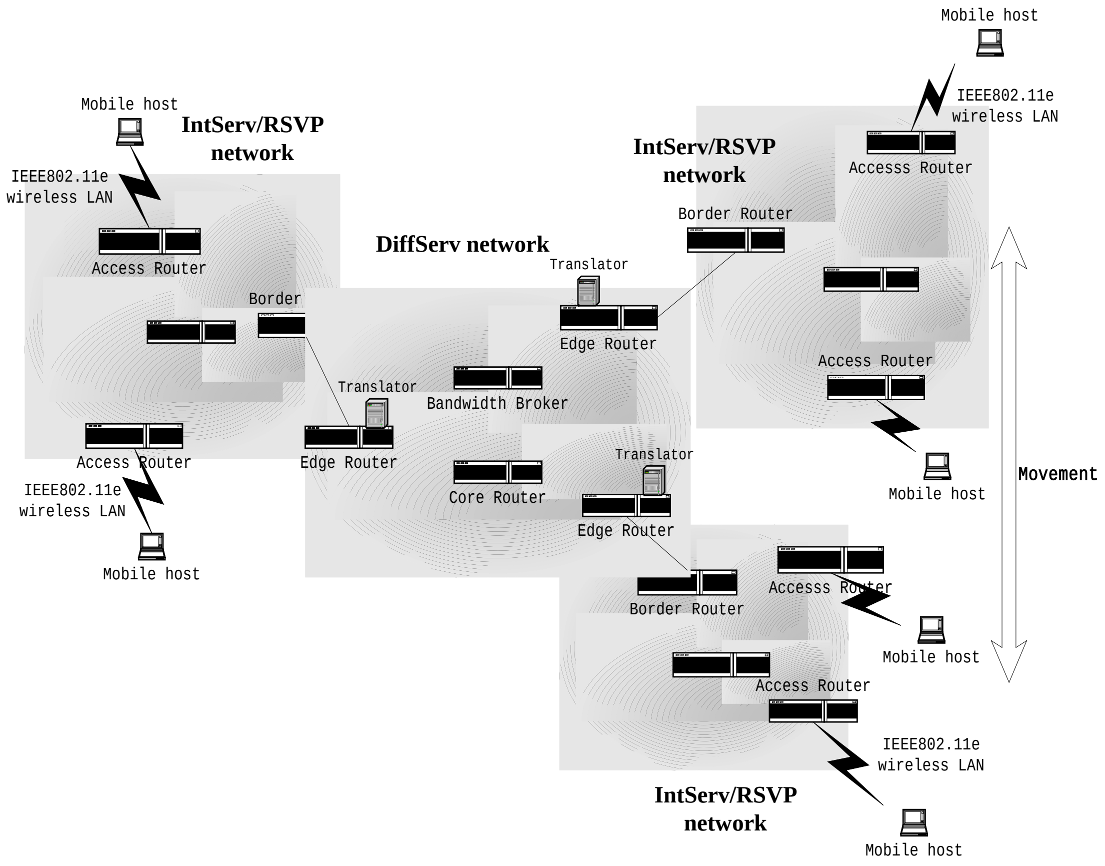


Figure 3.4 Mobile Host's Inter-Domain Mobility

Intra-domain mobility is shown in Figure 3.3 and inter-domain mobility in Figure 3.4. During intra-domain mobility the edge router through which the traffic goes into core network does not change while during inter-domain mobility the traffic goes into core network through other edge router.

IntServ and DiffServ can provide QoS in wired network. For wireless network, IEEE802.11e MAC protocol can provide QoS support for mobile host through assigning different flow different priority to transmit in MAC layer.

3.4 Services Mapping

The primary issue in integrating IntServ, DiffServ and IEEE802.11e is services mapping not only between IntServ and DiffServ but also between IntServ/DiffServ and

IEEE802.11e. In this subsection, we will firstly describe the services provided by IntServ and DiffServ. IEEE802.11e protocol does not provide explicit QoS services, it provides QoS guarantee by assigning different transmission priority to different traffic. After that, we will illustrate service mapping among IntServ, DiffServ, and IEEE802.11e.

3.4.1 Services Provided by IntServ

In addition to the best effort service, IntServ model provides two services: Guaranteed service (GS) and Controlled-load service (CL).

1) Guaranteed Service

Guaranteed Service (GS) guarantees that datagrams will arrive within the guaranteed delivery time and will not be discarded due to queue overflows, provided that the flow's traffic stays within its specified traffic parameters[36]. The service provides assured level of bandwidth or link capacity for the data flow. The GS service controls only the maximum delay, it does not control the minimum delay or minimize the jitter. The delay consists of the fixed delay and the queuing delay. The fixed delay includes transmission delay, propagation delay etc., which is a property of the chosen path and is determined not by GS but by the setup mechanism. Only queuing delay is determined by GS service. GS service imposes a strict upper bound on the end-to-end queuing delay as data flows through the network. The delay bound is usually set large enough even to accommodate cases of long queuing delays.

2) Controlled-load Service

Controlled-load (CL) service does not accept or make use of the specific QoS parameters such as packet loss and delay as control parameters. Instead, acceptance of a request for controlled-load service is defined to imply a commitment by the network elements to provide a service closely equivalent to that provided to uncontrolled (best effort) traffic under lightly loaded conditions[37]. The goal of this service is to provide the same QoS under heavy loads as under unloaded conditions. Although there is no specified strict bound on delay, it ensures that a very high percentage of packets do not experience delays much greater than the minimum transmission delay and do not experience congestion loss.

3.4.2 Services Provided by DiffServ

DiffServ provides Expedited Forwarding (EF) PHB and Assured Forwarding (AF) PHB besides Best Effort (BE).

1) Expedited Forwarding (EF) PHB

The EF PHB provides a low loss, low latency, low jitter, assured bandwidth, and end-to-end service. This service has also been described as Premium Service. Loss, latency and jitter are due to the queuing experienced by traffic while transiting the network. Therefore, providing low loss, latency and jitter for some traffic aggregate means that there are no queues (or very small queues) for the traffic aggregate[38]. In order to ensure that there is almost no queuing delay for these premium packets, the aggregate of the EF traffic's maximum arrival rate must be less than its configured minimum departure rate at every transit node.

When implementing EF PHB, some means must be included to limit the damage that EF traffic could inflict on other traffic. Packets exceeding this limit must be shaped or discarded by traffic conditioners to bring the traffic into conformance.

2) Assured Forwarding (AF) PHB

Assure Forwarding (AF) PHB group is a means to offer different levels of forwarding assurances for IP packets[39]. This service provides a reliable service for customers even during network congestion. The assured service traffic is considered *in*-profile if the traffic does not exceed the bit rate allocated for the service; otherwise, the excess packets are considered *out*-of-profile. The *in*-profile packets should be forwarded with high probability. However, the *out*-of-profile packets are delivered with lower priority than the *in*-profile packets[8].

Assured Forwarding PHB group provides forwarding of packets in four independently forwarded AF classes. Within each AF class, a packet is assigned one of different levels of drop precedence. Each class is allocated a configurable minimum amount of buffer space and bandwidth. In the case of network congestion, the drop precedence determines the relative importance of the packets within the AF classes.

There are no strict timing requirements (delay or delay jitter) associated with the forwarding of AF packets.

3.4.3 Services Mapping

1) Services Mapping between IntServ and DiffServ

IntServ service types are specified by a set of parameters known as TSpec (*Traffic Specification*) while DiffServ service types are specified by the DiffServ Code Points (DSCPs). When combining IntServ with DiffServ, IntServ services must be mapped into DiffServ network. The mapping procedures include [35]:

- Selecting the appropriate PHBs in the DiffServ domain for requested service in the IntServ domain (when the PHB has been selected for a particular IntServ flow, it is necessary to assign an appropriate DSCP to packets from this flow);
- Performing appropriate policing, shaping and marking at the edge router of the DiffServ domain;
- Taking into account the resource availability in the DiffServ domain, perform admission control for traffic coming from the IntServ domain.

When a PHB is selected for a particular IntServ flow specified by TSpec, it is necessary to assign an appropriate DSCP code to packets from this flow. To ensure that QoS can be achieved for IntServ flows when running over a DiffServ domain, appropriate service mapping should be selected.

Both IntServ and DiffServ define different services that can be used by different types of applications. As we discussed above, the EF service in DiffServ provides a low loss, low latency, low jitter and assured bandwidth end-to-end service, which is nearly equivalent to GS service in IntServ that offers strict assurance of both throughput and delay. These two kinds of services are suitable for real time applications such as Voice over IP (VoIP), which is called non-adaptive applications. On the other hand, the AF service in DiffServ is a means to offer different levels of forwarding assurances for IP packets. It could implement the function of CL service in IntServ that requires services tightly approximate to BE service under unloaded network conditions. AF service and CL service can support adaptive applications such as one-way voice or video, which request soft QoS guarantees for their operation, i.e. they may be tolerant in terms of delay bounds and jitter. Both Best Effort services in IntServ and DiffServ do not guarantee any bandwidth and only get the available bandwidth. They are associated with applications requiring no QoS like file transfer or e-mail. Therefore GS service is mapped to EF service, CL service is mapped to AF service and BE service is still BE service.

2) Services Mapping between IntServ/DiffServ and IEEE802.11e

The services in IntServ and DiffServ are not supported in wireless mobile environment. Therefore in order to guarantee end-to-end QoS for mobile host, these services have to be mapped to the ways that support different QoS requirements in MAC layer. In the IEEE802.11e MAC protocol, there are eight traffic categories (TCs) with different priority. MAC Service Data Units (MSDUs) are delivered through multiple backoff timers that are determined by TC-specific parameters. In order to satisfy different QoS requirements, the IntServ Guaranteed Service (GS) class is mapped to TCs of higher priority, Controlled Load (CL) class is mapped to TCs of medium priority and the Best Effort (BE) class is mapped to TCs of the lower priority.

Now we summarize services mapping in the following table.

Table 3-1 Services mapping

IntServ	**DiffServ**	**IEEE802.11e**
Guaranteed Service	Expedited Forwarding	TCs of higher Priority
Controlled-load Service	Assured Forwarding	TCs of medium Priority
Best effort Service	Best effort Service	TCs of lower Priority

3.5 Summary

In this chapter we proposed an end-to-end QoS architecture for mobile host, which combines IntServ/DiffServ, mobile IPv6, and IEEE802.11e. We introduced how its components interoperate and how the services that provided by every component are mapped. However, we only illustrate that this architecture could provide end-to-end QoS for mobile host from the concept. We will evaluate this QoS architecture through simulation in chapter 5.

Chapter 4

Improvement on Handoff QoS of the Architecture that Combines IntServ/DiffServ, MIPv6 and IEEE802.11e

4.1 Introduction

In the above description of the end-to-end QoS architecture for mobile hosts that combines IntServ/DiffServ with Mobile IPv6 and IEEE802.11e, we have illustrated that every segment of QoS “chain” in this architecture can provide QoS support, and so the total end-to-end QoS can be guaranteed. However we have not yet discussed the QoS performance of our proposed architecture during handoff. Handoff performance is a significant factor in evaluating wireless networks. As we mentioned in chapter 1, there have been several schemes to provide handoff QoS when mobile hosts move. Therefore we might combine these schemes with our QoS architecture to provide end-to-end QoS for mobile hosts no matter they stay within one subnet or roam to other subnets. However, these schemes provide handoff QoS by means of extending RSVP to mobile environment. RSVP is almost the most complicated QoS technology for applications and for network elements, therefore these handoff QoS schemes are also complicated. Implementing them in real wireless network is either difficult or costly. Here we consider the methods to improve handoff QoS performance from the viewpoint of mobile IPv6 itself. In this chapter, we propose an approach that assigns *router advertisement (RAD)* messages of mobile IPv6 higher transmission priority[22], which is carried out on MAC layer and supported by QoS feature of IEEE802.11e. Then we use this approach, and combine another approach that the MN sends *binding* *update (BU)* messages of mobile IPv6 to the previous base station that it just visited, to improve the handoff performance of our QoS architecture.

Firstly we will describe the factors that affect the handoff performance and introduce some existing approaches to improve the handoff performance in section 4.2. In section 4.3 we will describe the effect of losing *RADs* on handoff performance. Then we will provide a solution to eliminate the loss of *RADs* – by assigning *RADs* higher priority to transmit.

4.2 Handoff Procedure

When a mobile node performs a handoff between different IP subnets it must perform three basic steps: movement detection, forming a new care-of address and sending *BUs* to HA and CNs to redirect the data flow. In a nutshell, a mobile node detects that it has moved to a new subnet by analyzing *RADs* periodically broadcasted by the base station. The mobile node can also request the base station to send *RADs* by sending a *router solicitation*. After receiving *RADs*, the mobile node first needs to verify the uniqueness of its link-local address. It performs duplication address detection (DAD) on its link-local address. Then the mobile node may obtain a new care of address through stateless or stateful address autoconfiguration. Once it has obtained a new care of address, it may perform DAD for this new care of address. However, DAD takes quite long time with respect to the handoff latency. Therefore, the mobile node should perform DAD in parallel with its communications, or choose not to perform it [33]. Once the new care of address construction is done, the mobile node sends *BUs* to its HA and CNs to update its binding cache.

Handoff latency includes the time to discovery the new prefix on the new subnet, the time to form the new care of address, and the time to notify HA and CNs about the new locality of the MN. In order to reduce the handoff latency, it is very important to minimize the movement detection delay, CoA forming delay and flow redirection delay.

As to CoA forming delay, besides the time to perform DAD, it is said that the long random wait time for getting a care of address is another factor that leads to CoA forming delay. The random wait time is introduced to avoid synchronization. A shorter random wait times is recommended to decrease this delay [42].

When a MN gets a new CoA, it sends *BU* messages to HA and CNs to indicate its new CoA. If the HA and/or CNs are far from the MN, even if the MN's movement is small, *BU* messages have to travel across several IP networks to reach HA and CNs. This will lead to long handoff latency. Moreover, before the *BU* messages reach CNs, CNs do not know the change of MN's address, and so they still send packets to the previous address, which will cause these packets dropped. Here we adopt the idea of signaling and tunnel between previous access router (PAR) and new access router (NAR) in Fast Handovers for MIPv6[23] to solve this problem. The key operation of Fast Handovers protocol involves setting up a routing path between the PAR and the NAR to enable the MN to send and receive IP packets. After the PAR receives a Fast Binding Update message from the MN, it can forward the packets to MN on the tunnel between the PAR and the NAR. In this thesis, we also let the MN send *BUs* indicating the new CoA to the previous base station that it just visited to reduce the handoff latency and handoff packet drop. In the chapter 5, we will show how much improvement this method can be achieved quantitatively through simulation.

4.3 Improving Handoff QoS Performance by Assigning *RADs* Higher Priority

Now we focus on the first step of handoff procedure —movement detection. Base stations (BSs) sent *router advertisements (RADs)* message periodically. When a mobile node enters into the coverage area of a base station, at some point of time it receives a *RAD* message from this new base station. The mobile node maintains a base station (BS) list that includes the base stations from which it received *RADs*. The BS list needs to be updated. When a mobile node receives a *RAD* from a base station, it will check its BS list. If this base station is not in the BS list, the mobile node will begin to do handoff to attach itself to the new base station and add a entry for this new base station. If this base station is already in the BS list, the mobile node refreshes the lifetime entry of this base station. In this way, a handoff per each received *RADs* from a different base station while the mobile node moves within overlapping coverage area of two base stations is avoided[43].

The problem arises when some *RADs* are dropped due to congestion in the wireless channel. The loss of *RADs* will cause the problem that the mobile node cannot detect its movement to a new link instantly. Moreover, when the mobile node is in the overlap area of two base stations, the loss of some *RADs* could lead to the situation of unnecessary handoffs due to the lack of refreshment of the BS list. The mobile node may switch between two base stations repetitively, which is usually called *ping-pong* handoff. This ping-pong handoff between two base stations may lead to large handoff latency and affects the handoff performance seriously.

IEEE802.11e MAC protocol can transmit different traffic with different priority. In order to eliminate the loss of *RADs*, we use this QoS feature of IEEE802.11e to assign

RADs a higher priority to transmit. We will minimize or totally eliminate the loss of *RADs* by this way and improve the handoff performance.

Now we will investigate the drop of *RADs* with no priority and with higher priority. Here we mainly consider the *RADs* that are dropped in IFQ (Interface Queue) due to congestion. The data transmission rate of wireless channel is 1M, and the saturated bandwidth of wireless channel is around 0.8 ~ 0.9M.

Figure 3.5 shows the comparison of the probability of dropping *RADs* between with higher priority for *RADs* and with no priority for *RADs*.

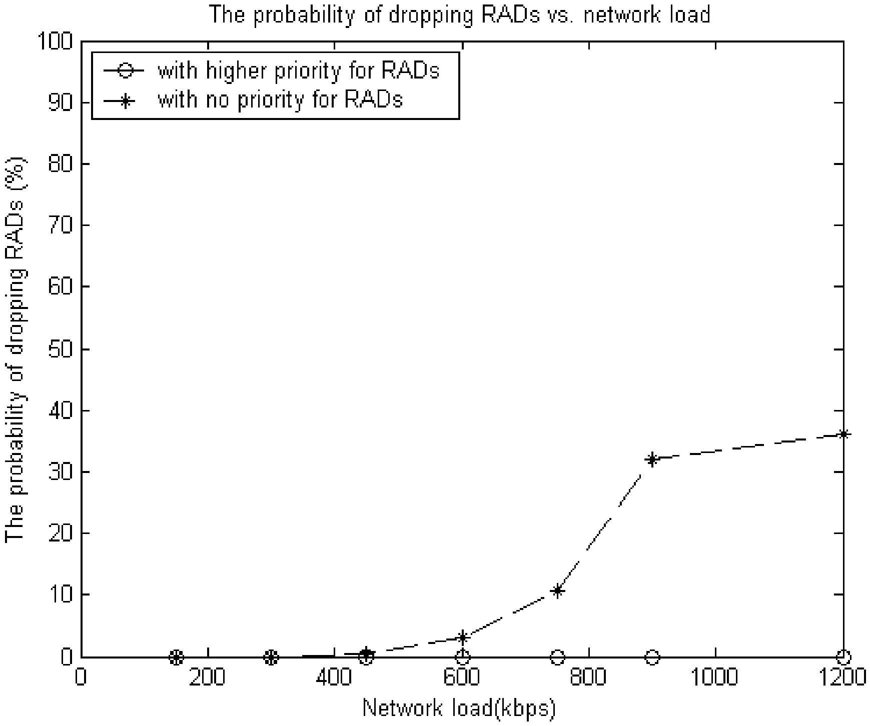


Figure 4.1 The probability of dropping *RADs* in IFQ varying with network load

From the figure we can see that if we do not assign the *RADs* priority, when the network is under medium load (less than 0.6M), the probability of dropping *RADs* is about 3%, i.e. the *RADs* are dropped slightly; when the network load is more than 0.6M, the probability of dropping *RADs* becomes larger. When the network load is in congestion, the probability of dropping *RADs* is more than 30%, which will worsen the handoff performance seriously. If we assign *RADs* message higher priority to transmit, there are not dropped *RADs* even though the network is under congestion.

In the next chapter, we will discuss the effect of assigning *RADs* higher priority on handoff performance in detail.

4.4 Summary

In this chapter, we proposed an approach to improve the handoff performance, which is assigning *router advertisement (RAD)* messages of mobile IPv6 higher transmission priority. We combined this approach with another approach that the MN sends *binding update (BU)* messages of mobile IPv6 to the previous base station to improve the handoff performance of our end-to-end QoS architecture proposed in chapter 3. We will examine the efficiency of these two approaches through simulation in next chapter.

Chapter 5

Simulation Studies

5.1 Introduction

In the previous chapters, we have already discussed the end-to-end QoS architecture that combines IntServ/DiffServ, mobile IPv6 and IEEE802.11e. We have also proposed an approach of assigning *RADs* higher priority and combined this approach with another approach of sending *BUs* to the previous base station to improve the handoff performance of our QoS architecture. In this chapter, we will conduct simulation to evaluate the end-to-end QoS provided by this architecture and to examine how much the handoff performance can be improved through these two approaches. The purposes of our simulation are:

1) Evaluate and compare the end-to-end QoS of the architecture that combines IntServ/DiffServ, MIPv6, and IEEE802.11 with that of the architecture that combines IntServ/DiffServ, MIPv6, and IEEE802.11e. The end-to-end QoS is studied through following cases: when mobile nodes stay in home network, after intra-domain handoff, and after inter-domain handoff[22].

2) Study and compare the handoff performance of the QoS architecture of no priority for *RADs* with that of higher priority for *RADs* supported by the IEEE802.11e MAC protocol[22];

   Study and compare the handoff performance of the QoS architecture of *BUs* sent to the previous base station that mobile node just visited with that of no *BUs* sent to the previous base station.

This chapter will begin with a description of simulation overview, including simulation tools, performance criteria, simulation scenarios, and simulation configurations etc. After that, we will do simulation on different scenarios and analyze the simulation results.

5.2 Simulation Overview

5.2.1 QoS Models in Simulation System

Our simulation is based on the Network Simulator 2 (NS2)[44], which is developed by University of California at Berkeley and Lawrence Berkeley National Laboratory and widely used by the networking community to analyze IP networks. Besides the basic NS2 system, our simulation system comprises RSVP model, DiffServ model, Mobile IPv6 model and IEEE802.11e model. In order to make simulation results closer to real situation and more persuasive, we are not intending to develop new models in this study, but to use off the shelf and robust models that have been verified by other researchers as reliable and stable.

- RSVP model for NS2 implementation [45] is taken from University of Bonn. This RSVP/NS model uses Weighted Fair Queuing (WFQ) to enforce bandwidth guarantee. It includes soft state with freely adjustable refresh intervals, FF reservation style, gathering of statistics in the links and nodes, and interfaces to parameter based and measurement based admission control. It also includes the possibility of reserving a portion of a link's bandwidth for RSVP messages to avoid the loss of RSVP messages[46].
- DiffServ support has been added to NS in 2000 [47]. The model has been developed by Nortel Networks since 1998. This model follows the DiffServ architecture in that a node can have Edge or Core capabilities. It includes

many of the popular PHBs, Meters, and Policers available at the IETF today. This model has been found to provide suitably similar results in comparison with real world DiffServ products.

- MobiWan is software package based on NS2 to simulate mobile IPv6 [48][49]. It has been developed by MOTOROLA Labs Paris in collaboration with INRIA PLANETE Team. MobiWan has developed a set of NS Agents that simulate the Mobile IPv6 and IPv6 protocols. It can support local mobility (within a single administrative domain) and global mobility (across domain boundaries). In MobiWan, NS addressing is brought from 3 levels to 4 levels, which makes configuration of the network easier.
- IEEE802.11e model we use is developed by He DaJiang and Shen Qi, from Institute for Infocomm Research of A*STAR in Singapore. So far, they have implemented EDCF in NS2. The exact values for the QoS parameters are used according to IEEE802.11e wireless LAN Standard.

The above four models are developed by different institutes under different versions of NS. They are not interoperable directly each other. Therefore we have to modify them to make them compatible. It is involved and time-consuming work. Here we just state the main modification and do not present the details one by one. The main modification include:

1) These four models work with different version of NS. Therefore, the first task is to make them workable with the same version of NS. In our simulation, we use NS 2.1b6.

2) Add translator in DiffServ domain, which is attached to edge router. Translator is responsible for integrating IntServ and DiffServ model, including translating messages and mapping services between IntServ domain and DiffServ domain.
3) Change the flat address format in wired network in NS2 to hierarchical address format in order to be compatible to MobiWan (simulation platform of mobile IPv6).

5.2.2 Performance Criteria

In this simulation we use goodput, end-to-end delay, delay jitter, packet drop ratio as QoS performance criteria and use handoff latency, handoff packet drop to evaluate the handoff performance. The performance of flows that require GS, CL and BE services (we call them GS flow, CL flow and BE flow respectively) will be compared and evaluated using these criteria.

- **Goodput (kbps)**: how much data the receiver receives per second.
- **Delay (s):** the time that a packet needs to travel through the network from the sender to the receiver.
- **Average flow delay (s):** the average of all packets delay in the flow;
- **Delay jitter (s):** the difference between the maximum delay and the minimum delay;
- **Packet drop ratio (%):** the rate at which packets are dropped while traversing through the network
- **Handoff latency (s):** the time that elapses between the last packet received by mobile node via the old base station and the first packet received by mobile node via the new base station after a handoff.
- **Handoff packet drop (pkt):** the number of packets lost due to handoff.

5.2.3 Simulation Scenario

In order to achieve the simulation objectives, our simulation will include three parts:

Part 1: Simulation study on the handoff performance of the QoS architecture

In this part we will investigate the improvement on handoff performance through two approaches: assigning *RADs* higher priority and sending *BUs* to the previous base station that the mobile node just visited. This part includes two scenarios:

- Scenario 1: The mobile node firstly moves from one cell to adjacent cell within the same subnet, and then moves to other subnet. In this case, we will compare handoff performance of the QoS architecture of *BUs* sent to the previous base station that the mobile node just visited with that of no *BUs* sent to the previous base station.
- Scenario 2: The mobile node moves from original location to adjacent cell within the same subnet. The interest of this case is to compare handoff performance of the QoS architecture of no priority for *RADs* with that of higher priority given to *RADs* supported by IEEE802.11e under different simulation conditions.

Part 2: Simulation study on the end-to-end QoS provided by the QoS architecture

In this part we will evaluate and compare the end-to-end QoS provided by the architecture of IntServ/DiffServ/MIPv6/IEEE802.11 with that provided by the architecture of IntServ/DiffServ/MIPv6/IEEE802.11e, We will also investigate the end-to-end QoS obtained by the mobile node after intra-domain handoff and inter-domain handoff. This part also includes two scenarios:

- Scenario1: The mobile node is connected to home agent and does not move. Here, we compare the QoS obtained by GS, CL and BE flows using a combination of IntServ/DiffServ/MIPv6/IEEE802.11 with the case of IntServ/DiffServ/MIPv6/IEEE802.11e;

- Scenario2: The mobile node firstly moves to the adjacent base station within one subnet, and then moves to other subnet. In this scenario, we study the end-to-end QoS obtained by the mobile node after intra-domain and inter-domain handoff.

Part 3: Simulation study on a more practical situation

In this part, we will simulate on a more practical scenario, which comprises more sources, more mobile nodes and different traffic types. We will study the average end-to-end QoS obtained by more mobile nodes, which includes average end-to-end QoS before movement, after intra-domain handoff and after inter-domain handoff.

5.2.4 Simulation Configuration and Parameters

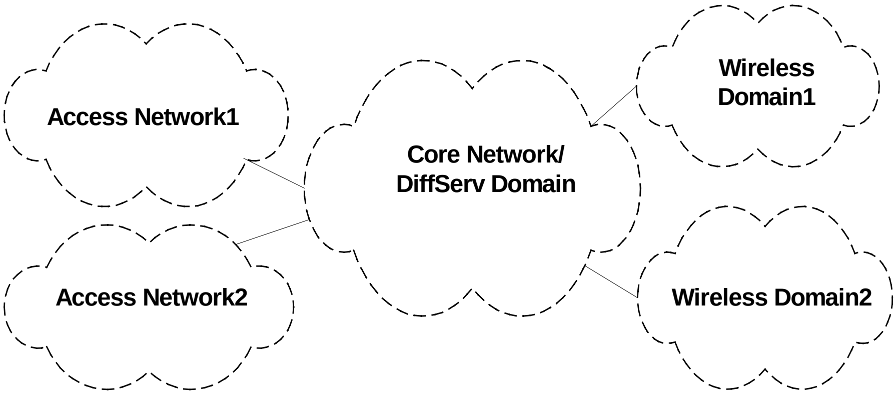


Figure 5.1 Basic simulation configuration

The simulation in this thesis is based on the basic simulation configuration shown in Figure 5.1, which comprises five domains: two access networks, one core network/DiffServ domain, and two wireless domains. In the following sections, simulation will be studied in three parts. In parts 1 and 2, access network1 is IntServ domain that includes two RSVP-sources, access network2 is a best effort domain that includes one best effort source. In part 3, two access networks are IntServ domains with more sources. The links inside IntServ Domain and the links that connect IntServ domain with DiffServ domain are RSVP-link on which the resources are reserved. The links inside DiffServ domain and the links that connect the DiffServ with wireless domain are DiffServ-link on which the packets are forwarded based on their DSCP (in DiffServ model in NS, the DiffServ-link is simplex-link).

During simulation, we assume sources are fixed nodes, and destinations are mobile nodes. Now we list the common parameters and traffic management in simulation. The individual parameters will be listed in every simulation scenario.

Table 5-1 The common parameters

Packet size	1000bytes
Link delay (in core network)	25ms
Link delay (in access network)	1ms
Bandwidth (in wired network)	1M
Bandwidth (in wireless LAN)	1M

Table 5-2 Traffic management

	IntServ/RSVP domain	DiffServ domain
Admission control	Param*	Token Bucket
Buffer management	DropTail	RED
Scheduler	WFQ	Priority

Param* — A parameter-based "Simple Sum" algorithm[45]

5.3 Simulation Study – Part 1: Simulation study on the handoff performance of the QoS architecture

In this part we will investigate the improvement on handoff performance through two approaches: assigning *RADs* higher priority and sending *BUs* to the previous base station that the mobile node just visited.

In order to investigate the efficiency of each approach and the ultimate handoff performance after two improvements, we will combine two approaches to study the handoff performance in the simulation study. That is, we compare the handoff performance of the QoS architecture of *BUs* sent to the previous base station with that of no *BUs* sent to the previous base station under the condition that *RADs* have been assigned higher priority. Additionally, we compare the handoff performance of the QoS architecture of higher priority for *RADs* with that of no priority for *RADs* under the condition that the *BUs* have been sent to the previous base station.

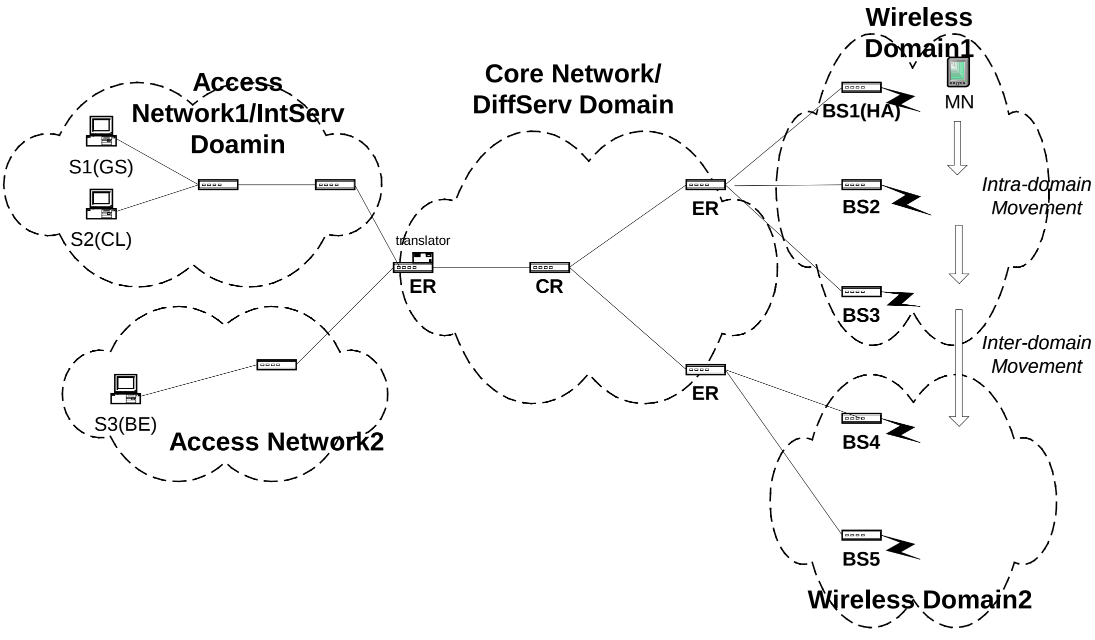


Figure 5.2 Simulation Configuration1

5.3.1 Scenario 1: Improvement on Handoff Performance by Sending *BUs* to the Previous Base Station

Improvement on handoff performance by sending *BUs* to the previous base station that the mobile node just visited will be studied in this scenario. Simulation configuration is shown in Figure 5.2.

The wireless domains are IEEE802.11e wireless LAN. The base station BS1 is the home agent of the mobile node. In this case, the mobile node is not originally connected to its home agent BS1, but connected to BS2. The mobile node firstly moves from BS2 to BS3, then to BS4.

Firstly, we investigate how many packets are tunneled from the previous base station that MN just visited to the current base station. Then we will compare the intra-domain handoff latency and inter-domain handoff latency of *BUs* sent to the previous base station with that of no *BUs* sent to the previous base station. After that we will evaluate the improvement on intra-domain handoff performance and inter-domain handoff performance by sending *BUs* to the previous base station.

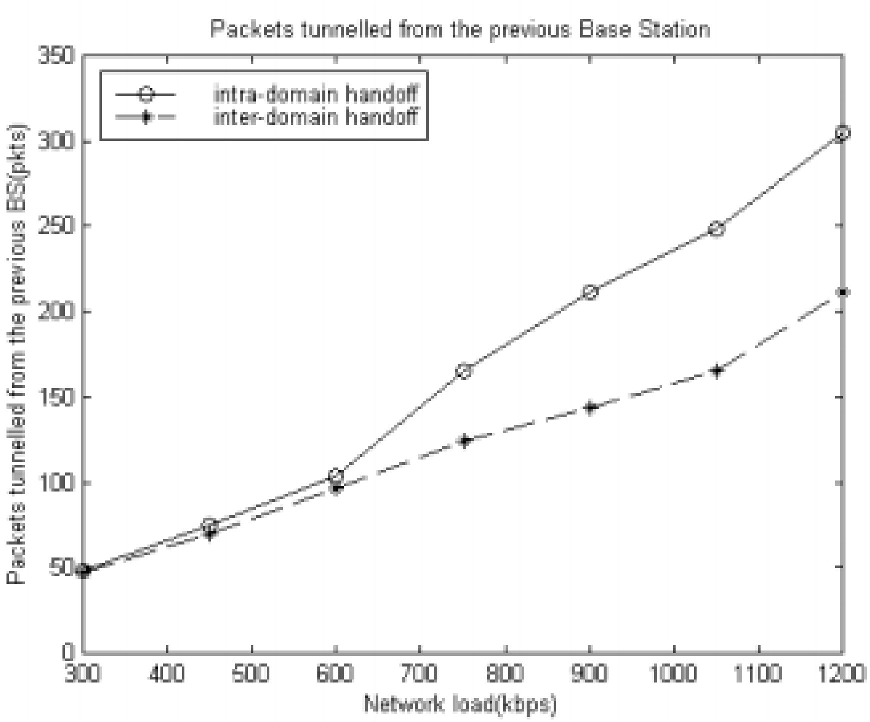


Figure 5.3 Packets tunneled from the previous base station that the mobile node just visited to the current base station

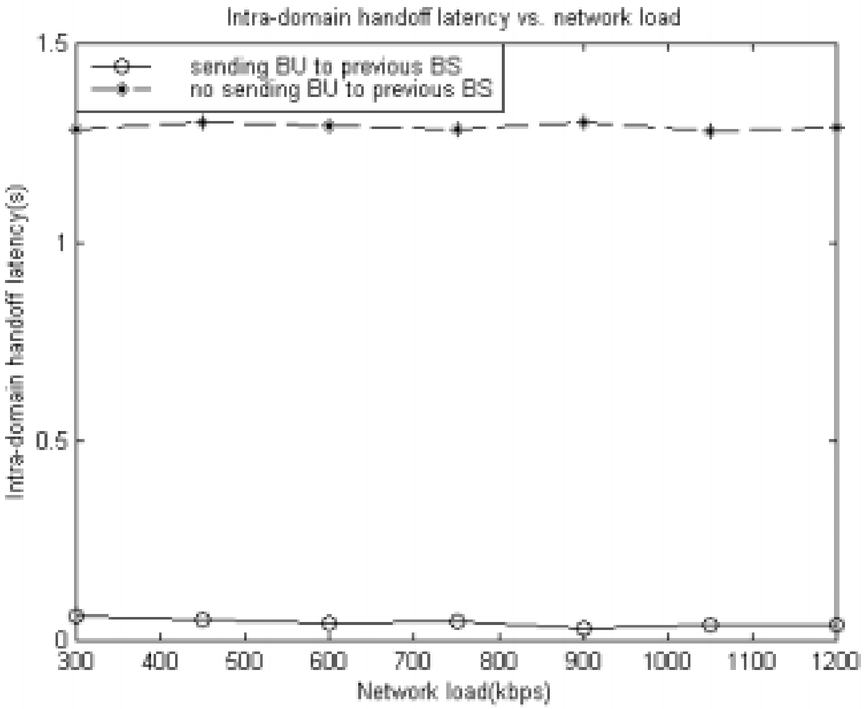


Figure 5.4 Intra-domain handoff latency varying with network load

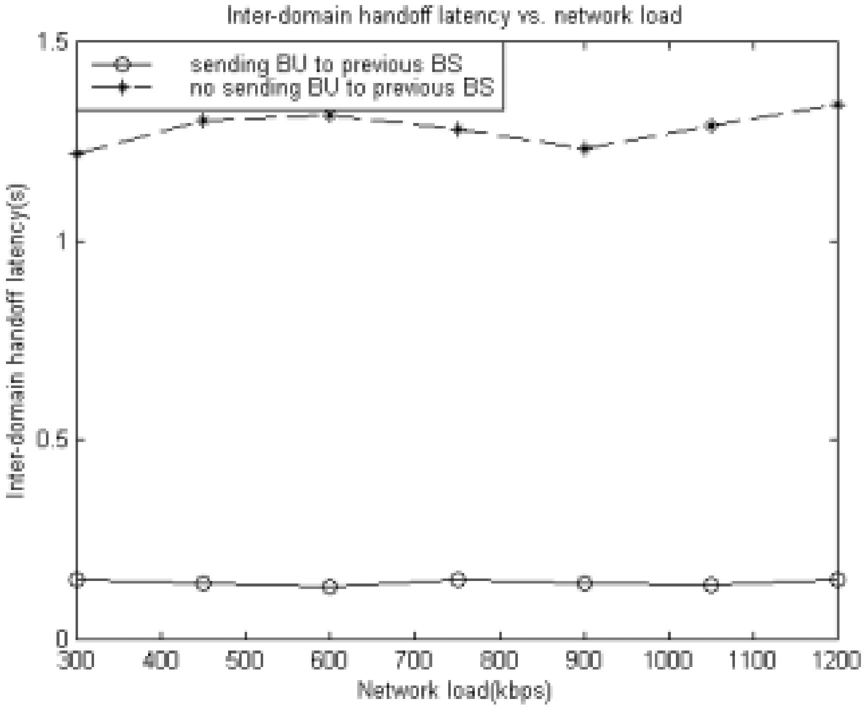


Figure 5.5 Inter-domain handoff latency varying with network load

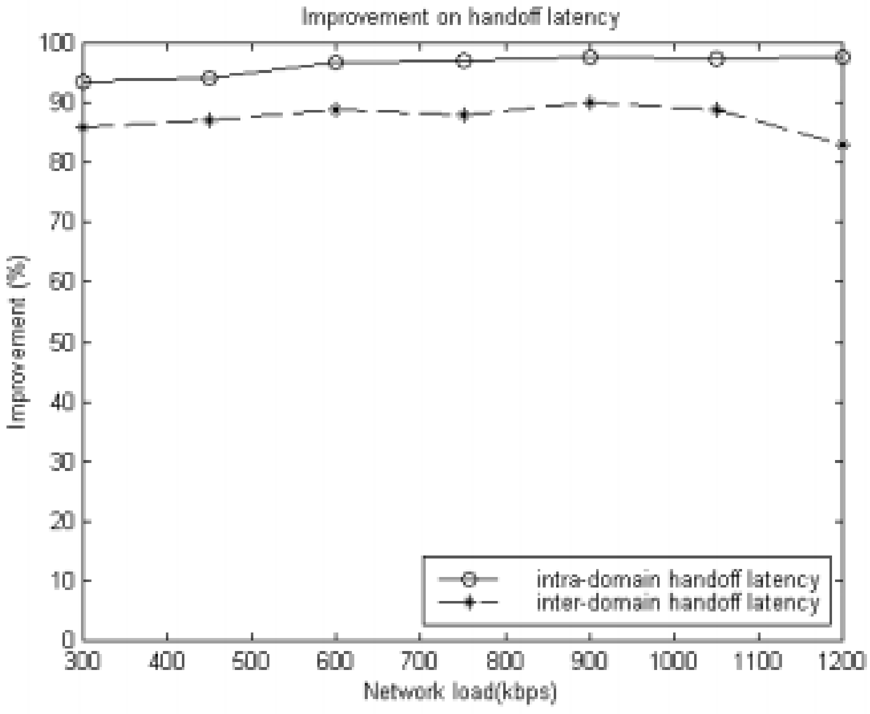


Figure 5.6 Improvement on intra-domain and inter-domain handoff latency with sending *BUs* to the previous base station

Figure 5.3 shows the number of packets tunneled from the previous base station that mobile node just visited to the current base station. If the *BUs* are not sent to the previous base station, these packets will be dropped.

Figure 5.4 and 5.5 show that the intra-domain and inter-domain handoff latency is reduced remarkably by sending *BUs* to the previous base station, comparing to the condition that no *BUs* is sent to the previous base station.

Figure 5.6 shows that the improvement on intra-domain handoff latency is more than 95% while the improvement on inter-domain handoff latency is less than 90%. The reason is that *BUs* will travel a very short way to the previous base station that mobile node just visited in case of intra-domain movement, while they will travel longer way in case of inter-domain movement.

From the above simulation results, it can be seen that the intra-domain and inter-domain handoff latency are notably reduced through sending *BUs* to the previous base station that mobile node just visited. Moreover, it was found that there is almost not packet dropped due to handoff when sending *BUs* to the previous base station, which tunnels the packets to the current base station. Therefore, we use handoff latency as criterion to evaluate the handoff performance in the following simulation.

5.3.2 Scenario 2: Improvement on Handoff Performance by Assigning *RADs* Higher Priority

In this scenario, we will compare handoff performance of the QoS architecture of no priority for *RADs* with that of higher priority given to *RADs* supported by IEEE802.11e under different simulation conditions. Simulation configuration is shown in Figure 5.2.

The wireless domain is IEEE802.11e wireless LAN. In this case, the mobile node is originally connected to its home agent BS1, and moves from BS1 to BS2 within the wireless domain1.

We will compare handoff performance of no priority for *RADs* with that of higher priority for *RADs* under the following conditions: different network load, different last wired link delay, different overlap between base stations, and different speed of the mobile node.

1. Handoff performance varying with network load

In this case, the speed of mobile node is 4m/s, the distance between base stations is 450m [43], and the last wired link delay is 1ms. Every source rate changes from 100kbps to 400kbps simultaneously, so the total source rate varies from 300kbps to 1200kbps.

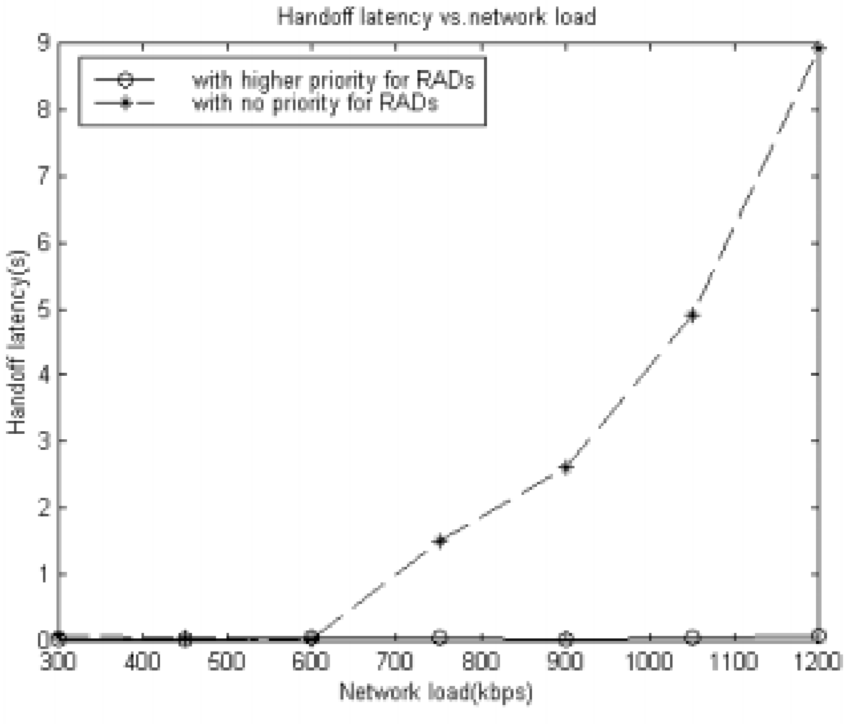


Figure 5.7 Handoff latency varying with network load

From Figure 5.7 it can be seen that the handoff latency increases remarkably from the medium network load (around 600kbps) if higher priority for *RADs* is not assigned. When assigning higher priority for *RADs*, the increase of handoff latency can almost be ignored (less than 100ms).

2. Handoff performance varying with last wired link delay

The last wired link delay represents the "distance" between base stations and between the CNs and MNs [43]. In this case, the speed of mobile node is 4m/s, and the distance between base stations is 450m. The three source rates are 400kbps respectively. The network is in heavy congestion.

The last wired link delay ranges from 1ms to 900ms.

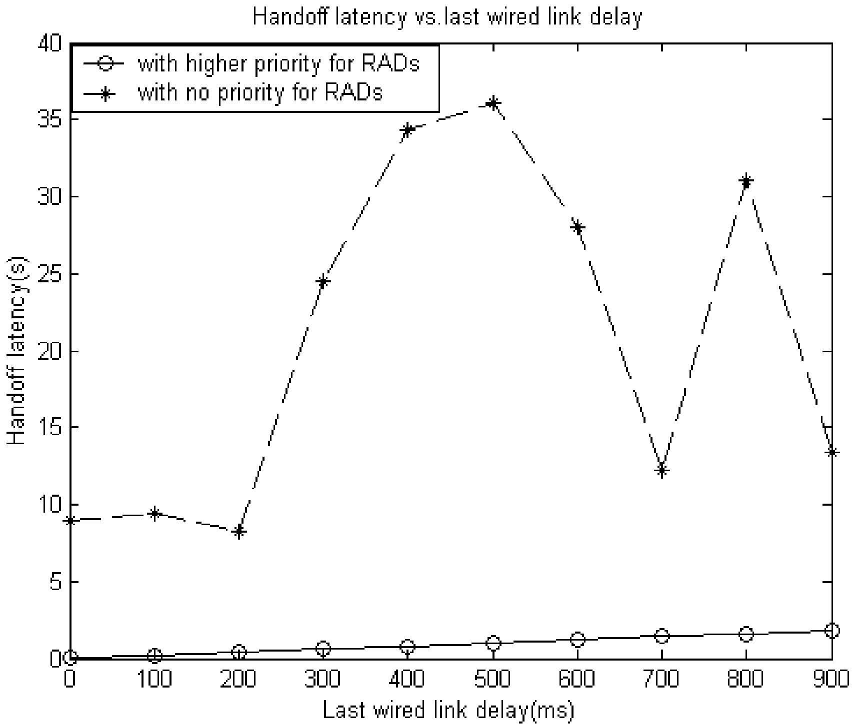


Figure 5.8 Handoff latency varying with last wired link delay

From Figure 5.8 we can see, with higher priority for *RADs*, the handoff latency increases with the last wired link delay steadily. The values of handoff latency with no priority for *RADs* are greatly larger than the values with higher priority for *RADs* in every point. Moreover, handoff latency with no priority for *RADs* is unpredictable. The handoff latency becomes very long when the *RADs* are dropped seriously at some point.

3. Handoff performance varying with the overlap between base stations BS1 and BS2

In this case, the speed of mobile node is 4m/s, the last wired link delay is 1ms, and the source rate is 400kbps for each source respectively. The network is in heavy congestion. The overlap of BS1 and BS2 ranges from 100m to –25m (25m away each other).

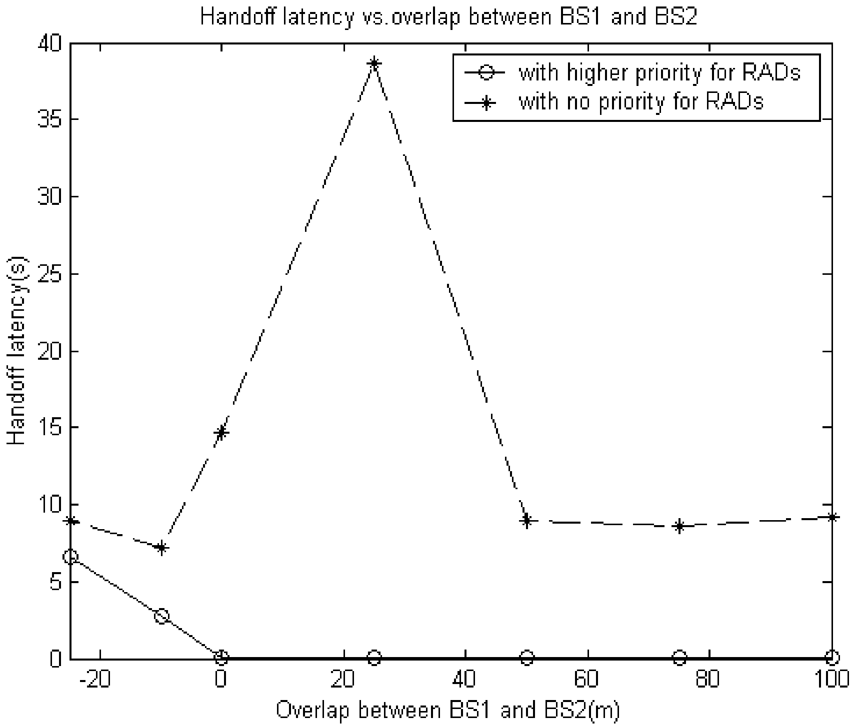


Figure 5.9 Handoff latency varying with overlap between base stations BS1 and BS2

Figure 5.9 shows that with higher priority for *RADs,* the handoff latency is quiet low (around 50ms) when two base stations overlap and the handoff latency almost does not change as the overlap increases. With no priority for *RADs*, the handoff latency is much longer. At some points, the handoff latency sharply increases. The reason for this phenomenon is that the *RADs* are dropped seriously in these points.

4. Handoff performance varying with the speed of the mobile node

In this case, the distance between base stations is 450m, the last wired link delay is 1ms, and the three source rates are 400kbps respectively. The network is in heavy congestion.

The speed of the mobile node ranges from 1m/s to 60m/s.

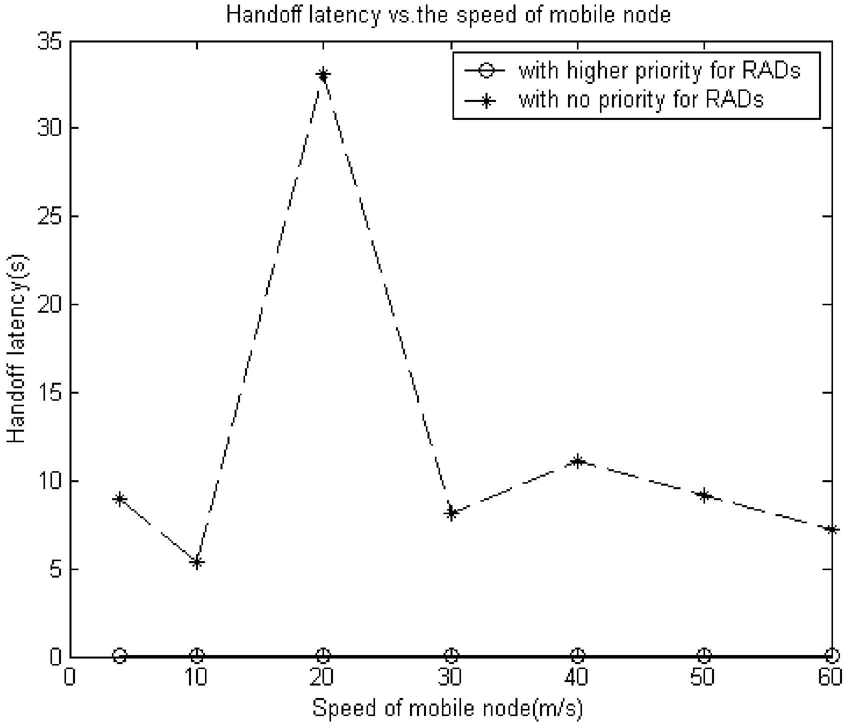


Figure 5.10 Handoff latency varying with the speed of the mobile node

From Figure 5.10 it can be seen that with higher priority for *RADs*, the handoff latency keep quiet low, and the handoff latency almost does not change with the change of speed. With no priority for *RADs*, the handoff latency becomes longer. Similarly, the handoff latency dramatically increases at some points.

Figure 5.7 to 5.10 show that the loss of *RADs* of MIPv6 affects the handoff performance seriously. The handoff latency with higher priority for *RADs* is much shorter than that with no priority for *RADs*. Moreover, it can be observed from Figure 5.8 to 5.10 that the trend of the curve with no priority for *RADs* is unpredictable, on which exists some jitter points. The reason is that *RADs* are dropped seriously in these points.

These figures show that handoff latency is remarkably improved and keeps quiet low when the loss of *RADs* is eliminated through assigning them higher priority.

5.4 Simulation Study – Part 2: Simulation study on the end-to-end QoS provided by the QoS architecture

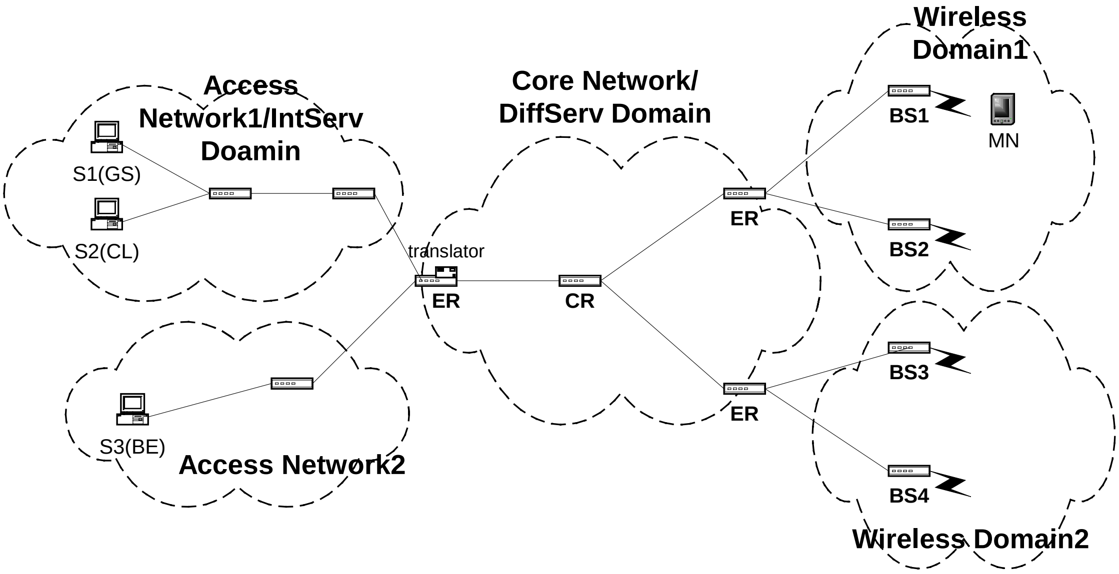


Figure 5.11 Simulation configuration 2

In this part we will evaluate and compare the end-to-end QoS provided by the architecture of IntServ/DiffServ/MIPv6/IEEE802.11 with that provided by the architecture of IntServ/DiffServ/MIPv6/IEEE802.11e. We will also investigate the end-to-end QoS obtained by the mobile node after intra-domain handoff and inter-domain handoff.

In this part, sources S1 and S2 are in Access Network1/IntServ domain, which will use RSVP to reserve resources for every flow. Source S3 is in Access Network2/best effort domain. S1 generates GS flow, S2 generates CL flow and S3 generates BE flow. This part also includes two scenarios.

5.4.1 Scenario 1: Compare the End-to-End QoS Provided by the Architecture of IntServ/DiffServ/MIPv6/IEEE802.11 with that Provided by the Architecture of IntServ/DiffServ/MIPv6/IEEE802.11e

In this scenario, one mobile node (MN) stays at its home network in wireless domain1. Firstly, We will set wireless networks as IEEE802.11 wireless LAN. Then we will change them to IEEE802.11e wireless LAN. We will compare the end-to-end QoS obtained by the GS, CL and BE flow in these two cases.

In this simulation, we increase three source rates from 100kbps to 400kbps simultaneously. The bandwidth is 1Mbps. The saturated bandwidth of wireless channel is about 0.8M with IEEE802.11, and it is about 0.9M with IEEE802.11e. The reason is that since EDCF in IEEE802.11e uses smaller Contention Window (CW) for high priority traffic, its average backoff time gets smaller and waiting time gets less, so the saturated bandwidth is larger than the one with IEEE802.11.

We will use goodput, average flow delay, packet drop rate, and delay jitter as QoS performance criteria.

From Figure 5.12, it can be seen that goodput of GS flow can be guaranteed with IEEE802.11e support while cannot be guaranteed with IEEE802.11 when the network is in congestion. The goodput of CL flow decreases slightly with IEEE802.11e while decreases largely with IEEE802.11 under congestion. The goodput of BE flow with IEEE802.11e deteriorates when the network is in congestion. Note that the goodput of BE flow with IEEE802.11e is still better than the one with IEEE802.11 when every source rate is 300Kbps (total is 900Kbps). The reason is that the saturated bandwidth of wireless channel is 0.9 M with IEEE802.11e, which is more than the one with IEEE802.11.

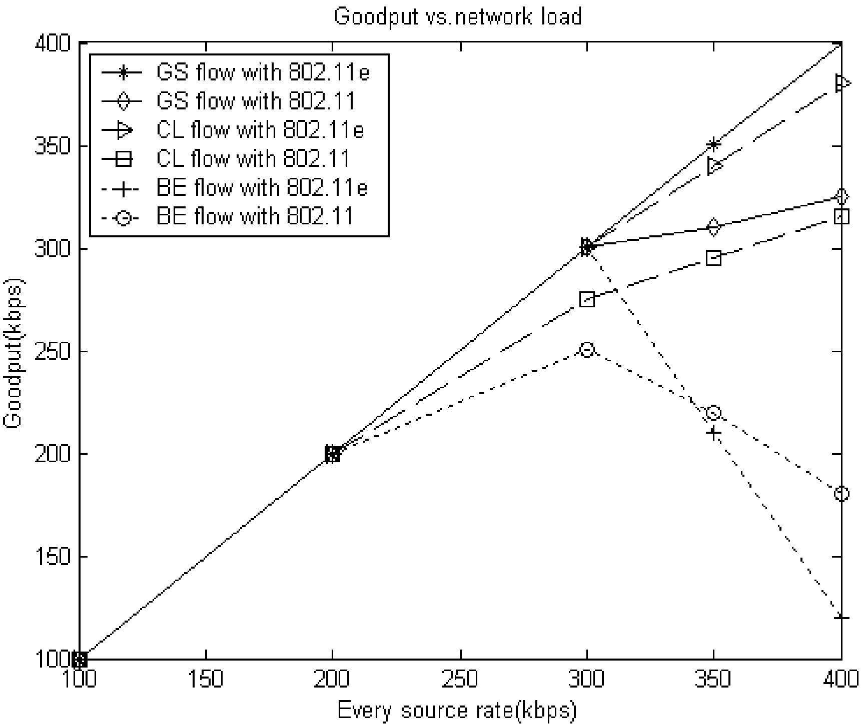


Figure 5.12 Goodput of GS, CL, BE flow vs. network load with IEEE 802.11 and with IEEE 802.11e

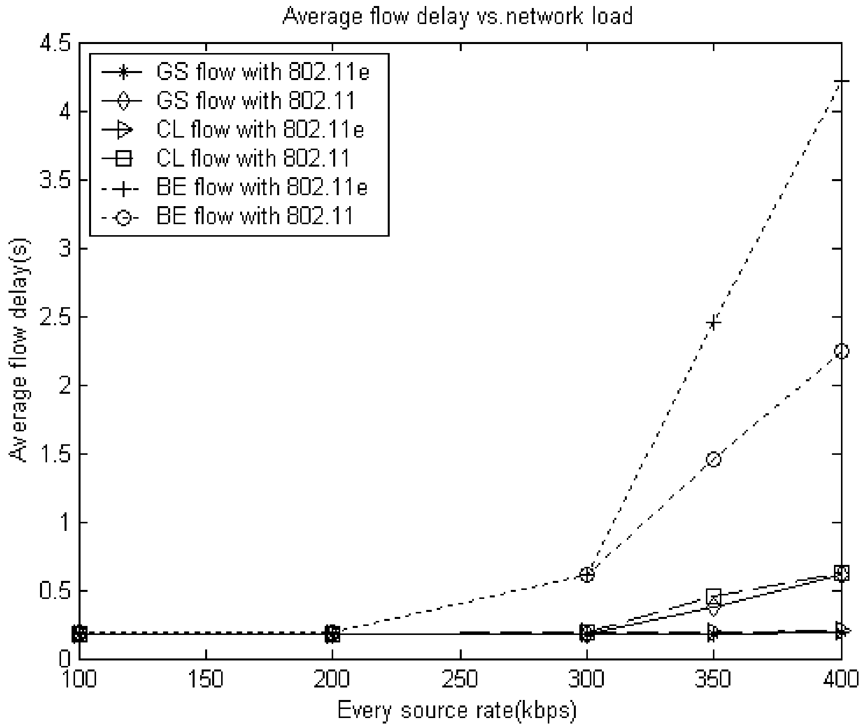


Figure 5.13 Average flow delay of GS, CL, BE flow vs. network load with IEEE802.11 and with IEEE802.11e

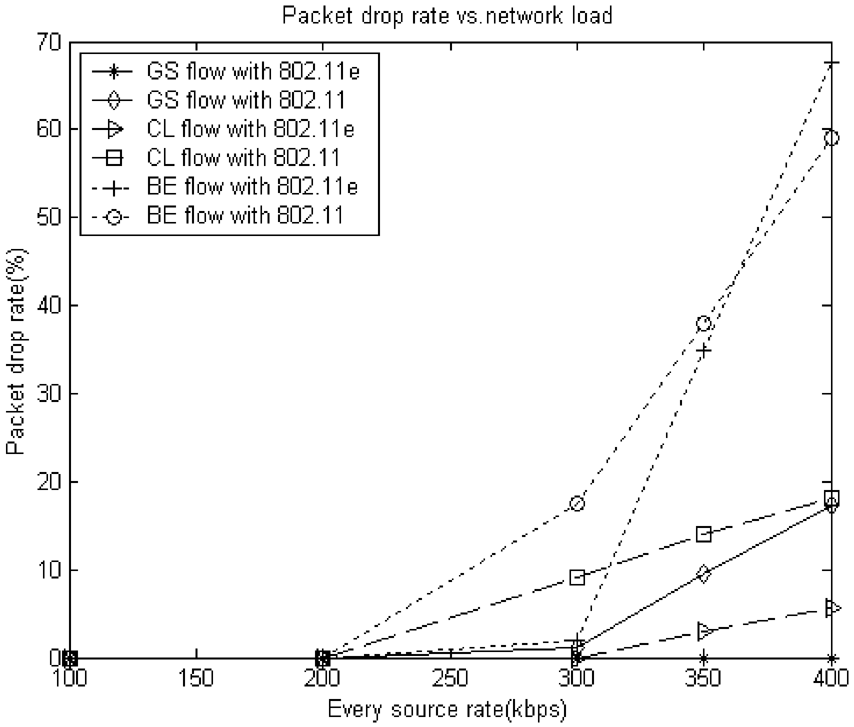


Figure 5.14 Packet drop rate of GS, CL, BE flow vs. network load with IEEE802.11 and with IEEE802.11e

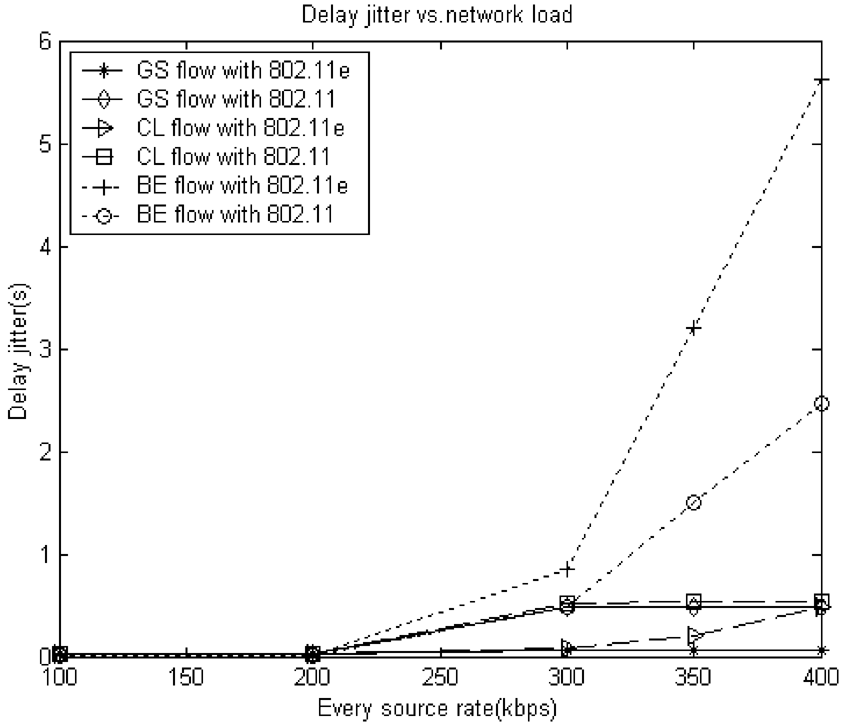


Figure 5.15 Delay jitter of GS, CL, BE flow vs. network load with IEEE802.11 and with IEEE802.11e

Figure 5.13 shows that the average flow delay of GS flow remains very small with IEEE802.11e while it deteriorates with IEEE802.11 when network is in congestion. The delay of CL flow also keeps very small with IEEE802.11e even when network is in congestion.

From Figure 5.14, we can see that the packet drop rate of GS flow is always zero with IEEE802.11e. The packet drop rate of CL flow keeps lower with IEEE802.11e than that with IEEE802.11 when network is in congestion.

Figure 5.15 shows that the delay jitter of GS flow remains very small (almost zero) with IEEE802.11e. The one of CL flow also remains lower with IEEE802.11e.

From the above figures, it can be seen that with the IEEE802.11e, the goodput of GS flow can be guaranteed and the packet drop rate of GS is zero. The delay and delay jitter of GS flow remain quiet low. Although we have not set a delay bound for GS flow, the delay of GS flow almost keeps unchanged when the network load varies from light load to heavy load. Therefore, we can claim that the strict requirement of GS flow on maximum delay can be achieved. The goodput, packet drop rate, delay and delay jitter of CL flow with IEEE802.11e change slightly when the network is in heavy congestion. The performance of CL flow under heavy network load is equivalent to the one under light network load. Therefore, it can be concluded that the end-to-end QoS obtained by GS and CL flow in the architecture that combines IntServ/DiffServ, MIPv6 and IEEE802.11e can be guaranteed.

5.4.2 Scenario 2: End-to-End QoS after Intra-Domain and Inter-Domain Handoff

In this scenario, the mobile node firstly moves to adjacent cell within the same subnet, and then moves to adjacent subnet. We will compare the end-to-end QoS obtained by GS, CL and BE flow after the intra-domain and inter-domain handoff in terms of goodput, packet drop rate, average flow delay and delay jitter.

The simulation results in this scenario are divided into two parts: the end-to-end QoS under light network load and under heavy network load. We set each source rate to 100k (total rate is 300k) when the network is under light load, and set every source rate to 400k (total rate is 1200k) when network is under heavy load (the network is in congestion).

1. End-to-end QoS after intra-domain and inter-domain handoff under light network load

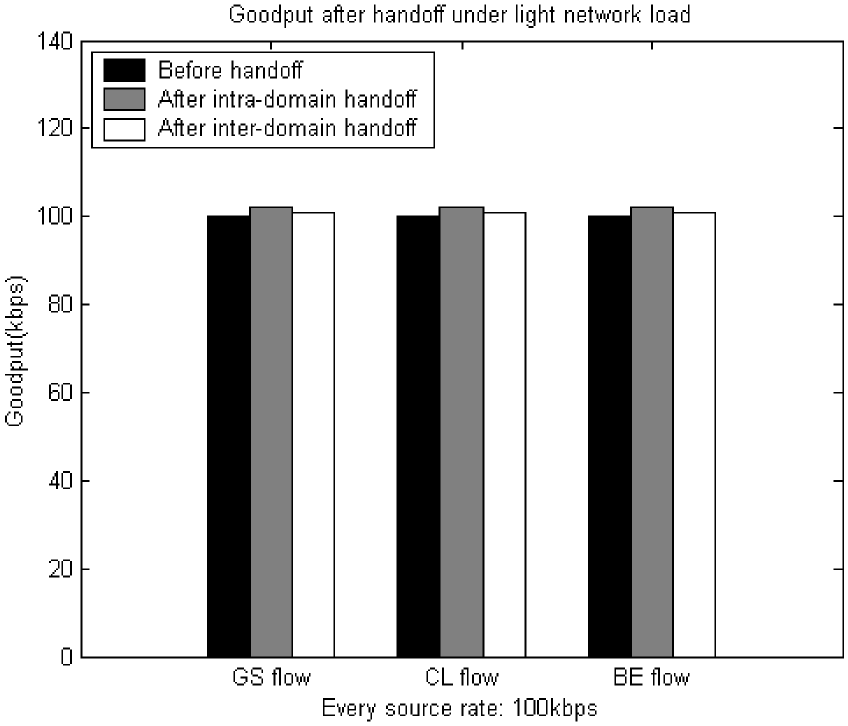


Figure 5.16 Goodput of GS, CL and BE flow after handoff under light network load

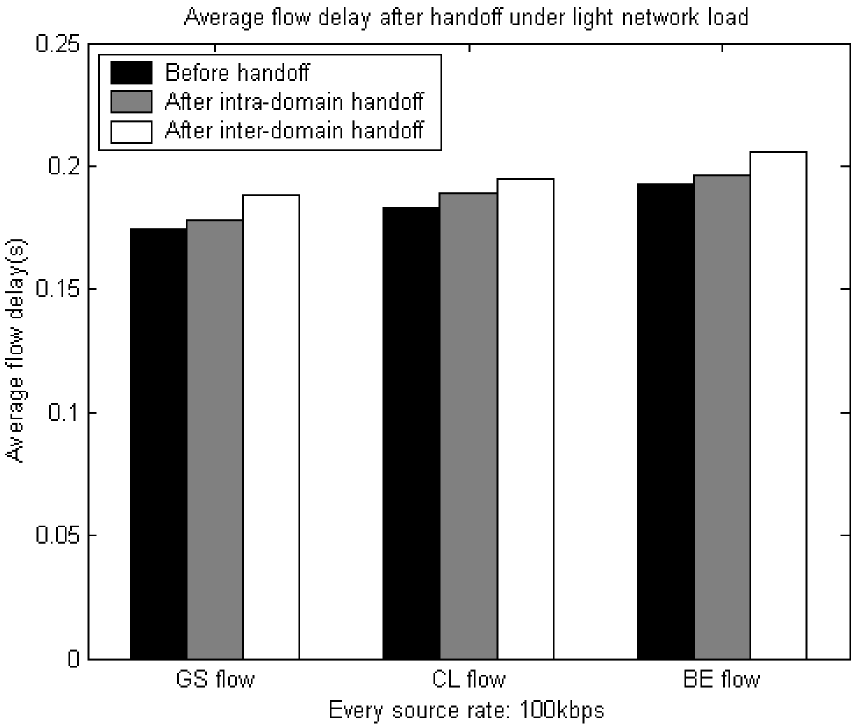


Figure 5.17 Average flow delay of GS, CL and BE flow after handoff under light network load

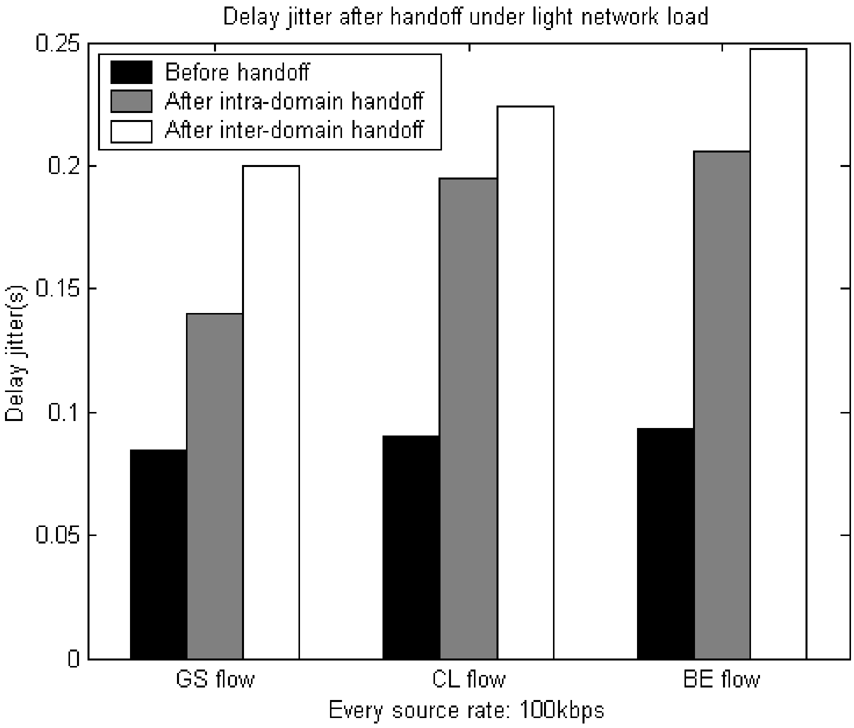


Figure 5.18 Delay jitter of GS, CL and BE flow after handoff under light network load

Packet drop rate of GS, CL and BE flow after handoff under light network load is zero respectively. Here, we will not figure them.

Figure 5.16 shows that the goodput of GS, CL and BE flow before handoff is equal to every source rate. After intra-domain handoff and inter-domain handoff, the goodput of these three flows increases slightly. The reason is that some packets are tunneled from the previous base station to the current base station after handoff.

Figure 5.17 and 5.18 show that the average flow delay increases slightly and delay jitter increases largely after intra-domain handoff and inter-domain handoff. It is because that the tunneled packets from the previous base station travel longer to MN.

2. End-to-end QoS after intra-domain and inter-domain handoff under heavy network load

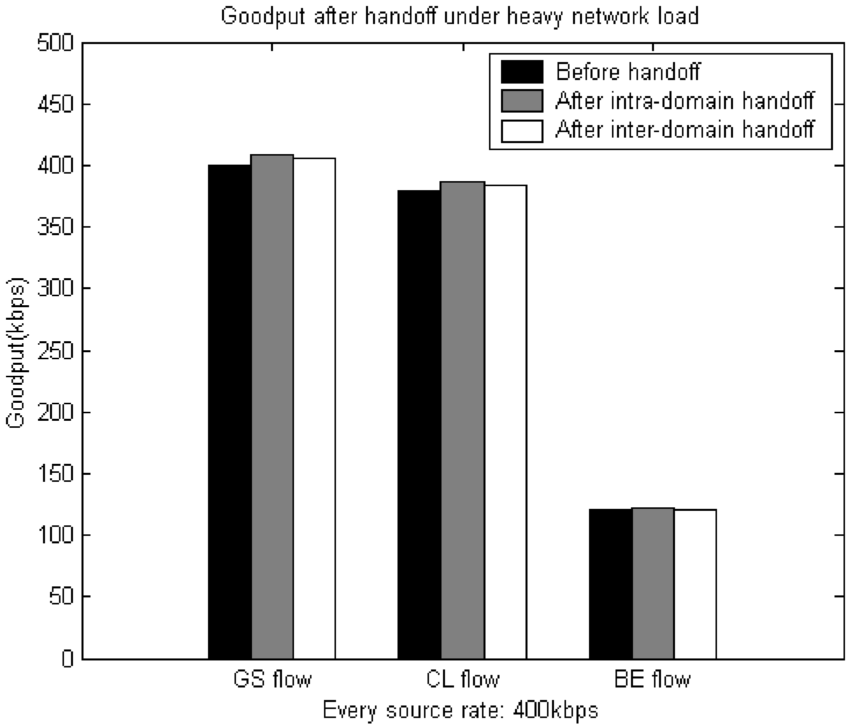


Figure 5.19 Goodput of GS, CL and BE flow after handoff under heavy network Load

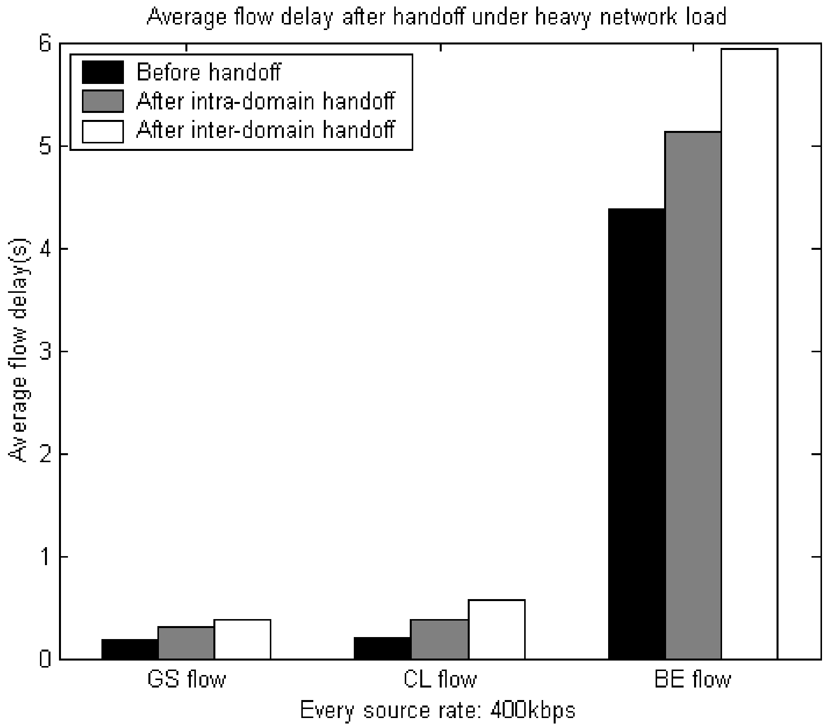


Figure 5.20 Average flow delay of GS, CL and BE flow after handoff under heavy network load

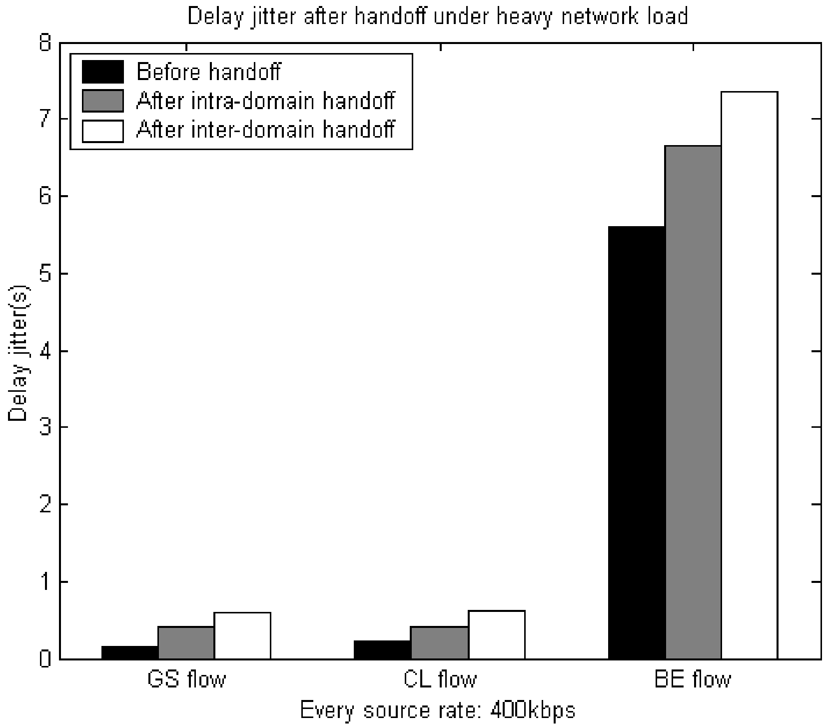


Figure 5.21 Delay jitter of GS, CL and BE flow after handoff under heavy network load

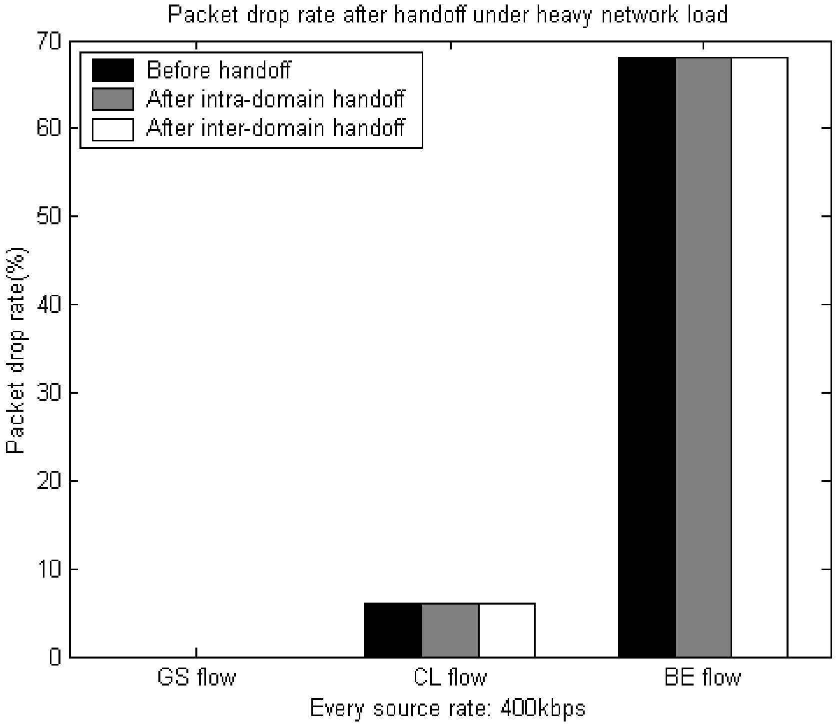


Figure 5.22 Packet drop rate of GS, CL and BE flow after handoff under heavy network load

Figure 5.19 ~ Figure 5.22 show the end-to-end QoS obtained by the mobile node under heavy network load. Before handoff, the goodput of GS flow is equal to source rate, the goodput of CL flow is slightly less than the source rate, and the goodput of BE flow decreases greatly. After intra-domain handoff and inter-domain handoff, the fact that some packets tunneled from the previous base station to the current base station results in the slight increase of goodput of these three flows. The average flow delay and delay jitter increase more under heavy network load than that under light network load after intra-domain handoff and inter-domain handoff. The packet drop rate remains unchanged after handoff.

5.5 Simulation Study – Part 3: Simulation Study on a More Practical Situation

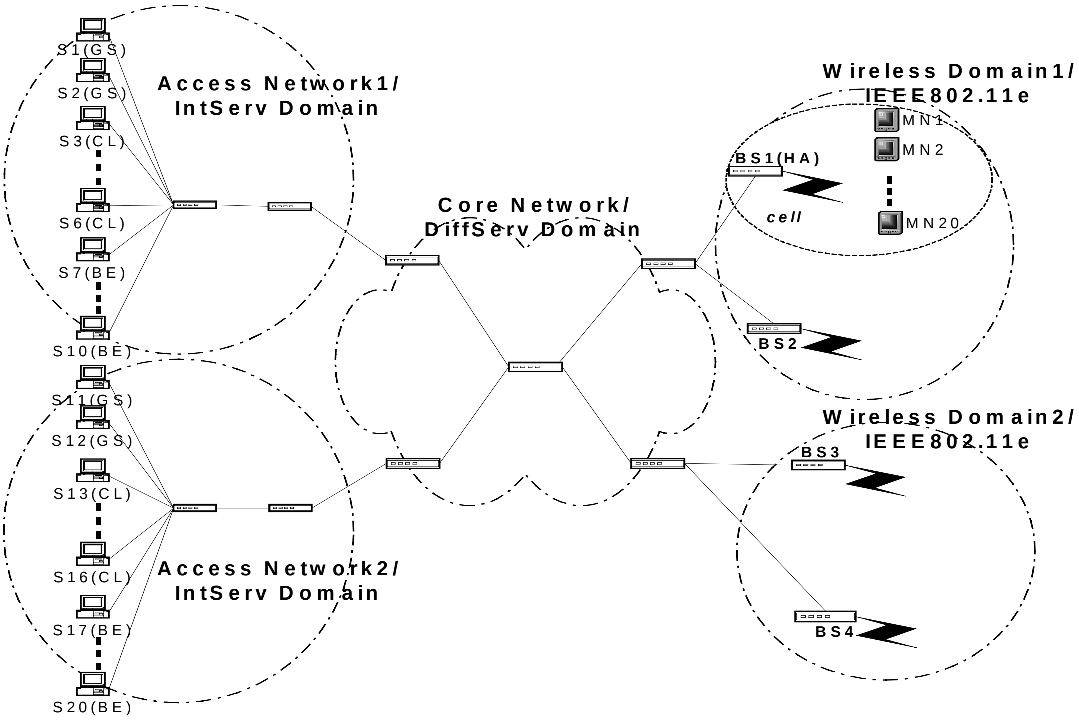


Figure 5.23 Simulation configuration 3 – more practical situation

In the previous simulation, we only set one mobile node and three CBR sources (one for GS, one for CL and one for BE flow), which is an ideal condition. Now we will simulate a more practical situation, i.e., more source nodes, more mobile hosts and different traffic types. We will study the average end-to-end QoS obtained by multiple mobile nodes.

Figure 5.23 shows a more practical situation. In this configuration, there are still five domains: two Access networks/IntServ networks, one core network/DiffServ network, and two wireless domains. Wireless domains are IEEE802.11e wireless LAN.

There are 10 sources in every access network respectively. 20% of sources generate GS flow, 40% generate CL flow, and 40% generate BE flow. Namely, in access network 1, sources S1, S2 generate GS flow, sources S3 ~ S6 generate CL flow, and sources S7 ~ S10 generate BE flow. In access network 2, sources S11, S12 generate GS flow, sources S13 ~ S16 generate CL flow and sources S17 ~ S20 generate BE flow. That is, altogether there are 4 sources that generate GS flow, 8 sources that generate CL flow and 8 sources that generate BE flow.

In wireless networks, every wireless subnet has two base stations. There are 20 mobile nodes, MN1, MN2, … , MN20. Source Si and mobile node MNi is a source-destination pair, i.e., S1 and MN1, S2 and MN2, … S20 and MN20 are source-destination pairs. The initial locations of all mobile nodes are in the cell of the base station BS1, i.e., BS1 is the home agent of these 20 mobile nodes.

We use CBR traffic as GS flow, Exponential (on/off) traffic as CL flow, and TCP traffic as BE flow [50].

Table 5-3 Different traffic type

GS flow (CBR traffic)	Every source rate = 64kbps
CL flow (EXP traffic)	T_{ON} = 500ms, T_{OFF} = 300ms; Every source rate = 64kbps
BE flow (TCP traffic)	Window size = 75

5.5.1 Scenario 1: Average End-to End QoS Obtained by the Mobile Nodes before Movement

In this scenario, mobile nodes stay in their home network and do not move. We will compare the average goodput, delay, and delay jitter obtained by GS, CL, and BE flows when the number of mobile nodes changes. As the number of mobile nodes increase from 5 to 20, the number of working sources also increases from 5 to 20 correspondently. Among working sources, the probability of different sources which generate GS, CL, BE flow is 20%, 40%, 40% respectively. The following table shows the detail that working sources vary with the number of mobile nodes.

Table 5-4 Source rate with different number of the mobile nodes

The number of MNs	The working sources	Source rate
5	S1, S3, S4, S7 and S8 work: S1 generates GS flow, S3 and S4 generate CL flow, S7 and S8 generate BE flow.	of GS flow: 64kbps of CL flows: 128kbps
10	S1, S2, S3, S4, S5, S6, S7, S8, S9 and S10 work: S1and S2 generate GS flow, S3, S4, S5 and S6 generate CL flow, S7, S8, S9 and S10 generate BE flow.	of GS flows: 128kbps of CL flows: 256kbps
15	S1 to S10, and S11, S13, S14, S17, S18 work: S1, S2, S11 generate GS flow, S3, S4, S5, S6, S13 and S14 generate CL flow, S7, S8, S9, S10, S17 and S18 generate BE flow.	of GS flows: 192kbps of CL flows: 384kbps
20	All sources S1 to S20 work: S1, S2, S11 and S12 generate GS flow, S3, S4, S5, S6, S13, S14, S15 and S16 generate CL flow, S7, S8, S9, S10, S17, S18, S19 and S20 generate BE flow.	of GS flows: 256kbps of CL flows: 512kbps

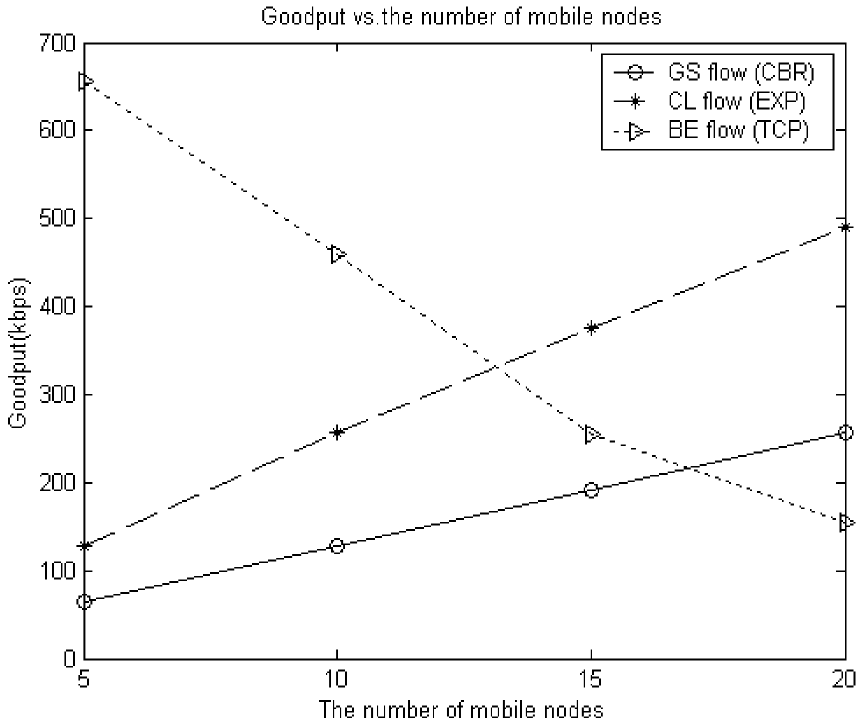


Figure 5.24 Average goodput of GS, CL and BE flows varying with the number of mobile nodes

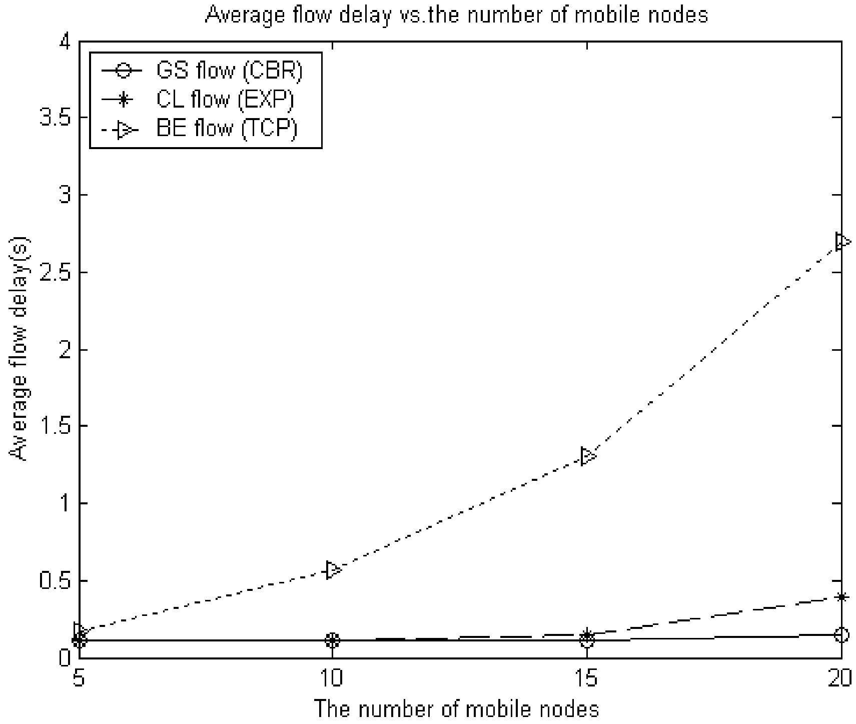


Figure 5.25 Average flow delay of GS, CL and BE flows varying with the number of mobile nodes

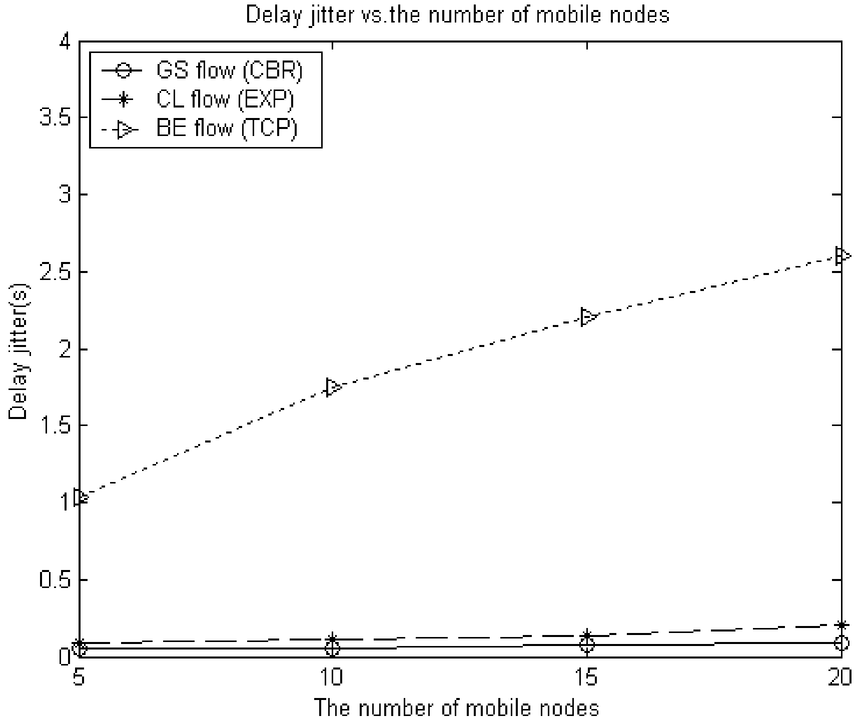


Figure 5.26 Average delay jitter of GS, CL and BE flows varying with the number of mobile nodes

Figure 5.24, 5.25, and 5.26 show the average goodput, delay and delay jitter obtained by different flows when the number of mobile nodes varies from 5 to 20. These figures show that the number of mobile nodes has little effect on the end-to-end QoS obtain by GS flows and CL flows when the number of mobile nodes is modest. When the number of mobile node increases to 20, the goodput of CL flows has a little loss and delay also increase slightly. The reason is that the total source rate is large at this time and the network has already been under heavy load. The goodput of BE flows (TCP traffic) decreases with the increase of the source rates of GS and CL flows. This is because that TCP traffic is adaptive traffic, which can adjust traffic according to the state of network. The delay and delay jitter of BE flows also increase notably as the network load increases.

5.5.2 Scenario 2: Average End-to-End QoS Obtained by the Mobile Nodes after Intra-Domain and Inter-Domain Handoff

In this scenario, we will study the end-to-end QoS obtained by GS, CL and BE flows after intra-domain and after inter-domain handoff. Firstly, mobile nodes do intra-domain handoff, move from the base station BS1 to BS2. Then mobile nodes do inter-domain handoff, move from the base station BS2 to BS3. Simulation is divided into two parts: under light network load and under heavy network load.

1. Average end-to-end QoS obtained by mobile nodes after intra-domain and inter-domain handoff under light network load

We set 10 mobile nodes in this case. Accordingly, we let 10 sources in access network1 work. Therefore, the total source rate of GS (CBR) flows is 128kbps, and that of CL (EXP) flows is 256kbps.

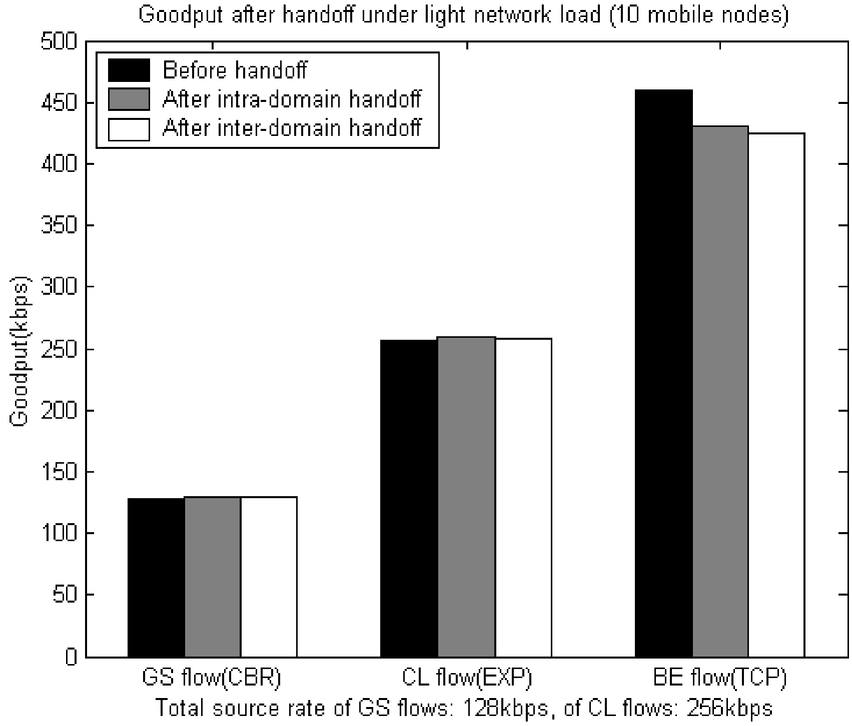


Figure 5.27 Average goodput of GS, CL and BE flow after handoff under light network load

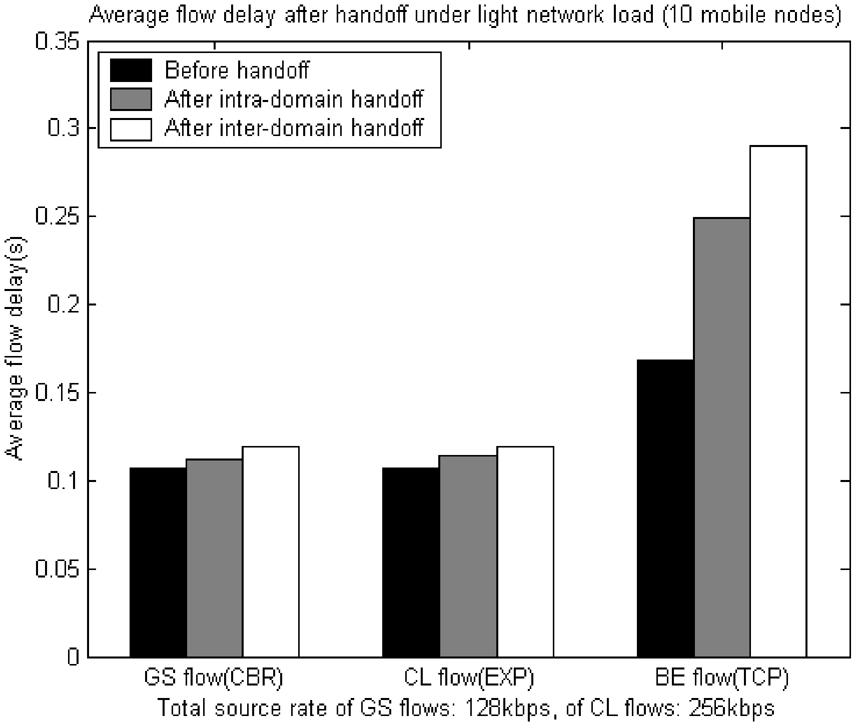


Figure 5.28 Average flow delay of GS, CL and BE flows after handoff under light network load

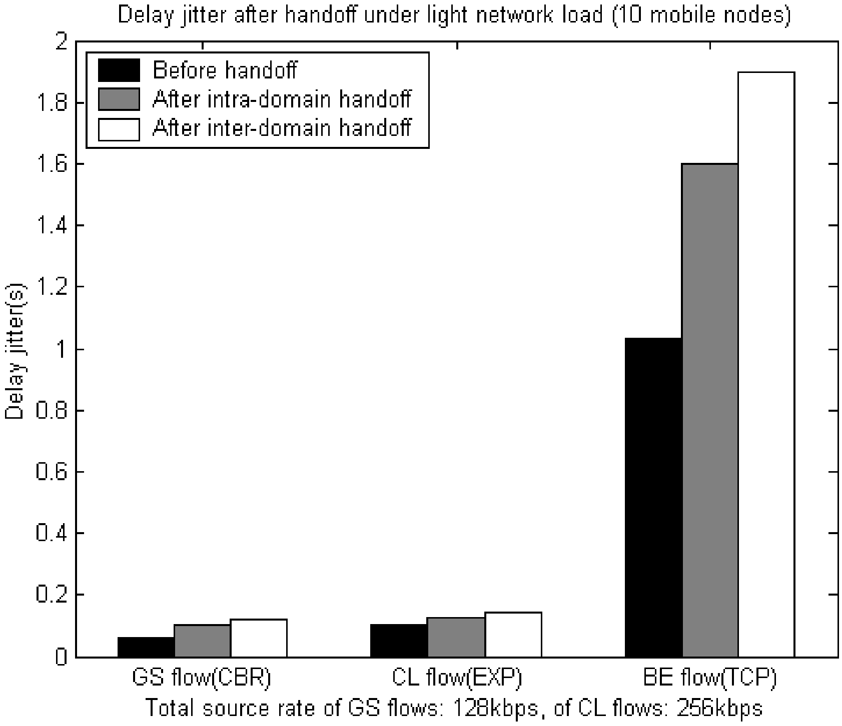


Figure 5.29 Average delay jitter of GS, CL and BE flows after handoff under light network load

Figure 5.27, 5.28 and 5.29 show that the average goodput of GS(CBR), CL(EXP) flows increases slightly after intra-domain handoff and inter-domain handoff. The reason is that some packets are tunneled from the previous base station to the current base station after handoff. The average flow delay increases slightly and average delay jitter also increases largely after intra-domain handoff and inter-domain handoff. The reason is that tunneled packets experience longer delay to mobile nodes. The decrease of average goodput of BE (TCP) flows after handoff is more than the decrease of that of GS and CL flows. The average flow delay and delay jitter of BE flows increase greatly after handoff. This is because the TCP traffic is adaptive traffic. It will adjust the window size and retransmit the packets when some packets are lost during handoff.

2. Average end-to-end QoS obtained by mobile nodes after intra-domain and inter-domain handoff under heavy network load

In this case we set 20 mobile nodes, and all 20 sources work. Therefore, the total source rate of GS (CBR) flows is 256kbps, and that of CL (EXP) flows is 512kbps.

Figure 5.30, 5.31, and 5.32 show that the average goodput of GS and CL flows increases slightly after intra-domain handoff and inter-domain handoff, which is similar to the situation under light network load. The average flow delay and average delay jitter under heavy network load increase more than that under light network load after intra-domain handoff and inter-domain handoff. The performance of BE flows deteriorates under heavy network load.

The simulation results in this scenario show that average end-to-end QoS obtained by GS and CL flows can also be guaranteed under the condition with more sources, more mobile nodes and different traffic types.

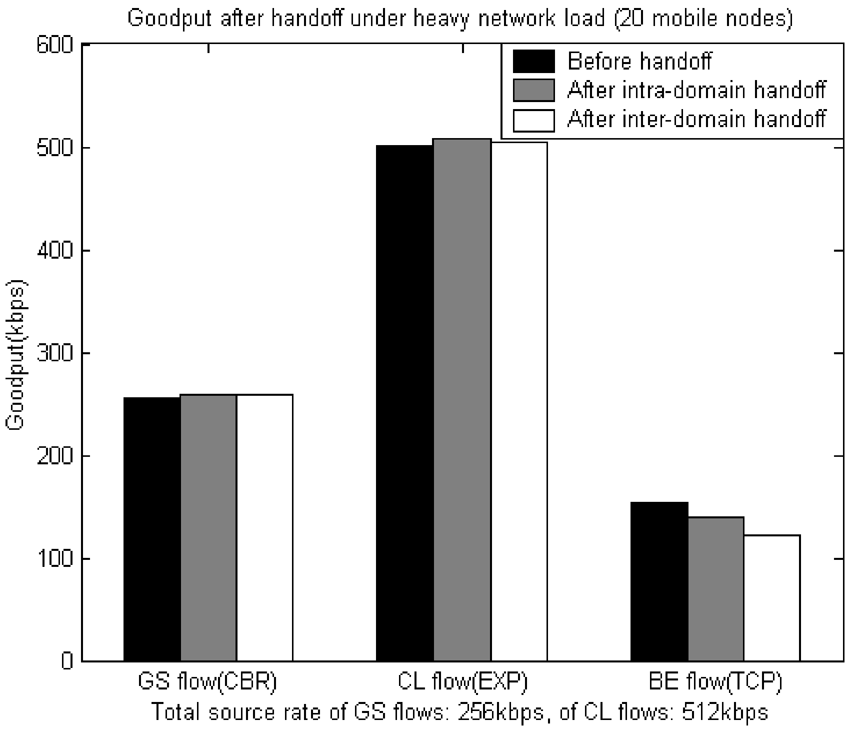


Figure 5.30 Average goodput of GS, CL and BE flows after handoff under heavy network load

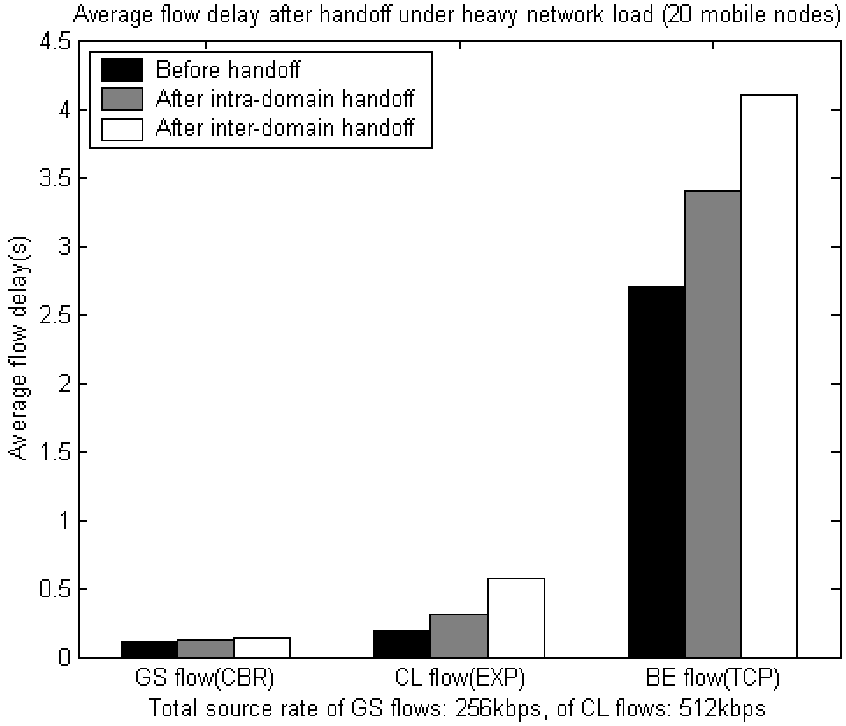


Figure 5.31 Average flow delay of GS, CL and BE flows after handoff under heavy network load

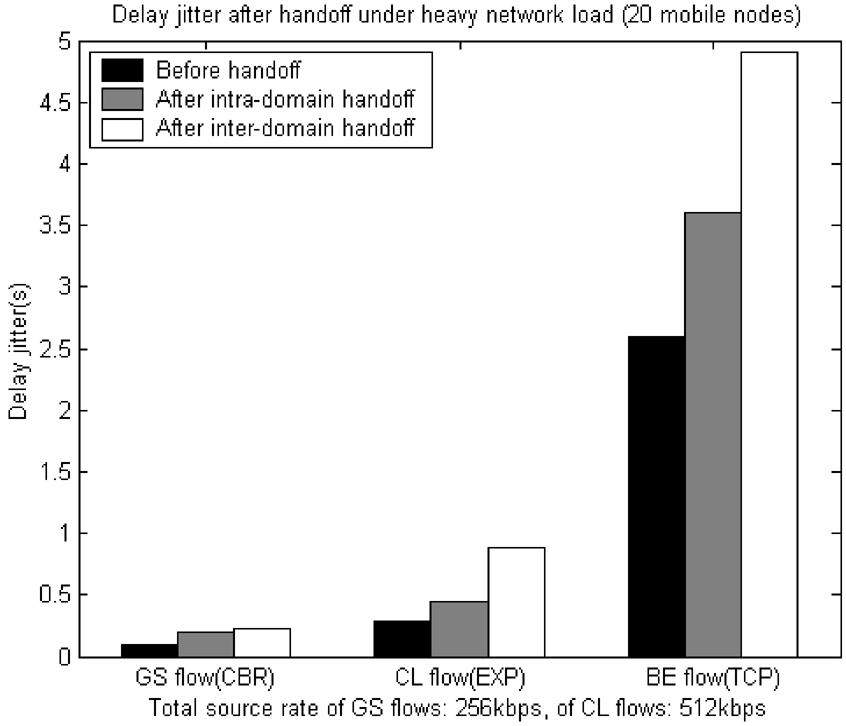


Figure 5.32 Average delay jitter of GS, CL and BE flows after handoff under heavy network load

5.6 Summary

In this chapter, we conducted simulations to evaluate the end-to-end QoS achievable by the architecture that combines IntServ/DiffServ, MIPv6, and IEEE802.11e wireless LAN standard. Simulation results show that the end-to-end QoS in terms of goodput, delay, and packet drop rate, and delay jitter can be guaranteed in the cases that mobile nodes are in their home network, after intra-domain handoff, and after inter-domain handoff.

We also designed simulations to examine the efficiency of the two approaches on improving handoff QoS performance — assigning *RADs* higher priority and sending *BUs* to the previous base station. Simulation results have shown that these two approaches can improve the handoff QoS performance effectively. By sending *BUs* to the previous base station, the packet drop due to handoff is almost eliminated, and the handoff latency is shortened by 85% ~98%. By assigning higher priority for *RADs*, the handoff latency decreases dramatically.

Chapter 6

Conclusion

6.1 Summary of Existing Work

The increasing popularity of real-time multimedia Internet applications and the rapid growth of mobile systems indicate that the future Internet architecture will have to support various applications with different QoS requirements, regardless of whether they are running on fixed or mobile terminals. The main contributions of this thesis work involve proposing an end-to-end QoS architecture for mobile hosts that combines IntServ/DiffServ, Mobile IPv6 and IEEE802.11e and proposing an approach to improve the handoff performance of this QoS architecture, which is assigning *RAD* messages of MIPv6 higher transmission priority.

Handoff performance is a significant factor in evaluating a QoS architecture for mobile hosts. Therefore, besides the approach of assigning *RADs* higher priority, we also made use of the approach of sending *BUs* to the previous base station that mobile nodes just visited, which come from the idea that signaling and tunnel between previous access router (PAR) and new access router (NAR) in Fast Handovers for MIPv6. We combined these two approaches to improve the handoff performance of our QoS architecture.

We conducted simulations to evaluate and compare the proposed QoS architecture that combines IntServ/DiffServ, Mobile IPv6 and IEEE802.11e with the existing QoS architecture for mobile hosts that combines IntServ/DiffServ, Mobile IPv6 and

IEEE802.11. We also examined how much the handoff performance was improved through these two approaches.

Simulation results conducted on five different scenarios show: 1) end-to-end QoS in terms of goodput, average flow delay, delay jitter and packet drop ratio cannot be guaranteed just from the use of IntServ over a DiffServ backbone with MIPv6 and IEEE802.11 at the wireless last hop, because IEEE802.11 does not provide QoS support. The QoS achieved at wired part is void at the wireless last hop. When IEEE802.11e is deployed at the last hop, the packet drop ratio is almost zero and end-to-end delay is very low for GS flow, the packet drop ratio and delay are also small for CL flow, but the performance of BE flow become worse as more resources are channeled to satisfy the QoS of the other two classes. Therefore, the end-to-end QoS for GS and CL flows are guaranteed. 2) end-to-end QoS could also be guaranteed after intra-domain and inter-domain handoff for the GS and CL flows. 3) Handoff latency is reduced by 85% ~ 98% and packet drop due to handoff is almost zero when sending *BUs* to the previous base station that mobile node just visited. 4) Handoff performance is also improved dramatically after *RADs* are given higher priority to transmit under heavy network load conditions.

6.2 Future Work

For our end-to-end QoS architecture for mobile hosts, we combine IntServ/DiffServ and IEEE802.11e with basic Mobile IPv6. Although the combination of two approaches that assigns *RADs* higher priority and sends *BUs* to the previous base station can improve the handoff performance efficiently, some work could be done to better improve handoff performance by incorporating other handoff optimization for MIPv6, such as FMIPv6 (Fast Handovers for Mobile IPv6)[23] and

HMIPv6 (Hierarchical Mobile IPv6)[51]. Therefore, we could integrate IntServ/DiffServ and IEEE802.11e with FMIPv6 and/or HMIPv6 to further improve the handoff performance of the end-to-end QoS architecture.

Publications

1. Wei Wu and Winston K.G. Seah, “Evaluation of End-to-end QoS Support for Mobile Hosts in IPv6 with IEEE802.11e”, IEEE 58th Vehicular Technology Conference (VTC Fall 2003), Orlando, USA, Oct. 2003
2. Wei Wu and Winston K.G. Seah, “End-to-end Internet Quality of Service with IntServ/DiffServ, Mobile IPv6 and IEEE802.11e”, submitted for publication.

Reference

[1]. XiPeng Xiao and Lionel Ni, "Internet QoS: A Big Picture", IEEE Network Magazine, March/April, pp. 8-18, 1999.

[2]. “White paper: QoS protocols and architecture”, white paper from Stardust.com and QoS Forum.com, July, 1999.

[3]. Braden,R., Clark, D.,Shenker, S., “Integrated Services in the Internet Architecture: An overview”, IETF RFC 1633, 1994

[4]. Blake, S., Black, D., Carison, M., Davies, E., Wang, Zh., Weiss, W., “An Architecture for Differentiated Services”, IETF RFC 2475, 1998.

[5]. Braden R., et.al, “Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification” RFC 2205, Sep. 1997.

[6]. Zhang L., et.al., “RSVP: a New Resource ReServation Protocol”, IEEE Network, vol.7, no.5, pp8-18, 1993.

[7]. W.Almesberger, S. Giordano, R.Mameli, S.Salsano, F.Salvatore, "A Prototype Implementation for the IntServ Operation over DiffServ Networks", IEEE Globecom 2000, S. Francisco, USA, November 2000.

[8]. Bai H., et.al “Running Integrated Services over Differentiated Service Networks: Quantitative Performance Measurements” Quality of Service over Next-Generation Internet, Boston, MA, July 29 - August 1, 2002.

[9]. Johnson D., Perkins C., “Mobility Support in IPv6”, Internet Draft, work in progress, Oct 2002

[10]. Talukdar A., et.al “MRSVP: A Reservation Protocol for an Integrated Services Packet Network with Mobile Hosts”, Technical Report DCS-TR-337, Department of Computer Science, Rutgers University, USA.

[11]. Chen W. T. and Huang L. C., “RSVP Mobility Support: A Signaling Protocol for Integrated Services Internet with Mobile Hosts,” INFOCOM 2000.

[12]. Terzis, Srivastava M., and Zhang L., “A Simple QoS Signaling Protocol for Integrated Services Internet with Mobile Hosts”, INFOCOM’99, vol.3, 1999, pp.1011-18.

[13]. Jain R., et.al., “Mobile IP with Location Registers (MIP-LR)”, Internet Draft, draft-jain-miplr-01.txt, July 2001.

[14]. Chiruvolu G., et.al., “Mobility and QoS Support for IPv6 Based Real-time Wireless Internet Traffic”, 1999 IEEE Int’l. Conf. Commun., vol.1, 1999, pp.334-38.

[15]. Shen Q. et.al., “On Providing Flow Transparent Mobility Support for IPv6-based Real-Time Services”, Proc. MoMuC 2000, Tokyo, Japan, Oct. 2000.

[16]. Bongkyo Moon and Hamid Aghvami, “RSVP Extensions for Real-Time Services in Wireless Mobile Networks” IEEE Communications Magazine, Dec. 2001.

[17]. Mahadevan I. and Sivalingam K., “Quality of Service Architectures for Wireless Networks: IntServ and DiffServ Models” IEEE ISPAN,pp420-425, 1999, http://www.computer.org/proceedings/ispan/0231/02310420abs.htm

[18]. Zhang G. and Mouftah H., “End-to-End QoS Guarantees Over DiffServ Networks”, Sixth IEEE Symposium on Computers and Communications (ISCC'01), Hammamet, Tunisia, July 03 - 05, 2001

[19]. Mahadevan I. and Sivalingam K., “Quality of Service in Wireless Networks Based on Differentiated Services Architecture” 8th International Conference on Computer Communications and Networks: Proceedings 11-13, pp548-553, Boston, Massachusetts, USA, Oct., 1999.

[20]. IEEE, “Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications”, IEEE Std. 802.11, 1999.

[21]. IEEE 802.11 TGe, “EDCF Proposed Draft Text”, TR-01/131r1, March 2001.

[22]. Wei Wu and Winston K.G. Seah, “Evaluation of End-to-end QoS Support for Mobile Hosts in IPv6 with IEEE802.11e”, IEEE 58th Vehicular Technology Conference (VTC Fall 2003), Orlando, USA, Oct. 2003.

[23]. Rajeev Koodli, “Fast Handovers for Mobile IPv6”, Internet Draft, draft-ietf-mobileip-fast-mipv6-05.txt, 30, September 2002.

[24]. Wang Zheng, “Internet QoS: Architectures and Mechanisms for Quality of Service”, Morgan Keufmann, 1st edition, California, USA, March 2001

[25]. Nichols, K., Jacobson, V., Zhang, L., “A two-bit Differentiated Services Architecture for the Internet”, IETF RFC 2638, 1999.

[26]. Shen Q., “On Providing Flow Transparent Mobility Support for IPv6-Based Wireless Real-Time Services”, M.Eng. thesis, National University of Singapore, 2001

[27]. Deering S. and Hinden R., “Internet Protocol, Version 6 (IPv6) Specification”, RFC 1884, Dec.1995

[28]. Thomson S. and Narten T., “IPv6 Stateless Address Autoconfiguration”, RFC 1971, Aug. 1996

[29]. Narten T., et.al., “Neighbor Discovery for IP Version 6 (IPv6)”, RFC 1970, Aug. 1996.

[30]. Mangold S., et.al, “IEEE 802.11e Wireless LAN for Quality of Service”, European Wireless 2002, Florence, Italy, Feb.2002

[31]. He D., Shen Q., “Simulation Study of IEEE802.11e EDCF”, IEEE 57th Vehicular Technology Conference (VTC Spring 2003), Jeju, Korea, Apr. 2003

[32]. Bernet Y., et.al "A Framework for Integrated Services Operation over Diffserv Networks" RFC2998, Nov. 2000

[33]. Montavont N. and Noel T., "Handover Management for Mobile Nodes in IPv6 Networks", IEEE Communications Magazine, pp38-43, Aug. 2002

[34]. Rexhepi V. and Heijenk G., "Interoperability of Integrated Services and Differentiated Services Architectures", CTIT technical report, TR-CTIT-00-30, October 2000.

[35]. Wroclawski J. and Charny A. "Integrated Service Mappings for Differentiated Services Networks", Internet Draft, work in progress, Feb. 2001

[36]. Shenker S., Partridge C. and Guerin R., "Specification of guaranteed quality of service" RFC2212, September 1997.

[37]. Wroclawski J., "Specification of the controlled-load network element service" RFC2211, September 1997.

[38]. Jacobson V., Nichols K., and Poduri K., "An expedited forwarding PHB" RFC2598, June 1999.

[39]. Heinanen J., Baker F., Weiss W., and Wroclawski J., "Assured forwarding PHB group" RFC2597, June 1999.

[40]. Pack S., Choi Y., "An End-to-End QoS Provisioning Architecture in Mobile Network", ISCIT 2001, Chiangmai, Tailand, pp5-8, Nov. 2001

[41]. Rexhepi V. et.al., "A Framework for QoS & Mobility in the Internet Next Generation", Technical Report University of Twente, TR-CTIT-00-10, June 2000, 52pp

[42]. Vatn J., "Long Random Wait Times for Getting a Care-of Address are a Danger to Mobile Multimedia, the 1999 IEEE International Workshop on Mobile

Multimedia Communications (MoMuC’99), San Diego, CA USA, 15-17 Nov. 1999.

[43]. Costa X., “A Simulation Study on the Performance of Mobile IPv6 in a WLAN-based cellular network” Computer Networks, No. 40, pp. 191-204, 2002.

[44]. “UCB/LBNL/VINT Network Simulator – ns (version2)”, http://www.isi.edu/nsnam/ns

[45]. “RSVP/ns,” http://titan.cs.uni-bonn.de/~greis/rsvpns/index.html

[46]. Greis M., “RSVP ns An Implementation of RSVP for the Network Simulator ns-2”, www.lab.dit.upm.es/~labrst/02-03/p.sim/rsvpns.pdf

[47]. “DiffServ/ns”, http://www7.nortel.com:8080/CTL

[48]. “MobiWan: Mobile IPv6/ns” http://www.inrialpes.fr/planete/mobiwan/

[49]. Thierry Ernst, “MobiWan: A NS-2.1b6 Simulation Platform for Mobile IPv6 in Wide Area Networks”, http://www.inrialpes.fr/planete/pub/mobiwan

[50]. Wu D. et.al., “On Implementation Architecture for Achieving QoS Provisioning in Integrated Services Networks” IEEE International Conference on Communications (ICC'99), Vancouver, Canada, June 6-10, 1999.

[51]. Soliman H. et al., “Hierarchical MIPv6 mobility management”, draft-ietf-mobileip-hmipv6-07.txt, work in progress, Oct. 2002.