QoS Support in Mobile/Wireless IP Networks using Differentiated Services and Fast Handoff Method

Suk-Un Yoon, Ji-Hoon Lee, Ki-Sun Lee, and Chul-Hee Kang
Broadband Telecommunication Laboratory
Department of Electronics Engineering, Korea University
5-1 Anam-Dong, Sungbuk-Ku, Seoul 136-701, Korea

***Abstract*-In mobile/wireless IP network where frequent mobility is encountered, it is required to accommodate Internet Quality of Service (QoS) mechanisms such as Differentiated Services (DiffServ) for mobile hosts. In this paper, we consider the issue of integration of mobile IP and DiffServ including handoff condition and propose the schemes with hierarchical agent architecture for supporting QoS in mobile/wireless IP networks. To support DiffServ for mobile hosts, we configure foreign agent to have pre-allocated resources in advance for static Service Level Agreement (SLA) case. In case foreign network has no pre-configured SLA or enough resources for mobile host, we configure new SLA among bandwidth brokers to be set up using dynamic SLA. Also, we propose the signaling and data procedures with DiffServ with respect to fast handoff.**

I. INTRODUCTION

With the rapid increase in the number of portable computer and the emergence of high-speed wireless access technology, mobile computing applications may become more practical. Real-time application such as Internet telephony, video conferencing, video-on-demand, audio library, and news-on-demand in mobile environments also should be realized. In mobile and wireless networks where real-time traffic carried, it is important concern how to provide Quality of Service guarantees for such traffic to mobile hosts.

QoS provisioning means that the real-time traffic should get predictable service from resource in the communication system. The proceeding QoS aspects will cause problems when including wireless environments, due to following major differences between wired and wireless networks: link characteristic and user mobility. To resolve these problems resource allocation and call admission control in wireless network have been studied. Also, to guarantee QoS in wired network Resource Reservation Setup Protocol (RSVP), Integrated Services (IntServ), and Differentiated Services (DiffServ) have been designed [2]. In this paper, we adopt DiffServ as a QoS provisioning method in IP networks and propose integration scheme of mobile IP and DiffServ.

Because mobile host freely move from place to place while continuing communications, it is another issue to handle user mobility in mobile network. To support user mobility in the Internet, the IETF (Internet Engineering Task Force) has designed a Mobile IP protocol. Mobile IP offers a mechanism for handoff, but it is difficult to support seamless QoS guaranteed handoff in mobile and wireless network environment. In order to support real-time traffic which has delay bounds during handoff, fast handoff method are proposed in [6]. The fast handoff method eliminates the service disruption period due to handoffs by anticipating the mobile host's (MH's) movement and using multiple bindings to direct the MH's traffic to multiple locations which it may move to. In new draft [7], multicasting to the multiple locations is replaced by bicasting to the previous FA and new FA. In this paper, we adopt fast handoff method to achieve seamless handoff and apply DiffServ to fast handoff with hierarchical agent.

II. RELATED WORKS

A. *Mobile IP & Route Optimization*

Mobile IP [1] is a modification to IP that allows hosts to continue to receive data no matter where they happen to attach to the Internet. In Mobile IP, the Home Agent (HA) in the home network of the mobile host intercepts packets destined for the mobile host using proxy ARP, and then delivers them to the mobile host's current attachment point to the Internet using tunneling. The current attachment point is defined by an IP address called Care-Of Address (COA). There are two different types of care-of address: a foreign agent care-of address is an address used by a mobile host as a tunnel exit-point when the mobile host is connected to a foreign link, and a co-located care-of address is an temporarily assigned address to one of a mobile host's interfaces. When a mobile host moves into a new foreign network, it receives the agent advertisement message including care-of address from new FA. At this time, the mobile host can realize that it moves into the new foreign network, and then sends the registration request message requiring registration of the new care-of address to the new FA. The new FA relays it to the HA of the mobile host so that the HA delivers the packets to the new FA instead of the previous FA.

Mobile IP suffers from a problem known as triangle routing, which refers to the path followed by a packet from a correspondent host to a mobile host which must first be routed via the mobile host's home agent. Triangle routing incurs potentially significant overheads in the delay and network resources consumed for communication with mobile hosts. Mobile IP with route optimization extension [9] avoids triangle routing as follows. When a mobile host's HA intercepts an IP packet, it informs the Correspondent Host (CH) of the mobile host's current COA, this is called a binding update message. The correspondent host can cache this information and send subsequent packets by tunneling them directly to the mobile host's COA. On the other hand, data in flight that had already been intercepted by the home agent and tunneled to the old care-of address are forwarded to mobile host's new care-of address by previous foreign agent which is reliably notified of mobile host's new mobility binding.

B. *Differentiated Services*

Differentiated Services [3,4] has recently become the preferred method to address QoS issues in IP networks. In DiffServ network packets are classified prior to entering the network via a packet marking mechanism, and the service that a router inside the network

This work is supported by Hyundai Electronics Industries Corporation.

0-7803-6596-8/00/$10.00 © 2000 IEEE

provides to a packet are only dependent on the packet's class. The QoS information is carried in band within the packet in the Type of Service (ToS) field of the IP header. An end-to-end service is obtained by concatenation of per-domain services and Service Level Agreement between adjoining domains along the path that traffic crosses in going from source to destination. Per domain services are realized by traffic conditioning at the edge and simple differentiated forwarding mechanisms at the core of the network. Currently, two Per-Hop Behaviors (PHBs) are under active discussion within the DiffServ working group: an Expedited Forwarding (EF) PHB and an Assured Forwarding (AF) PHB.

To receive differentiated services from Internet Service Provider (ISP), customers must have a SLA with its ISP. An SLA basically specifies the service classes supported and the amount of traffic allowed in each class. An SLA can be static or dynamic: static SLAs are negotiated on a regular basis such as monthly or yearly, dynamic SLAs must be negotiated using signaling protocol such as RSVP to request services on demand. In sharing of organization's (e.g. ISP's) Internet resource, the resource control can be done independently by individuals or can be done by agents that have some knowledge of the organization's priorities and polices and allocate resource with respect to those polices. This agent called Bandwidth Brokers (BB). BBs have two responsibilities: the primary one is to parcel out their region's marked traffic allocations and set up the leaf routers within the local domain, the other one is to manage the messages that are sent across boundaries to adjacent regions' BBs.

C. *Fast Handoff Method*

In order to support fast and seamless handoff without packet loss, one solution is to store the packets at the foreign agent (or base station) as a provision for the dropping of packets during the handoff event. Here, packet buffering methods are classified into two categories: a unicast-based buffering method and a multicast-based buffering method. In the former method, only the current foreign agent for the mobile host performs packet buffering and forwards the buffered data to the new foreign agent, to which mobile host is connected after handoff, immediately after the address of the new base station is informed to the previous foreign agent. On the other hand, in the latter method, all adjacent foreign agent of the current foreign agent perform packet buffering with anticipation of future handoff the mobile host. For this purpose, the packet destined for the mobile host are forwarded to all adjacent foreign agents in addition to the current foreign agent by using multicast routing. Therefore, there is no need to forward the buffered packets to the new foreign agent at the time when the handoff actually takes place. It thus makes the handoff latency be shorter than that of unicast-based buffering. However, for doing so, the overhead of buffering is increased and more complicated multicast routing is necessary, in the case of the multicast-based buffering.

Another solution is to make use of hierarchical agents and Mobile IP regional registration [8] to support seamless handoff, which solution is called fast handoff method. In early version of fast handoff method anticipating the mobile host's movement and using multiple bindings to direct the mobile host's traffic to the multiple locations, which it may move to, eliminate the service disruption period. In new version of fast handoffs, seamless Mobile IP handoff is done by bicasting to the previous FA and new FA while mobile host is moving between them. The anticipation of the mobile host's movement is attained by coupling with layer 2 functionality that is dependent on the type of access technology used.

III. System Design

A. *Network Architecture*

Hierarchical structure may be helpful to resource configuration in DiffServ as well as fast handoff in mobile environments. If the agents for mobile host are hierarchically grouped and BB are well allocated for the group (e.g., campus or company), new resource allocation request from mobile host after handoff are easily dealt with by BB in case of dynamic SLA. To support QoS in mobile wireless IP networks, we adopt three-level hierarchical structure of foreign agent. We assume that all foreign agents and home agents provide differentiated services acting as Leaf Router (LR) because these agents are suitable entity to identify flows of mobile hosts. In our system, several foreign agents can be clustered as an administrative domain (e.g., campus area) by the configuring agent which is called Hierarchical Proxy Agent (HPA). Sometime, this group can include a home agent. Also, HPA can be grouped as a local administrative domain (e.g., ISP domain) by the configuring agent which is called Hierarchical Domain Agent (HDA). These agents also support differentiated services. We configure each domain to have it's own BB. The proposed network architecture is shown in Figure 1.

B. *DiffServ for Mobile Host*

To support DiffServ in mobile environments, several problems should be considered. When a mobile host that communicates with enough resource moves to a new foreign network, it should be offered the requested resource by foreign network. We design the foreign networks to have two options to offer resource to the mobile host: providing pre-configured resource and newly configured resource. In the former case, service level agreements are negotiated between the foreign network and the home network or the other foreign network in advance to handle the mobile host. In the latter case, whether the foreign network cannot provide enough capacity to fulfill mobile host's reservation request or there are no pre-configured service level agreements, the mobile host have to request resource allocation to the local bandwidth broker using dynamic service level agreement. This new service level agreement should be coupled to fast handoffs that are in conjunction with layer 2 mobility. When the foreign agents have DiffServ functionality, they can identify and process flows of mobile host in performing Multi-Field (MF) classification. On the other hand, when the foreign agents don't support DiffServ functionality or a mobile host uses co-located care-of address, the mobile host notifies the service of its flow to the bandwidth broker which configures the first and border routers in foreign network.

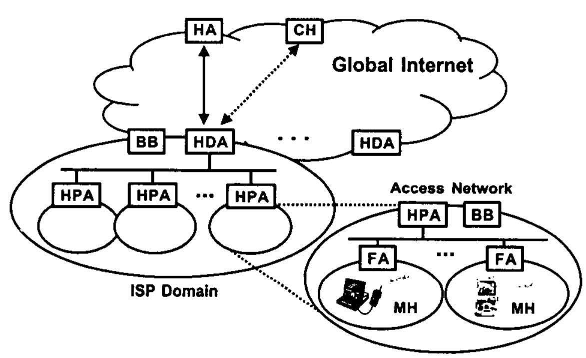

Fig. 1. Proposed Network Architecture.

1) *Static SLA for mobile host*: In static SLA all allocations should be statically pre-allocated through merely bilateral agreements between networks (or users). Because of no notification propagated to core networks, it is more simple and scalable than dynamic SLA. But it would not allow strict end-to-end guarantees. Boundary routers, in static SLA, should be manually configured with the classification, policing, shaping rules, that is, resources are allocated for each users. Unused resources, which can be shared by other users, also can be used by mobile hosts. But, in a QoS viewpoint, it is no reasonable for the mobile host to use only surplus resources. One possible solution is to separately allocate resources of the foreign agent for the mobile host in advance. That is, foreign agents have resources for mobile hosts. The foreign agent classifies mobile host's data based on DiffServ profile list which is made from visitor list.

- Sending from mobile host:

The delivery process of MH's data with static SLA is as follows.

a) MH request resources to local BB using signaling message (e.g., RSVP or a vendor-specific signaling mechanism) through FA.

b) If local BB grants the request, it will set the classification and shaping rule on FA.

c) MH sends packets to leaf router FA.

d) FA performs MF classification with respect to home address of MH in its DiffServ profile list. If the traffic of MH is non-conformant, FA will shape it.

e) Each intermediate router performs BA classifications from FA to DiffServ router in destination network that can be home network case of using reverse tunneling or other foreign network.

- Receiving on mobile host:

There are two possible cases in receiving data on MH.

In the first case, MH is just a receiver that don't need to control of incoming traffic. In this case MH merely receives the data from the source that performs MF classification with home address or COA of MH. The data can be delivered to MH by way of directly FA in case route optimization or via HA in case triangle routing with DiffServ forwarding mechanisms.

In the second case, MH wants to determine the transmission quality of incoming traffic. Because, in DiffServ, the sender commonly requests resources and it's data are processed in access networks, it is difficult to satisfy the mobile receiver. In [5], two receiver control types in DiffServ have been introduced. One is when the receiver is allowed to control which priority should set by the sender, in case a user is eager to get the result from the request delivered promptly. This type of receiver control can be solved at the session layer or application layer. The other is when the receiver needs to control the priority of the packets that come from the network on its access link. In this control, the semantics of the priority bits is changed across the receiver's access link and the access node will not grant priority according to TOS bits unless they are in agreement with the receiver's wishes.

In our system, the first case is considered in receiving on mobile host, that is, providing QoS for receiver is a portion of sender. But, we design a mobile receiver to be served existing service, if it wants, during handoff using receiver driven resource reservation in receiver's access network.

2) *Dynamic SLA for mobile host*: If the foreign network has no pre-configured SLA or enough resources for the mobile host and the mobile host wants strict end-to-end QoS, new SLAs among bandwidth brokers must be set up using dynamic SLA. The dynamic SLA can support strict end-to-end QoS, but on the other hand, there are the same scalability problems as IntServ signaling.

- Sending from mobile host:

The delivery process of MH's data with dynamic SLA is as follows.

a) MH request resources to local BB using signaling message through FA.

b) Local BB makes an admission decision, if the request is accepted; BB signaling message (e.g., RSVP PATH message and RESV message) is exchanged from local BB to BB in the receiver network.

c) During dynamic SLA, every DiffServ router (including mobility agent) is configured with the corresponding classification, policing, and shaping rules by BB.

d) MH sends packets to leaf router FA which performs MF classification with respect to home address of MH.

e) Each intermediate router performs BA classifications (and reshapes the traffic to make sure that the negotiated service are not violated in egress border router) from FA to DiffServ router in destination network.

- Receiving on mobile host:

To receive data on mobile host dynamic SLA should be configured from sender to receiver. The SLA configuration route can be varied according to delivery mode: triangle routing (sender network – home network – receiver network) or route optimization (sender network – receiver network). After SLA, data delivered to mobile host with DiffServ rules.

The end-to-end DiffServ data delivery from the source to the destination with static and dynamic SLA is shown in Figure 2 which including BB, hierarchical agent, and FA. In hierarchical structure, a packet will be encapsulated and decapsulated by regional tunnel management protocol. When the packet are encapsulated with a new care-of address, the DSCP of the new IP header are duplicated from that of original IP header.

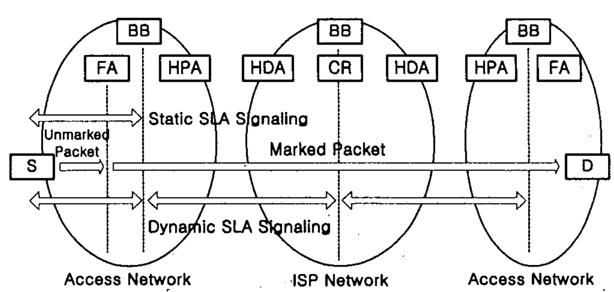

Fig. 2. End-to-End Data Delivery with SLA.

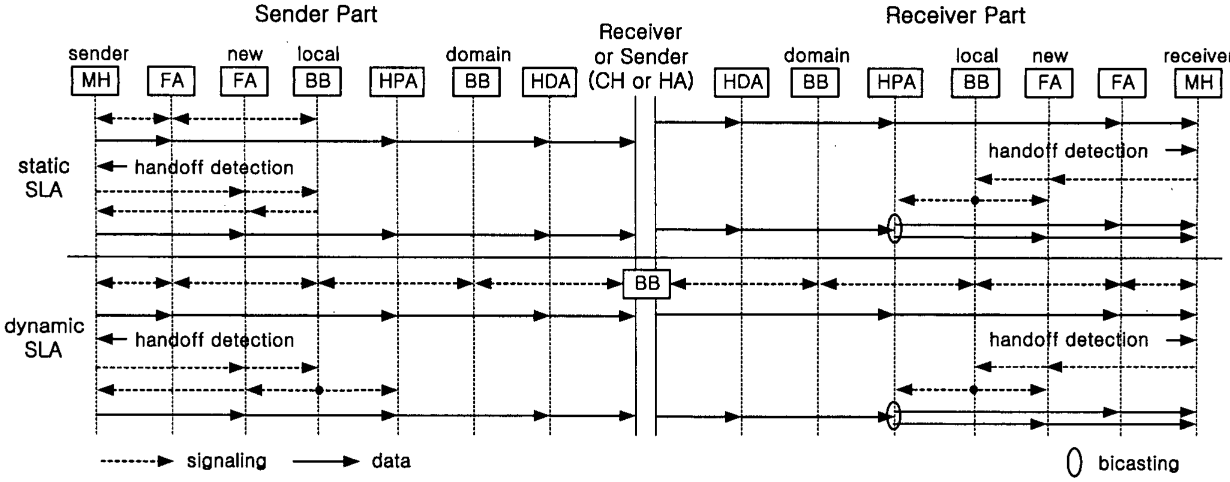

Fig. 3. Flow Chart of Inter-FA Handoff with SLAs.

C. *Fast Handoffs with SLA*

In our hierarchical system, three types of handoff (i.e., inter-FA handoff, inter-HPA handoff, and inter-HDA Handoff) are come into bring by the movement of hosts. Mobile hosts detect the change of attachment point by beacon signal of base stations. Using Layer 2 (L2) wireless technology, MH may solicit or hear the agent advertisement message periodically broadcasted from FA. Because of hierarchy of agents, the agent advertisement information includes not only FA address but also addresses of hierarchical agents.

For the handoff within the local domain (i.e., inter-FA and inter-HPA), hierarchical agent shields the mobility of MH and reduces service disruption and SLA time. To support DiffServ with fast handoff, MF classification by new static or dynamic SLAs have to be set on a new FA. When L2 handoff is detected, mobile sender requests a new SLA to a new local BB to maintain the transmission of its traffic. After handoff, information of MH is deleted from visitor list and DiffServ profile list of old FA by timer or L2 information. On the other hand, mobile receiver requests hierarchical agent to bicast traffic to the previous FA and new FA (or previous HPA and new HPA) using regional registration request. Also, if it wants to receive established service, mobile sender requests local BB to reserve necessary resource.

But, new mechanism is required for inter-domain handoff (i.e. inter-HDA) to support strict end-to-end QoS. One possible solution is a path extension, which uses both existing SLAs and a new SLA between adjacent two domains during a new end-to-end SLA and registration time. In this case, the data in flight to MH's previous domain are forwarded to MH's new domain after a SLA, requested by the MH, between adjacent domains. Also, sender's data are transmitted via earlier domain agent to make use of negotiated SLA during new end-to-end SLA. However, for not strict QoS service, mobile host only sends or receives data after registration to FA and HA or binding update to CH with static SLA.

The detail signaling and data flow with SLAs for intra-domain handoff are shown in Figure 3 and for inter-domain handoff are shown in Figure 4. Wireless link setup and mobile IP registration procedures are omitted for simple flow chart.

D. *Multicast Support*

When FAs and hierarchical agents are willing to support multicasting, our proposed system is adaptable to QoS guaranteed multicast. Because the hierarchy of multicast router should effectively deal with multicast group membership query and report. We propose multicast scheme in hierarchical structure based on the method of using foreign agent because of its several advantages such as optimal multicast routing, low delivery overhead, and low protocol overhead. Multicast operations based on foreign agent routing are similar to route optimization of unicast except for using home agent. A mobile sender sends data to multicast router through foreign agent employing a class D multicast address in destination field and a care-of-address in source field. A mobile receiver receives multicast data from multicast router through foreign agent, hierarchical proxy agent and hierarchical domain agent by IGMP information. Hierarchical agent enhances resource efficiency of public Internet by one time registration and transmission on behalf of several foreign agents in the same cluster and domain. MH's movement incurs a new registration by IGMP message which is processed by correspondent multicast router. Here various kinds of IGMP can be used. But join and leave function are required in IGMP for the resource efficiency of wireless networks.

There are several problems in the use of DiffServ for Mobile IP multicast service provisioning. For dynamic multicast group membership including host mobility, it is difficult to predict in advance the amount of network resources that may be consumed by multicast traffic originating from an upstream network for a particular group. Also, peering SLAs with downstream DS domains is not so easily performed due to the possibility of dynamic group membership at a DS ingress node. The first problem can be solved by dynamic resource renegotiation in hierarchical structure supported by bandwidth broker protocol through requests of hosts. Also, the second problem is easily relaxed in hierarchical network structure by sharing the load of ingress node with hierarchical agent. In our multicast system, fast handoffs are easily achieved. When a MH detects handoff by hearing beacon signal, to receive multicast data, it sends group join message to upper level hierarchical agent or next multicast router through FA.

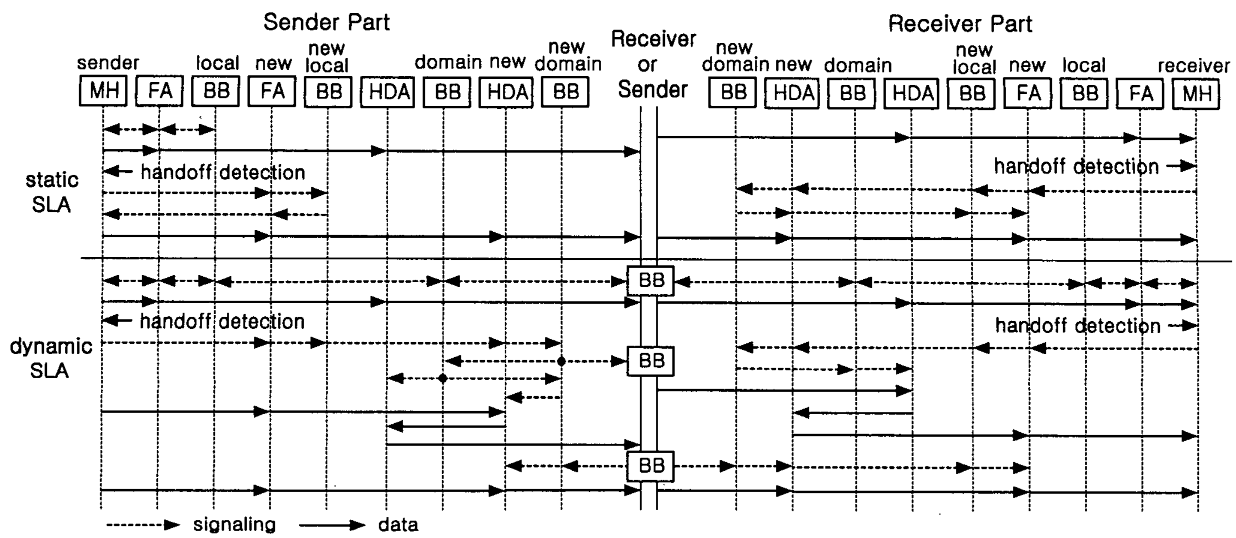

Fig. 4. Flow Chart of Inter-Domain Handoff with SLAs.

IV. Conclusion

We present integration scheme of Mobile IP and DiffServ including fast handoff with hierarchical agent structure. The proposed scheme considers both static SLA and dynamic SLA in DiffServ for mobile host. The static SLA, in which case foreign agents have separately pre-allocated resources for mobile host, have merits such as simple and scalable like original advantages of DiffServ. With the dynamic SLA mobile host can acquire strict end-to-end QoS though foreign network don't have resource for mobile host. For the handoff within the local domain, hierarchical agent shields the mobility of host and reduces service suspension and SLA time. The path extension for inter-domain handoff uses both existing SLAs and a new SLA between adjacent two domains during a new end-to-end SLA and registration time.

To apply our system to real networks such as IMT-2000, the static and dynamic SLA for mobile host must be well blended with each other. Also, resource management connected with wireless bandwidth for mobile host as well as Authentication, Authorization, and Accounting (AAA) function are considered with DiffServ.

References

[1] C.E. Perkins, Editor, "IP mobility support," Network Working Group, Request for Comments 2002, Oct. 1996

[2] Xipeng Xiao and Lionel M. Ni, "Internet QoS: a big picture," *IEEE Internet*, pp. 8-18, Mar./Apr. 1999

[3] S. Blake et al., "An architecture for differentiated services," Request for Comments 2475, Dec. 1998

[4] K. Nichols, V. Jacobson, and L. Zhang, "A two-bit differentiated services architecture for the Internet," Internet draft, draft-nichols-diff-svs-arch-02.pdf, Apr. 1999

[5] Borje Ohlman, "Receiver control in differentiated services," Internet draft, draft-ohlman-receiver-ctrl-diff-00.txt, Mar. 1998

[6] K. El Malki and N. A. Fikouras, "Fast handoff method for real-time traffic over scaleable mobile IP networks," draft-ietf-mobileip-fast-handoffs-00.txt, Mar. 1999

[7] Karim El Malki and Hesham Soliman, "Hierarchical mobile IPv4/v6 and fast handoffs," Internet draft, draft-elmalki-soliman-hmipv4v6-00.txt, Mar. 2000

[8] Eva Gustafsson, Annika Jonsson, and Charles E. Perkins, "Mobile IP regional registration," Internet draft, draft-ietf-mobileip-reg-tunnel-02.txt, Mar.2000

[9] Charles E. Perkins and David B. Johnson, "Route optimization in mobile IP," Internet draft, draft-ietf-mobileip-optim-09.txt, Feb.2000

[10] Charles E. Perkins and Kuang-Yeh Wang, " Optimized smooth handoffs in mobile IP," *IEEE International Symposium on Computers and Communications*, pp. 340-346, Jul. 1999

[11] Chu-Sing Yang et al., "Support an efficient connection for mobile IP," *The Ninth International Workshop on Database and Expert Systems Applications*, pp. 514 -519, Aug. 1998

[12] Cheng Lin Tan et al., "A fast handoff scheme for wireless networks," *The second ACM international workshop on Wireless mobile multimedia*, pp. 83-90, Aug.1999

[13] Suk-Un Yoon, Ji-Hoon Lee, and Chul-Hee Kang, "Efficient and QoS guaranteed multicast support in mobile/wireless networks," *IEEE International Conference on Third Generation Wireless Communications*, pp. 170-175, Jun. 2000

[14] Ji-Hoon Lee, Suk-Un Yoon, Sung-Kwan Youm, and Chul-Hee Kang, "Adaptive resource allocation mechanism including fast and reliable handoff for IP-based mobile networks," *IEEE International Conference on Third Generation Wireless Communications*, pp. 306-312, Jun. 2000