A Network Mobility Management Scheme for Fast QoS Handover

Cheng-Wei Lee  
Dept. of Information Management,  
National Taiwan University  
No.1, Sec. 4, Roosevelt Rd.,  
Taipei City 106, Taiwan  
+886-2-33661200  
jameslee@net.edu.tw

Meng Chang Chen  
Institute of Information Science,  
Academia Sinica  
128 Academia Road, Section 2,  
Nankang, Taipei 115, Taiwan  
+886-2-27883799#1802  
mcc@iis.sinica.edu.tw

Yeali S. Sun  
Dept. of Information Management,  
National Taiwan University  
No.1, Sec. 4, Roosevelt Rd.,  
Taipei City 106, Taiwan  
+886-2-33661200  
sunny@im.ntu.edu.tw

ABSTRACT
The evolution of wireless access technologies has led to a new era of Mobile Internet. Network mobility, which considers the mobility of an entire network, is particularly suitable for mobile environments. Efficient network mobility QoS-handover design is essential to meet QoS requirements for real-time communication. In this paper, we propose a cross-layer hierarchical network mobility architecture and protocol, called HiMIP-NEMO, which is designed for all-IP networks. HiMIP-NEMO optimizes the routing between a mobile network node and the correspondent node, and supports fast QoS provisioning in the network mobility service domain. The simulation results demonstrate that Hi-MIP-NEMO reduces handover latency and packet loss, and supports high velocity vehicles.

Categories and Subject Descriptors
C.2.1 [Network Architecture and Design]: Wireless communication; C.2.3 [Network Operations]: Network management; D.4.8 [Performance]: Simulation.

General Terms
Management, Performance, Design.

Keywords
Mobile IPv6, Network Mobility, QoS Handover

1. INTRODUCTION
A concept called network mobility has become popular for supporting such high-mobility mobile networks. Network mobility refers to the mobility of an entire network that changes its point of attachment to the Internet as one unit, and all data packets sent to and from the mobile network are transmitted via one or more designated mobile routers (MR). Figure 1 shows an example that the mobile hosts in the mobile network (called mobile network node, or MNN) are immobile inside a moving vehicle which is equipped with an MR and provides network access service to the MNN. The vehicle usually moves within the wide area wireless network coverage of a single operator, e.g., Operator 1, and rarely crosses the boundary. The wide area wireless network can be an IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMAX) network or an IP-based cellular system. The major advantage of network mobility is that it reduces the overhead of handovers as well as the power consumption of the MNN.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WICON’08, November 17-19, 2008, Maui, Hawaii, USA.  
Copyright 2008 ICST 978-963-9799-36-3…$5.00.
QoS-handover becomes a more important and complex process. It is essential that the resources needed for the existing services must be reserved in time when the mobile network is moving across the coverage area of BSs to ensure a successful QoS-handover, not just network connectivity re-establishment.

Traditional layered protocol reference model can no longer be efficient to support fast QoS-handover for network mobility service, where individual protocol layers are independently designed and functionally they do not cooperate. For example, during handover, the radio link attachment procedure to a BS in a new subnet may have been completed while the network-layer handover procedure is still waiting for the reconfiguration of mobile nodes’ IP addresses. This inefficiency dictates a combined cross-layer mobility management and resource allocation design to reduce latency and packet loss during handovers.

In this paper, we propose a QoS-integrated network mobility architecture and protocol, called HiMIP-NEMO, which is designed for all-IP networks. When a vehicle travels in the network mobility service domain, the hierarchical architecture and the cross-layer designed protocols provide fast QoS provisioning for the mobile network in the vehicle. The function of the MR is simplified, and an MNN can use its applications and even security mechanism such as IPSec without modification in the mobile network.

2. RELATED WORK

To achieve route optimization, the MR in Prefix Scope Binding Update (PSBU) [3] send aggregate binding updates (BUs) for the mobile network to every correspondent nodes (CN), while in MIRON [4] and ROB [6], the MR helps the MNN to configure a geographically meaningful care-of address (CoA).

The handover performance of NEMO-bs, PSBU, ROB, MIRON, SIP-NEMO [7] and SIP-NGM [5] depends on how fast the new care-of address/prefix configuration and re-registration processes of the MR and the MNNs can be completed. According to [8], duplicate address detection must be run before the newly configured CoA can be accepted, even when the address is provided by the DHCPC6v6 server. In addition, subsequent re-registrations are time-consuming processes that degrade the handover performance. On the network level, fast handover for Mobile IPv6 [9] is designed to operate with the assistance of the link layer for reducing handoff latency and packet loss, and can be applied on NEMO-bs (FNEMO-bs). However, FNEMO-bs suffers from handover trigger timing, multiple tunnels and packet transmission delay fluctuation problems.

Current resource reservation protocols, such as RSVP [10], MRSVP [11] and HMRSVP [12] can only be applied to end-to-end reservation. NEMOR [13] and the BSR scheme [14] are proposed to support QoS for NEMO-bs. NEMOR uses a generic signaling protocol called Next Step In Signaling (NSIS) [15] to exploit DiffServ on the bi-directional tunnel between the MR and the HA. The HA maintains the mapping between the prefix of the serving and target BSs, and the CNs. On the other hand, BSR adopts the concept of HMRSVP to manage network mobility with QoS guarantee. However, the MRs in NEMOR and BSR can only execute re-reservation process after regaining network connectivity under the BSs in new subnet, which degrade the handover performance. In addition, they cannot cooperate with FNEMO-bs that is designed for fast regaining network connectivity.

3. HiMIP-NEMO ARCHITECTURE AND PROTOCOL

Assume that an MR gets a new network prefix from its egress interface at a new location using prefix delegation protocol, and sends router advertisements (RA_MIP) to the mobile network. Consider an MNN adopting Mobile IPv6 and having an HoA. When attached to the MR, it listens for RA_MIP, and configures a geographically meaningful CoA which will be used to optimize the routing of packets. Once connected, the MNN may for example place a VoIP call.

3.1 System Overview

To achieve fast and QoS-guaranteed handover, we propose the use of Foreign Mobility Agent (FMA) in a hierarchical routing architecture which is typical in access networks. Using Figure 1 as an example, FMAs are functional modules residing in the routers of the network mobility service domain and are connected by a high speed wired network to facilitate their fast communications to support QoS-handover. FMA is responsible for reserving resources and allocate new resources for mobile networks once handover happened. A root-FMA is the edge router of the service domain. All other FMAs between the root-FMA and the leaf-FMA are called intermediate-FMAs. A leaf-FMA resides in access router (AR) having wired connections to several BSs, which can form one or many subnets. It is also responsible for managing radio channel resources of its child BSs.

The switching-FMA (SWF) refers to the cross-point of the old and new transmission paths of a handover instance. For example, if the MR in Figure 1 is in the subnet of BS1 and plans to move to the neighboring subnet of BS2, a handover occurs and the intermediate-FMA1 becomes the SWF. A main function of FMA is to maintain routing table for all the MRs attached to their child BSs.

In this paper, we focus on the handover procedure between BSs in different subnets. In the proposed HiMIP-NEMO architecture, a QoS-incorporated registration process is performed whenever an MR enters the network mobility service domain. Furthermore, we distinguish two types of handover in the HiMIP-NEMO operations: proactive and reactive to speed up the handover process. Two layer 2 messages from the wireless access network are used to trigger HiMIP-NEMO operations. When attaching to a BS, an MR will send a registration message, MRReg, containing its information to the BS. When the MR boots up in or enter the network mobility service domain, the MRReg will lead to a HiMIP-NEMO registration process.

The other layer 2 message is the handover indication (HOind) message which is sent by an MR when it is about to disconnect from the current serving BS and hand over to a target BS. An HOind will be sent only when the layer 2 pre-handover negotiations is successful. When receives the HOind message, a HiMIP-NEMO defined layer 3 message, MRHOnotify, will be sent by the serving BS to the target BS. During this message forwarding, based on the prefixes of the serving and target BSs, a route optimization is performed and the two BSs’ nearest common ancestor (i.e. the SWF) is found in the hierarchical Mobile IP backhaul network. This refers to HiMIPv6-NEMO proactive handover. In addition, when abnormal condition happened, which causes failed proactive handover, a HiMIPv6-NEMO reactive handover will take over to recover the on-going sessions.
The REG-REQ and MOB_HO-IND messages defined in the IEEE 802.16e standard [16] have similar functionalities as MRreg and HOind used here, respectively. But they do not provide similar functions here whereby the BSs and FMA in the hierarchical Mobile IP backhaul network cooperate to speed up optimal routing and resource allocation in efficient QoS-handover.

3.2 The QoS-incorporated Registration Process

The messages exchanged during the registration process are shown in Figure 2. In HiMIP-NEMO, during MR’s registration, an extra field is defined in the MRreg to carry the prefix of the mobile network (prefixMR). The MR then aggregates QoS requirements of the MNNs and sends the information to the BS. If the serving BS does not have QoS and handover related information of the MR, it sends a HiMIP-NEMO defined layer 3 message, MRinfo, to its leaf-FMA containing the MAC address (MAC_MR), the prefixMR, and the QoS parameters of the MR (QoSMR). The leaf-FMA first checks if there is enough radio channel resource on the BS. If not, a HiMIP-NEMO defined message, MRinfo_reply, is sent to the BS, rejecting the entrance of the MR. Otherwise, the leaf-FMA searches its MR list. If not found, the leaf-FMA creates a new record for the MR and records all the information in it including the QoS requirements. It then sends a HiMIP-NEMO defined message, newMRquery, to the root-FMA (step 2.) with the MAC_MR, prefixMR, and QoSMR. In the meantime, each intermediate-FMA on the path will also search its MR list for a record of the MR. If not found, a new one is created and the information in the newMRquery is copied into the record. When the root-FMA receives the message and does not find a record on its MR list, it replies a newMRreply with the MAC_MR, and the newMR and the QoS_reservation_confirm (QoSr) fields set to "true" to the leaf-FMA (step 3).

The REG-REQ and MOB_HO-IND messages defined in the IEEE 802.16e standard [16] have similar functionalities as MRreg and HOind used here, respectively. But they do not provide similar functions here whereby the BSs and FMA in the hierarchical Mobile IP backhaul network cooperate to speed up optimal routing and resource allocation in efficient QoS-handover.

3.3 QoS-handover protocols

Figure 3 shows the procedure and the messages exchanged during a proactive handover. A layer 2 message, HOind, is sent by an MR when it is about to disconnect from the serving BS and hand over to a target BS (step 1). An extra field is defined in the HOind message to carry the prefixMR. When receives the HOind, the serving BS sends an MRHO_notify to the target BS (step 2), carrying the MAC_MR, the prefixMR, and the QoSMR. Note that on the way of forwarding the MRHO_notify, the first FMA that finds both the serving and target BSs are its children, becomes the SWF (step 3). An FMA under the SWF on the old path, i.e., the ancestor of the serving BS but not the target BS, will execute an expiration function to release prior QoS reservations. The expiration function should be carefully calibrated to alleviate the ping-pong effect.

The SWF creates a routing rule for this handover instance in its routing table to redirect packets whose prefix is the (old) prefixMR to the target BS. In step 4, each FMA on the path from the SWF towards the target BS will create a new record, and a routing rule for the mobile network as the SWF does. In addition, the leaf-FMA checks whether there is enough radio channel resource of the target BS. If so, it simply forwards the MRHO_notify to the target BS. The BS uses the information of the mobile network to execute fast layer 2 network entry process that eliminates unnecessary information exchange between the BS and the MR (step 5). Once the process completes, all packets whose prefix is the (old) prefixMR are redirected to the mobile network through the target BS. The MNNs can now continue the on-going communications with their CNs (step 6). Afterwards, the MR receives RA from the leaf-FMA (step 7), gets a new prefixMR and broadcasts the prefixMR (step 8 and 9). The MNN then performs the normal Mobile IPv6 operations (steps 10). In the meantime, messages are sent by the SWF for the release of the resources at routers on the

Figure 3. The QoS-incorporated proactive handover
previous routing path. At this point, the routing rules inserted by the FMAs for special routing of the mobile network must be removed. The FMAs can start a timer when setting the routing rule and reset the timer when the rule is hit. After an interval when no packets are sent to the old prefix, which means the MNN has completed the re-registration processes, the routing rules can be removed. In this process, because of the simultaneous execution of finding an optimal new route and performing resource allocation, the handover delay is thus minimized.

In addition, HiMIP-NEMO supports an efficient reactive handover procedure. If the target BS does not receive an MRHOnotify, it sends an MRinfo. FMAs that do not have any information about the MR follow the same procedure as step 2 in Figure 2. Eventually, one of the FMAs will find a record of the MR, and it will become an SWF. The SWF stops the forwarding of newMRquery and performs the following works: a) making QoS reservations for the MR; b) adding a routing rule to its routing table for the (old) pre-fix; c) sending a newMRreply with the newMR field set to “false” and the QoS reservation confirm field set to “true” to the leaf-FMA; and d) sending an MRHOinform message to the root-FMA with the MAC_MRR prefix and QoS_MRR, indicating an abnormal condition. It then creates a newMRreply message and sends along the path back to leaf-FMA. Each intermediate FMA makes corresponding QoS reservation. Once the new QoS route is established, the communications of the MNNs proceed without interruption. In the meantime, messages are sent to release the resources allocated at the routers on the old routing path.

4. PERFORMANCE EVALUATION

We have implemented HiMIP-NEMO on a ns2 simulator that is patched with Mobile IPv6 and IEEE 802.16e modules. A serving BS will send MRHOnotify to the target BS when executing an 802.16e handover, which is suggested but not defined in the IEEE 802.16e standard. To compare the performance, we have implemented Mobile IPv6 Fast Handsovers [9] and also Mobile IPv6 Fast Handovers over IEEE 802.16e Networks [17] for NEMO-bs (FNEMO-bs) on ns2. We have implemented a buffer mechanism on the NAR in case that the tunneled packets arrive the NAR before receiving a UNA sent by the MR.

The radius of an IEEE 802.16e cell is 10 kilometers, and the overlap of two adjacent cells is 200 meters. A vehicle with moving speed: $V_{MR} = \alpha \text{ km/h}$ first starts up in the coverage of BS1 and then moves to the coverage of BS2. When crossing the boundary of BS1 and BS2, there are three signal strength thresholds for the MR to send MOB_SCN_REQ, MOB_MSHO_REQ and MOB_HO-IND, and they are mapped to three corresponding positions. In other words, the MR will send MOB_SCN_REQ when the distance between BS1 and the MR is 9870 meters, and send MOB_MSHO_REQ and MOB_HO-IND at 9930 and 9930+ $\alpha$ meters, respectively. Note that once the MR receives FBack, it will send MOB_HO-IND immediately; however, if the threshold of sending MOB_HO-IND is reached, the MR will send MOB_HO-IND despite that the FBack is not received.

A two-level FMA hierarchy is used in our simulation, and the link delay between leaf-FMA and the root-FMA is set to 10 ms to simulate multiple hops in real deployment. Leaf-FMA1 and leaf-FMA2 (when using FNEMO-bs, PAR and NAR, respectively) broadcast RAs containing different prefixes through BS1 and BS2 respectively. We assume that in NEMO-bs the home networks of the MR and the MNN are the same. When using NEMO-bs in our simulation, a binding update (BU) will be sent by the MR to the CN which is communicating with the MNN, resulting optimized route. In addition, the delay of duplicate address detection (DAD) at new access router (NAR) is set to 1 second when using FNEMO-bs. A G.711 session is established from a CN to a MNN, which is CBR over UDP with 64 kbps and 20 ms packet interval. Handover delay is defined as the interval between the time that the MNN receives the last packet and the time that the MNN receives the first packet when the MR is attached to BS1 and BS2 respectively. Assume that the IEEE 802.16e handover of the MR is successful.

![Figure 4. The handover delay.](image)
higher than that of the packet through optimal path. Both of these two phenomena will degrade the performance of the G.711 session. On the other hand, the routing path remains optimal in Hi-MIP-NEMO. The communication between the CN and MNN will not be affected when the mobile network moves.

5. CONCLUSION
In this paper, we present a QoS-integrated cross-layer hierarchical network mobility management architecture and the protocols, HiMIP-NEMO, which demonstrates the advantages of combining optimal route-redirecting, resource allocation with network mobility management. Simulation results show that HiMIP-NEMO reduces latency and packet loss for mobile network. Because the network side is responsible for all handover negotiations, including resource reservation, HiMIP-NEMO supports high velocity vehicles, and, in addition, the function of the MR is simplified which benefit network mobility service deployment. Furthermore, an MNN can use the network mobility service without modification, including its security mechanism, e.g. IPsec.

6. REFERENCES