To appear in Proc. of 16th IEEE Int'l Symp. on Personal Indoor and Mobile Radio Communications (PIMRC), Berlin, Germany, Sept 2005

Fast Seamless Handover Scheme and Cost Performance Optimization for Ping-Pong Type of Movement

Zongkai Yang*, **Yuming Wang***, **Dasheng Zhao***, **Jianhua He***, **Xiaoming Fu**†
*Department of Electronics and Information Engineering
Huazhong University of Science and Technology
Wuhan 430074, China
Email: dr.ymwang@163.com
†Institute for Informatics, University of Goettingen
37083 Goettingen, Germany

***Abstract*— The ping-pong type of movement is a typical motion manner in mobile IPv6 networks, which will bring frequent handovers and thus increase signaling burden. On the other hand, reducing handover delay in this case seems to be more significant. In this paper we propose a fast seamless handover scheme for the ping-pong type of movement as an extension to the hierarchical mobile IPv6. Based on the simulation results, it can be observed that, by setting the reservation active flag (RAF) and the offline count down timer (CDT), the scheme significantly reduces QoS signaling cost and handover delay. Furthermore, the simulations work out an optimized CDT for acquiring better cost performance of resource reservation.**

I. Introduction

Nowadays, mobile communication and Internet technology are two emerging developing technologies that tend to converge to form the fundamental elements for future information communication infrastructure. As the IETF proposed solution for mobility support in the Internet, Mobile IP has become a promising technology which is expected to be the primary implementing technology and important developing direction in the next generation mobile networks. With the development of IP-based wireless multimedia communications, Mobile IP technology has been standardized by the industry and investigated both in theory and practice. Though the IETF had released the first Mobile IPv4 protocol specification as early as 1996 [1] (a recent update is available in [2]), Mobile IPv6 specification [3] was just approved a few months ago as a proposed standard. Compared with the relatively more mature research status in the wired access networks, Mobile IP is still undergoing a number of issues, some being criticial, especially when in terms of large-scale deployment, such as fast handover of mobile nodes, end-to-end QoS guarantee, framework scalability and reliability, performance and security.

Ping-pong type of movement is a typical motion manner of the mobile node (MN). Moving back and forth between two access routers (AR) especially when the MN stays exactly at the boundary will result in frequent handovers. Today, ping-pong type of movement happens more frequently as the cells become smaller. An analysis of the influences of ping-pong type of movement on the performance of some typical mobile IP(v6) extensions can be found in [4]. When Quality-of-Service is concerned (where would be one of the key incentives of mobile communications), resources just reserved a moment ago need to be released because of the MN's departure; on the other hand, resources just released a moment ago need to be reserved again because of the MN's return to the previous network. As frequently reserving and releasing resources will greatly increase QoS signaling cost and handover delay, we believe this is an important issue needs to be addressed. Actually, in a recent work regarding next-generation Internet signaling protocols in mobile environments [5] (section 8.3), the ping-pong type of movement has been defined as an open issue. Usually, when an MN departs from an access network, resources along the old path should be released as quickly as possible to avoid waste of resources. It is inefficient to wait until the soft-state timer (which is typically used in most QoS and other IP signaling protocols) expires in the mobile access network where resources are scarce. However, immediate release of resources along the old path should be avoided in case of a ping-pong type of movement so that the old reservation can be reused after a very short period of time. Noticeably, the current IP QoS signaling protocol specification (QoS-NSLP) [6] specified by the IETF defines a REPLACE flag which can help to keep the reservation along the old path. However, there is still a lack of resource management scheme in such a ping-pong type of movement.

In this paper we study the MN's typical ping-pong type of movement under the architecture of hierarchical mobile IPv6, and propose a fast seamless handover scheme adapting to the hierarchical architecture, with a detailed analysis of its properties. Based on the simulation results, it is observed that, by setting the reservation active flag (RAF) and the offline count down timer (CDT), the scheme significantly reduces QoS signaling cost and handover delay, thus reduces network burden of signaling messages and probability of service interruption. Furthermore, the MN is allowed to find out a proper CDT according to its mobility characteristics to acquire better cost performance of resource reservation along several neighboring domains.

The rest of the paper is organized as follows. Section II describes the architecture of hierarchical mobile IPv6 and the general handover scheme based on the hierarchical architecture. Section III illustrates the fast seamless handover scheme, including the state diagram and the detailed handover procedure. In section IV, the scheme is evaluated by simulations, and cost performance optimization is also analyzed. Finally, we draw a conclusion in section V.

II. General Handover Scheme

In mobile networks, reducing handover delay is the key of service interruption avoidance. Furthermore, frequent handover will bring heavy signaling burden to the core network. In order to speed up the MN's handover and reduce the amount of singling messages, researchers have presented many mobility support schemes which are ordinarily called "IP micro-mobility protocols" [7].

A typical IP micro-mobility protocol, Hierarchical Mobile IPv6 (HMIPv6) [8], which is an extension of mobile IPv6, is designated in Fig. 1. HMIPv6 provides mobility management based on a hierarchical architecture, and a new network entity called Mobility Anchor Point (MAP) is introduced. When the MN moves into a foreign network, it uses the current MAP as its local home agent (HA), and acquires a global routable Regional Care-Of Address (RCoA) from the subnet that the MAP belongs to as its local home address. At the same time, the MN is configured with an on-Link Care-Of Address (LCoA) based on the prefix advertised by its default AR as its local care-of address. The MN's HA and correspondent nodes (CN) maintain binding information between the MN's home address "MN" and global care-of address "RCoA", and according to this, datagrams towards the MN will be transmitted by its HA or sent by its CN to the global care-of address "RCoA". On the other hand, the MAP maintains binding information between the MN's local home address "RCoA" and local care-of address "LCoA", and according to this, transmits the received datagrams to the local care-of address "LCoA".

When the MN switches between the ARs that belong to the same MAP domain, it keeps its global care-of address (also its local home address) "RCoA" unchanged, and only changes its local care-of address "LCoA", so it just needs to send a local Binding Update to register itself on the current MAP, that is, its local HA. When the MN finds itself moving into a new MAP domain and has acquired a new global care-of address "RCoA", it must send Binding Updates to its HA and CNs. In this way, intra-domain handover delay is reduced, and signaling burden brought by frequent handover is lightened.

Suppose the MN is performing an intra-domain handover from AR1 to AR2, the MN sends a Binding Update massage only to MAP1, and the binding registration will be finished as the Binding Acknowledgement message is returned by MAP1. When the MN is performing an inter-domain handover from

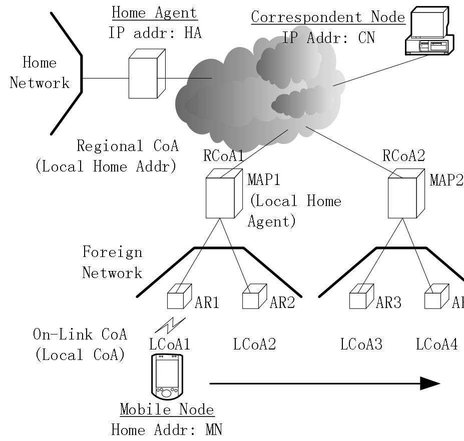

Fig. 1. Hierarchical mobile IPv6 architecture.

AR3 to AR2, the MN must send Binding Update messages both to its HA and CNs besides MAP1, and the binding registrations will not be finished until all the Binding Acknowledgement messages are returned. After all the binding registrations are finished, data can be exchanged directly between the MN and its CNs. In some situation, see Fig. 2, a path must be reserved between the MN and its CN before data exchange starts, in order to guarantee the quality of services. According to the current QoS-NSLP specification, the MN can send RESERVE massage directly, and it is sure that the path is reserved successfully once the right RESPONSE message is received. After that, the MN can send data packets to the CN. Similarly, the CN must send the RESERVE message and receive the right RESPONSE message before it sends data packets to the MN.

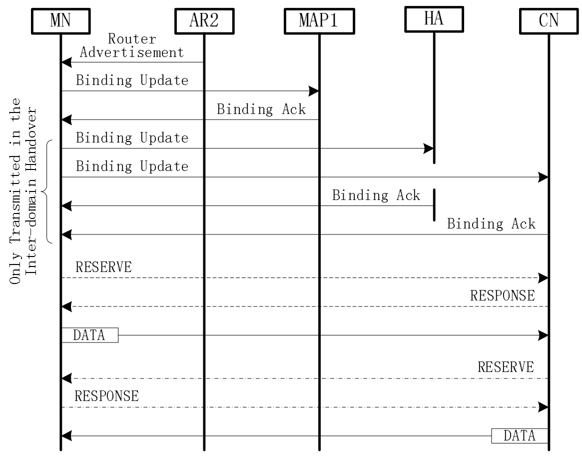

Fig. 2. General intra/inter-domain handover process.

Thus it can be seen that when the MN is making a handover to a new AR, no matter it is an intra-domain handover or an inter-domain handover, the MN and the CN must wait the right reserve response before the first data packet is sent in order to guarantee the quality of services. Hence, except for binding registration delay, QoS handover delay is also introduced. In ping-pong type of movements, QoS handover delay further increases the probability of service interruption, moreover, frequent handover brings up large numbers of QoS signaling messages and thus greatly increases network burden.

III. Fast Seamless Handover Scheme

A. State Diagram

A state diagram for resource management is explored to decrease the frequency of resource reserving and releasing in a case of ping-pong type of movement. As Fig. 3 has indicated, it consists of the following three states:

Online - The MN enters the current AR for the first time and has finished binding registrations and resource reservation. The MN may also enter the current AR again to try back to this state from "idle" state before the offline count down timer (CDT) is time out. The reservation active flag (RAF) is set and the MN is exchanging data with its CN.

Idle - The MN has handed over from the current AR to a new one. In order to get ready for the MN's momentary return, the RAF is cleared to de-active the MN's resource reservation through the current AR instead of releasing the reserved path completely. Besides, the CDT is started.

Offline - The CDT is time out, and it seems little probability of the MN's return to the current AR, thus the MN's related state information will be cleared and the reservation through the current AR will be released completely.

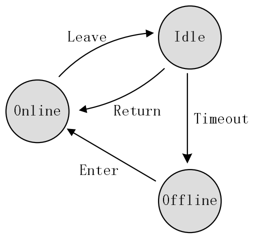

Fig. 3. State diagram of the fast seamless handover scheme.

The CDT is denoted by Δt with N levels ($0 < \Delta t_1 < \Delta t_2 < \cdots < \Delta t_N$) and initialized with a lesser value Δt_1 at the MN's first time access to the AR. Each time the MN returns to "online" state from "idle" state, the CDT will step into an upper level until it reaches the highest value Δt_N. In the ping-pong type of movement, the MN's frequent return to the original AR in a very short time leads the CDT to a higher and higher level, and will help to prevent reservation through the original AR from being released, thus conduces to a faster QoS handover and lower signaling cost. The MN can detect whether it is a ping-pong movement user according to its previous mobility characteristics. If not, for example, the MN moves linearly, then Δt_1 is set to zero to avoid waste of resources.

B. Handover Procedure

According to the above state machine description, the fast seamless handover scheme is realized by the control of the RAF and the CDT. Setting the RAF means the resources reserved through the current AR are in use, on the contrary, clearing the RAF means the resources reserved through the current AR are not in use. Once the reservation has been made, when the MN returns to the AR before the CDT is time out or departs from the AR to start the CDT, it needs only to set or clear the RAF instead of performing complete QoS signaling exchanging procedure to realize equivalent resource reserving and releasing. In this way, QoS signaling cost and handover delay are both greatly reduced.

As Fig. 4 has indicated, when the MN is making a handover at its first access to AR2, no matter it is an intra-domain handover or an inter-domain handover, besides the normal binding registration, a complete resource reservation procedure must be performed before the MN and the CN can exchange data. Different from the general handover scheme, in the fast seamless handover Scheme, in order to support a hierarchical architecture, MAP1 separates the end-to-end reserving path into two segments, thus, when the MN is moving in the same MAP domain, only the segment of the path between the MN and MAP1 needs to be updated, and this greatly reduces intra-domain QoS handover delay.

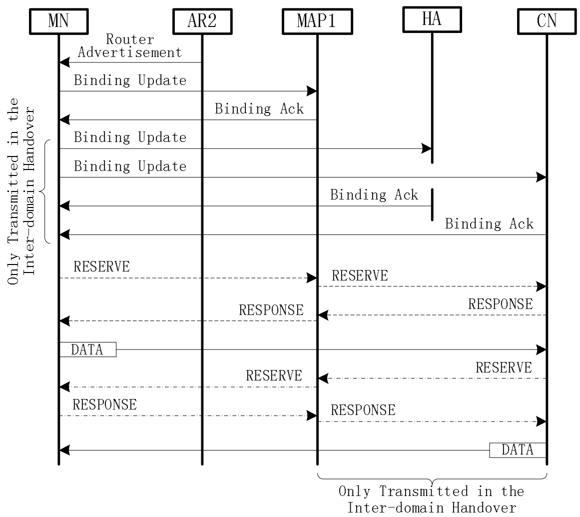

Fig. 4. Intra/inter-domain handover procedure at first time access to AR2.

As Fig. 5 has indicated, if the MN returns to AR2 in MAP1 domain in a very short time before the MN's CDT related with AR2 is time out, besides the normal binding registration, it need only send an ACTIVE message to set the RAF along the path instead of performing a complete resource reservation procedure, and the data packet can be sent without any waiting in the wake of the ACTIVE message, thus QoS handover delay is completely eliminated. In this case, a piggybacking ACTIVE message with the data packet might further reduce QoS signaling cost, but this might bring about some modifications of the data packet header, and will not be discussed here. When the MN is departing from AR2, it needs to send INACTIVE message to de-active the reserved resources.

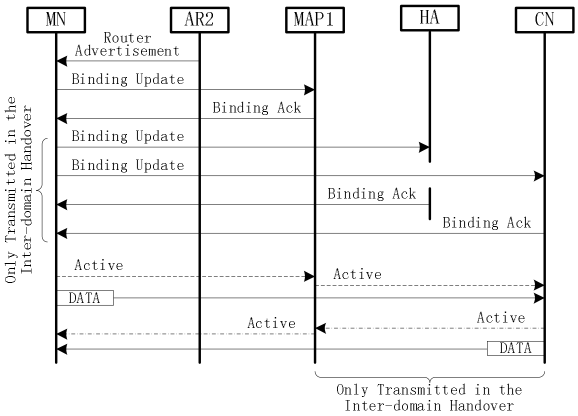

Fig. 5. Intra/inter-domain handover procedure at subsequent return to AR2.

IV. Simulation and Evaluation

The simulation topology is indicated in Fig. 6, and two MAP domains are defined in it. AR1, AR2 and MAP1 belong to the first domain (MAP1 domain), while AR3 and MAP2 belong to the second domain (MAP2 domain). The MN makes handover between AR1, AR2 and AR3 randomly, and stays in each AR for random time. No matter in intra-domain handover from AR1 to AR2 or in inter-domain handover from AR3 to AR2, both the comparison of signaling cost and the comparison of handover delay between the general handover scheme and the fast seamless handover scheme are made.

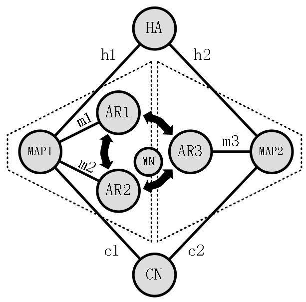

Fig. 6. Simulation topology.

"h_1/h_2", "c_1/c_2" and "$m_1/m_2/m_3$" in the simulation topology are numbers of hops between the MAP and the HA, CN, AR respectively. Distance between the MN and each AR is 1 hop. Considering Router Advertisement messages are sent by the AR periodically, and the MN just receives the unsolicited multicast messages passively, so signaling cost and delay related to the Router Advertisement message will not be taken into account for the moment. Suppose packet flow distribution is symmetrical over the network, and each router has an equal switching capability, for simplicity, assume that signaling cost and delay introduced by each hop is same, and is denoted by s_0, d_0 respectively, then the calculating expressions of signaling cost and delay in the general handover scheme and the fast seamless handover scheme are listed in TABLE I.

Detailed parameter settings in the simulations are listed in

TABLE I

Calculating expressions of signaling cost and delay.

Handover Scheme	Evaluation Items	Intra-domain Handover	Inter-domain Handover
General Handover	Signaling Cost S_g	$(6+6m+4c)\cdot s_0$	$(10+10m+2h+6c)\cdot s_0$
	Delay D_g	$(4+4m+2c)\cdot d_0$	$(6+6m+4c)\cdot d_0$
Fast Seamless Handover	Signaling Cost S_f	$(4+4m)\cdot s_0$	$(8+8m+2h+4c)\cdot s_0$
	Delay D_f	$(2+2m)\cdot d_0$	$(4+4m+2c)\cdot d_0$

TABLE II. Furthermore, the CDT is defined with 5 levels (10, 20, 40, 80, 120 seconds respectively), and the probability distribution of the MN's random stay time in a certain AR is listed in TABLE III, which defines a general case with part of ping-pong type of movement approximately.

TABLE II

Parameter settings.

h_1/h_2	c_1/c_2	$m_1/m_2/m_3$	s_0	d_0
10 Hops	12 Hops	2 Hops	1	0.1 Seconds

TABLE III

Probability distribution of the random stay time.

Stay time (s)	$0\sim5$	$5\sim10$	$10\sim20$	$20\sim40$	$40\sim60$	$60\sim\infty$
Probability	45%	20%	15%	10%	7.5%	2.5%

In the simulations, the MN made about 100000 random handovers between AR1, AR2 and AR3. Fig. 7 shows the change of cumulative signaling cost in all handovers (including both intra-domain and inter-domain handovers), and obviously, the fast seamless handover scheme greatly reduces the cumulative signaling cost than the general handover scheme.

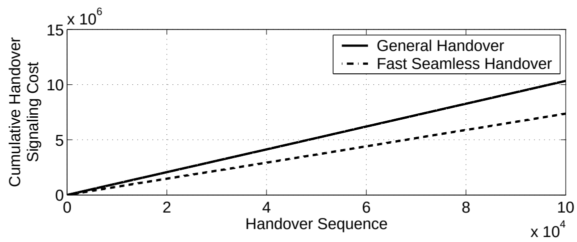

Fig. 7. Comparison of cumulative signaling cost.

Fig. 8 shows that, with the increase of handover number, compared with the general handover scheme, the fast seamless handover scheme greatly reduces the average intra/inter-domain handover delay.

It is just because of keeping reservation along the old path with an offline count down timer (CDT) that the signaling cost and handover delay in ping-pong type of movement are greatly reduced. Regarding the product of reserved bandwidth along the old path (B) and keeping time (T) as the paying cost (C_b), and the depression of handover signaling (ΔS) and

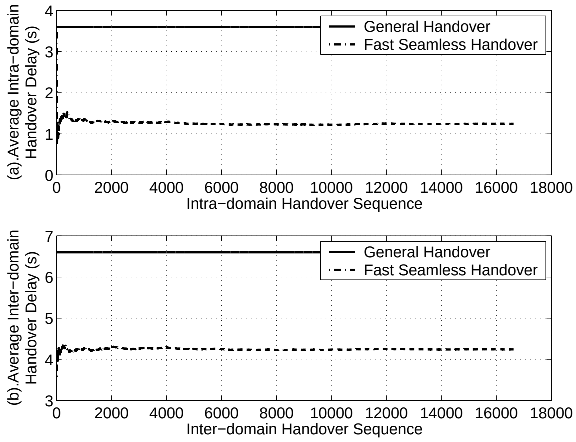

Fig. 8. Comparison of average intra/inter-domain handover delay.

delay (ΔD) as the achieved performance, cost performance for signaling and delay depression are defined as follows,

$$CP_s = \frac{\sum_i \Delta S}{\sum_j C_b} = \frac{\sum_i S_g - S_f}{\sum_j B \times T} \quad (1)$$

$$CP_d = \frac{\sum_i \Delta D}{\sum_j C_b} = \frac{\sum_i D_g - D_f}{\sum_j B \times T} \quad (2)$$

The best cost performance is achieved when CP_s and CP_d reach their maximum. For a determinate motion model of the MN, the value of CP_s and CP_d are influenced by the CDT (suppose its 5 levels are respectively defined as Δt, $2\Delta t$, $4\Delta t$, $8\Delta t$ and $12\Delta t$ seconds). If $\Delta t \to 0$, the MN cannot return to the original AR before the CDT is time out, as a result, CP_s and CP_d approach to zero. Along with Δt increasing, the probability that the MN returns before the CDT is time out increases, thus CP_s and CP_d will increase. As $\Delta t \to \infty$, reservation along the old path will not be released forever, and it is not an optimized manner in reality, so CP_s and CP_d won't reach their maximum. Accordingly, a proper Δt can be found to achieve the best cost performance.

Suppose the MN makes random handovers between the ARs according to the random stay time defined in TABLE III, and $\Delta t \in [1, 10]$, Fig. 9 illustrates the relationship between the cost performances (CP_s & CP_d) and the CDT interval Δt.

It can be observed that, when $\Delta t = 3$, namely, the CDT's 5 levels are respectively 3, 6, 12, 24 and 36 seconds, the best cost performances for signaling and delay depression is achieved. In the simulations, the Δt that CP_s reaches its maximum may be different with the one for CP_d, corresponding Δt can be selected according to the MN's emphasis. Besides, Δt is not limited to integer, here we just have a rough discussion about the existence of Δt for the best cost performance. Detailed Δt estimating method will be studied in the subsequent research work.

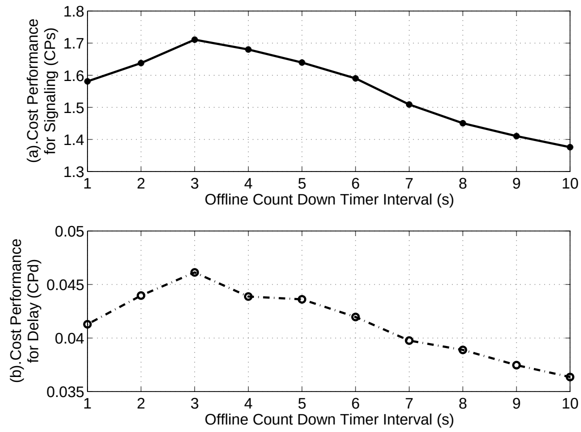

Fig. 9. CP_s, CP_d and Δt relationship curves.

V. Conclusion

Ping-pong type of movement is a typical motion manner of mobile node. In order to effectively provide users with an uninterrupted service which satisfies a certain quality, the handover problem, especially the QoS handover problem, should be well solved. The fast seamless handover scheme presented in this paper adapts to the hierarchical architecture, and well reduces QoS handover delay and signaling cost in the ping-pong type of movement. Further reducing of handover delay helps to decrease the probability of service interruption. At the same time, the use of the RAF effectively reduces network burden of signaling transmission when mobile node is making frequent handovers. Furthermore, mobile node is allowed to find out a proper CDT according to its mobility characteristics to acquire better cost performance of resource reservation along several neighboring domain.

References

[1] C. Perkins, "IP Mobility Support for IPv4", *IETF RFC2002*, Oct 1996.

[2] C. Perkins, "IP Mobility Support for IPv4, revised", *IETF internet draft*, draft-ietf-mip4-rfc3344bis-01 (work in progress), Jun 2004.

[3] D. Johnson, C. Perkins, and J. Arkko, "Mobility Support in IPv6", *IETF RFC3775*, Jun 2004.

[4] Robert Hsieh and Aruna Seneviratne, "A comparison of mechanisms for improving mobile IP handoff latency for end-to-end TCP", *Proceedings of the 9th annual international conference on mobile computing and networking (MobiCom'03)*, Sep 2003.

[5] S. Lee, S. Jeong, H. Tschofenig, X. Fu, J. Manner, "Applicability Statement of NSIS Protocols in Mobile Environments", *IETF internet draft*, draft-ietf-nsis-signaling-protocol-mobility-00 (work in progress), Oct 2004.

[6] S. Van den Bosch, G. Karagiannis, A. McDonald, "NSLP for Quality-of-Service signalling", *IETF internet draft*, draft-ietf-nsis-qos-nslp-05 (work in progress), Oct 2004.

[7] Pierre Reinbold and Olivier Bonaventure, "IP micro-mobility protocols", *IEEE communications surveys and tutorials*, Vol. 5, No. 1, pages 40-57, Mar 2003.

[8] H. Soliman, C. Catelluccia, K.E. Malki and L. Bellier, "Hierarchical Mobile IPv6 mobility management (HMIPv6)", *IETF internet draft*, draft-ietf-mipshop-hmipv6-02 (work in progress), Jun 2004.