Computing and Informatics, Vol. 23, 2004, 363–376

FORWARDING BASED DATA PARALLEL HANDOFF FOR REAL-TIME QOS IN MOBILE IPV6 NETWORKS*

Hyunseung Choo

School of Information and Communication Engineering
Sungkyunkwan University, Korea
e-mail: choo@ece.skku.ac.kr

JungHyun Han

Department of Computer Science and Engineering
Korea University, Seoul, Korea
e-mail: jhan@korea.ac.kr

Manuscript received 23 August 2004; revised 2 Decemer 2004
Communicated by Pavol Návrat

Abstract. Real time mobile applications with guaranteed quality of service (QoS) are expected to be popular due to drastic increase of mobile Internet users. Many Resource ReSerVation Protocol (RSVP) based handover schemes in MIPv4 were studied in the literature for those services. However, the buffering overhead is unavoidable for the latency to optimize the route in new path establishment. Even though the data forwarding based schemes minimize the data loss and provide faster handoff, there are still some overheads when forwarding them and limitation on MIPv4. In this paper we propose a novel handoff scheme in MIPv6 based on forwarding which balances route traffic and reduces the overhead. The comprehensive performance evaluation shows that the disruption time and the signaling overhead are significantly reduced up to about 62 % and 73 %, respectively, in comparison with well-known previous schemes discussed in [1, 2]. Furthermore, it is able to transmit data with the reduced latency and guarantee the fast and secure seamless services.

* Short version of this paper was published in ICCS 2004, Krakow (Poland). This work was supported in parts by BK21 and the Ministry of Information and Communication, Korea.

Keywords: Forwarding scheme, handoff, Mobile IPv6, quality of service (QoS), routing

1 INTRODUCTION

The growth of wired and wireless networking technologies promises a new era of mobile computing. The seamless access to the global information network provides Internet users advanced services whenever and wherever they want. At the same time, the mobile Internet services to support real-time traffic flows with guaranteed quality of service (QoS) are being fuelled by various wireless and portable computing devices such as laptops and personal digital assistants. Therefore, many mobility support protocols including Mobile IP(MIP) are proposed to allow such demands and being under development [3].

Mobile IP [4, 5, 6] has been designed by IETF to serve the increasing needs of mobile users who wish to connect to the Internet and to maintain communications as they move from place to place. In Mobile IPv4 (MIPv4), each MN is always identified by its home address regardless of its current point of attachment to the Internet. While away from its home IP subnet, the MN is also associated with a care of address (CoA), which indicates its current location. Therefore, Mobile IP enables to the mobile node when it is away from its home network. Due to drastic increase of mobile users, it is preferable to move to MIPv6 [7, 8].

RSVP [9, 10] for QoS which enables a destination node to reserve resources along a fixed path to a source node is quite remarkable. RSVP based hand over schemes in MIPv4 to provide faster services with QoS have been studied [11, 12]. However, the buffering overhead is unavoidable during the latency to optimize the route in path establishment on the movement of the mobile node(MN). If the MN moves frequently in a certain area like cell border, *zigzag* makes the significant overhead. Besides, there are some disruption times needed in the new path establishment which is not preferable for the seamless services. There are many researches to reduce the MIP handoff latency in the literature [1, 2, 11, 13].

We briefly discuss low latency handoff scheme with Neighbor Casting [1], the two path handoff scheme [2], and optimized handoff scheme [11] as related works. We observe the data loss caused by the disruption time and the bandwidth overhead due to RSVP channel. Even though the data forwarding based low latency handoff scheme minimizes the data loss and provides the faster handoff between old foreign agent(old FA) and new one(new FA), there are some overheads when forwarding data. It has a limitation on the employment of MIPv4. The motivation of our works is to mitigate the limitation.

In this paper we propose a novel handoff scheme in MIPv6 for the faster and more reliable services. This scheme minimizes the possible latency occurred in the route optimization based on forwarding, load-balances in routers involved, and thus reduces the overhead. This scheme performs the data forwarding and the route optimization in parallel. The performance evaluation shows that the disruption time and the signaling overhead are significantly reduced up to about 62 % and 73 %, respectively, in comparison with well-known previous schemes discussed in [1, 2]. Furthermore, it is able to transmit data with minimum latency and guarantees the fast and secure seamless services.

This paper is organized as follows. In the following section we first discuss previous works. In Section 3, we describe our proposed scheme. Then, in Section 4, we evaluate the performances of the new scheme and related works. We finally conclude this paper along with the future direction in the last section.

2 RELATED WORKS

2.1 The Low Latency Scheme

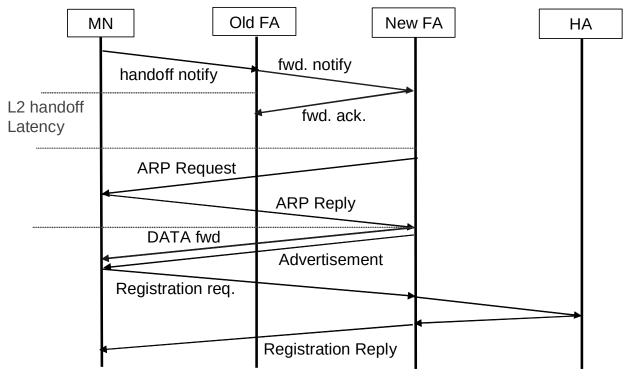

Fig. 1. The low latency scheme message flow

Figure 1 shows that the MN receives the data before the agent advertisement message. To minimize the latency of handoff, the low latency scheme which uses the data forwarding between the new and the old FA is presented in [1]. The low latency scheme is built upon MIPv4 and data forwarding procedure. First, Neighbor FA discovery mechanism means how FAs discover their neighbor FAs dynamically and forward data in mobile networking environment. Each FA knows its neighbor FAs by MNs since the MNs move around and eventually from neighbor FAs. That is, the MN keeps the address of the old FA and transmits this information to the new FA whenever it hands over from a FA to another one. Each FA maintains a table for its neighbor agents. The MN sends the registration message which includes the information for the old FA to the new one. Then the new FA recognizes the old FA as its neighbor agent and sends a neighbor FA Notification Message to the old FA. Therefore, the old FA is informed of the new FA and updates its neighbor FA table. The handoff mechanism is based on the L2 handoff including Neighbor FA discovery. When the MN moves from the old FA to the new one, the old FA is notified from the new FA by Neighbor FA discovery mechanism. If CN sends the data, the old FA can start forwarding the data to the new FA. Therefore, the MN receives the data shortly after L2 handoff without establishing a new path between the old FA and the new one when it hands over.

2.2 The Two Path Scheme

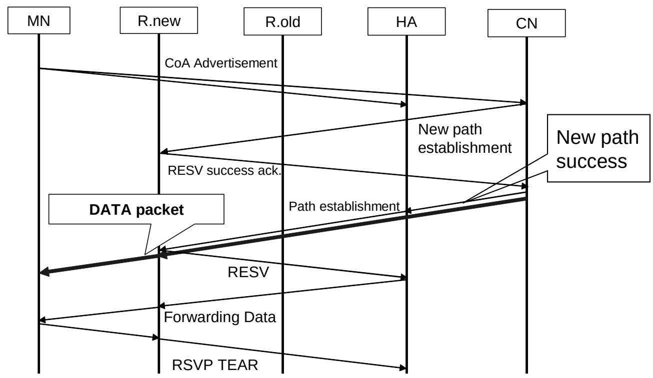

Fig. 2. The two path scheme message flow

In MIPv4, incoming packets from CN are received by the MN through the HA. As a result, it takes too much time to support real-time traffic flows with guaranteed QoS. However, the two path scheme is available to support large-scale mobility with faster handoff. This mechanism requires new procedures and signaling messages as shown in Figure 2. When the MN moves into a new subnetwork, the MN's new CoA is sent to both the HA and the CN. Then, the CN which receives the MN's new CoA sends a PATH message to the MN to establish a new mobile path. Since the mobile path which is a direct RSVP path does not depend on forwarded packets from the HA, it is more efficient; however, incoming packets may be lost or may arrive out of sequence.

Therefore, this scheme alternately uses the home path, i.e., from the CN through the HA to the MN. Since the MN maintains a connection with the HA, it is able to establish Home path whenever needed. When this approach uses the home path, it ensures correct and ordered forwarding of incoming packets, and reduces the packet loss.

2.3 Optimized Smooth Handoff Scheme

In [4], several methods to alleviate triangle routing and to reduce data loss are proposed. First, to reduce data loss during a handoff, it presents a buffering scheme at the FAs and smooth handoff scheme. The FA buffers any data packets forwarded to a MN. When a handoff occurs, the MN includes a handover request in its registration, and the new FA in turn requests the old FA handoff the buffered packets to the new location. To reduce duplicates, the MN buffers the identification and source address fields in the IP headers of packets. It receives and includes them in the buffer handoff request so that the old FA does not need to transmit those packets which have already been received by the MN. Also, Smooth handoff uses binding update to reduce packet loss during a handoff. Including FA buffering mechanism, it decapsulates the tunnelled packets and delivers them directly to the MN.

Second, hierarchical FA management reduces the administrative overhead of frequent local handoffs, using an extension of the Mobile IP registration process so that security can be maintained. The FAs in a domain are organized into a hierarchical tree to handle local movements of the MNs within the domain. Meanwhile, when a packet for the MN arrives at its home network, the HA tunnels it to the root of the FA hierarchy. When an FA receives such a tunnelled packet, it re-tunnels the packet to its next lower-level FA. Finally the lowest-level FA delivers it directly to the MN. When a handoff occurs, the MN compares the new vector of CoA with the old one. It chooses the lowest-level FA that appears in both vectors, and sends a Regional Registration Request to that FA. Any higher-level agent need not be informed of this movement since the other end of it forwarding tunnel still points to the current location of the MN.

3 THE PROPOSED SCHEME

3.1 The Operational Mechanism

In the new forwarding based data parallel handoff technique, the CN has the option of using two different paths to reach the MN. The first path is called the R.new path, which is a direct path to any new location of the MN. Since the R.new path does not depend on the packet from HA, it is the more efficient and preferred path to use. However, while the MN is moving to a new location, incoming packet may be lost. Therefore we have to maintain the R.old path, i.e. from the CN to the R.old to the MN, to ensure uniform and fast forwarding of incoming packets, and to reduce the packet latency and the disruption.

The new technique requires a new procedure and signaling messages to coordinate the use of the paths. These are illustrated in Figures 3, 4 and 5. We describe each operation along with the figure as shown below:

1. **Figure 3 (1-1, 1-2):** The MN located at R.old is communicating through an existing session with the CN. When the MN moves into a new subnetwork of R.new, the MN's new CoA, the handoff notification, is sent to both the R.new and the R.old.
2. **Figure 3 (2-1, 2-2, 2-3):** When the R.old receives the MN's new CoA, it begins to forward the data packet to the R.new. R.old's new path request message

is sent to the CN. When R.new receives the MN's new CoA, R.new sends the CoA to the HA to register MN's new location (binding update).

3. **Figure 4 (3):** When the CN receives the R.old's *new path request*, the CN sends a *new path establishment* message to the R.new to establish a new optimized path.
4. **Figure 4 (4):** If the R.new path attempt is successful, the R.new sends *new path acknowledgement* message to R.old.
5. **Figure 4 (5):** When the R.old receives acknowledgement message from R.new, R.old sends *the request for tear down of the old path* to CN.
6. **Figure 4 (6):** When this happens, the CN stops sending data packets to the R.old and R.old explicitly tears down the forwarding path to R.new after it sends the final data packet to R.new.
7. **Figure 4 (7):** When the CN switches to the R.new path from the R.old path, the new path can be used. The R.new buffers incoming packets for the MN from the CN until the forwarding path is destroyed.

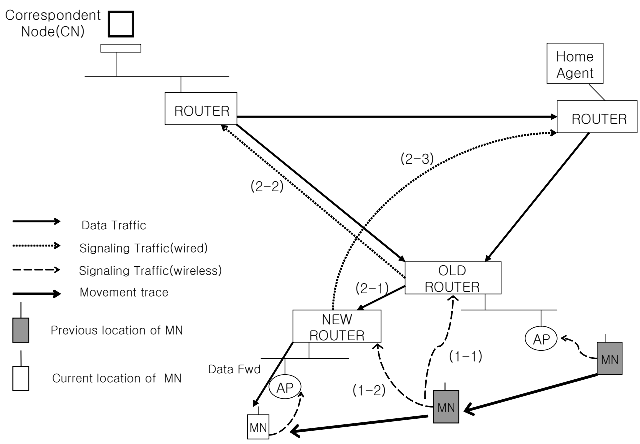

Fig. 3. Before route establishment

3.2 Characteristics of the New Scheme

In the design of this scheme, we applied the following approaches.

- Mobile IPv6 compatibility: Mobile IPv6 is the proposed by IETF, and this scheme is designed to be compatible with it.
- Resource preallocation: the handoff process requests a resource preallocation to guarantee that MNs obtain the requested resources right away after they enter the new cell/network.

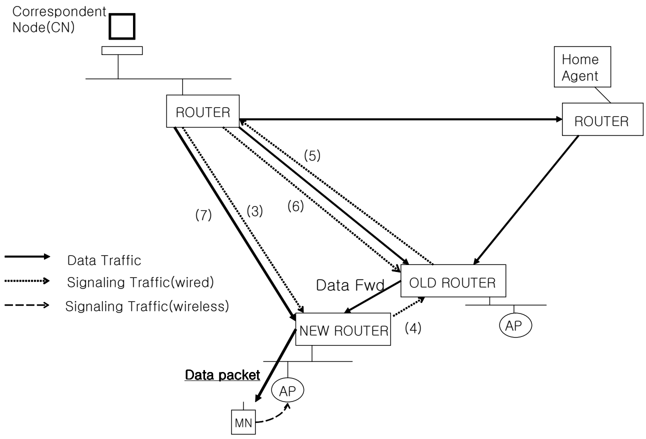

Fig. 4. After route establishment

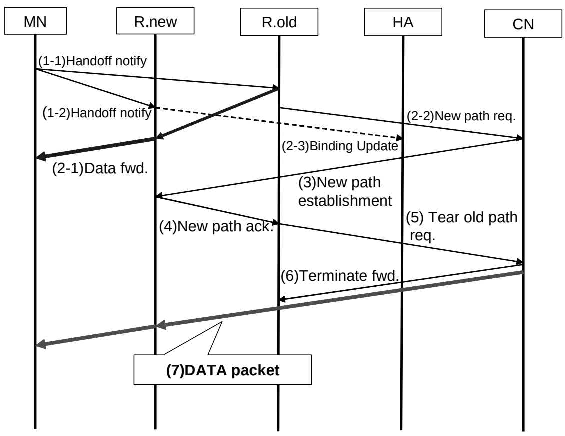

Fig. 5. Message flow of the proposed scheme

- Buffering: To avoid data loss and unordered packets, the R.new is instructed to buffer packets for the incoming MN.
- Forwarding to an adjacent router: When a MN handoffs to adjacent router, its packets can be forwarded to the R.new by the R.old.
- Less location update only with HA: It is a frequent case that a MN moves to another networks or cells. This scheme does not make frequent HA updates when a handoff is performed. It contacts the HA for the update of the information for the care-of address.

- Triangular message exchange: It has to exchange messages for the smooth hand-off. When a router sends a message to next router, it utilizes a next path resource to complete the message exchange.

This scheme introduces the disruption time while the R.old, R.new, and the CN exchange messages. Such problems are ignored for wired networks since wired links are much more reliable and the disruption time is short in duration. In wireless network, we assume that routers have MN's location table. They perform fast forwarding to the MN by sending *MN's location* packets to each other. In the next section, we describe a framework to measure the disruption time and total overhead use for the new technique.

4 PERFORMANCE EVALUATION

4.1 The Signaling Time

The time to send a message over one hop, M, can be calculated as

$$M = \alpha + \beta + \gamma \tag{1}$$

where α is the transmission time, β is the propagation time, and γ is the processing time. The transmission time α, is computed by b/B, where b is the size of the control message in bits, and B is the bit rate of the link on which the message is sent. B has two kinds of values which are for wired and wireless links. The propagation time, β, has two kinds, too. Suppose that the processing time, γ, always has a fixed value when the message is sent. The signaling time which is the total time needed to transmit and process the messages between router i and j is computed by

$$T_{i,j} = T_{sig} \times D_{i,j} \tag{2}$$

where T_{sig} is the signaling time over one hop and $D_{i,j}$ is the node distance between router i and j. T_{sig} is computed as follows. Wired link ($T_{RSVP} = M \times \frac{1+p}{1-p}$) and wireless link($T_{RSVP} = M \times \frac{1+q}{1-q}$) where q is the probability of wireless link failure. As mentioned, $D_{i,j}$ is the number of hops and computed by

$$D_{i,j} = N_{R_i,R_j} = N_{R_i,R_{join}} + N_{R_{join},R_j}. \tag{3}$$

N_{R_i,R_j} is the number of hops from R_i to R_j and R_{join} is the router placed at MAP (merged access point). The number of hops over routers on which the signaling passes are calculated by summing these two values. The RSVP signaling time is computed by

$$T_{i,j} = T_{RSVP} \times D_{i,j}, \tag{4}$$

where T_{RSVP} is the time to perform RSVP signaling over one hop. Wired link ($T_{RSVP} = M \times \frac{1+p}{1-p}$) and wireless link ($T_{RSVP} = M \times \frac{1+q}{1-q}$), here p is the probability of resource assignment failure on RSVP. When the resource failure occurs, the RSVP signaling is assumed to be repeated until either a session is successfully established, or the MN moves to a new router location.

4.2 The Disruption Time and Total Signaling Cost

The disruption time is to measure the time that a node cannot get packets. The shorter this time, the better the performance. The proposed scheme does not get data when the MN receives forwarded data through New.FA from Old.FA while data forwarding occurs. This time is computed by the equation (1). Namely, if the disruption time is short, we can say that data flow is continuous and performance is good.

Signaling overhead is due to sending and receiving signaling from neighbor FA and increases in proportion to the signaling sending time. It is proportional to generated signal in one cell. Then if signaling overhead is larger, the resource use rate becomes larger. Accordingly, small signaling overhead is more efficient. Consider the case that each signal source generates UDP packets at constant signal rate c continuously. Then the amount of generated signal during a handoff of mobile node is $n{\cdot}\tau{\cdot}c\ \ bytes$. n is the number of neighboring FAs which MN sends signal. (Suppose that each cell is in hexagonal form.)

In our paper signal is transferred by one cell which MN moves; then n is 1. τ is the total signaling period during which all signaling and forwarding processes finish for receiving datas by the new path over one handoff process. The time which takes the parallel process is included in total signaling period. This is because we need to calculate all signaling time for the entire signaling process. The total amount of generated signal from a cell for a time unit period is $n \cdot \tau \cdot c \cdot \lambda$ where λ is the handoff rate. The total amount of signal sent by the sources to the MNs in the cell for the same time unit period is $m \cdot c$ where m is the total number of active MNs in the cell. The handoff rate, λ, is computed by

$$\lambda = \frac{\rho v L}{\pi} \tag{5}$$

where λ is the handoff rate or cell boundary crossing rate (1/sec), ρ is the active MN density in 1 m^2, v is the MN moving speed (m/sec), and L is the cell perimeter (m). The signaling overhead ratio is defined as the number of bytes generated from a cell divided by total number of bytes sent by the source, i.e. $\xi = \frac{n\tau c\lambda}{mc} = \frac{n\lambda\tau}{m} = \frac{\rho v L n\tau}{\pi m}$. Denoting r as the cell radius for a hexagonal cell scenario and $L = 6r$, ρ is computed by

$$\rho = \frac{m}{cellarea} = \frac{m}{\frac{3\sqrt{3}}{2}r^2}. \tag{6}$$

λ is computed based on previous two equations (λ, ρ), and we get the following: $\lambda = \frac{\rho v L}{\pi} = \frac{4mv}{\sqrt{3}r\pi}$. We can substitute this values to ξ, and then get the result below:

$$\xi = \frac{4vm\tau}{\sqrt{3}r\pi} \tag{7}$$

where τ is the total signaling period. This is the sum of the time taken for signaling processes including parallel signaling process based on Figure 5. $\tau_1 = time\_handoff\_notify(step1 - 1) + time\_path\_req(step2 - 2) + time\_path\_establish(step3) + time\_path\_ack(step4) + time\_stop\_transfer(step5) + time\_stop\_fwd(step6)$ where we disregard handoff latency for handoff process. τ_1 is the time which takes fundamental signaling process. As mentioned above, the MN notifies handoff to R.new, then R.new begins binding updates. These processes are performed in parallel through signals during the handoff. So it must be contained in signaling time. $\tau_2 = time\_handoff\_notify(step1 - 2) + time\_binding\_update(step2 - 3)$. Total signaling time is the sum of the above two time periods.

$$\tau = \tau_1 + \tau_2 \tag{8}$$

This is computed based on equation (1) in Section 4.1. We use 200m-radius hexagonal cell topology and assume that the mobile speed is fixed at 5 m/sec.

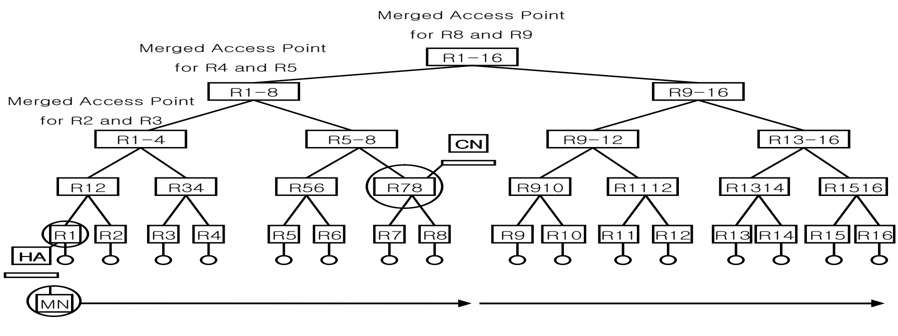

Fig. 6. Hierarchical router configuration

4.3 Numerical Results

Table 1 shows the system parameters [14, 15] that were used to compare the proposed scheme and the previous scheme [1, 2].

The Disruption Time: Figure 7 shows the comparison of the proposed scheme with the related ones [1, 2] with respect to the disruption time. The results are computed based on the first two equations($T_{i,j}$). As shown in Figure 7, disruption time for the two path scheme is greater than other schemes applying

SystemParameters	*Values*
Wireless propagation time, $\beta 1$	2 msec
Wired propagation time, $\beta 2$	0.5 msec
Signaling processing time, γ	0.05 msec
RSVP processing time, γ_{RSVP}	0.5 msec
Packet size, S	50 Bytes
Probability of link failure, q	0.5
Probability of resource denial, p	0.5
Wireless bit rate, B1	144 Kbps
Wired bit rate, B2	155 Mbps

Table 1. Parameters

forwarding. This is due to the character of forwarding method which allows to get data while the new path is established. The proposed scheme has better performance than the low latency scheme using forwarding. For each MN location, the average disruption time for the proposed scheme is 53 % and smaller by 3 % than the two path scheme and the the low latency scheme, respectively.

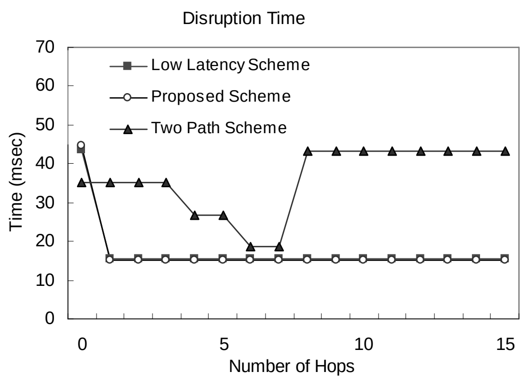

Fig. 7. The disruption time

The Total Signaling Overhead: Figure 8 is the graph which shows the comparison of the proposed scheme with [1, 2] with respect to the total signaling overhead. As shown in Figure 8, the proposed scheme has the best performance compared with the previous schemes. The signaling overhead of the low latency scheme based on MIPv4 is computed only for the signaling related to the forwarding, hence it has a fixed value. However the proposed scheme and the two path scheme are affected by the time for path establishment from CN to R.New, so signaling traffic is changed according to the model. Therefore results of two schemes have similar trends. In Figure 8, average total signaling overhead for the proposed scheme is 48 % and smaller by 59 % than the two path scheme and the low latency scheme, respectively.

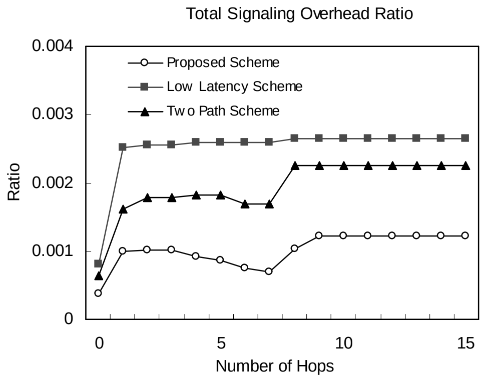

Fig. 8. Total signaling overhead

5 CONCLUSION

In this paper, we propose the new scheme with guaranteed QoS whose data transmission is faster and data loss is smaller than the two path scheme based on MIPv6 and the data forwarding scheme based on MIPv4. The proposed scheme based on MIPv6 has the merits of the two related schemes. Namely it receives data fast using data forwarding, and makes the new path from CN to R.new while forwarding data. Besides, R.old is charged with new path request message from CN to R.new instead of R.new, and one step in path setup is saved compared to the low latency scheme. The proposed scheme is efficient in terms of the disruption time and total signaling overhead as discussed in performance evaluation and offers services of more improved quality in mobile communication environment where the real-time transmission is very important. We are investigating other aspects of the proposed scheme.

REFERENCES

[1] Shin, E.—Wei, H.-Y.—Chang, Y.—Grtlin, R. D.: Low Latency Handoff for Wireless IP QOS with Neighbor Casting. IEEE ICC, Vol. 5, 2002, pp. 3245–3249.

[2] McNair, J.—Akyldiz, I. F.—Bender, M. D.: Handoffs for Real-Time Traffic in Mobile IP Version 6 Networks. IEEE GLOBECOM, Vol. 6, 2001, pp. 3463–3467.

[3] Akyldiz, I. F.—McNair, J.—Ho, J. S. M.—Uzunalioglu, H.—Wang, W.: Mobility Management in Next Generation Wireless Systems. IEEE, Vol. 87, 1999, No. 8, pp. 1347–1384.

[4] Perkins, C.: Mobile IP. IEEE Communications Magazine, Vol. 35, 1997, No. 5, pp. 84–99.

[5] Perkins, C. E.: IP Mobility Support. IETF RFC 2002.

[6] Perkins, C. E.: Mobile Networking in the Internet. ACM/Baltzer Mobile Networks and Applications (MONET) Journal, Vol. 3, 1998, pp. 319–334.

[7] David, B.—Perkins, C.—Arkko, J.: Mobility Support in IPv6. IETF draft, Internet Draft `draft-ietf-mobileip-ipv6-17.txt`, May 2002, work in progress.

[8] Okajima, I.—Umeda, N.—Yamao, Y.: Architecture and Mobile IPv6 Extensions Supporting Mobile Networks in Mobile Communications. IEEE VTC, Vol. 4, 2001, pp. 2533–2537.

[9] Braden, R.—Zhang, L.: Resource Reservation Protocol (RSVP) – Version 1 Message Processing Rules. RFC 2209, September 1997.

[10] Chen, W.-T.—Huang, L.-C.: RSVP Mobility Support: A Signaling Protocol for Integrated Services Internet with Mobile Hosts. IEEE INFOCOM, Vol. 3, pp. 1283–1292, 2000.

[11] Perkins, C. E.—Wang, K.-Y.: Optimized Smooth Handoffs in Mobile IP. IEEE ISCC, pp. 340–346, 1999.

[12] Terzis, A.—Srivastava, M.—Zhang, L.: A Simple QoS Signaling Protocol for Mobile Hosts in the Integrated Services Internet. INFOCOMM, May 1999.

[13] Talukdar, A. K.—Badrinath, B. R.—Acharya, A.: On Accommodating Mobile Hosts in an Integrated Services Packet Network. IEEE INFOCOM, Vol. 3, pp. 1046–1053, 1997.

[14] McNair, J.—Akyldiz, I. F.—Bender, M. D.: An Inter-System Handoff Technique for the IMT-2000 System. IEEE INFOCOM, Vol. 1, pp. 208–216, 2000.

[15] Xie, J.—Akyldiz, I. F.: An Optimal Location Management Scheme for Minimizing Signaling Cost in Mobile IP," IEEE ICC, Vol. 5, pp. 3313–3317, 2002.

Hyunseung Choo has received BS in mathematics from Sungkyunkwan University, Korea in 1988, MS in computer science from the University of Texas at Dallas, USA in 1990, and PhD in computer science from the University of Texas at Arlington (UTA), USA in 1996. From 1997 to 1998, he was a Patent Examiner at Korean Industrial Property Office. He is currently Associate Professor of the School of Information and Communication Engineering at Sungkyunkwan University. His research interests include switching networks, mobile computing, grid computing, and optical networking. He has published over 80 papers in international journals and refereed conferences. Dr. Choo is a member of IEEE.

JungHyun Han is an associate professor of the Department of Computer Science and Engineering at Korea University, where he directs the Interactive Media Laboratory. Prior to joining Korea University, he worked at the School of Information and Communications Engineering of Sungkyunkwan University, and at the Manufacturing Systems Integration Division of the US Department of Commerce National Institute of Standards and Technology (NIST). Dr. Han received a BS in computer engineering at Seoul National University, an MS degree in computer science at the University of Cincinnati, and a PhD degree in computer science at the University of Southern California. His research interests include mobile entertainment, and 3D game design and development.