Mobile IP route optimization method for a carrier-scale IP network

Takeshi Ihara, Hiroyuki Ohnishi, and Yasushi Takagi
NTT Network Service Systems Laboratories
3-9-11 Midori-cho, Musashino-shi, Tokyo 180-8585, Japan
Phone: +81-422-59-3458   Fax: +81-422-59-3494
E-mail: ihara.takeshi@lab.ntt.co.jp

Abstract

The current Mobile IP specification forces all packets forwarded to a mobile terminal (MT) to be routed via that terminal's Home Agent (HA), which often leads to triangular routing, which in turn causes data transmission delay and wastes network resources. IETF (Internet Engineering Task Force) is studying a solution that can optimize the route length, but this solution requires the implementation of additional functions on all correspondent nodes, so it will not be realistic.

In this paper, we propose the use of Mobile IP border gateways (MBGs) to optimize the routing in a carrier-scale network and reduce the processing burden on the HA without adding more functions to the correspondent nodes, and show the effect of reducing the route length and the HA's load. We also propose a hierarchical arrangement of Foreign Agents (FAs) to reduce the processing burden on MBGs

1. Introduction

In next-generation mobile data networks, mobility management that enables mobility at the IP level will be needed to improve terminal mobility, provide seamless roaming between fixed and mobile networks, and allow new IP applications to be developed.

Mobile IP [1], which is the most promising technology studied by IETF, can provide IP level mobility. Moreover, 3GPP (Third Generation Partnership Project) [2] and 3GPP2 [3], which are aimed at establishing regional standards, are beginning to study the application of Mobile IP to the IMT-2000 core network as shown in Figure 1 [4].

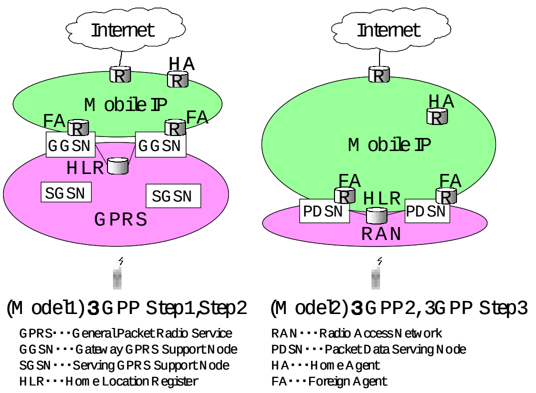

Figure 1. Next-generation mobile data network applying Mobile IP

However, Mobile IP has not been applied to any carrier-scale network, and the triangular routing in Mobile IP causes an additional load on the Home Agent (HA), wastes network resources, and increases the delay time.

IETF is studying a way of optimizing the triangular path and there have been several reports on route optimization by the IETF method [5]. However, this method requires that specific functions be implemented on all correspondent nodes, such as thousands of WWW servers in the Internet, so it is not easy to achieve route optimization by this method.

In this paper, we propose a method for a large network that enables route optimization without modifying existing WWW servers [6].

This paper is organized as follows. Section 2 overviews the current Mobile IP protocol and its problem concerning the route optimization procedure. Section 3 describes our route optimization method which uses Mobile IP border gateways (MBGs). Section 4 evaluates the effect of route optimization and the reduction of the HA load achieved by using MBGs. Section 5 calculates the burden on an MBG when it is applied to a commercial Mobile IP network and a hierarchical architecture is proposed to reduce that burden. Section 6 concludes this paper and discusses technical issues in future.

2. Overview of Mobile IP

2.1 Basic procedure

Figure 2 (a) shows the location registration procedure in a Mobile IP network.

(1) A Mobile Terminal (MT) moves from its home network to another network (R1).

(2) The FA belonging to the visited network assigns the care-of address to the MT (R2).

(3) The MT reports its care-of address to its HA (R3).

Figure 2 (b) shows the IP packet uplink/downlink forwarding procedure in the Mobile IP network.

・Uplink forwarding (from MT to correspondent node)

IP packets are forwarded directly to the correspondent node (U1).

・Downlink forwarding (from correspondent node to MT)

(1) IP packets are transferred to the home address of the MT (D1).

(2) The HA intercepts the packets destined for the MT that is away from the home network and tunnels them to the MT's care-of address (D2).

(3) Packets are detunneled and delivered to the MT (D3).

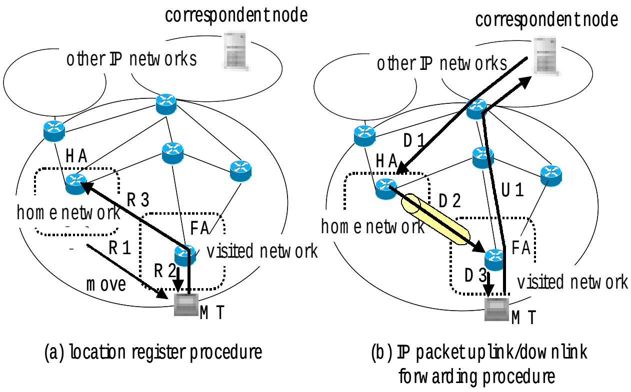

Figure 2. Mobile IP procedure

2.2 Route optimization proposed by IETF

The basic procedure illustrated in Figure 2 forces all downlink forwarded packets to be routed via the MT's HA. In other words, such packets cannot be forwarded to the MT directly; instead, a triangular route path is generated. This increases the concentrated load on the HA, wastes network resources, and lengthens the delay time.

To solve this problem, IETF has proposed an extension, shown in Figure 3, called "Route Optimization in Mobile IP"

(1) After receiving several packets destined for the MT, the HA performs a tunneling operation to the associated care-of address, and also indicates the MT's care-of address to the correspondent node when the terminal is away from its home network (K1).

(2) The correspondent node can subsequently send the traffic intended for the MT directly to its care-of address by means of a tunneling mechanism (K2).

This extension optimizes the routing, but requires that specific functions be implemented on all the correspondent nodes, such as WWW servers in the Internet. It is not practical to take such implementation within the next two or three years.

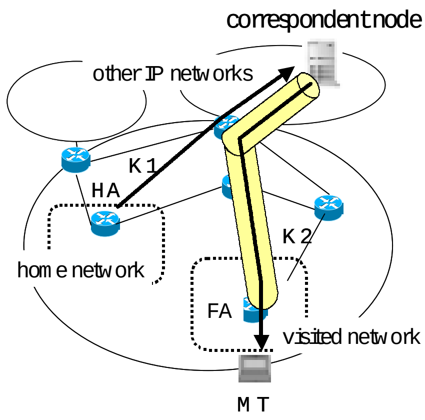

Figure 3. IETF route optimization in Mobile IP

3. Route optimization using MBGs

We propose a route optimization method that enables a Mobile IP network to optimize the downstream route by using its Mobile IP Border Gateway (MBG) routers as shown in Figure 4. In this method, HAs notify the MBG of the care-of address and home address of each MT after a few packets are transferred through the HAs, then the MBG keeps the binding information as these two addresses. Subsequent packets arriving from other networks can be routed directly to the care-of address according to the binding information. This method enables all the packets arriving from other networks to be routed directly to the MT.

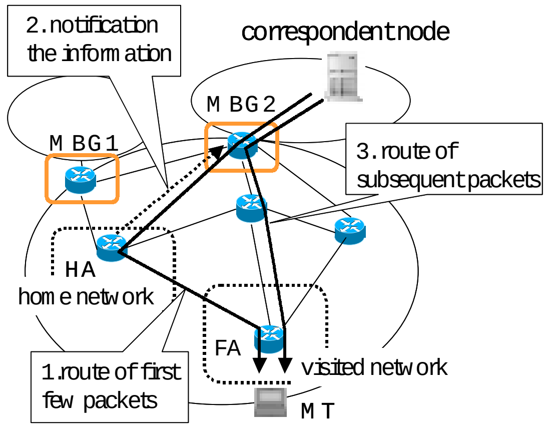

Figure 4. Mobile IP border gateway (MBG)

For the packets arriving from the MT within the Mobile IP network, the existing IETF method described in Section 2 can be applied because MTs within the networks have the required Mobile IP functions. Usually IP networks have few border gateway routers (BG), so the number of MBGs that need to have these specific functions can be limited. In addition, most MTs communicate with WWW servers in other networks more frequently, so it is very effective to optimize the downstream arriving from the other networks.

An HA located at a border gateway (BG) can also forward the packets to the MT directly (Figure 5). This model is called the HA/BG model in this paper. However, the MBG model is better than the HA/BG model in two points.

First, the HA/BG model can't optimize the route when a Mobile IP network connects to several other IP networks. Because the HA/BG node doesn't know the location of an MT that doesn't belong to the HA, the packets are not forwarded to the MT directly. On the other hand, an MBG can forward packets to the MT's location because each MBG has location information concerning the MT.

Second, the MBG model (Figure 6) enables route optimization in the Mobile IP network and another IP network when the other IP network can select the nearest BG from the correspondent node.

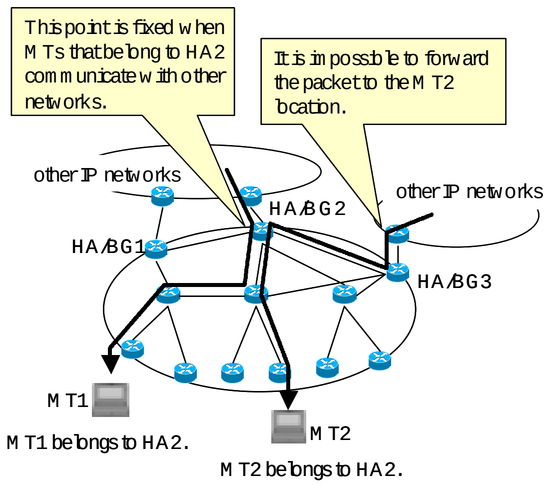

Figure 5. HA/BG model

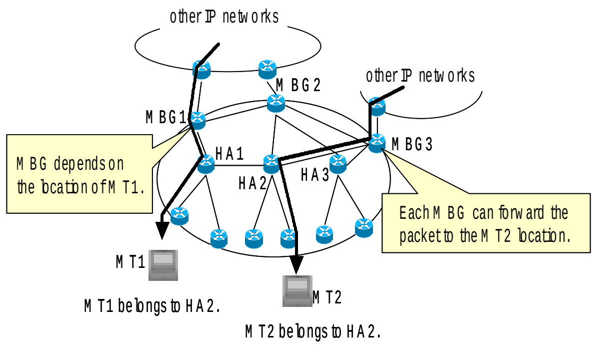

Figure 6. MBG model

3.1 Procedure outline

Figure 7 shows the detailed route optimization procedures.

a) First few packets sent to MT
After sending the first few packets to the MT, MBG2 sends a message to the HA. The message includes the relationship between the MBG ID number and the packet's source IP address to inform the HA of which MBG handled the packet (M1).
b) Subsequent packets
(1) The HA informs MBG2 of the MT's care-of address (M2).
(2) MBG2 can subsequently send packets intended for the MT directly to its care-of address by means of a tunneling mechanism (M3).

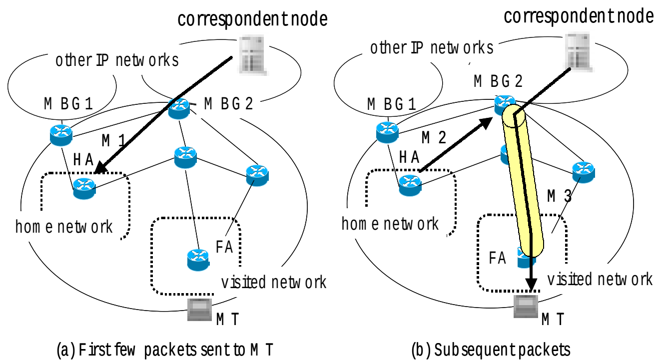

Figure 7. Route optimization using MBGs

4. Evaluation

The MBG method can reduce the route length and the HA load. In this section, we evaluate its effects by applying MBG.

4.1 Efficiency of route length reduction

We evaluated the effects of route optimization using the MBG method from the viewpoints of network topology, network scale, and the HA and FA arrangement schemes. We use mean route length as the indication for the evaluation of the proposed method. In calculating the route length, we consider the calculation of shortest-path used in OSPF, which considers the bandwidth and link length in addition to the number of hops. In this paper, we want to do the evaluation on a simple condition as the first step. If we consider the bandwidth, use's traffic types (e.g., bandwidth, call attempt rate) should be variables and the evaluation might be very complicated. Therefore, we consider only the length of the link for the simplicity.

4.1.1 Network model

We considered the two types of three-layer hierarchical network model shown in figure 8. In Model 1, the first and second layers both had a ring topology. In Model 2, the first layer had a ring topology and the second layer had a star topology. In both models, third layer had a star topology. The link lengths of the first, second, and third layers were a, b and c respectively.

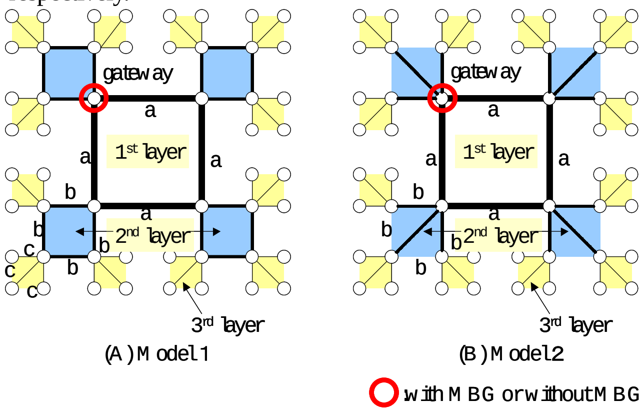

Figure 8. Hierarchical network model

We examined two arrangements for each model to determine the better HA and FA arrangement .
Case 1: The HAs are located in the first layer and the FAs are located in the third layer.
Case 2: The HAs are located in the second layer and the FAs are located in the third layer.

4.1.2 Calculation and Results

Figure 9 shows the relationship between the ratio (R) between the scale of the first and second of layers and the average route length,
where R=b/a, a is the network scale of the first layer, b is that of the second layer (0.1a to a), and the network scale of the third layer is assumed to be c = 0.1a.
Other assumptions made for this evaluation were as follows:
- the packet is transferred so as to minimize the route length in the ring topology network.
- the distance between each layer is 0.
Under these assumptions, the average route length (L) is calculated as

$$L_{\text{without MBG}} = \sum_{i=1}^{n} \sum_{j=1}^{m} \{\ \text{min(route length from the gateway node to the HAi)} + \text{min(route length from HAi to FAj)}\ \}/(n \times m), \quad (3)$$

$$L_{\text{with MBG}} = \text{min(route length from the gateway node to the FAj)}\ /m, \quad (4)$$

where HAi is the HA to which the MT belongs, FAj is the FA visited by the MT, n is the number of HAs, and m is the number of FAs.

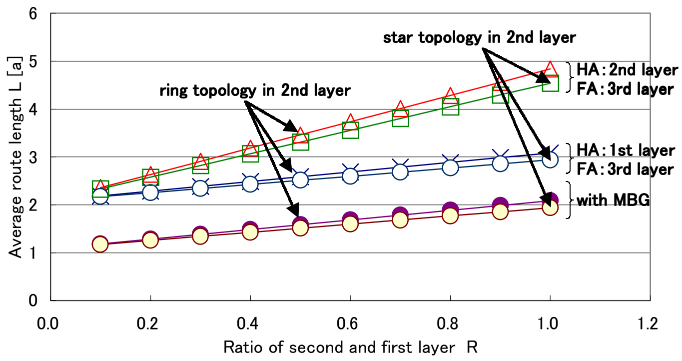

Figure 9. Ratio (R) between the scale of layers 1 and 2 and average route length

The route length in case 1 is shorter than that in case 2 when MBGs are not used. On the other hand, when MBGs are used, the route is the same for both models. So if MBGs are used, the case 2 arrangement will reduce the burden on the HA without lengthening the route.

If we show this result using the more understandable indication, we can calculate the link cost by a*(link cost per unit length).

4.2 Efficiency of HA load reduction

In the early stage of the introduction of Mobile IP, the processing capacity of HA may be not so high, and multiple HAs should be deployed in the network. On such condition, many triangular route would be generated, and the load of HA would be high.

Using MBGs can reduce the load of HA by providing a direct route to the MT. In this section, we evaluate the relationship between the number of HAs and effectiveness of the load reduction of HAs by introducing MBG.

4.2.1 Network model

We assumed two-layer hierarchical networks model as shown in figure 10.

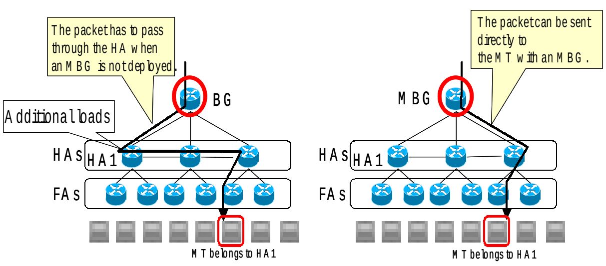

Figure 10. Network models for evaluation of load reduction

4.2.2 Calculation and Results

Figure 11 shows the relationship between the number of HAs and the normalized HA load (HA load with an MBG / HA load without an MBG).

Assumptions

-Each FA belongs to only one HA.

-An FA has only one link.

-All HAs connect to each other.

Under these assumptions the HA load is calculated as,

Load of HAi without MBG=(traffic generated by users within the HAi area) + (traffic generated by users who belong to the HAi but are not in the home area),

Load of HAi with MBG=(traffic generated by users within the HAi area),

where the total area of the network is 620 km2, the number of terminals is 300,000, the MT velocity is 10 km/h, and the ratio of traffic to and from the external network to all traffic is 0.7.

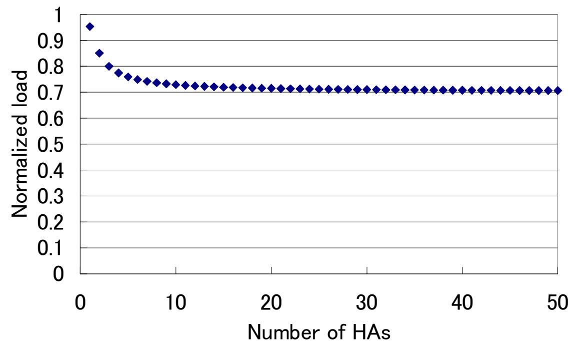

Figure 11. Normalized load

The normalized load decreased as the number of HAs increased when the number of HAs was small, but reached a constant value when the number of HAs was large because most of the users would not be in the home area.

5. Reducing the number of registrations to the MBGs

All MTs communicating with other networks register their location information with an MBG, but the number of registrations should be reduced to decrease the burden on the MBGs.

The hierarchical FA architecture can solve this problem if we add Gateway Foreign Agents (GFAs). Figure 12 shows a hierarchical FA architecture with GFAs . In this architecture, an MT registers only when it moves to a new GFA area. This reduces the number of registrations to the MBGs.

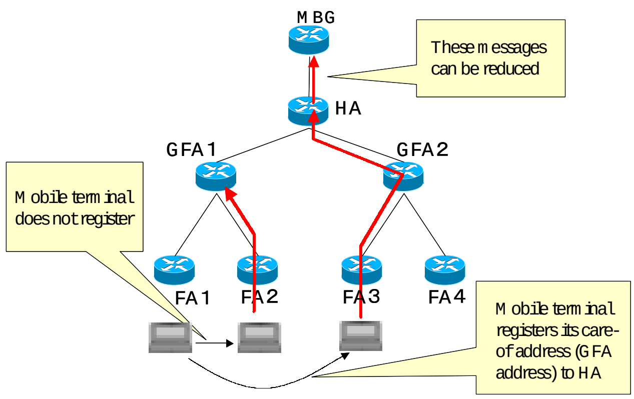

Figure12. Hierarchical FA architecture

By applying the hierarchical architecture with GFAs, the area of coverage of FA can become virtually wider. The number of registrations to one HA is inversely proportional to squared root of the number of FAs connected to one GFA. On the other hand, GFA should terminate the registration request message and forward it if necessary. Therefore, the processing load on GFA could be more than that on transit router, which performs only

the relay of the messages. From the viewpoint of the processing load on the total network, the effectiveness by introducing the GFA is trade off between the additional load of GFA and the reduction of HA, however, considering that the concentration of load on the HA is the main technical issue for the application of Mobile IP for a carrier-scale network, the hierarchical architecture with GFAs is effective solution because it can perform the load distribution.

6. Conclusion

We have proposed a route optimization mechanism featuring Mobile IP border gateways (MBGs).

Using the MBGs can reduce the route length and the HA load regardless of the HA's location. The number of registrations can be reduced by enlarging the area covered by an FA and applying a hierarchical FA architecture.

In future, the performance of the Mobile IP router and the scalability of a network under actual geographical and traffic conditions, based on a commercial mobile network, should be evaluated. The MBG procedure should also needs to be investigated in detail.

References

[1] C. Perkins, "IETF RFC 2002, IP Mobility Support", 1996.

[2] 3GPP home page (http://www.3gpp.org/)

[3] 3GPP2 home page (http://www.3gpp2.org/)

[4] "3G TR 23.923 v1.2.0, Combined GSM and Mobile IP Mobility Handling in UMTS IP CN"

[5] C. Perkins and D. B Johnson, "Route Optimization in Mobile IP" (IETF internet-draft)

[6] H. Ohnishi, T. Ihara, and Y. Takagi, "Proposed scheme for route optimization on the Mobile IP network", IEICE Society Conference, B-7-166, pp.259, 2000. (in Japanese)

[7] W. L. Tan, T. Chah and M. C. Chuah, "Mobile IP Hierarchical Mobility Management", APSITT'99, 1999.