(3) IP² Mobility Management

Masami Yabusaki, Shoichi Hirata,
Takeshi Ihara and Hiroshi Kawakami

The aim of mobility management in IP-based IMT network Platform (IP²) is to improve the network's performance and the network's transmission quality. It addresses the network's performance by taking advantage of mobility characteristics, choosing optimal paths, efficiently using network resources, etc., and tackles the network's transmission quality by realizing lossless, low-latency transport. This article also reviews the elemental technologies currently subject to research in concrete terms.

1. Introduction

In the phase "Beyond IMT-2000", which is subsequent to 3rd Generation (3G) mobile communications referring to as International Mobile Telecommunications-2000 (IMT-2000), most traffic is expected to be based on Internet Protocol (IP), and a network will be required for efficiently transferring IP traffic. Network Laboratories advocates the IP-based IMT network Platform (IP²) as a network architecture meeting such requirements [1].

This article reviews the objectives of the mobility management technologies required for implementing IP², and the elemental technologies for its implementation.

2. IP² Mobility Management Technologies Objectives

The IP² mobility management technologies primarily aim to improve performance and transmission quality by reducing the redundancies and waste in the network itself, in order to provide advanced high quality services to users at affordable prices. Improvements in the network's performance is addressed by taking advantage of mobility characteristics, choosing optimal paths, efficiently using network resources, etc., and the network's transmission quality is tackled by realizing lossless, low-latency transport, as described below in concrete terms.

2.1 Network Performance Improvements

(1) Taking Advantage of Mobility Characteristics

Conventional mobile communication systems applied uniform control methods with respect to all mobile terminals, for location control and handover control of mobile terminals. In future mobile communication services, as the term "ubiquitous" signifies, anything and everything might become a user of mobile communication services —not only human beings and means of transportation but also pets and mail. Accordingly, the number of mobile terminals is expected to explosively increase. It will be necessary to execute mobility control adapted to mobile terminals on an individual basis, which are wide in variety, as it will be inefficient to apply the conventional, uniform control methods to these mobile terminal for the reasons explained below.

The ways in which mobility characteristics are harnessed are as follows, with reference to location control as an example. Location control is concerned with controlling an inactive mobile terminal in standby mode, which consists of location registration and paging technologies. Location control in conventional mobile communication systems fixed the size and shape of the location registration area with respect to all mobile terminals. While this made it possible to simplify network control, it imposed a heavier processing burden on mobile terminals and network control depending on the mobility characteristics and the call termination frequency due to increased location registration traffic, paging traffic, etc., resulting in poorer service performance.

IP^2 aims to support mobility control adapted to the mobility characteristics and the communication characteristics of mobile terminals on an individual basis. Moreover, IP^2 is expected to lead to the new services creation using location information through the management of mobility characteristics of mobile terminals on an individual basis.

(2) Route Optimization

In IMT-2000 packet-switched networks, the handover control method is concerned with fixing the anchor point (i.e. route-switching point). Although this method can avert the loss of packets, it becomes incapable of presenting the shortest route as the mobile terminal moves away from the anchor point due to the generation of a redundant route commonly referred to as the trombone path. Consequently, it leads to problems such as transmission delays and the inefficient use of network resources.

IP^2 aims to implement routing control that is capable of continually providing the optimal paths even if the mobile terminal is repeatedly subject to handover.

(3) Sharing of Network Resources

One of the most important services in future mobile communication will be information delivery to a large number of specific mobile terminals such as the information distribution on traffic jams to car navigation systems, downloads for upgrading mobile-terminal software, and notices to users about delays in public transport. The information delivery services over conventional mobile communication networks have been implemented by setting the number of point-to-point connections the same as the number of receivers. For example, if the information had to be distributed to 10,000 people, the distributing server was required to set 10,000 connections for the receivers. As for IP technologies, multicast techniques have been examined to suppress such wasteful use of network resources. IP multicast is a technique in which the router copies and distributes data only in the direction required by using a specific address. In IP^2, we are examining a method of implementing distribution services through the efficient use of network resources based on IP multicast technology. In mobile communication networks, the sender and the receiver of the delivered information may move, and the movement may frequently give rise to packet loss and duplication. The crucial challenge in mobile communications is to implement multicast services that optimize the efficiency of network resource utilization considering of these conditions.

2.2 Transmission Quality Improvements over Network

IP^2 aims to assure a quality level over the network that does not affect services in which real-time properties are required, such as Voice over IP (VoIP) and streaming services, by minimizing the transmission loss and transmission delays.

3. Elemental Technologies for Mobility Management

3.1 Technologies harnessing Mobility Characteristics

(1) Adaptive Location Control

Conventional mobile communication systems used both location registration and paging on the grounds that the total control load can be minimized by combining them. In general, there is a tradeoff between the number of control signals of location registration and paging, and an optimal size of the location registration area does exist, at which the total number of control signals of location registration and paging is minimized depending on the mobility characteristics of the mobile terminal (such as the speed and direction in which it moving) and the paging frequency. However, not all terminals have the same characteristics; the movement speed obviously varies between pedestrians and means of transportation such as trains and automobiles, and there are some items that hardly move, such as vending machines and home electric appliances. As far as the means of transportation are concerned, the scope of movement is limited (e.g. automobiles are confined to roads and trains are limited to railroad tracks), as is the movement direction (the travling direction is regulated in highways and railroad tracks). If the location registration area is set uniformly with respect to all mobile terminals under such circumstances, it may not be optimized for certain mobile terminals.

Adaptive location control is a technology that assigns a location registration area that is sized and shaped according to the mobility characteristics of each mobile terminal and the paging frequency (**Figure 1**). For example, in cases where the object moves slowly (e.g. pedestrians), unnecessary paging can be reduced by assigning a small location registration area. In cases where the object moves fast such as a bullet train, the location registration counts can be reduced and the scope of paging can be narrowed down by assigning a large location registration area shaped according to the railroad track. In cases where the mobility characteristics change (say, the individual stays in all night after coming back home), it is necessary to consider reassigning the location registration area. Figure 1 shows the various location registration areas assigned, which are differentiated by coloring.

As for the control method in concrete, when the mobile terminal is detected to have moved from a location registration area to another location registration area, or stayed in the same location registration area for a certain period of time, the mobile terminal sends a location registration signal to the network consisting of information on the movement speed and direction, paging frequency, and so on. Based on the information received, the network calculates the size and shape of the optimal location registration area and informs the mobile terminal to that effect. The network executes broadcasting and paging according to the assigned location registration area, and the mobile terminal executes location registration according to the location registration area.

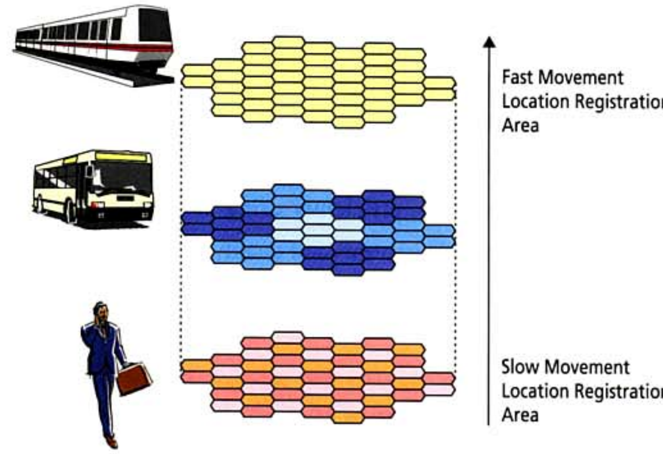

Figure 1 Adaptive Location Control

Of note, there are a number of ways for the mobile terminal to detect the movement speed, such as embedding a Global Positioning System (GPS) function in the mobile terminal, and using the information on the base station's coordinates broadcast by the base station.

(2) Concatenated Location Registration Control

Concatenated location registration control is a technique applied to cases in which many mobile terminals move in groups: the concatenated relation between the individual mobile terminals constituting the group and the mobile terminal representing the group are registered in the network in advance, and the visited location information is updated at the network only by the mobile terminal representing the group. This arrangement allows fewer location registration signals than cases in which location registration is executed individually by mobile terminals.

This concatenation is not limited to two stages. The ubiquitous environment in the future is presumed to constitute the so-called Personal Area Network (PAN), in which users carry more than one wearable device —not only a mobile phone but also a PC with mobile communication functions and a Personal

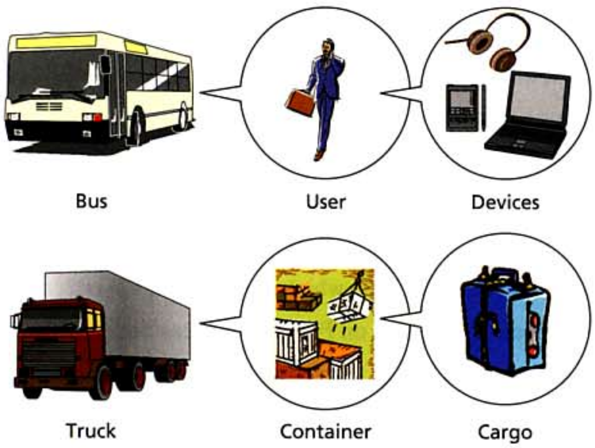

Figure 2 Usage Patterns of Concatenated Location Registration Control

Digital Assistant (PDA). An example is a case in which the user travels by bus (**Figure 2**). Firstly, when the user picks up the devices, the first concatenation is formed with respect to the user's mobile phone (this example assumes that the mobile phone is the mobile terminal representing the user group). This concatenation makes it possible to collectively manage the location information of the devices by using the location information of the mobile phone. When the user hops on a bus, the next concatenation is formed, which similarly makes it possible to manage the location information of the mobile phone by using the location information of the bus. Consequently, the location information of the devices can be managed by using the location information of the bus. Another example is a case in which cargo is loaded into a container, and the container is then transported by truck (Figure 2). The approach to location information management is the same as in the previous example.

The future challenge is to determine what kind of cases in which such multi-phase concatenations are valid in real life, and how to establish and release such concatenated relations. Another vital challenge is how to improve the efficiency of handover control with respect to an active terminal in communication mode in a concatenated state using the concatenated relations.

The aforementioned adaptive location control and concatenated location registration control are deemed to be applicable to 2nd Generation (2G) and 3G systems as well, and further improvements in functionality are being sought.

3.2 Route Optimization

(1) Anchor Relocation Handover Control

In order to prevent the generation of redundant routes upon handover as described in 2.1(2), it is effective to relocate the anchor point to an adequate position according to the mobile terminal's movement, rather than fixing it.

The problem in hierarchical Mobile IP [2], which is the existing handover method that relocates the anchor point, is that it requires more time for the handover process and gives rise to the loss of packets because the process of relocating the anchor point is carried out during the execution of handover. IP^2 executes the process of relocating the anchor point prior to the handover commencement, thereby shortening the duration between the commencement and completion of handover [3].

The following is a concrete description of the anchor relocation control procedures in IP^2, with reference to **Figure 3**.

① The mobile terminal moves to the territory of the Old Access Router (O_AR), which is an Access Router (AR), and then transmits a Binding Update (BU) message to

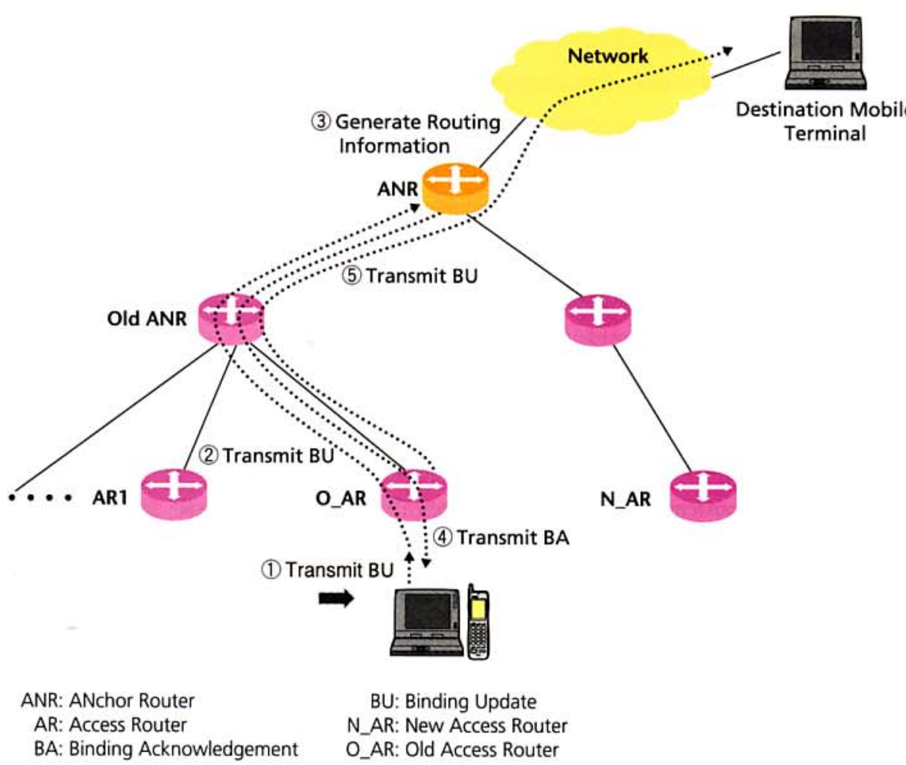

Figure 3 Anchor Relocation Control

O_AR for the purpose of making an anchor relocation request.

② O_AR relays the received BU to the ANchor Router (ANR), which is the new anchor point (assuming that O_AR already knows the address of ANR).

③ ANR receives BU, and then newly generates routing information on the mobile terminal, that is, information that the packets will be transmitted to the mobile terminal via O_AR, as the anchor point.

④ ANR transmits Binding Acknowledgement (BA) to the mobile terminal, as a message in response to BU.

⑤ O_AR transmits a message (BU) to notify the destination mobile terminal that the anchor point has been relocated.

⑥ The destination mobile terminal receives BU, and then transmits packets via ANR thereafter.

By setting the position of the anchor point at the intersection of paths from the destination terminal to O_AR and O_AR's adjacent ARs, the shortest route can be sustained wherever the mobile terminal goes.

(2) Mobile Multicast Techniques

The existing methods of multicast control in mobile communications are: Foreign Multicast [4], which involves the reconstruction of the multicast tree every time the mobile terminal moves; and Bidirectional Tunneling [4], which involves the transmission of multicast packets via the Home Agent (HA) based on IP tunneling, using Mobile IP techniques. The former focuses on the effective use of network resources, whereas the latter focuses on the reduction of the load incurred in multicast-tree updating. However, both methods have their respective merits and demerits, facing the following technical challenges.

Foreign Multicast:

① If the receiver of the multicast service moves, the receiver must participate in the multicast tree at the destination, which leads to increased handover processing time.

② If the sender of the multicast service moves over the multicast tree with the shortest route, the entire multicast tree must be reconstructed per movement, which increases the load on the network. Another challenge is that a new means to inform the receiver of the sender's movement is required.

Bidirectional Tunneling:

③ As packets are transferred between the mobile terminal and HA on a unicast basis with the use of IP tunneling, the number of packet forwarding must always be as many as the number of terminals, which undermines the resources consolidation effects of multicast.

④ As both uplink and downlink communications always transit HA, the challenge is that it is impossible to avert such problems as path redundancy and concentrated traffic in HA.

As indicated by the challenges referred to in ③ and ④, in the case of multicast based on Bidirectional Tunneling, mobile communication networks cannot fully appreciate the effects of multicast communications, which are aimed at the effective use of resources. IP^2 is based on Foreign Multicast; in order to solve the challenges referred to in ① and ② above, methods are being studied to reduce the frequency of multicast-tree reconstruction by forwarding/receiving multicast packets via a proxy in the network and use resources more effectively by informing the receiver of the sender's movements by multicast.

3.3 Technologies to Raise Transmission Quality

(1) Buffering Handover Control

If you buffer the downlink packets at the anchor point upon handover and resume the transfer of buffered packets after the completion of handover, you can achieve lossless packet forwarding, even though packet forwarding will temporarily be suspended during handover.

The existing method called smooth handover [5] buffers the packets once at the Old Access Router (O_AR) as the anchor point, and transfers them to the New Access Router (N_AR) after the handover completion. However, the order in which packets arrive might be inversed between packets transferred from O_AR and packets transmitted from the destination mobile terminal after the handover completion without going through O_AR, which results in lower throughput due to the order correction process. This is attributable to whether the packets transmitted from the destination mobile terminal goes through O_AR or not, i.e. the difference in the path length in the paths before and after handover.

IP^2 designates the router located at the paths intersection from the destination terminal to O_AR and O_AR's adjacent ARs as the router in charge of buffering to prevent the adverse effects from arising due to the inversed packet arrival order [6].

The buffering control procedures in IP^2 are explained below in concrete terms, with reference to **Figure 4**.

① Triggered by the radio link layer, etc., the mobile terminal predicts the occurrence of handover and transmits a message called Buffering Indication (BuI) to O_AR for the commencement of buffering.

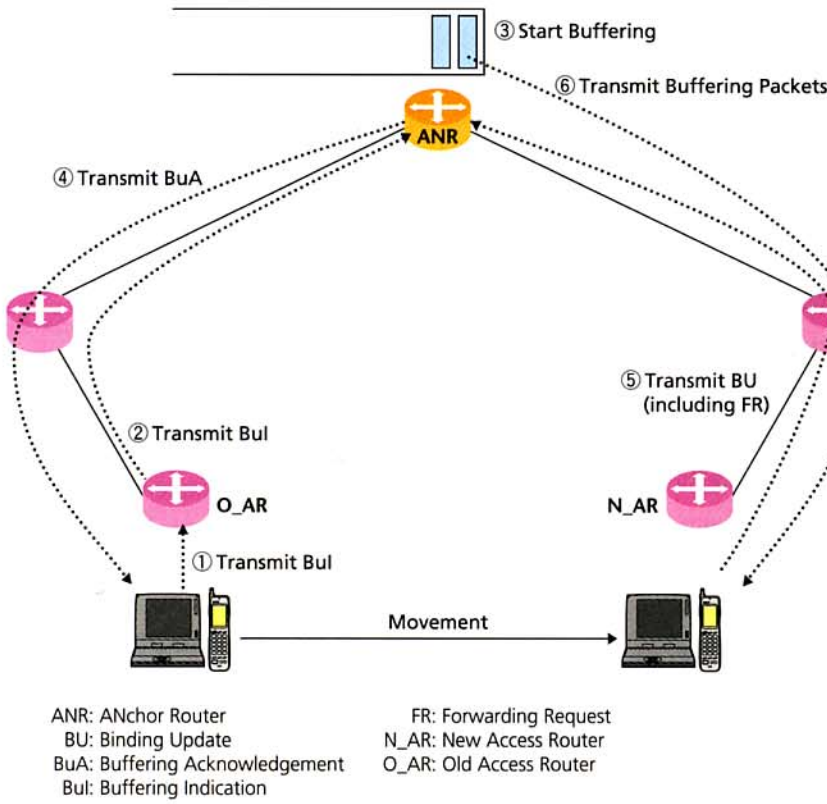

Figure 4 Buffering Handover Control

② O_AR relays the received BuI to the anchor point, which is at ANR (assuming that O_AR already knows the address of ANR).

③ ANR receives BuI, and then starts buffering.

④ ANR transmits a BuI acknowledgement message called Buffering Acknowledgement (BuA) to the mobile terminal.

⑤ When the mobile terminal moves into the territory of N_AR, it sets the Forwarding Request (FR) information on the packet subject to buffering at ANR in the BU message to start buffering, and transmits it to ANR.

⑥ ANR receives BU, and then transmits the buffered packets to the mobile terminal via N_AR.

In step ① above, if BuI is transmitted too quickly, the buffer size will be larger than what's needed at ANR, whereas if BuI is transmitted too slowly, the packets transmission will continue even after the weakening of the radio link on the O_AR side, resulting in the packets loss. The future challenge is to determine what is appropriate as the trigger by the radio link layer for reference.

Of note, IP² implements handover with limited transmission delay and absolutely no transmission loss by combining the aforementioned method with anchor relocation control referred to in 3.2 (1).

(2) Multipath Handover Control

If the cell radius is reduced to implement high-speed wireless access, fast public transport, it creates a situation in which handover is repeated in a short period of time. In a state of fast handover, packets might frequently be transmitted to a cell in which the terminal was located prior to its movement, even if the terminal has already moved to another cell, raising concerns about poorer communication quality and lower throughput. Buffering handover, which is a technique to prevent packet loss, tolerates delays associated with Transmission Control Protocol (TCP) and is effective in applications that are sensitive to packet losses. However, delays due to buffering in real-time communications cause deterioration in the quality of voice, video, etc. Therefore, it is necessary to suppress the loss of packets without giving rise to delays. IMT-2000 is capable of executing handover without packet loss, as the mobile terminal can simultaneously communicate with multiple base stations thanks to the diversity handover technology. As diversity handover is a handover method dependent on the wireless access technology, IP² —which accommodates various wireless access systems— must support handover control with limited packet loss also for wireless access systems under no diversity handover environment. For this reason, we are developing multipath handover control, which supports the simultaneous transmission of packets through multiple base stations even to terminals with no diversity handover functions. To implement the multipath handover, it is important to satisfy the following two requirements to use resources effectively upon handover and to reduce packet forwarding jitters:

- Adequate selection of base stations at which the path should be set simultaneously; and
- Multipath settings based on the shortest route.

In IP², handover methods based on the application of IP multicast techniques and methods of setting optimal multicast points are currently being examined.

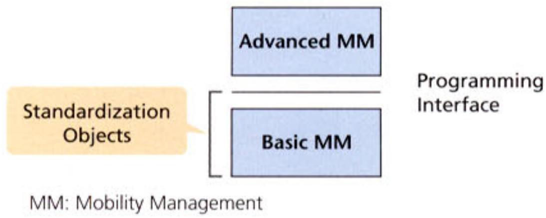

Figure 5 Standardization of Mobility Management Control Capabilities

4. Standardization Method

IP² aims to implement advanced location control and handover control as stated in Chapter 3. Upon standardization, however, it would be preferable to assure flexibility with respect to each carrier that satisfies the aforementioned requirements, considering that control is unnecessary for certain carriers and that control should be executed in regard to each carrier individually. This means that only the basic control capabilities and the programming interface for using them should be standardized, rather than standardizing all the advanced control capabilities (**Figure 5**) [7]. By doing so, each carrier will be able to introduce the advanced functions it wishes to provide on a plug-in basis, making flexible networking possible.

5. Conclusion

This article reviewed the objectives of mobility management technologies deemed necessary in IP², which satisfies the requirements for "Beyond IMT-2000", and discussed the elemental technologies for solving them in a systematically. We plan to conduct studies to incorporate the aforementioned elemental technologies for mobility management in the IP² architecture to operate them in a coordinated system as an integrated system.

REFERENCES

[1] H.Yumiba, K.Imai, and M.Yabusaki, "IP-Based IMT Network Platform," IEEE Personal Communication Magazine, Oct.2001.

[2] H.Soliman, C.Castellucia, K.El-Malki, L.Bellier, "Hierarchical MIPv6 mobility management," IETF draft-ietf-mobileip-hmipv6-06.txt, Jul. 2002.

[3] Iwasaki, Igarashi, Isobe, Koshimizu and Yabusaki: Proactive Region Reconfiguration Handover, Technical Report of IEICE NS2001-203, Dec. 2001.

[4] David B.Johnson, et al, "Mobility Support in IPv6", IETF draft-ietf-mobileip-ipv6-18.txt, Jun.2002.

[5] G.Krishnamurthi, R.Chalmers., C.E.Perkins, IETF draft-govind-seamoby-buffer6-00.txt, Feb. 2001.

[6] Isobe, Koshimizu, Hayashi and Yabusaki: A proposal of Lossless-Fast HO with buffering scheme at cross-over-route, IEICE Communication Society, Sep. 2001.

[7] Sawada, Ihara and Yabusaki: Mobility Management Modularization in IP², IEICE General Conference, Mar. 2002.

GLOSSARY

ANR: ANchor Router
AR: Access Router
BA: Binding Acknowledgement
BU: Binding Update
BuA: Buffering Acknowledgement
BuI: Buffering Indication
FR: Forwarding Request
GPS: Global Positioning System
HA: Home Agent
IMT-2000: International Mobile Telecommunications-2000
IP: Internet Protocol
IP²: IP-based IMT network Platform
MM: Mobility Management
N_AR: New Access Router
O_AR: Old Access Router
PAN: Personal Area Network
PDA: Personal Digital Assistant
TCP: Transmission Control Protocol
VoIP: Voice over IP