Invited Paper

Emerging Standards for Mobility Management in Next-Generation All-IP Networks

Hidetoshi Yokota Akira Idoue

KDDI R&D Laboratories, Inc.
2-1-15 Ohara, Fujimino, Saitama, 356–8502, Japan

Abstract

With the rapid growth of mobile users and bandwidth-rich applications, demands for higher capacity in the wireless networks are becoming more pronounced, and these ever increasing demands are driving the evolution of radio access technologies at a faster pace than ever before. Towards future mobile networks, voice and data services are converging and the existing mobile networks are facing the challenge of providing a new architecture that is optimized to IP-based communications. With an eye on Beyond 3G, further enhancement of the existing mobile core network is an imminent issue for many mobile operators. Various industry fora and standard bodies such as 3GPP, 3GPP2, WiMAX and IETF are actively developing new protocols, architectures and standards. While they are apparently developing their own standards in an independent way, they are nevertheless interworking and harmonizing with each other in many facets for the next-generation wireless IP network infrastructure.

In this paper, we introduce the current status of the 3G standardization organizations towards the evolution of existing 3G systems and present a new mobility management protocol that is under developing to meet the requirements for the next generation mobile networks. Finally, current issues that we are facing and future directions are addressed.

Keywords: Mobile IP, micro-mobility, beyond 3G, All-IP networks

1 Introduction

Continuously evolving wireless technologies are producing advanced radio access systems and these technologies are not only boosting the speed of the same wireless technologies such as IEEE802.11a/b/g/n, but also creating new wireless network systems such as Mobile WiMAX. Also, in the cellular networks, 3G systems are becoming deployed worldwide and the 3G standardization organizations, i.e. 3GPP and 3GPP2, have started developing next generation cellular systems. From the perspective of network services, IP-based applications are becoming dominant, so the mobile network including both the radio access systems and the core network should be optimized for packet-switched communications. Since it is forecast that diverse wireless access technologies will coexist for the next decade, it is of particular importance for the mobile core network to have the capability to accommodate not only one type of access system, but also multiple access technologies so that the future mobile network will be able to interwork these networks in a more integrated way.

In this paper, the current trends in the 3G standardization work on all-IP networks are introduced and an emerging mobility management protocol for the next generation mobile networks that is under development in IETF is presented. Before the conclusions, current issues and future directions are also discussed.

2 Trends in the Network Evolutions of 3G Infrastructure

2.1 3GPP Evolution Trends

3G systems are being deployed worldwide and the number of subscribers to 3G services is also increasing rapidly. 3G network operators are anticipating that even higher capacity will soon become necessary in order to accommodate more users and bandwidth-rich applications. 3GPP started a feasibility study called "UTRA & UTRAN Long Term Evolution (LTE)" in 2004 to develop a framework for the evolution of the 3GPP radio access technology towards a high data rate, low latency and packet-optimized radio access technology. The target peak rate was set to 100 Mbps for downlink and 50 Mbps for uplink within a 20 MHz downlink spectrum allocation. The target user-plane latency was set to less than 5 ms in unload condition for small IP packets.

It was also projected that in the future 3G networks, IP-based communications and applications would become dominant; therefore, a new feasibility study on an IP-based core network started as AIPN (All-IP Network)[1] in 2004 and was completed in 2006. The AIPN is an evolution of the 3GPP system to meet the increasing demands of the mobile telecommunications market. A general overview of the AIPN is depicted in Figure 1. One of the key features of the AIPN is the adoption of a common IP-based network that provides IP-based network control and IP transport across and within multiple access systems. This includes the provision of IP-based mobility control that is not dependent upon specific access or transport technologies. Another key feature is to provide seamless mobility between heterogeneous access systems. The AIPN is expected to provide continuous user experience even if the user moves across a variety of different access systems. In AIPN, three types of mobility are defined:

- End-user mobility: the ability for the subscriber to communicate using the device or devices of his/her choice.
- Terminal mobility: the ability for the same UE (User Equipment) to communicate while changing its point of attachment to the 3GPP System. This includes both handovers within the same access system, and handover from one access system to another.
- Session mobility: the ability for a communication session to be moved from one device to another under the control of the user.

End-user mobility can be realized by the 3GPP security framework (e.g. USIM) or an end-user's service profile. Terminal mobility can be realized by a mobility management mechanism (e.g. Mobile IP) that enables data to be routed to the user device wherever it goes. Session mobility can be realized by controlling the session (e.g. with SIP) and adapting the quality of service to the device capability. The AIPN also interworks with external IP networks (e.g. Internet) and existing networks (e.g. PSTN) for smooth migration.

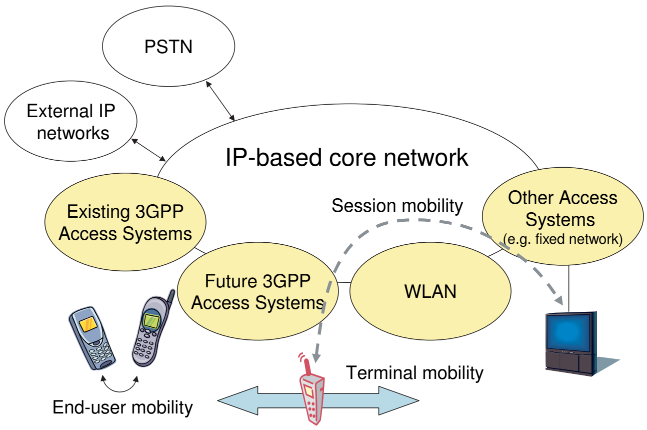

Figure 1: General overview of the AIPN

To develop a new system architecture for an evolved 3GPP system that accommodates and migrates to the LTE, a new feasibility study was started under the name "System Architecture Evolution (SAE)". Based on the requirements identified from the AIPN work, it became clear that access to the 3GPP network will be not only via UTRAN or GERAN but by WiFi, WiMAX, or even fixed broadband access. Therefore, the focus of this work is to study a variety of IP connectivity access networks. Functional components and possible interfaces are depicted in Figure 2 [1]. The key functional components in the evolved packet core are as follows:

[1]This figure is based on Ref.[2] and is subject to change according to further discussion.

Mobility Management Entity (MME): manages and stores UE context. It generates temporary identities and allocates them to UEs. It checks the authorization to ascertain whether the UE may camp on the TA or on the PLMN. It also authenticates the user.

User Plane Entity (UPE): terminates for idle state UEs the downlink data path and triggers/initiates paging when downlink data arrive for the UE. It manages and stores UE contexts, e.g. parameters of the IP bearer service or network internal routing information.

3GPP Anchor: is a functional entity that anchors the user plane for mobility between the 2G/3G access system and the LTE access system.

SAE Anchor: is a functional entity that anchors the user plane for mobility between 3GPP access systems and non-3GPP access systems.

The locations of these functional components and the interfaces between them are still under discussion.

Since GPRS (General Packet Radio Service) uses GTP (GPRS Tunneling Protocol), S3 and S4 interfaces are likely to be GTP; however, S1, S2 and S5a/b interfaces may be good candidates for IP-based mobility management protocols.

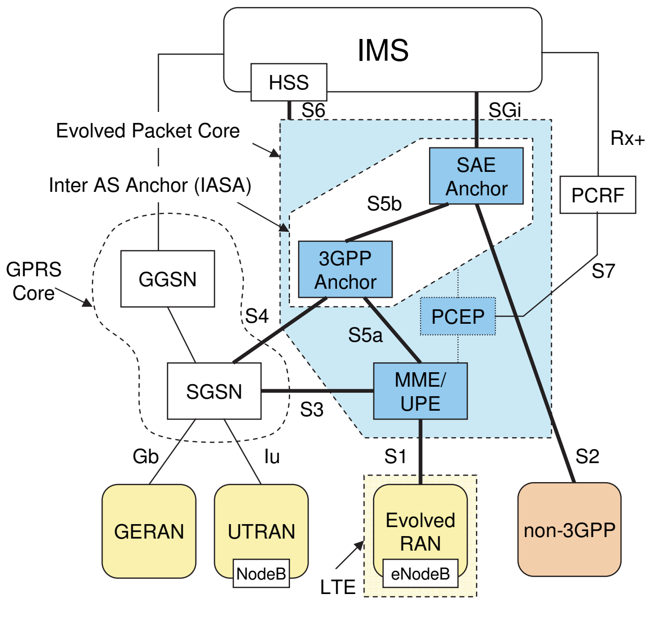

Figure 2: Baseline high level architecture of SAE

2.2 3GPP2 Evolution Trends

In 3GPP2 as well, a new work topic on the evolved network architecture got underway as "Packet Data Network Evolution (PDANE)" in 2006. The general architecture of PDANE is shown in Figure 3 [2]. As a mobility agent, the Home Agent (HA), which resides in the Home IP network, and the Local Mobility Home Agent (LMHA) are defined. The LMHA delegates an IP address to the Mobile Node (MN) and controls its mobility. The Controlling Access Point (CAP) has Radio Resource Management (RRM) functions and acts as the Access Gateway, that is, the first hop router to the MN. On the lowest level, the Base Transceiver System (BTS) functions are defined. Depending on the needs of a given deployment, some of these entities may not be deployed. If all of them are deployed, two levels of mobility management are required i.e. between the HA and the LMHAs and between the LMHAs and the CAPs. If some entities are omitted or collocated, only a single level of mobility management is required.

One of the goals for the PDANE is to streamline the network entities to enhance system performance. For example: reduce signaling latency, reduce media transmission delay, improve data transmission efficiency, and improve mobility management. Both in the wireless and wireline environments the trend is to converge and reduce the number of transport and signaling protocols and the number of network elements required to support end user applications. To this end, the PDANE leverages IP-based protocols and minimizes the 3GPP2 specific network entities compared with the current configurations. The evolved new architecture will be explored so that it can efficiently support the ongoing proliferation of new applications that use IP as the delivery protocol (i.e. VoIP, IPTV). This trend is being accelerated by the emergence of new broadband air interface technologies which offer significant throughput improvements to the end user.

3 Protocol trends for mobility management in all-IP networks

For both 3GPP and 3GPP2 network architectures, the mobility management protocol is one of the key issues and an efficient mechanism that can accommodate various access technologies is required. Micro-mobility management is an effective way of reducing the number of location registration messages with the home agent and shortening the round-trip time required while the MN moves locally [4]. It is anticipated that this mechanism will be implemented particularly between the LMHAs and CAPs in 3GPP2 PDANE since the HA resides in the home IP network, which could be remote from the LMHA. To conceal local movement of MNs from the HA, most micro-mobility management schemes define special administrative domains that are physically or logically separated from the global and Mobile IP enabled network. Within these domains, which we call micro-mobility networks, any type of routing scheme can be applied and there are mainly two approaches to the realization of micro-mobility[4]:

[2] This figure is based on Ref.[3] and is subject to change depending on the outcome of further discussion.

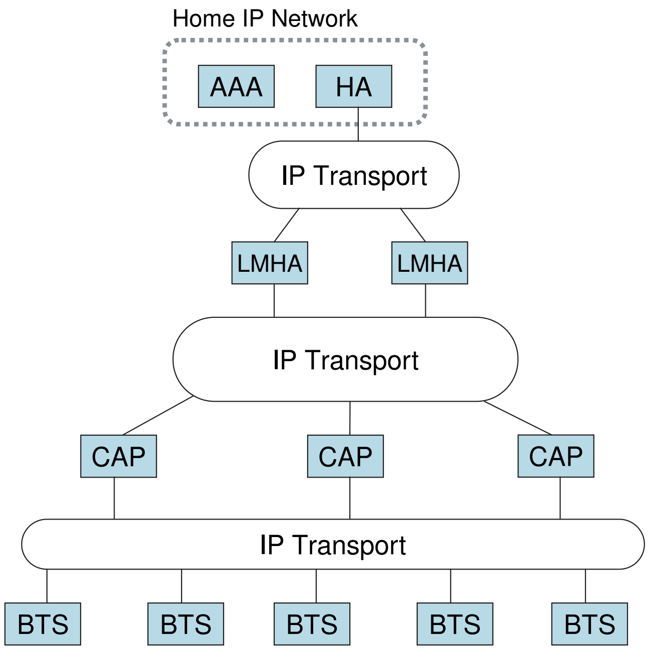

Figure 3: General architecture of Packet Data Network Evolution

Tunnel based approach: Mobility agents establish tunnels to forward packets whose destination address does not belong to the network. As long as the tunnel endpoints can support the protocol, intermediate nodes need not be aware of it. Regional Registration[5] or Hierarchical MIPv6[6] fall under this category. Ref [7] is also an example of this approach although it uses MPLS.

Host-routing based approach: Mobility agents in the micro-mobility network maintain the next hop for the MN and packets destined for the MN are relayed by these agents. Although there is no tunnel overhead, all the nodes need to be aware of the protocol. Cellular IP[8] and HAWAII[9] fall under this category.

When the PDANE in 3GPP2 is considered, the existing IP network can be used between the LMHA and the CAP, so the tunnel-based approach would be more suitable. If Mobile IP is employed for the mobility management protocol, not only the network but also the terminals need to support it. Even if 3GPP2 defines Mobile IP as the mobility management protocol, there are many other terminals that do not support Mobile IP such as the ones in 3GPP networks. An evolved network is required to migrate from the existing network seamlessly and to interwork with multiple access technologies; therefore, it needs to support mobile terminals that do not support Mobile IP. To meet this requirement, a proxy-based Mobile IP mechanism is a promising candidate. A proxy-based Mobile IP handles location registration of the mobile nodes on behalf of these nodes and they can change their subnets without disrupting communications. In the following section, we introduce and discuss a new proxy-based mobility management protocol that is currently under development in IETF.

3.1 Network-based Localized Mobility Management Protocol

In IETF, a new working group was formed in 2006 to develop network-based localized mobility management. This protocol is called NetLMM and it allows a mobile node to move around in an access network attaching to various points of the access network while maintaining an IP layer configuration that does not change as the mobile node's points of attachment change. It is intended to support packet forwarding to mobile nodes that change their point of attachment to the network without the use of any protocol support at the IP layer on the mobile node to support that mobility. A general overview of the NetLMM domain and the key components involved are shown in Fig 4.

The principal functional entities in a NetLMM infrastructure are the Mobile Access Gateway (MAG) and the Local Mobility Anchor (LMA). Their functionalities are as follows:

Mobile Access Gateway (MAG): a router that a mobile node is attached to as the first hop router in the NetLMM infrastructure. The MAG is connected to the mobile node over some specific link provided by a link layer, but the NetLMM infrastructure is agnostic about the link layer technology that is used. Each MAG has its own identifier used in NetLMM protocol messaging between the MAG and the LMA. There are multiple MAGs in a NetLMM infrastructure.

Local Mobility Anchor (LMA): a router that maintains reachability to a mobile node's address while the mobile node moves around within the NetLMM infrastructure. It is used to maintain forwarding information for the mobile nodes. This includes a set of mappings to associate mobile nodes by their identifiers with their address information, associating the mobile nodes with their serving MAGs and maintaining the relationship between the LMA and the MAGs. There may be one or more LMAs in a NetLMM infrastructure.

While NetLMM defines the protocol between the MAGs and the LMAs, there are other entities that will make up a mobile access network and that will be used to support various kinds of functionality (mobile nodes, AAA, routing, DNS, etc.), whose basic functionality may be used by the MAG and the LMA, but whose operation is not changed in any way for the proper operation of the NetLMM protocol.

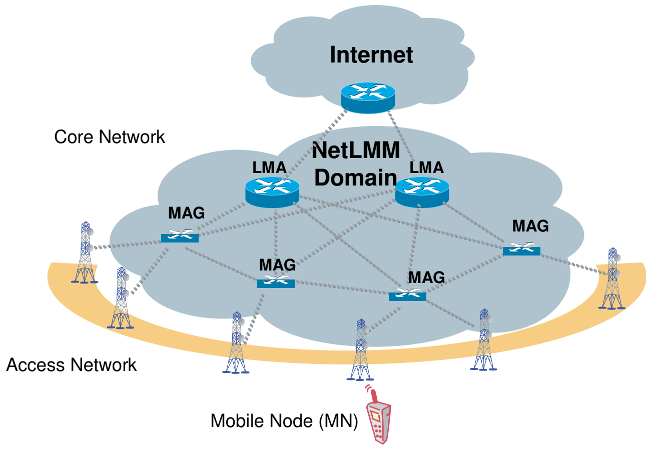

Figure 4: General overview of NetLMM domain

3.1.1 Protocol Overview[3]

The protocol consists of two major phases, an initiation phase and an operational phase. During the initiation phase, the MAGs and an LMA (or multiple LMAs) establish connectivity between each other. During the operational phase the MAGs and LMAs provide a mobile connectivity service to mobiles nodes that are attaching to the infrastructure, leaving the infrastructure, and moving around within the infrastructure. It is not assumed that a MAG is associated with only a single LMA. If multiple LMAs exist in a NetLMM domain, each MAG would most likely be associated with, and potentially communicate with all the LMAs rather than only a single LMA. On the other hand, each MN is always associated with one LMA wherever it moves within the NetLMM domain.

In the initiation phase, the NetLMM currently defines 3 pairs of messages to establish and maintain associations between the MAGs and the LMAs:

- Association Request / Association Reply,
- Disassociation Request / Disassociation Reply and
- Heartbeat / Heartbeat Ack.

A MAG associates itself with an LMA by sending an Association Request message that includes its MAG ID (e.g. MAG's MAC address or IP address) and the supported data forwarding modes (e.g. IPv6-in-IPv6). When the LMA receives this message, it creates an association with the MAG and populates state information about the association. The LMA then returns an Association Reply with its LMA ID (e.g. LMA's MAC address or IP

[3] This is based on [10] and is subject to change depending on the outcome of further discussion.

address) and an agreed upon data forwarding mode to the MAG. The MAG can uncouple the relationship with the LMA through sending a Disassociation Request, to which the LMA responds with a Disassociation Reply. Heartbeat messages are sent between the MAG and LMA to determine the current status of the reachability of the other entity. All of these messages may be sent optionally over an IPsec connection if additional security is desired.

In the operational phase, the NetLMM defines 2 pairs of primary messages and 5 pairs of additional messages to manage the attachment, departure, mobility, and other activities of MNs within the infrastructure:

- Location Registration / Location Registration Ack,
- Location Deregistration /Location Deregistration Ack,
- LMA Allocation Request / LMA Allocation Reply,
- MN Address Setup / MN Address Setup Ack,
- MN Address Remove /MN Address Remove Ack,
- Routing Setup / Routing Setup Ack and
- Routing Remove / Routing Remove Ack.

The Location Registration message and the Location Deregistration message are sent from the MAG to the LMA when the MAG detects an attachment or detachment of the MN, respectively. Notice that the MN may or may not have its IP address at this point. The LMA Allocation Request message is optionally sent to the LMA to authorize service for a particular MN by specifying its MN ID (e.g. MN's MAC address or Network Access Identifier). This message may come from various sources (e.g. the AAA server) prior to the location registration procedure. The MN could at any time configure a new IP address for itself using stateless address autoconfiguration. When the MN issues a new address DAD (Duplicate Address Detection) request, the MAG sends an MN Address Setup Request to the LMA with the MN ID and the new IP address. The LMA validates the usability of the address, updates the entry for the MN in the Routing Cache and acknowledges the request with the MN Address Setup Ack message to the serving MAG, which then replies to the MN DAD operation. On the other hand, the Routing Setup message is sent from the LMA to the MAG to inform the MAG of the MN's IP address. This situation happens when the MN moves to a new MAG. Since the IP address for the MN does not change and is stored in the LMA, the LMA can inform the new MAG of this information. The MAG then acknowledges this information with a Routing Setup Ack message.

3.1.2 Network Attachment Procedure

A proxy-based mobility management requires 1) a trigger to detect network attachment of the MN and 2) an identifier of the MN to register its location on behalf of the MN. This trigger and identifier can be obtained in various ways. One possible way is to use the authentication phase of the mobile node at the time of network attachment. In cellular networks, a mobile terminal is always authenticated when it attaches to the network for security and accounting. Also, in WiFi networks, IEEE802.1X[11] provides a dynamic authentication mechanism compared to a static WEP key. By leveraging these mechanisms, the NetLMM protocol can obtain the timing and necessary information for the location registration. Furthermore, the MAG needs to know the corresponding LMA for the MN that has attached to the network before the MAG sends the Location Registration message. To this end, for example, the AAA server can be leveraged. If this is the case, the AAA server stores the IP address or prefix for the MN and the corresponding LMA. Fig. 5 shows the initial network attachment procedure, where IEEE802.1X is used as the trigger for detection of network attachment. In this case, the MAG plays a role of both an AP and the first hop router for the MN. Note that this is just one use case and does not exclude other possibilities.

When a new MN attaches to the network, it sends an EAP-Start message (1). Then the MAG sends back an EAP-Request with its MAG_ID (2). When the MN in turn sends its MN_ID (e.g. NAI) to the MAG (3), it is conveyed to the AAA server with, for example, RADIUS Access-Request (4). Depending on the type of authentication (e.g. EAP-TLS), authentication information (e.g. credentials) is exchanged between the MN and the AAA server (5). If the authentication is successful, an Access-Accept with the corresponding LMA_ID is sent to the MAG (6). The AAA server also provides either a prefix or a full IP address for the MN with this message. In this use case, if a prefix is assigned, it must be unique to the MN (per-MN prefix) in order for the LMA to be able route the packets destined for the MN only by looking at this prefix. Detailed discussion is presented in Section 3.1.4. When the MAG receives this message, it takes out and stores the LMA_ID and sends the Registration Request to the corresponding LMA (7). The LMA then creates an entry for the MN and stores the received MAG_ID. If the location registration is successfully performed, the LMA sends back a Location Registration Ack message to the MAG (8). The MAG finally sends EAP-Success to the MN to indicate that the authentication is completed (9).

For the address assignment, there are two ways to do it: the stateless address auto-configuration and stateful address auto-configuration. In the case of the stateless auto-configuration, the MAG sends a unicast Router Advertisement to the MN with the obtained prefix since this prefix may be unique to this MN. In the case of the stateful auto-configuration, the MAG can act as the DHCP Server since it already obtains the prefix or the IP address for the MN although it can still be the DHCP Relay Agent. The MAG then assigns the IP address to the MN.

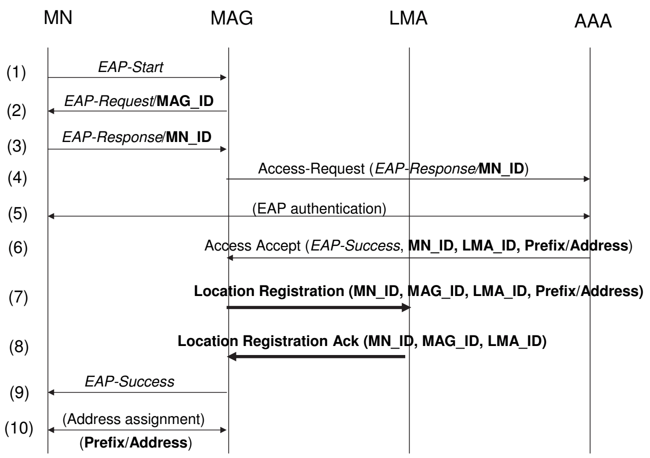

Figure 5: Initial network attachment

3.1.3 Inter MAG Handover Procedure

When a MN hands over from one MAG to another, the new MAG may not know if the event is a handover or an initial network attachment. This is because the interface between MAGs is not currently within the scope of the NetLMM protocol. The Inter-MAG handover procedure is shown in Fig. 6. When the MN attaches to the new MAG, it initiates the authentication phase. Steps from (1) to (5) are the same as the initial attachment in Fig. 5 although the authentication procedures may be different or simplified. When the new MAG receives the Access-Accept message from the AAA server (6), it creates a new entry for the MN and stores the LMA_ID and sends the Location Registration message to the corresponding LMA (7). When the LMA receives a Location Registration message, it checks in its Routing Cache if an entry for the MN ID carried in the message is already present. If an entry is already present in the Routing Cache, the LMA updates the entry for the MN with the new MAG_ID. The LMA then sends the Location Deregistration message to the old MAG to remove the entry for the MN (8). When the old MAG returns the Location Deregistration Ack to the LMA (9), the LMA sends the Location Registration Ack message back to the MAG (10). The MAG finally sends the EAP-Success message to the MN (11) to indicate that the authentication is completed. The MN already has its IP address, which does not change before and after the handover, so the address assignment procedure can be skipped.

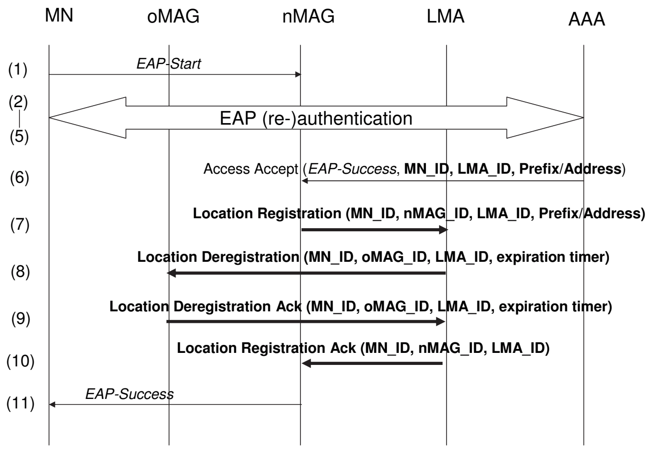

Figure 6: Inter MAG handover

3.1.4 Issues and Future Directions

- In the previous section, when the prefix is assigned to the MN, each MN is assigned a unique prefix, which is called “per-MN prefix assignment”. On the other hand, it is also possible to assign the same prefix to multiple MNs, and if each MN has a unique interface ID, it can construct a unique IP address. This is called “shared prefix assignment”. In the case of shared prefix assignment, the MN needs to configure its IP address with the stateless auto-configuration mode and it comes after the MN receives the Router Advertisement, which will be performed in step (10) as shown in Fig. 5. The LMA also needs to know the configured IP address for the MN to route packets destined for it since the LMA cannot disambiguate them by only the prefix, which is shared by multiple MNs. This means that another message is required after the MN has configured its IP address. The MN Address setup message, which is not used in the previous section, is defined for this purpose. When the stateless auto-configuration is performed, the possible trigger to send the MN Address Setup message is the DAD procedure. The DAD procedure is performed on the MAG link, but the MAG needs to verify the address uniqueness at the LMA; for this purpose as well, the MAG needs to send the MN Address Setup message to the LMA. Also, in the case of the stateful auto-configuration, the address assignment by DHCP usually happens after the authentication, that is, after the location registration; therefore, the MN Address Setup message is also needed in this case. There have been discussions on this issue and for now, the per-MN prefix mode, which is also recommended in [12], has been prioritized for consideration.

- Currently, NetLMM does not define the interface between MAGs; however, this interface can make handover more efficient by transferring the entry for the MN between old and new MAGs and/or forwarding the packets destined for the MN from the old MAG to the new MAG. This feature should be studied to further enhance the NetLMM protocol. As a reference, fast handover for Mobile IPv6 in 3G networks is discussed in [13]. Also, a possible solution for scalable localized mobility management can be found in [14][15].

- NetLMM primarily focuses on the IPv6 infrastructure and support of IPv6 nodes due to its flexibility, but it is envisioned that with modifications the protocol could be also used with an IPv4 infrastructure or to support IPv4 nodes. Once the specification for the IPv6 is complete, one for the IPv4 should also be studied.

4 Conclusion

In this paper, we showed the trends of the next generation mobile networks and presented the current status of the 3G standardization organizations. Since IP-based applications such as VoIP, e-mail, Web, video streaming are attracting more and more mobile users, it is expected that IP traffic will soon become dominant in the mobile environment. For signaling protocols as well, IP-based protocols will be adopted more broadly as we have already seen in IMS (IP Multimedia Subsystem)/MMD (Multimedia Domain). IETF is therefore playing a crucial role in providing protocols for the evolution of the 3G network. We presented a new proxy-based mobility management protocol called NetLMM, which is currently under development in IETF as a candidate for local mobility management in the evolved 3G networks. As shown in this paper, the evolved networks and terminals are expected to support multiple access technologies; therefore, corresponding standardization organizations are becoming more closely related with each other and it is therefore increasingly important to understand the relationship between these standards development organizations (SDOs) and to monitor their activities in parallel.

REFERENCES

[1] "All-IP Network (AIPN) Stage 1", 3GPP TS22.258 (version 7.0.0), December 2005.

[2] "3GPP System Architecture Evolution: Report on Technical Options and Conclusions", 3GPP TR23.882 (version 1.3.0), July 2006.

[3] "3GPP2 Packet Data Network Evolution", 3GPP2 X.R0046 v0.3, June 2006.

[4] A. T. Campbell, *et al.* "Comparison of IP Micromobility Protocols" IEEE Wireless Communications Magazine, Vol. 9, No. 1, pp.72–82, February 2002.

[5] E. Gustafsson, *et al.* "Mobile IPv4 Regional Registration" draft-ietf-mip4-reg-tunnel-00, IETF, November 2004.

[6] H. Soliman, *et al.* "Hierarchical MIPv6 Mobility Management" RFC4140, IETF, August 2005.

[7] F. M. Chiussi, *et al.* "A Network Architecture for MPLS-Based Micro-Mobility" Proc. IEEE WCNC'02, March 2002.

[8] A. Valkó "Cellular IP: A New Approach to Internet Host Mobility" ACM SIGCOMM Computer and Communication Review, Vol. 29, No.1, pp. 50–65, January 1999.

[9] R. Ramjee, *et al.* "HAWAII: A Domain-Based Approach for Supporting Mobility in Wide-area Wireless Networks" Proc. IEEE Int'l Conf. Network Protocols, November 1999.

[10] G. Giaretta, *et al.* "NetLMM Protocol" draft-giaretta-netlmm-dt-protocol-00, IETF, December 2006.

[11] "Standard for Port based Network Access Control", P802.1X/D10, IEEE, January 2001.

[12] M. Wasserman, "Recommendations for IPv6 in Third Generation Partnership Project (3GPP) Standards" RFC3314, IETF, September 2002.

[13] H. Yokota *et al.*, "Mobile IPv6 Fast Handovers for 3G CDMA Networks" draft-ietf-mipshop-3gfh-00, IETF, May 2006.

[14] M. Inoue *et al.*, "Scalable Mobile Core Network Architecture for All-IP Wireless Access", WPMC03, Vol. 1, pp. 18-22, October 2003.

[15] H. Yokota *et al.*, "Decentralized Micro-mobility Management for Large-scale Mobile Networks" Proc. WPMC03, Vol.2, pp.352-356, October 2003.