INTEGRATED MOBILE IP AND SIP APPROACH FOR ADVANCED LOCATION MANAGEMENT

Q Wang and M A Abu-Rgheff

University of Plymouth, UK

ABSTRACT

The vision of ubiquitous and continuous communications for a mobile user entails several emerging mobility types, which pose new requirements for location management. In this paper, we present an integrated approach for advanced location management towards supporting all kinds of mobility types. After a justification of our approach, we describe our architectural considerations on a uniform network architecture accommodating various mobility scenarios. We propose integrated mobility servers for Mobile IP and SIP (Session Initiation Protocol) to take advantage of their complementary features in supporting mobility. We also discuss the two major tasks of location management: location updates and paging using optimised signalling to reduce overheads and delay.

1. INTRODUCTION

All Kinds of Mobility

Traditional mobility management is limited to *terminal mobility*, which enables a moving mobile host (MH) to continue an ongoing call/session or initiate/receive a call regardless of its point of attachment. In the next generation systems (3G and beyond), a mobile user would access to heterogeneous networks (including ad hoc networks) for various services and multimedia sessions via a set of personal devices. Therefore, several new mobility types are emerging and fall into two categories [Wang and Abu-Rgheff (1)]. One is device-centric low-level mobility, including *ad hoc mobility* (routable in an ad hoc network) and *mode mobility* (switch between the infrastructure mode and the ad hoc mode) in addition to the terminal mobility. The other is user-centric high-level mobility, including *session mobility* (switch a session to another terminal), *personal mobility* (globally reachable user) and *service mobility* (maintain subscribed and personalised services regardless of service providers).

Why SIP + MIP?

SIP has been proposed to serve as the major call/session control protocol by principal 3G standardisation organisations and forums like 3GPP and 3GPP2. Meanwhile, there is a trend to explore SIP [Rosenberg et al (2)] for a complete mobility support since it has inherent support for personal mobility and can be extended for session, service and terminal mobility [Schulzrinne and Wedlund (3)]. However, we believe that Mobile IP (MIP) [Perkins (4), Johnson et al (5)] complements SIP in the following aspects:

- MIP can hide IP address changes from applications.
- MIP supports TCP sessions inherently while SIP finds much more difficult.
- MIP supports better terminal mobility in many cases [Kwon et al (6)].
- MIP has been incorporated into 3G systems, e.g. 3GPP2 uses MIP for IP mobility and 3GPP considers MIP as an option [3GPP (7)].
- MIP is more suitable to deal with low-level mobility types, esp. those related to ad hoc network.

Why Integrated Rather Than Hybrid?

Hybrid approach tends to apply both MIP and SIP in an overlapping way. This may easily result in redundant signalling for similar functions. Furthermore, the redundancy may cause conflicts and make system unstable. We thus argue that an *integrated*, rather than *hybrid*, approach would be more powerful and efficient through cross-layer design. We have proposed a multi-layer mobility management architecture [Wang and Abu-Rgheff (8)] to take full advantage of each layer's contribution of the protocol stack by applying cross-layer signalling schemes.

Focus and Structure of the Paper

To handle location management in all kinds of mobility types is a huge problem. This paper focuses on the following location management scenarios: mobility in case of personal area network (PAN) and cellular infrastructure network. Location management includes two major tasks: location updates and paging. This applies to most of the mobility types. Thus, the paper will discuss the mentioned mobility types from this perspective and highlight them wherever convenient.

The rest of the paper is organised as follows: Section 2 surveys related work. Section 3 presents a uniform network architecture with servers integrating MIP and SIP mobility functions. Section 4 describes our approach to integrated location updates and paging. Section 6 concludes the paper.

2. RELATED WORK

MIP-Based Schemes

MIP introduces two mobility agents, home and foreign agents, to handle location updates and to route traffic. The basic MIP suffers from the well-known triangular routing while MIP with Route Optimisation [Perkins and Johnson (9)] alleviates this problem by allowing correspondent hosts (CHs) to cache a dynamic binding of the MH's home and care-of IP addresses. Notably, Route Optimisation is an add-on to MIPv4 (4) whereas it is an integral part of the MIPv6 (5). Moreover, in MIPv6 packets can be forwarded with no tunnelling between the MH and the CHs, with the help of an additional routing header. Since MIP was initially proposed for mobile Internet rather than cellular networks, it lacks support for paging. Thus, MIP-based paging schemes were proposed, e.g., Zhang et al (10) and Ramjee et al (11). MIP can also be utilised in ad hoc networks. Pei and Gerla (12) applied the home agent notion to location management in hierarchical ad hoc networks. Jönsson et al (13) studied the combination of MIP and ad hoc routing when an ad hoc network interacts with an infrastructure network.

SIP-Based Schemes

Schemes in this category are best represented by (3). Moh et al (14) and Turányi et al (15) also discussed location management by implementing modified MIP and SIP respectively.

Largely, SIP is suitable for high-level mobility whereas MIP-based schemes are more appropriate for low-level mobility. Clearly, a careful design in an integrated approach should allow SIP and MIP co-operate in a complementary rather than competing way.

3. UNIFORM NETWORK ARCHITECTURE WITH INTEGRATED MOBILITY SERVERS

3.1 Architectural Considerations and Advantages

Considering various mobility types and the common functionalities of MIP and SIP mobility management, we propose a unified network architecture, on which our integrated location management framework is based. This architecture, shown in Figure 1, reflects our proposed considerations to facilitate different mobility scenarios.

First, both MIP and SIP signalling and data flows are handled through home/foreign mobility servers (HMS/FMS) for home/foreign network respectively. For the moment, HMS/FMS can be deemed as the traditional home/foreign agent (HA/FA) for MIP while for SIP they are a combination of home/foreign SIP proxy or redirect server, SIP location server, and SIP registrar.

Second, added benefits can be achieved by integrating the mobility servers (HMS/FMS). In our approach, all the components are integrated into the HMS/FMS (discussed in Section 3.2), which is capable enough to cope with both MIP and SIP, and optimised to minimise, if not eliminate, any possible functionality redundancy or information duplication. Thus complexity can be reduced when compared with an overlapping approach. On the other hand, some other important related components are also accommodated either for extended functionality or deployment/presentation convenience, as described in Section 3.2.

Third, mutual learning and benefits can be facilitated between MIP and SIP in this uniform architecture. For example, by following the MIP mobility model and

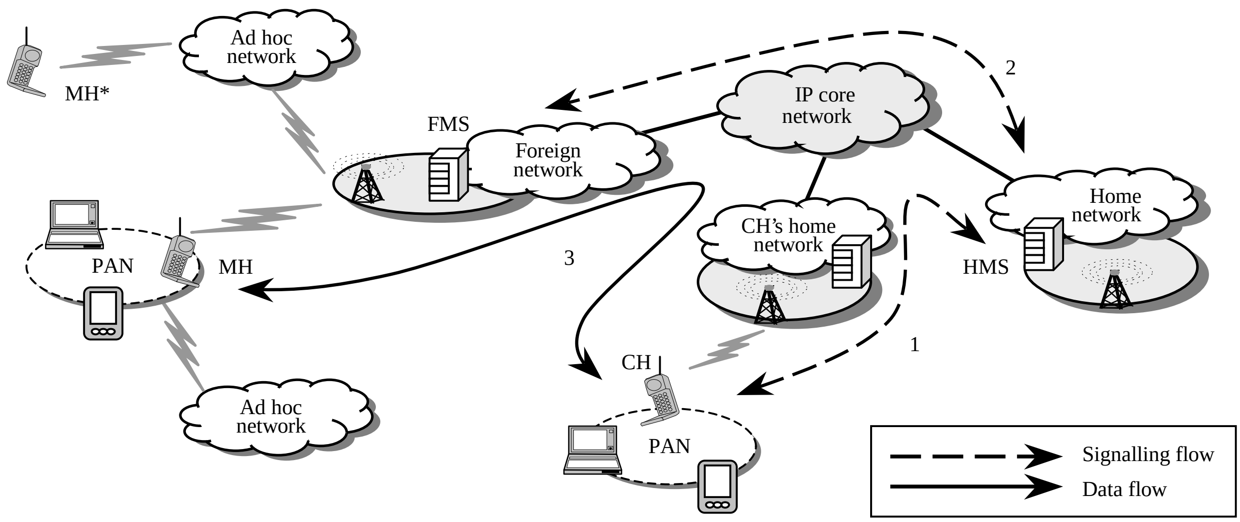

Figure 1. A unified network architecture for advanced location management

specifying the HMS/FMS, this architecture facilitates a solution to the simultaneous moving problem in SIP (both ends move to its foreign network at the same time). With their HMSs acting as the anchor points, both the MH and the CH contact their HMSs respectively to locate its counterpart.

Fourth, triangular routing is eliminated largely and thus data flows take place directly between both end users. For MIP, routing based on MIP with Route Optimisation is applied (see 1 in Figure 1). For SIP, the HMS can act as a redirect server (like MIP with Route Optimisation) or proxy server (see 1-2 in Figure 1). More details are discussed in Section 4.

Last, ad hoc networks, when incorporated into the architecture, interact with the infrastructure by means of connecting an access router (base station) directly or via another mobile host. In this paper, we are particularly interested in the PANs as a special form of ad hoc networks. Since it will be quite common in future that a mobile user uses a set of devices, which comprise a PAN, the case that both communicating ends are PANs will be the more generic scenario than both are single mobile hosts in most of the current network models. In such a PAN, at least one device has a connection with the infrastructure, acting as a gateway (or serving MH/CH) for other devices in the PAN.

3.2 The Main Functional Elements of HMS/FMS

The core part of HMS derives from the notions like the HLR (Home Location Register) in 2G cellular systems and HA in MIP enabled networks, but it is much more powerful and comprehensive. Clearly, to deal with both MIP and SIP signalling and data, all the functionalities of both architectures should be included while optimisation entails that related entities are integrated. The methodology for our optimisation and integration is to decompose the MIP and SIP entities to independent functional elements first and then merge the same or similar elements and retain the distinguished one intact, if enhancement is unnecessary. The integrated elements interface each other to fulfil the traditional tasks and maybe more. Through this methodology, e.g. HA could be de-coupled to a Registration Server and a Forwarding Agent. Consequently, the SIP Registrar is enhanced to handle MIP registration as well.

Like the HSS (Home Subscriber Server) in 3G cellular systems (UMTS), HMS could incorporate other mobility-related modules such as AAA (Authentication, Authorisation and Accounting) server and User Profile. AAA server in a foreign network grants or rejects an access request from a roaming user. In the remaining paper, we assume the user can successfully pass the AAA check and this procedure is omitted for simplicity. The User Profile stores user-specific information, such as subscribed and personalised services, and thus plays an important role in service mobility support.

Moreover, optional elements could be integrated into the HMS. The System Profile provides a roaming user with the major physical/link property information (static typical values or periodically updated values) to facilitate possible adaptation to QoS (Quality of Service). A Policy Table is also useful when traffic-engineering or other policies are applied. The most common policy proposed is to use SIP for UDP traffic while MIP for TCP although policies depending on mobility types are also possible. A Paging Initiator would be needed if the mobility server initiates the paging (discussed in Section 4.3).

The construction of a FMS follows the same methodology used for HMS and results in a similar, if not identical, structure. An IP Address Distributor (IPAD) is needed to assign a care-of-address (COA) or a new IP address to a MIP and SIP MH respectively. In IPv4, the IPAD can be a DHCP server. In IPv6, it is the stateless/stateful address auto-configuration server. Notably, a HMS may need an IPAD as well. In addition, there are some minor differences in capabilities of some elements. For example, the user profile in a FMS can be a simplified version of the one in the HMS. Figure 2 shows the components of a mobility server (HMS/FMS).

In sum, the proposed network architecture, with carefully designed mobility servers, should allow back compatibility of both MIP and SIP, reduced complexity with redundant entities/functions emerged, optimised performance with optimised messages and reduced signalling overheads and delay and enhanced functionality.

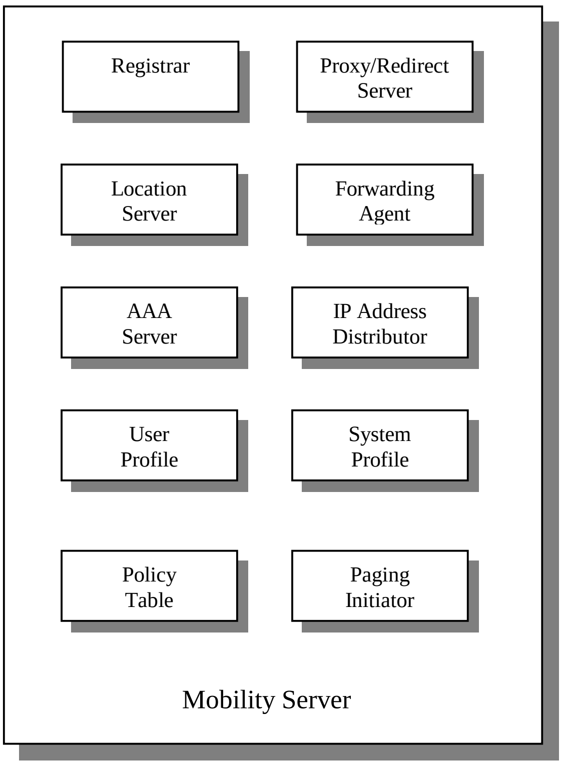

Figure 2. The components of a mobility server

4. INTEGRATED LOCATION MANAGEMENT FRAMEWORK

4.1 General Considerations of Location Management

We notice that different mobility in different scenarios may indicate different requirements for location management. For instance, a typical session handoff that takes place within a PAN usually does not trigger a location update procedure. However, when combined with an inter-system terminal handoff, it is clear that a location update is needed. Service mobility, indicating a change of service provider, may result in an inter-domain location update, or no location update at all in an overlay-networking environment. Ad hoc and mode mobility pose new problems for location management. For example, it may be more efficient to maintain rather than delete the roaming information for a MH in a FMA that is only temporally switches to the ad hoc mobility and will switch back to the infrastructure mode soon. We need enhanced methods than those of the traditional terminal-centric location management to deal with these contexts directly. One the other hand, notably, it may not be necessary to apply any additional location management in a "flat" ad hoc network, where all the MHs are routable. For a "hierarchical" ad hoc network, similar principle of its infrastructure counterpart would be also applicable.

Anyway, the basic location management notion applies to most, if not all, of the interested mobility types although they have their special contributions to or impacts on a specific location management protocol.

Based on the described network architecture and entities, therefore, we focus on an infrastructure-centric scenario when studying the two major tasks of location management: location updates and paging, although the former is emphasised. In designing both protocols, although specific messages are still under investigation, we propose the following criteria for message formation:

- Combine MIP and SIP advantages wherever possible
- Reuse SIP or MIP messages whenever cost-effective
- Weigh between signalling capability and message size although capabilities could be preferred

4.2 Location Updates (Tracking)

Service discovery. The roaming user needs to discover the FMS when entering a foreign network. Generally, a MH can passively or actively carry out this job by the MIP- or SIP-type. In MIP, the FA periodically advertises its existence while in SIP, normally the MH multicasts for searching the local SIP server. (It is possible that MIP/SIP uses the alternative way.) In our proposal, the MIP-type is preferred, esp. the beacons of an intersystem FMS could accommodate rich system-specific information, generated from the System Profile, and thus intersystem handoffs could be facilitated. Once the server has been identified, the following steps are proceeded as shown in Figure 3. Notably, the basic signalling and data flows may resemble some MIP- or SIP-based schemes but the message formats here are integrated ones.

New IP address. The serving MH obtains a new IP address or a COA from the IPAD (line 1 in Figure 3). It is also possible to obtain multiple IP addresses for other devices of the PAN.

Registration. The basic signalling flows during the registration are illustrated by line 2 to line 7. The serving CH then contacts the MH-PAN's HMS and fetches serving MH's current address (line 8) using MIP-RO or SIP redirect server function in the HMS, so that it can send packets to the MH directly (bypass the MH's home network, line 9 and 12). If the CH enquires the HMS before the MH's registration arrives, the packets would be sent to the last foreign network the MH has just left (line 10), and the last FMS forwards the incoming packets to the current FMS (line 11 and 12) upon receiving the registration notice (line 3).

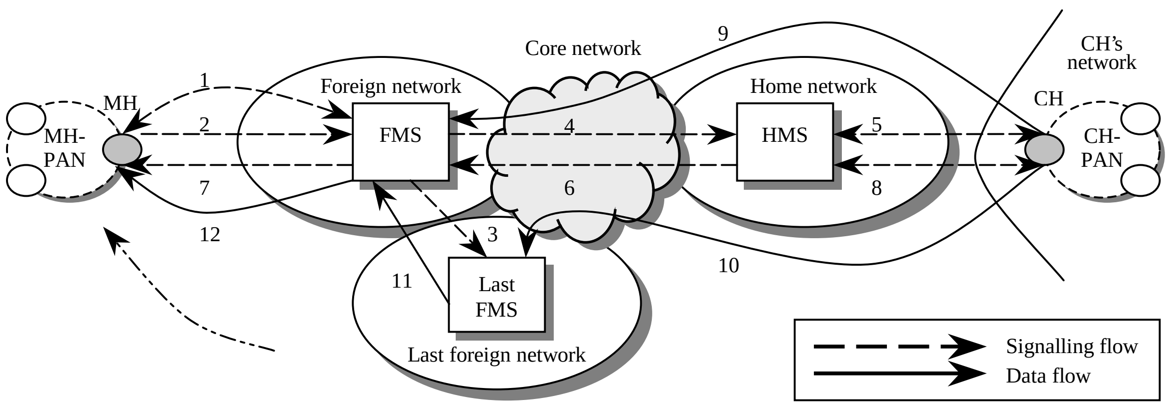

Figure 3. Location updates signalling and data flow

For registration, the rich SIP registration fields can be utilised to provide information more than just IP address. To gain separate IP addresses for each member of the MH-PAN, an aggregate message, rather than separate ones, should be generated from the serving MH on behalf of the PAN. Subsequent session handoffs would then be fulfilled by SIP re-inviting the CH-PAN. (It seems possible to avoid the end-to-end signalling with co-operation between FMS and the MH-PAN.) Sometimes, it is not desired to obtain an IP address for each member in a PAN, and the members of the MH-PAN share an IP address and thus the PAN appears as a single node to the rest of the system. In this case, a seamless session handoff is possible since there is no need for the SIP end-to-end re-invitation messaging.

4.3 Locating or Paging

Locating. In the mobile Internet context, each end host could be precisely identified by SIP personal mobility support feature, and thus paging may not be needed.

Paging. An IP-level paging scheme would be beneficial for large-scale wireless networks. Most IP-paging schemes are MIP-based and can thus be categorised as home agent based and foreign agent based. Clearly, in most cases, the latter has higher cost-effectiveness. The same principle applies to our architecture as well and thus we prefer an FMS-based routing protocol. It is the Paging Initiator who starts the paging procedure. Notably, paging algorithms are basically independent of the paging protocol. In addition, a domain distributed router based scheme is also possible as the HAWAII paging protocol shows (11).

5. CONCLUSION

Mobile IP and SIP together will play important roles in supporting mobile multimedia services in the next-generation wireless systems. They complement each other for a complete solution to all kinds of mobility types. An integrated approach is thus favoured to efficiently combine the capabilities of both protocols. This paper focuses on advanced location management issues when this integrated approach adopted. Architectural considerations were presented to accommodate various mobility scenarios. In particular, the framework of integrated mobility servers capable of both SIP and MIP was discussed. We also addressed the two major tasks of location management in case of PANs interfacing an infrastructure network. Our discussions will serve as a framework towards the ultimate goal of future pervasive personal communications.

Future work will be carried out to detail all the discussed messages and to interface the components in the mobility servers for more mobility scenarios.

REFERENCE

1. Wang Q and Abu-Rgheff M A, 2003, "Next-generation mobility support", IEE Communications Engineer, 14, 16-19

2. Rosenberg J, Schulzrinne H, Camarillo G, Johnston A, Peterson J, Sparks R and et al, 2002, "SIP: session initiation protocol", RFC 3261

3. Schulzrinne H and Wedlund E, 2000, "Application-layer mobility using SIP", MC^2R, 4, 47-57

4. Perkins C, 2002, "IP mobility support for IPv4", RFC 3220

5. Johnson D B, Perkins C and Arkko J, 2003, "Mobility support in IPv6", Internet draft, draft-ietf-mobileip-ipv6-21.txt (work in progress)

6. Kwon T T, Gerla M, Das S, and Das S, 2002, "Mobility management for VoIP service: Mobile IP vs. SIP", IEEE Wireless Commun, 9, 2-11

7. 3G TR 23.923 V.3.0.0, 2000, "Combined GSM and Mobile IP mobility handling in UMTS IP CN"

8. Wang Q and Abu-Rgheff M A, 2003, "A multi-layer mobility management architecture using cross-layer signalling interactions", Proc. IEE EPMCC'03, 237-241

9. Perkins C. and Johnson D B, 2002, "Route optimisation in Mobile IP", Internet draft, draft-ietf-mobileip-optim-12.txt (work in progress)

10. Zhang X, Castellanos J G, and Campbell A T, 2002, "P-MIP: paging extensions for Mobile IP", MONET, 7, 127-141

11. Ramjee R, Li L, Porta T L And Kasera S, 2002, "IP paging service for mobile hosts", WINET, 8, 427–441

12. Pei G and Gerla M, 1999, "Mobility management in hierarchical multi-hop mobile wireless networks", Proc. IEEE ICCCN'99, 452-456.

13. Jönsson U, Alriksson F, Larsson T, Johansson P, and Maguire Jr G Q, 2000, "MIPMANET - Mobile IP for mobile ad hoc networks", Proc. IEEE/ACM MobiHoc'00, 75-85

14. Moh M, Berquin G and Chen Y, 1999, "Mobile IP telephony: mobility support of SIP", Proc. IEEE ICCCN'99, 554 -559

15. Turányi Z R, Szabó C, Kail E and Valkó A G, 2000, "Global internet roaming with ROAMIP", MC^2R, 4, 58-68