A GENERALIZED MOBILITY SOLUTION USING A DYNAMIC TUNNELING AGENT

Subir Das, Archan Misra, Anthony Mcauley, Ashutosh Dutta
Telcordia Technologies, 445 South Street, Morristown, NJ 07960,USA
{subir,archan,mcauley,adutta}@research.telcordia.com

Sajal K. Das
Center for Research in Wireless Mobility and Networking (CReWMaN)
Department of Computer Science and Engineering
The University of Texas at Arlington
P.O. Box 19015, Arlington, TX 76019-0015, USA
das@cse.uta.edu

Abstract

This paper presents a tunneling agent-based mobility architecture which supports routing of datagrams to mobile nodes inside a mobility domain. The novelty of this architecture is that it allows users the flexibility to choose their own mobility protocols. It essentially uses the recently proposed base TeleMIP's two-level management framework with a new logical entity, called the Dynamic Tunneling Agent (DTA). *The architecture manages mobility by using two care-of addresses. While the global care-of address identifies the mobile's current domain, the local care-of address identifies the mobile's current point of attachment. By using DTA dynamically, this approach also offers several advantages over existing approaches.*

I. Introduction

The functionality needed in an IP-based network to support user roaming varies, depending on the network, movement, and application characteristics. Three basic functions that are typically required: registration, configuration and dynamic address binding. Registration allows roaming users to rapidly and automatically register their presence and requirements with the network. Configuration allows a network to automatically configure roaming users with the appropriate network-specific configuration information, such as temporary addresses and server identities. Dynamic binding allows corresponding nodes to locate roaming users and to allow continuous communication as the user moves among networks. Depending upon the application and network characteristics, these functions can be achieved either by a single protocol or by multiple protocols. In networking literature, these three operations are collectively known as *mobility management.*

Developing such an IP-based mobility management scheme for next (3^{rd} and 4^{th}) generation cellular environments is important to enable seamless integration with conventional IP-based data networks. By performing the location management functionality at the IP layer one can make the management scheme independent of the underlying link layer as well as (largely proprietary) physical layer technologies. This ensures that a uniform mobility management scheme can be deployed and maintained, even if the underlying layers are progressively upgraded. There are two key features that distinguish 3G/4G cellular architectures from currently deployed schemes:

- Packet-based services for all communications will not just be restricted to typical connectionless data-delivery services (e.g., Cellular Digital Packet Data system (CDPD) [1] or General Packet Radio Service (GPRS) [2]) but will involve real-time multimedia applications as well.
- Unlike current networks, where voice and data are managed using different entities such as Mobile Switching Centers (MSCs) and Inter-Working Functions (IWFs), a unified architecture will be used for network management, including the management of user mobility.

A. Existing Mobility Solutions

The standard IP-based mobility management scheme, Mobile IP [3] was designed primarily for the transparent support of non-real time data applications, without requiring changes in non-mobile nodes. In 3G/4G networks, however, voice (and possibly video), both with real-time requirements, are expected to be important traffic types. Research has shown why the simple Mobile IP scheme [3] is inappropriate for providing real-time support for several real-time multimedia IP-based applications, such as Voice-over-IP [4], [5], [6]. Also, the signaling overhead associated with Mobile IP-based location management techniques can be quite high, especially in cellular networks where the number of mobile nodes is quite large. In [7], it has been shown why a naive Mobile IP-based mobility management architecture may not easily support additional user management objectives in commercial telecommunication networks.

Besides reducing the latency as well as the frequency of IP-based location updates, it is also important to consider the network management and security issues in a multi-vendor environment. Important operational and architectural changes are thus required in Mobile IP [3] before it can be deployed for real-time mobility support in a commercial network. Recently, modifications to the base Mobile IP schemes [5], [6], [8], [9] have been proposed. Such modifications introduce the notion of management hierarchy; such a hierarchical approach enables support for fast intra-domain handoffs in cellular environments. Global signaling traffic is generated only when the mobile changes domains.

Since Mobile IP employs a flat mobility management architecture, only one care-of address (an address that specifies the mobile's current point of attachment) is adequate to ensure proper packet redirection. HAWAII [5] and Cellular IP [6] use a single care-of address to support the management hierarchy but they establish and maintain host-specific routes within a domain. Although Regional Registration [8] uses two different care-of address to manage intra-domain mobility, it is based on extensions to the basic Mobile IP.

On the other hand, Session Initiation Protocol (SIP) [10] is gaining more and more popularity in the IP world. SIP has been standardized within the Internet Engineering Task Force (IETF) as the protocol for inviting users to multicast conferences and Internet telephony calls. The basic SIP specification can establish, modify and terminate multimedia sessions or calls; SIP also supports personal mobility [4]. For real-time traffic, such as voice or video, SIP uses Real-Time Transport Protocol (RTP) over UDP. Researchers [4], [11] are presently trying to extend SIP for handling terminal mobility. Being an application layer approach, SIP offers a simple and scalable solution whereby users can utilize knowledge about the traffic and application profiles to make intelligent decisions on how to handle mobility in different situations. However, SIP mobility in its present form [4], [11] does not support TCP connections without Mobile IP or tunneling.

In view of the above limitations, we propose in this paper a generalized mobility solution using a new logical entity called *Dynamic Tunneling Agent (DTA)* for next generation wireless networks. The Mobility Agent (MA) based TeleMIP [7] architecture can be functionally viewed as a subset of DTA. More precisely, DTA can be viewed as a generalization of IDMP's (Intra-Domain Mobility Management Protocol) [12] Global Co-located (GC) mode of operation. The proposed architecture allows users the flexibility to choose their own mobility binding protocols. This approach is motivated by our desire to permit an user the flexibility of using Mobile IP, SIP (or indeed any other binding protocol) to bind their attachment to the network. The rest of the paper is organized as follows: Section II describes the mobility architecture; Section III discusses the protocol operation and its advantages over existing approaches. Finally, Section IV concludes the paper with some future issues.

II. Mobility Architecture using DTA

The generalized mobility architecture separates the network registration and configuration from binding and provides the flexibility to the network as well as users to adapt or interwork with different binding protocols. It also uses the strength of the TeleMIP [7] architecture by introducing a 2-layer mobility management hierarchy. By separating the local mobility management from global location updates, this architecture interworks with any global management schemes such as Mobile IP and SIP etc.

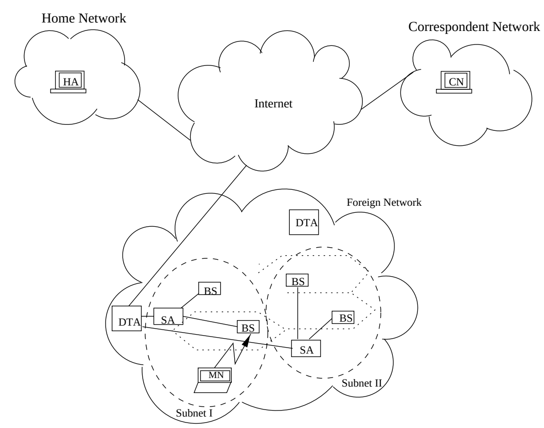

Fig. 1. Functional Mobility Architecture

The architecture uses the base TeleMIP architecture as reference and a new logical entity, called the *Dynamic Tunneling Agent (DTA)* (Figure 1). DTA not only provides a single stable point of attachment to all Mobile Nodes (MNs) under its serving area but also can assign unique globally reachable address to each and every MN from its global pool of addresses. While the DTA is functionally similar to Home Agents (HAs), it is located at a higher level in the network hierarchy than the subnet-specific agents (e.g., FA) and always tunnels the packets directly to the MN's local CoA. For intra-domain mobility, this scheme allows for the use of private or locally-scoped addresses. For example, it may assign private local addresses to users with Mobile IP for intra-domain roaming and thereby reduce the consumption of public addresses. The architecture also supports the use of multiple DTAs with various functionalities for an efficient use of public address space, redundancy and load balancing.

The architecture assumes that each subnet has at least one Subnet Agent (similar to FA [3]) or DHCP/DRCP server [13], [14]; these entities are similar to those described in [7]. Each of these elements must be associated with at least one DTA in that domain. When an MN registers in a new domain, it will receive two care-of addresses (CoAs): a domain-specific CoA (which is nothing but DTA's address) and a subnet-specific CoA for roaming in a particular subnet. Once the MN is attached to the local network, it will register with the DTA. Depending upon the need, DTA may assign the MN a new unique globally reachable domain-specific care-of address. MN will then replace the old DTA's address with the new one and update the location servers (e.g., SIP proxy for SIP users or HA for Mobile IP users). As long as the mobile roams within the domain, all future correspondence from CNs will be directed towards the domain-specific address; this assumes that CNs have the ability to bind the MN's global care-of address (GCoA) and use it in the destination field in the IP header. A DTA intercepts the packet destined to the MN and subsequently forwards the packet via tunneling (for IPv4) or via routing header modification (for IPv6). Therefore, the motion of the mobile between different subnets inside a domain will be transparent to the corresponding node.

III. Protocol Description

As described earlier, when an MN enters a DTA-based mobility domain, it receives two care-of addresses from its first subnet-level anchor point: one is for local connectivity (local care-of address, LCoA), and the other one is for global connectivity (global care-of address, GCoA). While the LCoA identifies the subnet which the MN is currently attached, GCoA specifies the DTA's interface address. There are two common ways to assign the addresses:

- Subnet Agent (SA) mode: the address is assigned using Agent Advertisements (similar to Mobile IP's FA mode). Hearing this beacon, MN sends a registration request to the SA. SA then assigns two addresses via registration reply message. The LCoA is usually SA's interface address and GCoA is DTA's interface address. In this mode, SA is responsible for demultiplexing all packets tunneled to this care-of address and then forwarding them to the individual mobile using layer-2 mechanisms. Nodes detect their subnet change via Agent Advertisements that are periodically broadcast by the SAs.
- Co-located (CO) mode: the address is assigned using additional subnet configuration protocols, such as DHCP [13] or DRCP [14]. This mode is identical to the co-located mode of Mobile IP. DHCP/DRCP servers assign two addresses with the OFFER message in response to the DISCOVER message from MNs. In this mode, an MN is responsible for decapsulating packets destined to it; also, in the absence of Agent Advertisements, additional layer-2 or layer-3 triggering support is needed to provide the MN the notification of subnet change.

Once the MN obtains configuration parameters (either from SA or DHCP/DRCP server), it registers with the DTA. Depending upon the application or based on QoS-related considerations, the DTA may choose to assign a new unique globally reachable domain-specific care-of address (GCoA) to the MN from its pool of addresses via the Registration Reply. The MN then replaces the old DTA's address with the new one and updates the location servers (e.g., SIP proxy for SIP users or HA for Mobile IP users).

A. Protocol Flow

While Figure 2 represents the message flow in Subnet Agent mode, Figure 3 describes the Co-located mode of operation. 'UPDATE' to location servers is necessary only when MN changes its GCoA. Figure 4 shows the packet re-direction header format for both cases. Solid lines from MN and SA to DTA represent using SA-mode with subnet-level private LCOA and dashed lines from MN to DTA and CN represent using CO-mode of operation as well as domain level private LCoA.

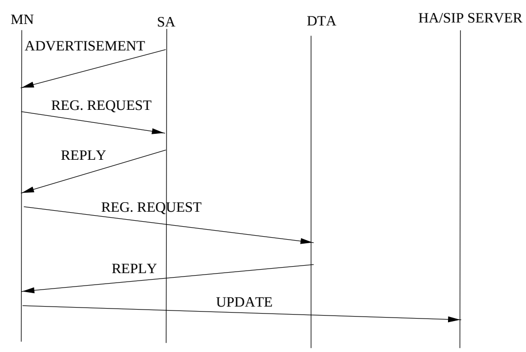

Fig. 2. Message Flow in Subnet Agent Mode

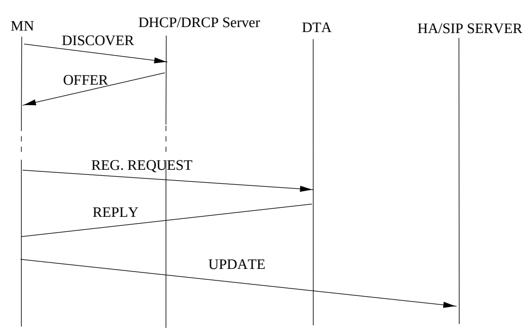

Fig. 3. Message Flow in Co-located Mode

B. Advantages

Unlike existing approaches, packets from external nodes (for example, from Mobile IP HAs) need not be tunneled to DTA's global care-of address. This is useful in certain mobility architectures [4] where users are not identified by their permanent home address or they may

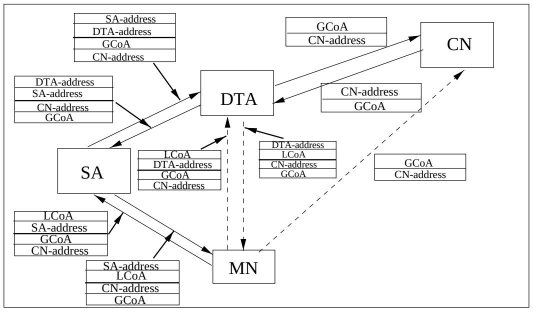

Fig. 4. Packet Header Format

not have permanent IP address. Furthermore, due to security and other QoS related issues many operators may not like the tunneled mode of operation. By introducing DTAs dynamically in the network, this architecture overcomes the following existing limitations:

- *Tunneling:* Tunneling is not necessary from the Correspondent Nodes (CNs) to DTA during intra-domain roaming (assuming the node is not using basic Mobile IP).
- *Home Address:* There is no need for permanent home address or home network for MN.
- *Global Update:* Since the architecture does not use permanent home address or network, no global update is necessary during intra-domain roaming.
- *Update Latency:* Since all binding updates are local (except inter-domain direct updates to CNs during live session), the update latency during subnet hand-offs is typically very low.
- *Routing:* Unlike approaches such as HAWAII [5], Cellular IP [6], no host-based routing is involved inside the domain. This architecture uses local tunneling from DTA to route the packets destined to the MN.

Thus the architecture alleviates many shortcomings over existing solutions. However, it requires every domain, in particular DTAs, to have a sufficiently large public address pool to allocate a separate global address to each and every registered MN. While address pool exhaustion could become an issue in IPv4, it hardly matters for IPv6 domains.

IV. Conclusion

This paper presented a generalized mobility architecture for supporting seamless mobility to roaming users in next generation networks. This approach is novel since it does not integrate binding as an integral part of the protocol and thereby offers the choice to users. By separating registration and configuration from binding, this solution provides the flexibility to the network as well as users to adapt or interwork with multiple binding protocols. Moreover, many applications, e.g., Web-browsing, do not require binding protocols whereas they need registration and configuration parameters. The *Dynamic Tunneling Agent (DTA)* is a generalized version of *Mobility Agent (MA)* described in TeleMIP [7]. DTA not only provides a single point of attachment to all MNs under its serving area but also has the ability to assign unique globally reachable address to each and every MN from its global pool of addresses. We outlined two modes of operation with the detailed protocol flow and provided a snapshot of the packet header format.

Although we highlighted several advantages over existing solutions, it would be interesting to practically implement and demonstrate these features. The most critical part is to deploy DTAs dynamically and introduce load balancing. It is worthwhile to mention here that we are currently working on several such issues [15], including fast handoff and paging support [16] in cellular environments. Moreover, we have a linux-based TeleMIP implementation in our test-bed [17].

References

[1] CDPD Forum. Cellular digital packet data system specifications. Release 1.1, January 1995.

[2] P. Rysavy, "General packet radio service (GPRS)", *GSM Data Today online Journal*, September 1998.

[3] C. Perkins, "IP mobility support", RFC 2002, Internet, October 1996.

[4] E. Wedlund and H. Schulzrinne, "Mobility support using SIP", *Proc. The Second ACM International Workshop on Wireless Mobile Multimedia*, ACM/IEEE, August 1999, pp76–82.

[5] R. Ramjee, T. La Porta, S. Thuel and K. Varadhan, "IP micro-mobility support using HAWAII", draft-ietf-mobileip-hawaii-01.txt, IETF, July 2000, Work in Progress.

[6] A. Campbell, J. Gomez, C-Y. Wan, S. Kim, Z. Turanyi and A. Valko, "Cellular IP", Draft draft-valko-cellularip-00.txt, IETF, July 2000, Work in Progress.

[7] S. Das, A. Misra, P. Agrawal and S. K. Das, "TeleMIP: Telecommunication enhanced mobile Ip architecture for fast intra-domain mobility, *IEEE Personal Communications*, August 2000, pp 50-58.

[8] E. Gustafsson, A. Jonsson and C. Perkins, "Mobile IP regional registration", draft-ietf-mobileip-reg-tunnel-03.txt, IETF, July 2000, Work in Progress.

[9] K. El Malki and H. Soliman, "Hierarchical mobile IPv4/v6 and fast handoffs", draft-elmalki-soliman-hmipv4v6-00.txt, IETF, March 2000, Work in Progress.

[10] M. Handley, H. Schulzrinne, E. Schooler and J. Rosenberg, "SIP: Session Initiation Protocol", RFC 2543, IETF, March 1999.

[11] F. Vakil, A. Dutta, J. C. Chen, S. Baba and Y. Shobatake, "Host mobility management protocol (HMMP)", Draft-ietf-itsumo-hmmp-00.txt, IETF, October 1999, Work in Progress.

[12] A. Misra, S. Das, A. Mcauley, A. Dutta and S. K. Das, "IDMP:an intra-domain mobility management protocol using mobility agents", draft-ietf-misra-idmp-00.txt, IETF, July 2000, Work in Progress.

[13] R. Droms, "Dynamic Host Configuration Protocol (DHCP)", RFC 2131, IETF, March 1997.

[14] A. Mcauley, S. Das, S. Baba and Y. Shobatake, "Dynamic Registration and Configuration Protocol (DRCP)", Draft-ietf-itsumo-drcp-01.txt, IETF, July 2000, Work in Progress.

[15] A. Misra, S. Das, A. Mcauley, A. Dutta and S. K. Das, "Introducing QoS Support in TeleMIP's mobility architecture", to appear in proceedings of ICPWC, December 2000.

[16] A. Misra, S. Das, A. Dutta, and S. K. Das, "Supporting fast intra-domain handoffs with TeleMIP in cellular environments", submitted to IEEE INFOCOM, 2001.

[17] K. Chakraborty, A. Misra, S. Das, A. Mcauley, A. Dutta, and S. K. Das, "Implementation and performance evaluation of TeleMIP", submitted to ICC, 2001.