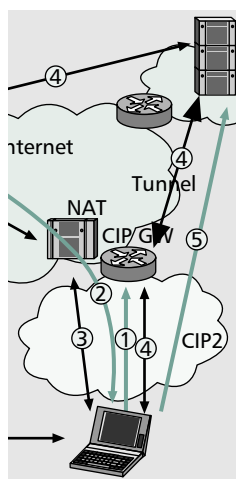


# HYBRID MULTILAYER MOBILITY MANAGEMENT WITH AAA CONTEXT TRANSFER CAPABILITIES FOR ALL-IP NETWORKS

CHRISTOS POLITIS, KAR ANN CHEW, NADEEM AKHTAR, MICHAEL GEORGIADIS, AND RAHIM TAFAZOLLI, UNIVERSITY OF SURREY  
TASOS DAGIUKLAS, AEGEAN UNIVERSITY



This article presents a multilayer mobility management scheme for All-IP Networks where local mobility movements (micromobility) are handled separately from global movements (macromobility). A hybrid scheme is also proposed to handle macromobility.

## ABSTRACT

This article presents a multilayer mobility management scheme for All-IP networks where local mobility movements (micro-mobility) are handled separately from global movements (macro-mobility). Furthermore, a hybrid scheme is proposed to handle macro-mobility (Mobile IP for non-real-time services and SIP for real-time services). The interworking between micro-mobility and macro-mobility is implemented at an entity called the enhanced mobility gateway. Both qualitative and quantitative results have demonstrated that the performance of the proposed mobility management is better than existing schemes. Furthermore, a context transfer solution for AAA is proposed to enhance the multilayer mobility management scheme by avoiding the additional delay introduced by AAA security procedures.

## INTRODUCTION

Today, there are two major technological forces that drive the communication era: wireless evolutionary systems and the Internet [1, 2]. As these forces converge, the demand for new services, increasing bandwidth, and ubiquitous connectivity continuously grows. Next-generation mobile systems will be based solely (or to a great extent) on the IP protocol.

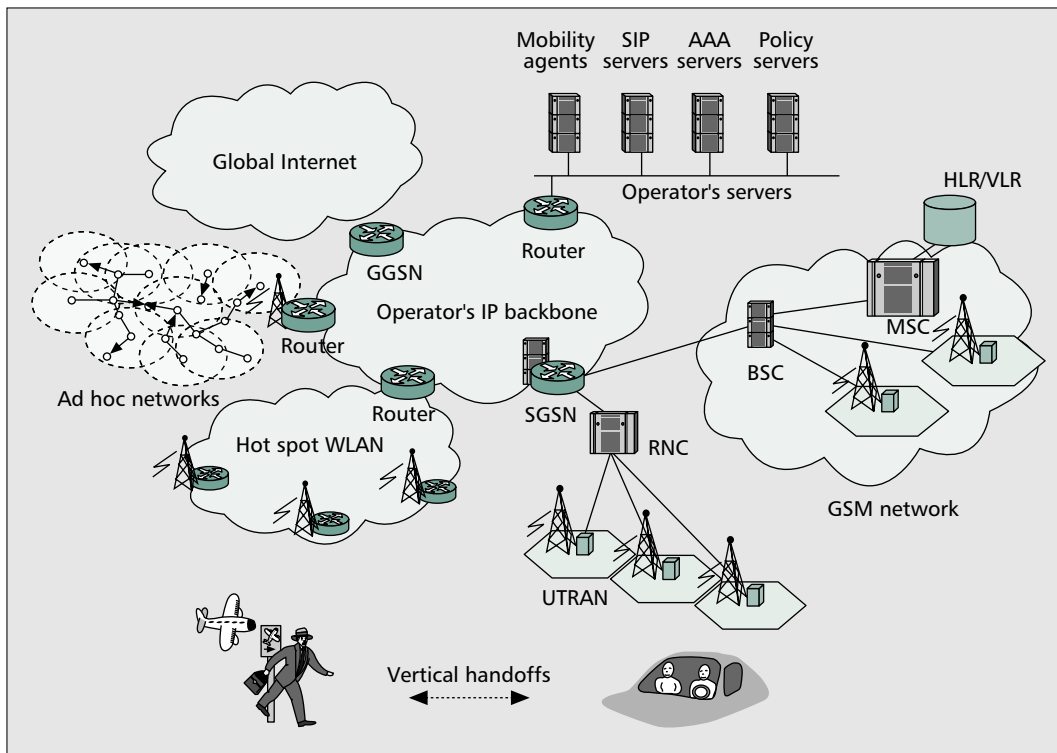
The aim of this convergence is to offer seamless multimedia services to mobile/wireless IP-based hosts across a variety of heterogeneous access technologies — Universal Mobile Telecommunications System (UMTS) radio access network (UTRAN), wireless LANs (WLANs), and emerging fourth-generation (4G) systems, for example — meeting the demands of both enterprise and public environments, anywhere and anytime [3]. As an effect, the all-IP network concept leads these developments and investigates their impact on the provisioning of real-time and non-real-time multimedia services,

such as mobile telephony, multimedia conferencing, or mobile Web access in ubiquitous environments [1, 2].

Not too long ago, communications meant voice and mobility meant cellular. But today we see that subscribers are increasingly relying on diverse communications solutions for a complex array of voice, data, and multimedia needs, many of which are being addressed by Internet/intranet connected network (e.g. in offices, homes, shopping areas, and transport facilities). What is missing is an overlying strategy for integration of these disparate solutions into what from the customers' perspective would appear as a single fabric. The core components for this integration strategy include cross-network and cross-service solutions for mobility, authentication, subscriber administration, and consolidated accounting and billing. These are all elements that today's cellular world offers better than anybody, but only for itself. The opportunity for the cellular community is to broaden its focus and associations by extending these core services to enterprise networks, Internet service providers (ISPs), public access hot spots such as airports and shopping malls, and private hot spots, such as home networks. It is foreseen that a family of seamless mobility handsets operating simultaneously in both cellular and WLAN environments will appear in the market.

Mobile IP (MIP) is the current standard for supporting the mobility of mobile users [4]. However, MIP exhibits several drawbacks. It is not efficient to handle local mobility of users in limited areas such as subnets [5, 6]. MIP also struggles with the problem of triangular routing, which introduces delay to the traffic toward a mobile host (MH), but not to that originated from the MH. For delay-sensitive traffic (e.g., VoIP or streaming multimedia services) this is not acceptable, due to the high latency in the network. Route optimization solves this problem, but on the other hand requires modifications to the IP stack of the end hosts.

With 3G systems just beginning to be deployed, it is necessary to consider how they will evolve to include a much wider range of users, applications, and economic development.



■ Figure 1. All-IP network architecture.

Furthermore, the IP-in-IP encapsulation used in MIP adds significant overhead, especially to real-time multimedia services. Apart from MIP, there are other outstanding solutions that support domain-based mobility [6]. The Session Initiation Protocol (SIP) [7] is an emerging protocol designed to provide basic call control and application-layer signaling for voice and multimedia sessions in a packet-switched network. Several wireless technical fora, such as the Third Generation Partnership Project (3GPP), 3GPP2, and Mobile Wireless Internet Forum (MWIF), have agreed on SIP utilization to provide session management and means of personal as well as service mobility. The main characteristics of SIP are its simplicity and extensibility, scalability and efficiency, and its inherent support of personal mobility.

This article proposes a solution to handle macro-mobility using a hybrid scheme (SIP for real-time and MIP for non-real-time mobile communications), while mobility within the sub-net area is handled by a micro-mobility protocol. The proposed mobility management scheme is verified and validated through both qualitative and quantitative (simulation results) analysis. Furthermore, a context scheme is proposed to support context transfer for AAA information in order to enhance the multilayer mobility management scheme. The introduction of a context transfer scheme aims to alleviate the additional delay introduced by authorization, authentication, and accounting (AAA) security procedures.

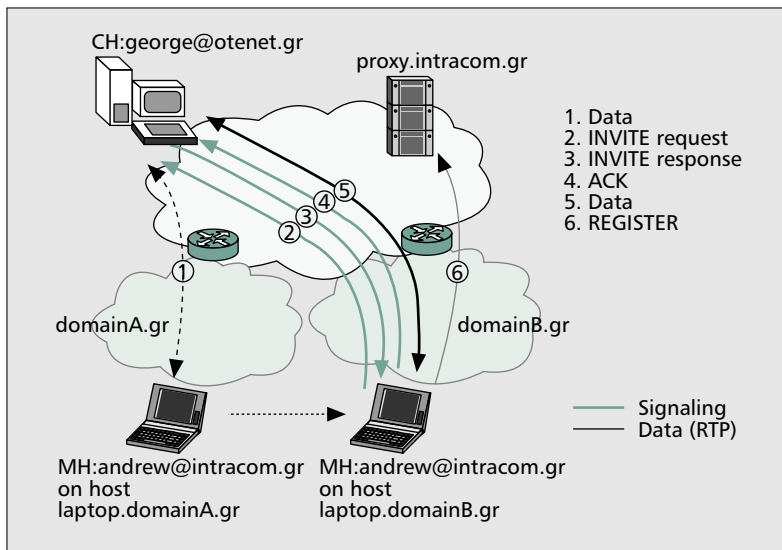
This article is organized as follows. We illustrate all-IP network architecture characteristics, and outline the requirements for mobility management in all-IP networks; we present the general characteristics of the proposed mobility

management framework associated with functional operation, qualitative evaluation, and performance evaluation through simulation; we present the context transfer scheme for conveying AAA information while the user hands off horizontally and vertically, and we draw conclusions.

## ALL-IP NETWORK ARCHITECTURES

With 3G systems just beginning to be deployed, it is necessary to consider how they will evolve to include a much wider range of users, applications, and economic development. There is no industry consensus on what systems beyond 3G will look like but, as far as the next-generation networks are concerned, ideas and concepts include [2, 3, 8]:

- Transition to an all-IP network infrastructure
- Support of heterogeneous wireless access technologies, such as UTRAN, WLANs, wireless personal area networks (WPANs), and 3G.
- Seamless handovers across both homogeneous and heterogeneous wireless technologies
- Mobility and quality of service (QoS) support at or above the IP layer
- Deployment of new protocols for services such as AAA (e.g., DIAMETER) and their interworking with existing technologies, such as the home/visited location register (HLR/VLR) and Mobile Application Part (MAP) in Global System for Mobile Communications (GSM), and Remote Authentication Dial-In User Service (RADIUS) in the Internet
- Support of different types of mobility (terminal, session and personal mobility)
- Mechanisms to support service roaming and service access while users move



■ **Figure 2.** SIP terminal mobility.

- Use of policy-based mechanisms in order to determine QoS, accounting and billing for multimedia services, while the users roam across different service providers.
- OSA/Parlay Gateway

Figure 1 illustrates a possible future all-IP network architecture. The main argument that leads toward such an integrated approach is twofold: support of innovative applications to generate new revenues, and a common transport technology to reduce costs. Considering that the entire telecommunications industry is funded out of the end user's pocket, it is straightforward that the future growth of this industry must ensure that end user's satisfaction is enriched with new services, that can be offered "anytime, anywhere, and anyhow."

## MOBILITY MANAGEMENT REQUIREMENTS IN ALL-IP NETWORKS

The most important requirements regarding mobility management of next-generation all-IP networks are the following [5, 9]:

### Support of different types of mobility:

- **Terminal mobility:** An end user's ability to use her/his own terminal regardless of location and the ability of the network to maintain the user's ongoing communication as she/he roams across heterogeneous radio cells in either the same or different administrative domains.
- **Service mobility:** The end user's ability to maintain ongoing sessions and obtain services transparently. The service mobility includes the ability of the home service provider to either maintain control of the services it provides to the user in the visited network, or transfer their control to the visited network.
- **Personal mobility:** The ability of end users to initiate and receive calls, and access subscribed network services on any terminal at any location in a transparent manner, and the ability of the network to identify end users as they move across administrative domains.

**Hierarchical topology:** This delineates the

separation of global from local mobility. The term *global mobility* refers to the movement of mobile hosts across different networks/domains (interdomain mobility), whereas the term *local mobility* is used to describe the movement within a specific subnet. For local (intradomain) mobility, it is important to differentiate active from idle systems to improve performance and scalability.

**Multilayer mobility:** Currently, the Internet Engineering Task Force (IETF) is standardizing the Mobile IP protocol to support dynamic mobility across Internet domains for MHs. There are two variations of MIP for IPv4 and IPv6 networks, respectively [10, 11]. However, MIP is struggling with the problem of triangular routing, which adds delay to the traffic toward an MH, but not that from the MH. For delay-sensitive traffic (e.g., voice or multimedia over IP) this is not acceptable, due to the high latency in the network. MIP route optimization solves this problem, but on the other hand requires modifications in the IP stack of the end hosts [12]. To alleviate the problems associated with MIP, application layer mobility is proposed, by using the Session Initiation Protocol (SIP). SIP [7] is an emerging protocol, designed to provide basic call control and application-layer signaling for voice and multimedia sessions in a packet-switched network. Supporting both real-time and non-real-time multimedia applications in a mobile environment may require mobility awareness on or above the IP layer in order to utilize knowledge about the traffic type. Several wireless technical fora (e.g., 3GPP, 3GPP2, MWIF) have agreed on SIP utilization to provide session management and a means of personal as well as service mobility.

The main features of SIP-based terminal mobility (Fig. 2) are the following:

- No permanent addresses are required for MHs. Collocated care-of addresses on each link are used during the registration process at each domain (e.g., through DHCP server).
- There is no modification of the IP protocol stack of the end hosts.
- It alleviates the problem of triangular routing.
- It inherently supports personal mobility.

However, SIP-based mobility management applies only to real-time communications over User Datagram Protocol (UDP); it breaks TCP connections (transparent terminal mobility is not supported). Supporting terminal mobility for TCP with SIP requires a tracking agent on every MH that maintains a record of its ongoing TCP connections, as well as IP encapsulation capabilities on each correspondent host (CH) [13].

This multilayer approach can be a complete mobility management framework by employing protocols such as HAWAII, Cellular IP, or Hierarchical MIP to support local or micro-mobility.

## THE PROPOSED MOBILITY MANAGEMENT SCHEME

### GENERAL CHARACTERISTICS

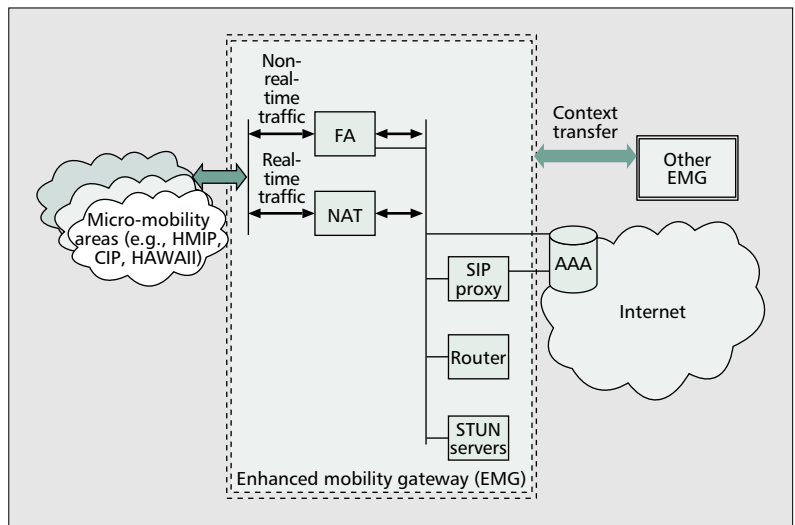
This article proposes a multilayer mobility management scheme where macro-mobility is separated and handled differently than

micro-mobility. The macromobility is based on the use of a hybrid and multiprotocol scheme, by following a multilayer and multiprotocol approach. Since there is no macromobility protocol that can meet the requirements of different types of services (i.e., loss for non-real-time services such as FTP or HTTP and delay for real-time services such as voice and multimedia), a hybrid solution is proposed that is based on the synergy, and extraction of the advantages, of existing protocols such as MIP and SIP. In the proposed scheme, SIP signaling is used to support interdomain mobility for real-time traffic (mainly Real-Time Transport Protocol, RTP, over UDP), while MIP applies to non-real-time traffic. The synergy between the two protocols is accomplished at the edge between the core and access networks at an entity called the Enhanced Mobility Gateway (EMG). Traffic from/toward an MH is separated on the domain edge routers, as illustrated in Fig. 3.

Network Address Translation (NAT) is required on the EMG, since MHs can be identified by their private home addresses within micro-mobility areas, depending on the micro-mobility protocol used (i.e., candidate protocols such as Cellular IP, HAWAII, and Hierarchical MIP). Following this approach, the IP encapsulation requirement on the end hosts is avoided. The employment of MIP foreign agents (FAs) at EMGs requires seamless interworking of MIP FAs with any micro-mobility modules (i.e., Cellular IP-CIP gateway, Hierarchical MIP-MAP). As mentioned previously, MIP supports macromobility for non-real-time traffic (e.g., TCP), which bypasses the NAT. This traffic is routed toward the MH through its home network using tunneling. The overhead introduced by IP-in-IP encapsulation is not as critical for this type of application.

However, macro-mobility for real-time traffic (i.e., real-time services) is supported with SIP signaling, as described in the previous section. The use of NAT for real-time traffic introduces problems involving blocking of the end-to-end communication because the voice and video devices behind the NAT have private IP addresses that are not routable outside their local domain or on the public Internet. This is due to the embodiment of the IP address of the calling end host in voice over IP (VoIP) signaling messages (e.g., H.323/H.225-H.245, MGCP, SIP) [14]. The problem of SIP NAT traversal is under investigation within the IETF. A number of solutions have been proposed: Simple Traversal of UDP through NAT (STUN), Traversal Using Relay NAT (TURN), SIP application layer gateways, MIDCOM protocol, and SDP extensions for NAT. Without loss of generality, the STUN approach has been chosen due to its simplicity in terms of implementation and its design methodology that does not require any modifications to the SIP servers. STUN allows entities (i.e., SIP clients) behind a NAT to first discover the presence and type of NAT, and then learn the address bindings allocated by the NAT [15].

On the other hand, mobility within a subnet area can be supported by a candidate micro-mobility protocol. Several candidate micro-mobility protocols have been proposed within



**Figure 3.** The enhanced mobility gateway separating micro-mobility from hybrid macro-mobility protocols.

the IETF (i.e., Cellular IP, HAWAII, Hierarchical Mobile IP, etc.) [16]. The integration between macro-mobility protocols (SIP and MIP) and micro-mobility protocols is accomplished through the EMG.

In the IETF the Seamoby working group (WG) is currently designing protocols that will allow real-time services to work with minimal disruption across heterogeneous wireless and wireline access technologies. Furthermore, IETF Seamoby WG is currently standardizing protocols for:

- Transferring state information between edge mobility devices (context transfer)
- Discovering candidate access routers upon handoff (handoff candidate discovery)
- Supporting IP paging (dormant mode host alerting)

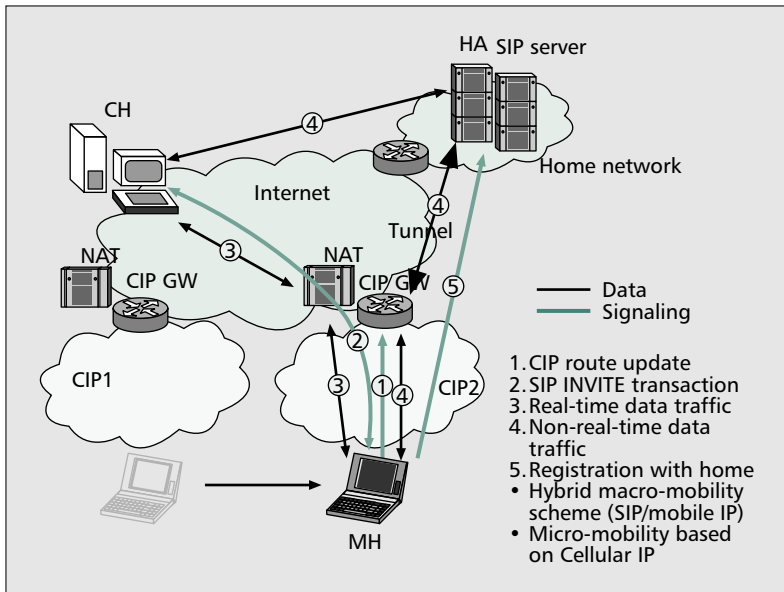
The main functionalities that could be supported by an IP micro-mobility protocol (as defined by IETF Seamoby WG) include but are not limited to IP paging, context transfer, and soft handoffs. For more information about the different micro-mobility protocols, and their comparisons can be found at [16].

Without loss of generality, this article is focused on the Cellular IP micromobility protocol. Cellular IP (CIP) has been proposed as an efficient protocol to support intradomain mobility to an MIP-enabled Internet. Some of the features of CIP include host-based routing, passive connectivity, paging, distributed location database, and use of home address to identify MHs [16, 17]. Moreover, CIP has been further enhanced with IPv6 capabilities and mechanisms that make the handover seamless and more efficient. Such capabilities include the use of IPv6 extension headers for transferring control information, route optimization, and the use of the IPv6 stateless autoconfiguration feature [18].

## FUNCTIONAL OPERATION

Traffic (i.e., TCP for non-real-time traffic and UDP for real-time traffic) from/toward an MH is separated on the EMG. The MHs are identified with their private home addresses within the





■ **Figure 4.** *The hybrid SIP/MIP approach.*

micro-mobility areas. This provides location privacy and application transparency, while the MH roams across a visited domain.

The seamless interworking of MIP with the CIP is accomplished within the EMG. As mentioned previously, MIP supports macro-mobility for non-real-time traffic (e.g., TCP), which bypasses the NAT. This traffic will be routed toward the MH through the use of tunneling. The smooth interworking between SIP and CIP necessitates the definition of a new type of control information in CIP route/paging-update packets: the SIP user identifier. It is an email-like address of the form `user@host`, where `user` is a user name and `host` is a domain name or numerical address. This information is inserted in the payload of the first route-update packet after handoff and may be repeated in a few subsequent route/paging-update packets for reliability. Upon receiving the first route-update packet, the CIP gateway (CIP GW) performs admission control to register the corresponding host at its caches (local registration).

If the mobile host moves during a session, the SIP user agent (UA) sends a SIP `re-INVITE` request message to each of its CHs. In this message, the MH includes its original SIP user identifier in the `from` field of the SIP header. It also includes the GW's address in:

- The `Contact` field of the SIP header, in order to inform the CH where it wants to receive future SIP messages
- The `c` (connection) field of the SDP header that contains a description of the session, in order to redirect the data traffic flow toward its new location

The CIP GW acts as the point where SIP responses will reach. Furthermore, the CIP GW is equipped with a SIP message tracking agent in order to forward the SIP responses to their original destination, the CIP GW. This agent checks the CIP GW's binding caches to determine whether a SIP response message must be forwarded toward a registered MH. This forwarding is accomplished after the destination IP address has been modified. This ensures that the aforementioned SIP `re-INVITE` transaction is correctly completed upon handoff. Upon reception of the SIP `INVITE` message by the CH, IP encapsulation is used to forward data information toward the MH. The encapsulated data packets are captured by the CIP GW, which in turn decapsulates and forwards them to the recipient MH, following the CIP routing scheme. Data traffic from the MH is regularly routed without the use of tunneling. The MH completes the handoff by sending a SIP `Register` message toward a SIP server on its home network (home registration) (Fig. 4).

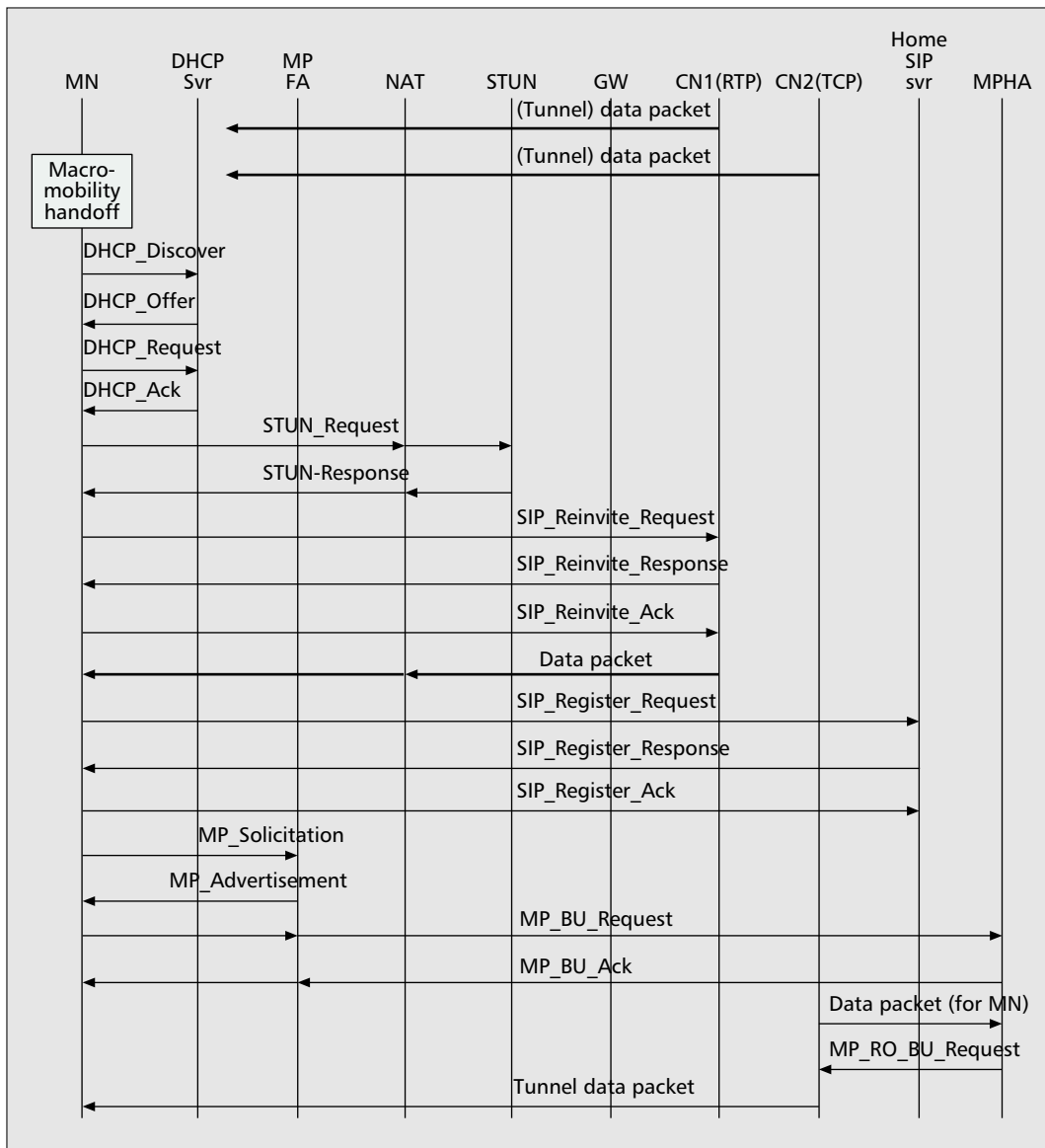
Figure 5 illustrates the signaling exchanged between two mobile nodes using the hybrid proposed nobility management scheme; one that supports real-time traffic and one that supports non-real-time traffic. Furthermore, in this scenario we use the NAT/STUN functionalities, which increase the overall signaling load. The main advantages of the proposed hybrid mobility management include the removal of tunneling (MIP features), which is quite vital for real-time services.

### QUALITY EVALUATION

Table 1 illustrates the main characteristics of the proposed schemes against other existing mobility management approaches. Similar to MIP, the proposed scheme supports transparent terminal mobility. However, unlike MIP, this scheme supports optimized routing, personal and session mobility, fast handoffs, and paging. The only difference between the proposed scheme and SIP-based mobility management is that the SIP-based mobility management architecture requires modifications in the IP stack of end hosts to support IP-in-IP encapsulation, and thus, similar to the MIP route optimization option, it will experience problems gaining wide acceptance.

	Optimized routing	Transparent terminal mobility	Personal mobility	No modifications to IP stack	No single point of failure	Seamless handoff support	Paging support
MIP		√		√			
MIP with route optimization	√	√			√		
SIP	√		√	√	√		
Hybrid SIP/Mobile IP	√	√	√	√	√	√	√

■ **Table 1.** *Comparisons of the proposed mobility management against other existing approaches.*

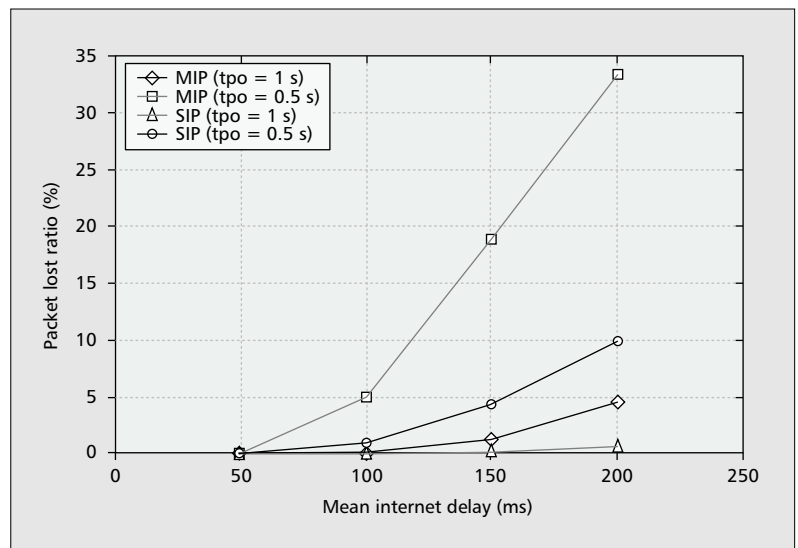


■ **Figure 5.** MSC flow for the proposed hybrid SIP/MIP mobility management scheme.

The proposed mobility management is evaluated using a simulation environment. The main emphasis has been placed on the hybrid macro-mobility scheme (MIP vs. SIP), which greatly affects the overall performance. The performance and comparisons of different micro-mobility schemes is outside the scope of this article. More information about the performance of micro-mobility protocols can be found at [16, 19].

**The Simulation Environment** — In order to evaluate the proposed mobility management architecture, we selected a set of simulation scenarios that consider the performance of real-time and non-real-time applications using the proposed mobility management against standardized solutions (MIP). Therefore, we split the mobility domains in two: the first is SIP-based, where real-time applications are treated; the second is MIP-based, where non-real-time applications are targeted.

The scenario in Table 2 was considered for this simulation set.



■ **Figure 6.** PLR for RTP session running over SIP and MIP versus delay between home and visited network (MN is static).

<p><b>Simulation setup (real-time)</b></p> <p>No fragmentation at HA  MN is in visited/foreign network  CN is another network (not MN's home network)  CN send UDP packets to MN  50 packets/s for SIP, CN sends packets directly to MN for MIP, CN sends packets to HA which then tunnel packets to MN  No handoff took place  Internet delay = 100 ms, exponentially distributed</p>	<p><b>Application configuration</b></p> <p>Start time = constant (20)  On time = constant (40)  Off time = constant (0)  Interarrival = constant (0.02)  Packet size = exp (1024) bytes</p>
<p><b>Simulation variables</b></p> <p>Internet delay  Playout delay (tpo)  Maximum playout delay set at application layer of receiver (MN).  Playout delay = 1000 ms (for streaming application) and 500 ms (for speech or voice)  Delay at the wireless link (twl): The delay as wireless link was set to 20 ms and 50 ms. This represents delay at wireless link level.</p>	<p><b>Other information</b></p> <p>If packet reach receiver (MN) at delay &gt; tpo, packet is discarded  MIP gives higher packet lost ratio because packets always take 2 steps (first to HA and then FA) to reach MN; higher jitter  Packet lost ratio = (pkt_discarded/ total_pkt_generated) × 100%</p>
<p><b>Simulation setup (non-real-time)</b></p> <p>FTP traffic is considered  MN is in visited/foreign network  Internet delays are constant  SACK version of TCP was used; packets to MN</p>	<p><b>IP configuration (non-real-time)</b></p> <p>MN: Address = 128.0.0.1, Mask=255.0.0.0  FTP server: Address = 128.0.0.2, Mask = 255.255.0.0</p>

■ **Table 2.** Simulation parameters.

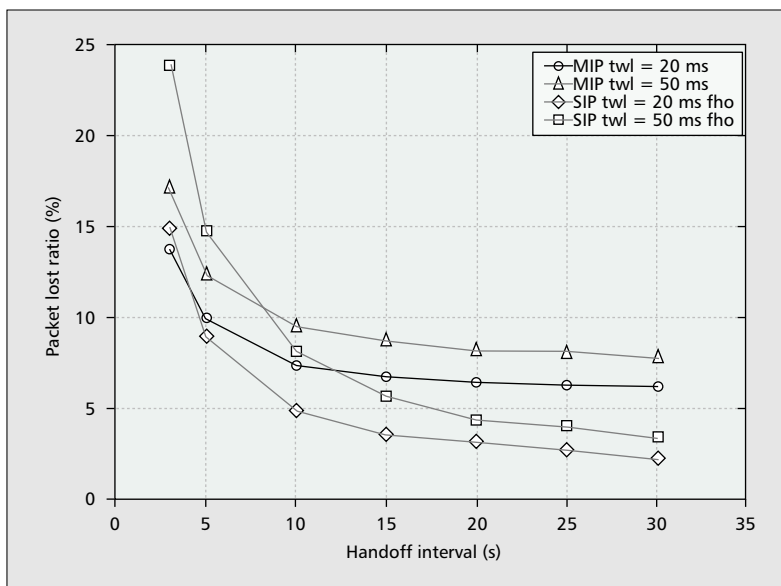
**The Impact of Mobility Management Schemes on Real-Time Services** — Figure 6 illustrates the packet loss ratio (PLR) for RTP packets for MIP and SIP mobility for both schemes vs. the mean interval delay time between the home and visited network domains. In the above scenario, the MH is relatively static. As is shown, the PLR for MIP is higher than that of SIP for two different types of multimedia applications (streaming application with tpo = 1 s and voice application tpo = 0.5 s). However, it must be emphasized that for real-time applications per-

formance can be affected by other factors such as wireless link delay, DHCP, security, and AAA operation.

In Fig. 7, as the handoff interval increases (the MH performs handoff regularly), the PLR for MIP marginally increases only as handoff frequency increases. On the other hand, SIP is more sensitive to packet loss increases for both full hand off and half hand off scenarios. In the half handoff scenario, RTP is resumed at the SIP MH when the SIP OK message is received from the CH, without waiting for reception of the SIP-Re-register-OK signaling flow from the home registrar. In the full handoff, the SIP MH resumes RTP after the SIP OK is received from the CH and home registrar. From these two figures, it is shown that SIP gives better performance for static MHs, while MIP gives better performance for MHs with frequent movement.

Figure 8 illustrates the impact of mean Internet delay on the service disruption (RTP stop time) due to handoff for both full and half handoff. It is shown that disruption time is lower for MIP than SIP for both half and full handoff. This is due to the separate address acquisition during SIP-based handoff. There are several messages exchanged during this procedure, and the MH starts to send handoff messages only after completing this procedure.

**The Impact of Mobility Management Schemes on the Performance of Non-Real-Time Services** — Figures 9 and 10 illustrate the TCP congestion window and received TCP segment numbers for both MIP and SIP mobility management schemes. Because of triangular routing in MIP, SIP performs better than MIP. However, both MIP and SIP



■ **Figure 7.** The impact of handoff interval on the PLR of multimedia services using both MIP and SIP.

employ tunneling. The former uses tunneling from the home agent (HA) to the MH, while the latter employs tunneling from the CH to the MH.

The above have been observed by varying the parameters of non-real-time applications, as shown in Table 3 and Fig. 11.

The main reasons for SIP's better performance are the following:

- All CHs interpret SIP messages.
- All CHs perform IP encapsulation.
- Support from TCP tracking (to allow TCP session to continue).
- Support from the application layer.
- Cross-layer cooperation is needed between SIP, TCP, IP, and perhaps the OS of the MH.

MIP performs better when the triangular problem is treated. Therefore, a complete and highly efficient mobility approach/framework is to split mobility support: use MIP for non-real-time services and SIP for real-time applications.

## AAA CONTEXT TRANSFER FOR SEAMLESS AND SECURE HANDOVERS

From a handoff performance perspective, one of the key issues in the development of the proposed multilayer mobility management scheme is the minimization of handoff delay when an MH roams across homogeneous/heterogeneous networks. The introduction of AAA functionalities adds an undesired delay component while the user requests network access and at the same time hands off. The delay time introduced by AAA transactions adds to the handoff latency and consequently affects ongoing sessions. During handoff, interactions between the MH and AAA servers must be avoided or at the very least reduced. Therefore, one of the main objectives is to minimize and if possible eliminate the additional delay introduced by AAA procedures.

Context transfer could facilitate this by forwarding the AAA pre-established information to the new access router (as shown in Fig. 12). In this article a context transfer solution is proposed for transferring AAA state information from the old access router to the new access router (nAR). The motivation for this stems from the benefits of avoiding re-establishment of AAA and providing an interoperable solution that works for any layer 2 radio access technology. This solution contributes to the seamless operation of application streams, minimizes packet loss, reduces delay, saves on bandwidth over the radio link, and reduces errors.

## AN OVERVIEW OF AUTHENTICATION IN WIRELESS NETWORKS

WLANs authenticate mobile users according to the IEEE 802.1x standards [20]. These standards specify how to run the Extensible Authentication Protocol (EAP). EAP provides a mechanism for supporting various authentication methods over wireline and wireless networks. An access point that supports an EAP AAA client is not required to have an under-

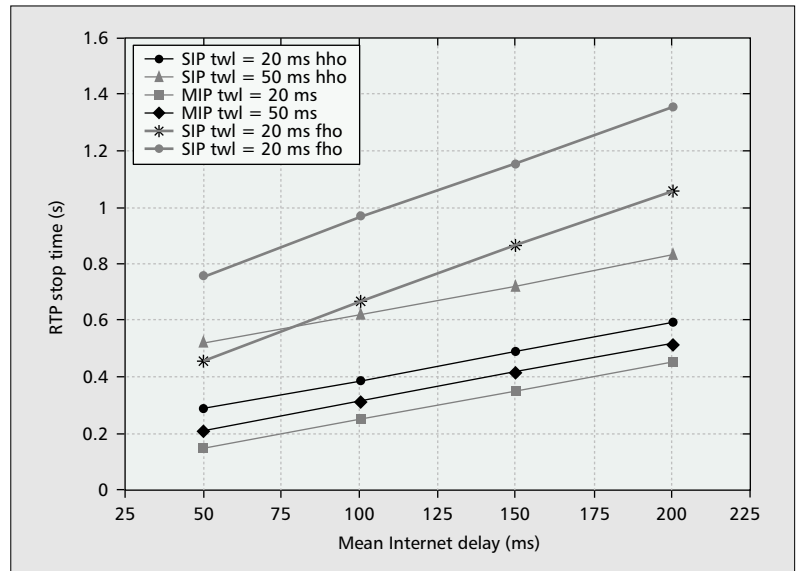


Figure 8. The impact of Internet delay on the disruption of multimedia service.

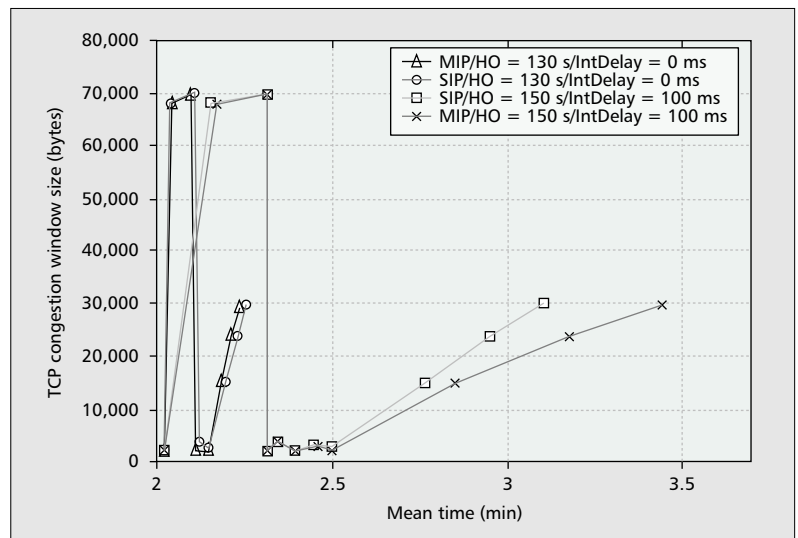


Figure 9. The impact of MIP and SIP mobility management schemes on the TCP congestion window.

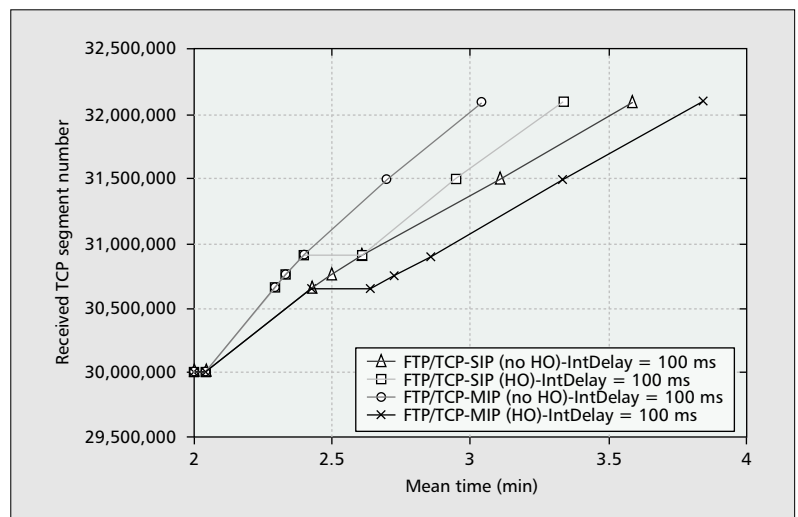
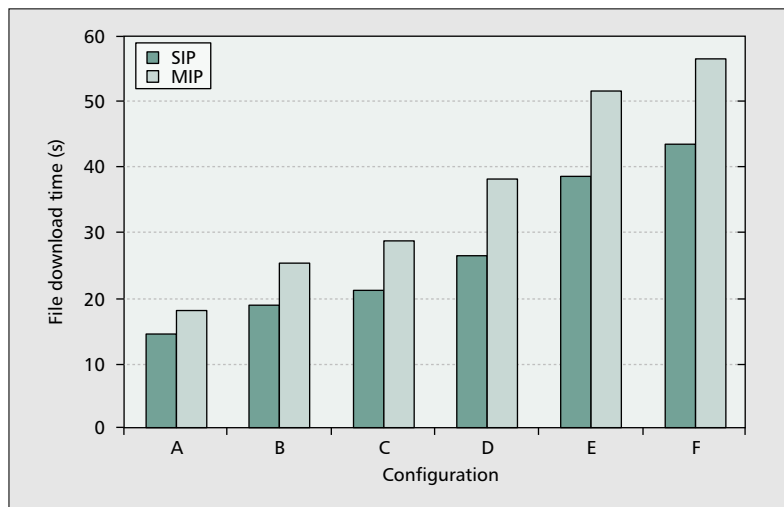


Figure 10. The impact of MIP and SIP mobility management schemes on the received TCP segment number.



Case	Setting	MIP (s)	SIP (s)
A	File size/transfer = 2 Mbytes MTU = 1500 Internet delay = 50 ms Wireless link delay = 20 ms TCP setting: SACK Buffer size = 32,768 MSS = Ethernet Nagle's SWS avoidance = enabled	18.1678	14.4211
B	Same as case A, except handoff at 129 s	25.4107	18.7808
C	Same as case A, except Internet delay = 100 ms	28.6107	21.3432
D	Same as case C, except handoff at 135 se	38.3842	26.6087
E	Same as case A, except buffer size = 16,384	52.0107	38.9429
F	Same as case E, except handoff at 144 s	56.6107	43.6087

■ **Table 3.** FTP download time for TCP settings.



■ **Figure 11.** FTP download time for different TCP settings.

standing of the specific EAP type used in the EAP authentication process. The AAA client is aware only of when the EAP authentication process starts and ends. EAP is a transport protocol used by different authentication types known as EAP methods. Among the EAP methods developed specifically for wireless networks, a family of methods based on public key certificates and the TLS protocol is included. These are EAP-TLS, EAP-TTLS, and PEAPP. Without loss of generality, the EAP-TLS method is considered [21]. There are two main elements associated with the EAP-TLS protocols:

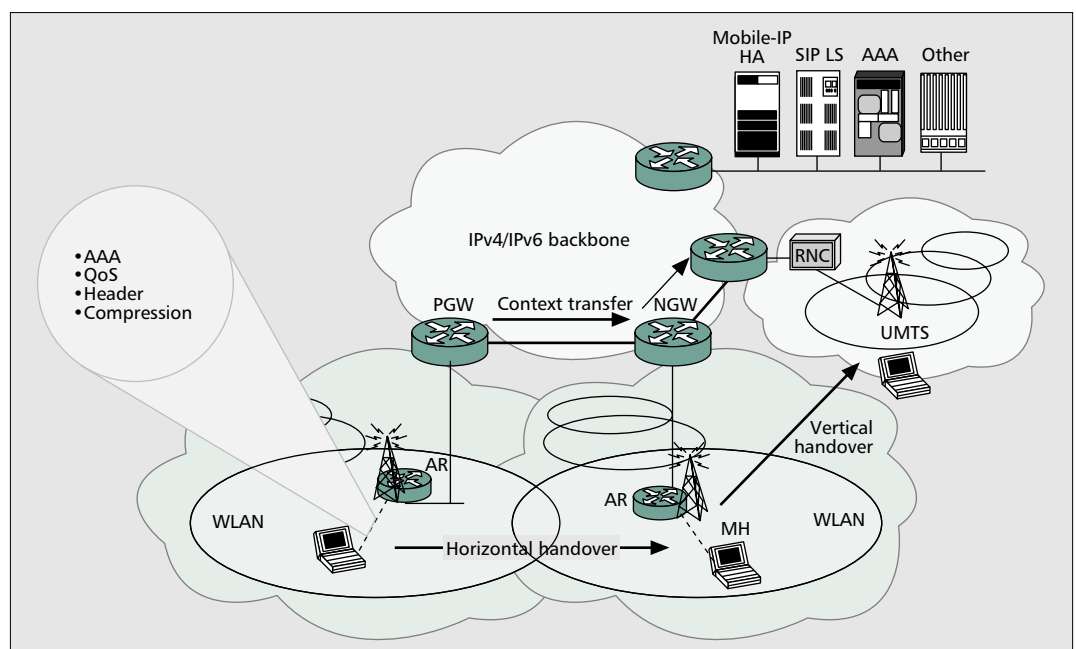
- EAP [21] allows wireless client adapters to communicate with different back-end servers, such as RADIUS or DIAMETER. Without loss of generality, RADIUS has been chosen [22].
- IEEE 802.1x, a standard for port-based network access control. EAP-TLS was the chosen protocol as it is 802.1x/EAP compliant, and supports mutual authentication and dynamic Wired Equivalent Privacy (WEP), which are essential for WLAN networks [23].

In brief the following steps take place in the EAP-TLS protocol:

- A mobile client requires a valid certificate to authenticate to a mobile WLAN network.
- The AAA server requires a server certificate to validate its identity to the clients.
- The certificate-authority-server infrastructure issues certificates to the AAA server(s) and clients.

The components involved in the 802.1x/EAP authentication process are the following:

- Supplicant (mobile user)
- Authenticator (access point)
- Authentication server (RADIUS server)
- The authenticator must support 802.1x/EAP authentication, and the supplicant and authentication server must support EAP/TLS authentication.



■ **Figure 12.** Context transfer in all-IP infrastructures.

## CELLULAR IP ENHANCEMENTS FOR TRANSFERRING AAA CONTEXT

Within a CIP domain, during a handoff from one base station (BS) or access point (AP) to another, CIP control packets could be used to initiate and transfer authorized context from the CIP-GW to the new base station (NBS). The context information will be stored at the CIP-GW, with a copy of this context (state information) forwarded to the NBS [23] (Fig. 13).

One of the main advantages of using CIP is the distinction it makes between idle and active users. This separation allows the network to follow a MH in active state from BS to BS and deliver packets without searching for the MH. By separating the caches for active and idle MHs, only a smaller cache need be searched for most of the packets, which results in faster lookups and better scalability. This CIP advantage of separating active hosts from idle MHs is also a benefit to the context transfer mechanism since it also targets active MHs.

In order to incorporate the above context transfer mechanism into the CIP protocol the following enhancements have been made [24]:

- Introduction of a context-update (CU) packet
- Introduction of a context cache at each CIP leaf node and the gateway
- Augment the cellular-IP route-update packet with a flag to indicate handoff

For interdomain handoff:

- Introduction of a context-update request (CU-Req) packet
- Introduction of a context-update reply (CU-Rep) packet

### THE AAA CONTEXT TRANSFER SOLUTION

As mentioned earlier, this article proposes an AAA context transfer solution for transferring AAA state information stored at the CIP-GW to the MH's NBS after handoff takes place. Figure 14 shows a signaling flow diagram of the EAP-TLS message exchanges after handoff between the MH, the NBS, and the RADIUS server in the absence of the context transfer solution. As shown, multiple message exchanges are required between these entities before the MH is authorized to access the network. This delay could be very large, especially if the RADIUS server resides far away from the NBS; hence, it must be avoided.

Figure 15 shows the messaging flow when the AAA context transfer solution is used. It clearly shows how the number of message exchanges is reduced and how communication with the RADIUS server is avoided.

### INTRADOMAIN CONTEXT TRANSFER SETUP

The performance of the proposed context transfer scheme was evaluated in a wireless networking testbed (WNT). Figure 16 illustrates the hardware configuration for the intradomain setup in the testbed. The EMG is running on a Linux PC. There are two CIP nodes running on Linux laptops in the setup. We are using the open source CIP implementation from Columbia University [25]. They have two network interfaces, one wired and one wireless. The wired

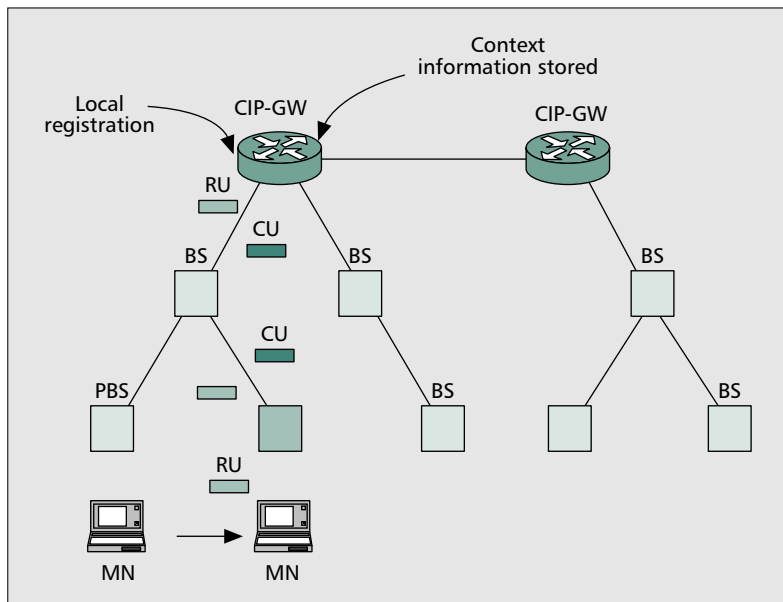


Figure 13. Context transfer enhancements of Cellular IP.

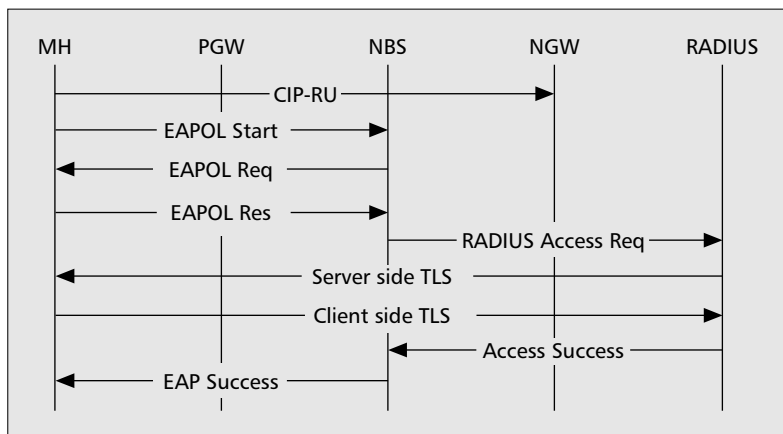


Figure 14. The signaling flow of AAA messages.

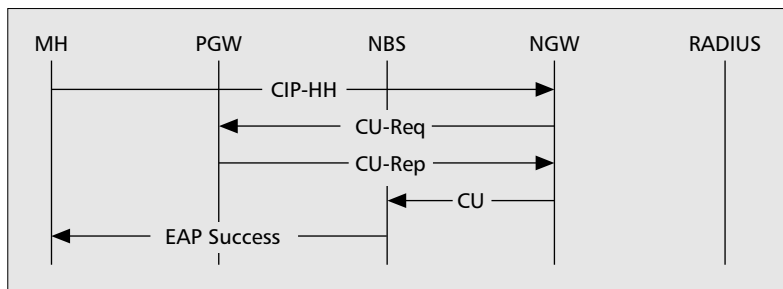
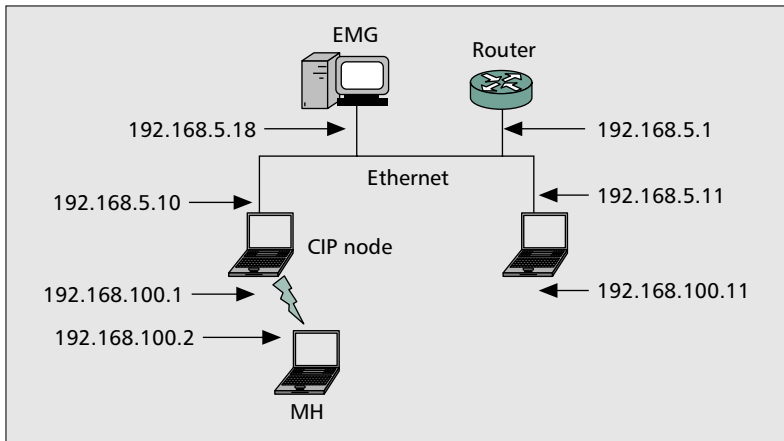


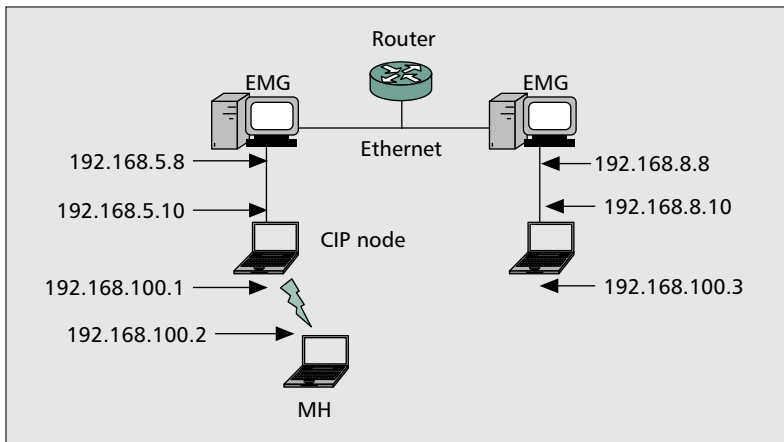
Figure 15. Signaling flow for AAA using context transfer.

interface is used to connect to the gateway, while the wireless interface serves as the AP. The AP is simply a Linksys WPC11 wireless card with the open source hostAP driver (hostap-0.0.3). The MH is a Linux laptop equipped with a wireless network interface card. The gateway is connected to the backbone via a Cisco Catalyst switch.

Handoff is initiated from the mobile node by sending a route-update packet toward the EMG (see Fig. 16). When an active node connects to an NBS, it transmits a route-update packet to



■ Figure 16. Intradomain context transfer setup.



■ Figure 17. Interdomain context transfer set-up.

Msg	Time (s)	Source	Destination	Protocol	Info
1	48.304	MH	CIP-GW	CIP	Route Update
2	50.738	MH	AP2	EAPOL	Start
3	50.74	AP2	MH	EAP	Request
4	50.748	MH	AP2	EAP	Response
5	50.753	AP2	MH	EAP	Request
6	51.538	MH	AP2	EAP	Response
7	51.739	MH	RADIUS	TLS	Client Hello
8	51.756	AP2	MH	EAP	Request
9	52.999	MH	AP2	EAP	Response
10	53.01	RADIUS	MH	TLS	Server Hello
11	54.265	MH	AP2	EAP	Response
12	54.275	AP2	MH	EAP	Request
13	55.257	MH	RADIUS	TLS	Handshake
14	55.276	RADIUS	MH	TLS	Handshake
15	56.519	MH	AP2	EAP	Response
16	56.523	AP2	MH	EAP	Success

Handoff delay = 56.523 – 48.304 = 8.219 s

■ Table 4. EAP/TLS signaling exchange (AAA context transfer disabled).

the EMG. The route-update packet will update route caches in nodes along the way from the NBS to the EMG. A new flag is introduced, called the H (handoff) flag, in the route-update packet. When the route-update packet reaches the EMG, if the H flag is enabled, the EMG will send a CU packet toward the mobile node. The CU packet, carrying the feature contexts, will be routed along the reverse path on a hop-by-hop basis toward the mobile node. When the CU arrives at the NBS, the NBS stores the context data in its context cache and discards the packet. Having received the AAA related data corresponding to the MH, the BS then authenticates the MH straightaway without waiting for the host to initiate the complete EAP-TLS messages exchange.

### INTERDOMAIN CONTEXT TRANSFER SETUP

The setup for interdomain context transfer scheme evaluation is shown in Fig. 17. In this case, there are two EMGs running on Linux PCs. Furthermore, each EMG is connected to a CIP node on the downlink. Also, the gateway is connected to the backbone through a Cisco Catalyst switch. The leaf nodes are the same as the ones used for the intradomain case. Because this setup is to be used for handoffs between domains, the two gateways are on different subnets, as illustrated in Fig. 17.

This is similar to the intradomain process with additional messages to request and forward the desired context information from the previous EMG (PEMG) to the new EMG (NEMG). When an active node connects to an NBS, it transmits a route-update packet to NEMG. When the route-update packet reaches the NEMG, if the H flag is enabled and identifies the MH as a newcomer to its domain, it requests the context information from the PEMG by sending a CT-Req packet. On reception of the CT-Req, the PEMG forwards the context information to the NEMG using a CT-Rep message. The NEMG in turn stores the context at the context cache and creates a CU packet containing the context. The CU packet, carrying the feature contexts, will be routed along the reverse path on a hop-by-hop basis toward the mobile node. When the CU arrives at the NBS, the NBS stores the context data in its context cache and discards the packet. As for intradomain handoff, the new BS authenticates the MH on the basis of the received context and hence avoids the delay due to the EAP-TLS message exchange.

### PERFORMANCE EVALUATION [26]

Table 4 shows the EAP/TLS packets captured at the MH during the authentication procedure when an interdomain handoff takes place. For this set of observations, the context transfer has been disabled; therefore, full re-authentication is required. The handoff is initiated by the CIP route-update packet with the H flag set (packet 1 in the figure). The re-authentication process is initiated with an EAPOL Start message sent by the MH to the new AP (AP2), while successful authentication is indicated by the EAPOL Success message. Using the timestamps associated with these two messages, we can find out the

time taken for a successful authentication. The time difference between the CIP route-update packet and the EAP Success packet is used to determine the time taken for the handoff from one BS to another and the subsequent re-authentication.

Altogether the handoff delay is about 8 s, which demonstrates that the EAP/TLS exchange is a significant delay component in this scenario. In contrast, Table 5 shows the handoff delay resulting when the context transfer mechanism is enabled. For this scenario the MH moves from AP2 back to AP1. As can be seen in the table, the handoff delay has been significantly reduced to only approximately 0.4 ses. In this case, again the route-update (with H flag set) indicates the handoff, and then the context transfer takes place between the new and previous gateways, followed by the “reduced” re-authentication procedure based on the received context.

Finally, the NBS (AP1) informs the MH that it has been successfully authenticated by sending the EAP Success message as indicated in Table 5. It is important to note that the improvement is almost 10 times. We have repeated this test a number of times, and it has been observed that although the actual times vary, context transfer enabled handoff is much faster than that without context transfer.

## CONCLUSIONS

This article presents a mobility management framework that can efficiently handle real-time services in all-IP networks. The presented framework proposes use of a hybrid scheme to handle macro-mobility (MIP for non-real-time services and SIP for real-time services). Interworking between micro-mobility and macro-mobility is implemented at an entity called the enhanced mobility gateway. Simulation results have shown that SIP is not worse than MIP for handling macro-mobility. For real-time applications (RTP/UDP-based) SIP performs better than MIP, while for non-real-time applications (TCP-based), SIP performs better under the following conditions: all CHs interpret SIP messages; all CHs perform IP encapsulation; support from TCP tracking (to allow a TCP session to continue), support from the application layer, and cross-layer cooperation are needed between SIP, TCP, IP, and perhaps the OS of the MH. Therefore, a hybrid solution is proposed to handle global mobility (SIP and MIP) together with the candidate local mobility protocol (CIP).

Furthermore, an innovative context transfer solution has been proposed to complement multilayer mobility management with the objective of avoiding the additional delay introduced by AAA operation. For this solution existing messages of CIP were used as triggers, and additional messages were introduced to carry the AAA context information to the appropriate base station. Based on the results shown here, the proposed AAA context transfer solution reduces the overall handoff delay by a factor of 20. This is because the full EAP/TLS procedure is avoided by transferring the AAA context to the new BS, thus enabling it to re-authenticate the MH without contacting the AAA server. This work

Msg	Time (s)	Source	Destination	Protocol	Info
1	59.786	MH	CIP-GW	CIP	Route update
2	60.167	AP1	MH	EAP	Success

Handoff delay = 60.167-59.786 = 0.381 s

■ **Table 5.** EAP/TLS signaling exchange (AAA context transfer enabled).

demonstrates how the context transfer mechanism improves the overall handoff performance and hence aids in realizing seamless and secure mobility management in all IP infrastructures.

## ACKNOWLEDGMENTS

This work has been performed in the framework of the IST-2001-32449 project EVOLUTE, which is partly funded by the European Union. The authors would like to acknowledge the contribution of their colleagues from Intracom, FhG Fokus, Alcatel-SEL, Motorola UK, University of Surrey, and Cefriel. The authors would also like to acknowledge the contributions of their colleagues Dionisios Gatzounas and Dimitrios Theofilatos.

## REFERENCES

- [1] R. Berezdivin, R. Breining, and R. Topp, “Next-Generation Wireless Communications Concepts and Technologies,” *IEEE Commun. Mag.*, vol. 40, Mar. 2002, pp. 108–16.
- [2] F. Sestini, J. Da Silva, and J. Fernandez, “Expanding the Wireless Universe: EU Research on the Move,” *IEEE Commun. Mag.*, vol. 40, June 2003, pp. 132–40.
- [3] T. Oda *et al.*, “White Paper on Cooperative Networks (CoNet) Research Challenges and Priorities,” *World Wide Research Forum*, 9th mtg., Zurich, Switzerland, July 2003.
- [4] H.-Y. Lach and C. Janneteau, “Network Mobility in Beyond-3G Systems,” *IEEE Commun. Mag.* vol. 41, July 2003, pp. 52–57.
- [5] D. Theofilatos, T. Dagiuklas, and D. Gatzounas, “Mobility Management for Supporting Multimedia Services to All-IP Network Infrastructures,” *5th Euro. Pers. Mobile Commun. Conf.*, Glasgow, Scotland, Apr. 2003.
- [6] EVOLUTE D2.1 Deliverable, “Architecture Specification,” Aug. 2002.
- [7] “SIP: Session Initiation Protocol,” IETF RFC 3261, June 2002.
- [8] D. Wisely *et al.*, *IP for 3G: Networking Technologies for Mobile Communications*, John Wiley, 2002.
- [9] T. Dagiuklas, D. Theofilatos, and D. Gatzounas, “Supporting Multimedia Services for Mobile Users in All-IP Based Networks: Requirements, Functions and Issues,” *Wireless World Research Forum (WWRF)*, London, U.K., June 2002.
- [10] “IP Mobility Support for IPv4,” IETF RFC-3220, Jan. 2002.
- [11] C. Perkins and D. Johnson, “Mobility Support in IPv6,” Internet draft, Mar. 2002.
- [12] C. Perkins and D. Johnson, “Route Optimization in Mobile IP,” draft-ietf-mobileip-optim-11.txt, work in progress, Sept. 2001.
- [13] C. E. Wedlund and H. Schulzrinne, “Mobility Support using SIP,” *2nd ACM/IEEE Int’l. Conf. Wireless and Mobile Multimedia*, Seattle, WA, Aug. 1999.
- [14] J. Rosenberg *et al.*, “SIP Traversal over NATs Scenarios,” Internet draft, work in progress, Mar. 2002.
- [15] “STUN — Simple Traversal of UDP over NATs,” IETF RFC 3489, Mar. 2003.
- [16] A. Campbell *et al.*, “Comparison of IP Micromobility Protocols,” *IEEE Wireless Commun.*, no. 35, Apr. 2002, pp. 72–82.
- [17] A. Campbell *et al.*, “Cellular IP,” IETF Internet draft, work in progress, Jan. 2000.
- [18] Z. Shelby, D. Gatzounas and A. Campbell, “Cellular IPv6,” Internet draft, work in progress, Nov. 2000.
- [19] K. A. Chew, C. Politis, and R. Tafazolli, “Performance Evaluation of Micro-mobility Protocols for All-IP Based Infrastructures,” *WWRF*, Eindhoven, THE Netherlands, Dec. 2002.
- [20] IEEE 802.11 Working Group, <http://grouper.ieee.org/groups/802/11/index.html>

This work demonstrates how the context transfer mechanism improves the overall handoff performance and hence aids in realizing seamless and secure mobility management in all IP infrastructures.

- [21] Cisco Systems, Inc., "Extensible Authentication Protocol Transport Layer Security Deployment Guide for Wireless LAN Networks," White Paper, 2002.
- [22] "RADIUS," IETF RFC 2058, Jan. 1997.
- [23] "PPP EAP TLS Authentication Protocol," IETF RFC 2716, Oct. 1999.
- [24] M. Georgiades, C. Politis, and R. Tafazolli, "Context Transfer Extension to Cellular-IP," Internet draft, work in progress, Dec. 2003.
- [25] [http://www.comet.columbia.edu/cellularip/linux\\_src\\_code.htm](http://www.comet.columbia.edu/cellularip/linux_src_code.htm)
- [26] M. Georgiades et al., "AAA Context Transfer for Seamless and Secure Multimedia Services over All-IP Networks," 5th Euro. Wireless Conf., Barcelona, Spain, Feb. 24–27, 2004.

### BIOGRAPHIES

CHRISTOS POLITIS [M] (c.politis@surrey.ac.uk) received his engineering degree from the Technological University of Athens, Greece, in 1996, his M.Sc. in mobile and satellite communications from the University of Surrey (UNIS), United Kingdom, in 1999, and his Ph.D. in mobile networking from the Center for Communication Systems Research (CCSR) at the same university in 2004. Past positions include telecommunications engineer with INTRACOM SA, Athens, IT engineer at AMSAT, and wireless communications engineer at Hellenic Air Force General Staff. He is currently a senior researcher in CCSR, UNIS. He is involved in the EU-funded IST EVOLUTE, STRIKE, and MAGNET projects. He is a patent holder and has published more than 30 papers in international journals and conferences. His research interests include wireless LANs, mobility management, security and QoS frameworks for IP-based infrastructures, and personal area networks.

TASOS DAGIUKLAS [M] (ntan@aegean.gr) received his Engineering degree from the University of Patras, Greece, in 1989, an M.Sc. from the University of Manchester, United Kingdom, in 1991, and a Ph.D. from the University of Essex, United Kingdom, in 1995, all in electrical engineering. Currently, he is employed as teaching staff at the University of Aegean, Department of Information and Communications Systems Engineering. Past positions include senior posts at INTRACOM and OTE, Greece. He has been involved in more than 12 EC R&D research projects in the fields of all-IP network and next-generation services. His research interests include all-IP networks, systems beyond 3G, and multimedia services over fixed and mobile networks. He has published more than 50 papers in international journals and conferences in the above fields. He is a member of the Technical Chamber of Greece.

KAR ANN CHEW (k.chew@surrey.ac.uk) received his Ph.D. in mobile communication from UNIS in 2004. He was awarded a B.Eng. in electronics and electrical engineering from the same university in 1998, and a Diploma in electronics and communication from Singapore Polytechnic in 1996. He has been a research fellow at CCSR, UNIS since 2000, working in research programs supported by Mobile Virtual Center of Excellence (VCE), United Kingdom. From 2000 to 2002 he led a team of researchers working in the area of mobility management in all-IP networks. He is now leading another team investigating various issues in interworking of mobile and broadcasting networks.

NADEEM AKHTAR (n.akhtar@surrey.ac.uk) received a B.Sc. (Eng.) degree in electronic engineering from Aligarh Muslim University, India, in 1998 and an M.E. degree in telecommunications from the Indian Institute of Science, Bangalore, in 2000. He worked as a software engineer in Motorola India Electronics Ltd., India from 2000 to 2001. Since 2001 he has been working as a research fellow at CCSR, UNIS, and also studying for a doctorate degree. His current research interests include mobility management in all-IP networks and routing in mobile ad hoc networks.

MICHAEL GEORGIADES (m.georgiades@surrey.ac.uk) received a B.Eng. degree in communications and radio engineering from King's College London in 2000 and an M.Sc. degree in telecommunications from University College London in 2001. In 2002 he joined INSIG Ltd. as a systems development engineer for wireless Internet solutions. In the same year, he joined CCSR, UNIS as a research fellow, where he is also studying for a Ph.D. degree. His current research interests include mobility management and security issues in IP-based networks. He has been working for IST EVOLUTE and Ambient Networks projects.

RAHIM TAFAZOLLI (r.tafazolli@surrey.ac.uk) head of the Mobile Communications Research Group in CCSR, UNIS. His research activities are on optimization techniques for mobile multimedia networks, mainly on advance resource management, mobility management, and media access control. He has published more than 300 research papers in refereed journals and international conferences, and as an invited speaker. He currently has more than 15 patents in the field of mobile communications. He is advisor and consultant to a number of mobile companies. He is the founder and past Chairman of the International Conference on 3G Mobile Technologies. He is also a member of the IEE Committee on the UK Regulations on Information Technology & Telecommunications, a member of the WWRF Vision Committee, and past Chairman of the New Technologies group of WWRF, and academic coordinator of the U.K. Mobile VCE.