Context Transfer Support for Mobility Management in All-IP Networks

Michael Georgiades

Submitted for the Degree of
Doctor of Philosophy
from the
University of Surrey

Centre for Communication Systems Research
School of Electronics and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, U.K.

May 2008

Summary

This thesis is a description of the research undertaken in the course of the PhD and evolves around a context transfer protocol which aims to complement and support mobility management in next generation mobile networks. Based on the literature review, it was identified that there is more to mobility management than handover management and the successful change of routing paths. Supportive mechanisms like fast handover, candidate access router discovery and context transfer can significantly contribute towards achieving seamless handover which is especially important in the case of real time services. The work focused on context transfer motivated by the fact that it could offer great benefits to session re-establishment during the handover operation of a mobile user and preliminary testbed observations illustrated the need for achieving this.

Context transfer aims to minimize the impact of certain transport, routing, security-related services on the handover performance. When a mobile node (MN) moves to a new subnet it needs to continue such services that have already been established at the previous subnet. Examples of such services include AAA profile, IPsec state, header compression, QoS policy etc. Re-establishing these services at the new subnet will require a considerable amount of time for the protocol exchanges and as a result time-sensitive real-time traffic will suffer during this time. By transferring state to the new domain candidate services will be quickly re-established. This would also contribute to the seamless operation of application streams and could reduce susceptibility to errors. Furthermore, re-initiation to and from the mobile node will be avoided hence wireless bandwidth efficiency will be conserved.

In this research an extension to mobility protocols was proposed for supporting state forwarding capabilities. The idea of forwarding states was also explored for remotely reconfiguring middleboxes to avoid any interruption of a mobile users' sessions or services. Finally a context transfer module was proposed to facilitate the integration of such a mechanism in next generation architectures. The proposals were evaluated analytically, via simulations or via testbed implementation depending on the scenario investigated. The results demonstrated that the proposed solutions can minimize the impact of security services like authentication, authorization and firewalls on a mobile user's multimedia sessions and thus improving the overall handover performance.

Acknowledgements

I would like to thank my supervisors Dr. Klaus Moessner and Prof. Rahim Tafazolli for their valuable advise and support during the course of this work. I am also grateful to former and present members of CCSR for their friendship, encouragement, discussions and support. Finally I would like to thank my family and my friends for being there for me.

Contents

List of Figures

List of Tables

ABBREVIATIONS

2G	2nd Generation
3G	3rd Generation
3GPP	3rd Generation Partnership Project
3GPP2	3rd Generation Partnership Project 2
AAA	Authentication, Authorization and Accounting
ACK	Acknowledgement
AODV	Adhoc On-Demand Distance Vector Routing
AR	Access Router
ARP	Address Resolution Protocol
BCMP	BRAIN Candidate Mobility Management Protocol
BRAIN	Broadband Radio Access for IP Based Networks
BU	Binding Update
CaAR	Context anchor Access Router
CAR	Candidate Access Router
CARD	Candidate Access Router Discovery
CBR	Constant Bit Rate
CGA	Cryptographically Generated Address
CIP	Cellular IP
CIPv6	Cellular IP Version 6
CN	Corresponding Node
CoA	Care of Address
CT	Context Transfer
CTP	Context Transfer Protocol
CTCS	Context Transfer Candidate Service
CU	Context Update
DHCP	Dynamic Host Confiduration Protocol
DRCP	Dynamic Registration and Configuration Protocol
DSR	Dynamic Source Routing
FA	Foreign Agent
FMIPv6	Fast Handovers for Mobile IPv6
FTP	File Transfer Protocol
GCoA	Global Care of Address
GFA	Gateway Foreign Agent
GGSN	Gateway GPRS Support Node

GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
GTP	GPRS Tunnelling Protocol
GW	Gateway
HA	Home Agent
HAWAII	Handoff-Aware Wireless Access Internet Infrastructure
HIP	Host Identity Protocol
HLR	Home Location Register
HMIP	Hierarchical Mobile IP
HSE	Host-Specific Entry
HTTP	Hyper Text Transport Protocol
ICMP	Internet Control Message Protocol
IDMP	Intra-Domain Mobility Management Protocol
IEEE	Institute of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
ISP	Internet Service Provider
LA	Location Area
LAN	Local Area Network
LCoA	Locar Care of Address
LCP	Link Control Protocol
LLC	Logical Link Control
LTE	Long Term Evolution
MAC	Medium Access Control
MANET	Mobile Adhoc NETwork
MAP	Mobility Anchor Point
MARP	Mobility-Aware Routing Protocol
Mbps	Megabits Per Second
MERTORA	Mobile Enhanced Routed - Temporally Ordered Routing Algorithm
MIP	Mobile IP
MIPv4	Mobile IP Version 4
MIPv6	Mobile IP Version 6
MMP	Mobility Management Protocol
MN	Mobile Node
MPC	Mobile Policy Table
MS	Mobile Station
NAP	New Access Point
NAR	New Access Router
NetLMM	Network-based Local Mobility Management
OAR	Old Access Router

PAN	Personal Area Network
PAP	Previous Access Point
PAR	Previous Access Router
PDA	Personal Digital Assistant
PDN	Packet Data Network
PDP	Packet Data Protocol
PPP	Point-to-Point Protocol
PS	Packet Switched
QoS	Quality of service
RA	Router Area
RAI	Routing Area Identifier
RAU	Routing Area Update
RCT	Reactive Context Transfer
PCT	Predictive Context Transfer
GCT	Gateway supported Context Transfer
RFC	Request For Comments
RNC	Radio Network Controller
RR	Resource Record
RSVP	Resource reSerVation Protocol
RTP	Realtime Transport Protocol
RU	Route Update
SAE	System Architecture Evolution
SDP	Session Description Protocol
SGSN	Serving GPRS Support Node
SIP	Session Initiation Protocol
SIP UA	SIP User Agent
SLA	Service Level Agreement
SS7	Signaling System 7
STM	State Transfer Module
TCP	Transmission Control Protocol
TEXT	Time Efficient conteXt Transfer
TeleMIP	Telecommunications-Enhanced Mobile IP
TORA	Temporally-Ordered Routing Algorithm
TTCP	Test TCP
TTL	Time To Live
UA	User Agent
UDP	User Datagram Protocol
UMTS	Univeral Mobile Telecommunications System
UTRAN	UMTS Terrestrial Radio Access Network
VLR	Visitor Location Register
WLAN	Wireless Local Area Network
WNT	Wireless Network Testbed

Chapter 1

Introduction

As the demand for multimedia and mobility services increases, the tremendous growth in data traffic has forced the wireless industry to evolve towards All-IP networking [1], [2]. The idea is to develop an All-IP core network that will accommodate inter-working between the various heterogeneous access technologies [3]. Currently, different wireless access technologies exist for providing network access, meeting different requirements of the mobile users' needs and services. WLANs provide high-data rate for local area access, current and next generation cellular networks provide voice and data for wide-area communication, and satellite networks have been used extensively for worldwide coverage. Regarding cellular telecommunications, the absence of having global standards when developing cellular networks from the beginning have resulted in regional standardisation, e.g. TDMA and CDMA developed by TIA in North America and GSM by ETSI in Europe. With the evolution towards third generation (3G) wireless and the need for global interoperability, two new partnership projects were formed to address this. 3rd Generation Partnership Project (3GPP) [4], which is developing 3G standards for GSM-based systems and 3rd Generation Partnership Project 2 (3GPP2) [5] which is developing 3G standards for IS-95-based CDMA systems. Due to the di-

rection in which the two systems are evolving the creation of a combined single system should be possible in the near future due to the common network layer used. Both 3GPP and 3GPP2 have been designed with this convergence in mind aiming towards a global IP-based mobile telecommunication network [6]. The two protocol architectures would differ in the underlying networks but they will work very similarly from the network layer and above. In 3GPP, a project known as Long-Term Evolution (LTE) aims to improve the UMTS mobile phone standard to cope with future requirements and define an evolved access system with primary focus to cope with the rapid growth in IP traffic [7], [8]. Furthermore it is expected by the research community that IP based 3GPP services will be provided through a variety of access technologies. Support for seamless mobility between 3GPP access systems, WLAN, WiMAX and other heterogeneous access networks will be one of the main aims of future network evolution. Moreover LTE/SAE aims to address any requirements stemming from the work in SA1 on an All-IP Network [2]. SA1 is the service requirements working group of 3GPP and is responsible for defining service and feature requirements, a framework (architecture) for services, specification of services, specification of service capabilities, identification of technical and operation issues to next market requirements as well as charging and accounting requirements for All-IP Networks. Another objective in scope of supporting mobility between heterogeneous access technologies is how to maintain and support the same capabilities of access control (authentication, authorization, privacy and charging) when moving between different radio access technologies.

1.1 Motivation

One of the main research challenges for next-generation all-IP based systems is the design of intelligent mobility protocols and infrastructure support that will take advantage of IP-based technologies to achieve global roaming among various access technolo-

gies. The Internet Engineering Task Force (IETF) is developing a suite of protocols to achieve such mobility. Besides successful location area and routing area updating, achieving seamless handover is also one of the key aims in mobility management. It is important to note that in the next few years, the majority of terminals will be mobile and the majority of traffic will originate from IP-based applications offering more and more real-time services. The quality of real-time services like IP telephony and video-on-demand will depend greatly on the ability to minimize the impact of the handover, hence traffic redirection of ongoing sessions. This research began by examining proposed protocols to handle IP mobility management such as Mobile IP [9], SIP [10], tunnel-based protocols like Hierarchical Mobile IP [11] and host-specific protocols like Cellular IP [12] (see Chapter 2 for more details). With fundamental aim to achieve seamless and secure handover in an all-IP network environment, it was identified that there is more to mobility management than the problem of sending packets to the correct access router and making the handover procedure as seamless as possible.

This has lead the research to issues associated with IP Mobility Management such as Context Transfer [13] and the impact of Security on Mobility [14]. Context Transfer aims to contribute to the enhancement in handover performance. When a mobile node (MN) moves to a new subnet it needs to continue certain transport- or routing-related services that have already been established at the previous subnet. Such services are referred to as 'context transfer candidate services' because they can be supported with the context transfer protocol which will be used to forward service related state information and thus minimize renegotiations at the new point of attachment. Examples include header compression, QoS policy, Authorization, Authentication and Accounting (AAA) profile and IP security (IPsec) state (More information on these can be found in Chapter 3). Re-establishing these services at the new subnet will require a considerable amount of time for the protocol exchanges and as a result time-sensitive real-time traffic will suffer during this time. Alternatively, context transfer candidate

services' state information can be transferred, for example, from the previous access router to the new access router so that the services can be quickly re-established. A layer-3 context transfer protocol will result in a quick re-establishment of context transfer candidate services at the new domain. Layer 2 protocols may also define their own solutions for context transfer. These solutions primarily aim to facilitate the transfer of Layer 2 related context between two radio access networks or two radio access points. However a layer 2 solutions that supports multiple radio interfaces will be difficult to achieve. On the other hand, operating at layer-3 ensures interoperability among layer-2 radio access technologies. This would also contribute to the seamless operation of application streams and could reduce susceptibility to errors. Furthermore, re-initiation to and from the mobile node will be avoided hence wireless bandwidth efficiency will be conserved.

1.2 Research Issues

This section summarizes the different research issues related to context transfer. More details about these can be found in chapter 3. Not all listed issues have been researched in this thesis but all could definitely be considered for future work.

- Identify different possible services that can be supported by this protocol. Different services like AAA, QoS, Header Compression, IPsec and Multicast Membership have been identified in [13] but this is definitely not an extensive list.

- Evaluate the impact these services can have on the overall handover performance and identify possible gain improvements the context transfer protocol can provide. Depending on the service supported either the total perceived handover delay can be improved or the overall quality of the service can improved.

- Research and compare different possible options of transferring the desired context information. Options can include: assigning the mobile node responsibility for storing and sending context; making the context transfer candidate services to be responsible for transferring their own state; enhancing the mobility management protocols to be responsible for transferring the state; utilising or introducing a central entity within the network to store a copy of the state; propose a dedicated protocol for transferring any service context.

- How and when to initiate context transfer is very important in order to get the timing right and also synchronization with the other mobility management protocols. Choosing how context transfer should be triggered is also crucial in achieving exactly this. As mentioned in [13] the context transfer solution must define the characteristics of these trigger mechanisms used to initiate context transfer.

- Another issue researched was how a context transfer solution could be added or integrated in future architectures. In this work, context transfer was integrated in a hybrid multilayer mobility management [17] and in ambient network proposed architectures [92]. The different constraints and requirements resulted in the proposal of a modular approach for context transfer.

A number of issues identified but not covered in this thesis include interoperability with other Layer 2 context transfer protocols and the use of context transfer in 3GPP LTE architecture. Layer 2 radio protocols may also define their own solutions for context transfer. These solutions primarily aim to facilitate the transfer of Layer 2 related context between two radio access networks or two radio access points. However a layer 2 protocol that supports multiple radio interfaces will be difficult to achieve. More information on the different research issues and scope of this work can be found in Chapter 3.

1.3 Contribution and Achievements

This section gives a summary of the different achievements and research contributions made in this work. The three main contributions can be summarised as:

- Context Transfer extensions to mobility management protocols
- Context Transfer support for middleboxes and firewalls
- A modular context transfer approach for next generation networks

The first part proposes possible ways of extending the mobility management protocols to support the forwarding of context information during the handover operation. Context transfer protocol extensions were proposed for Mobile IPv4, Hierarchical Mobile IP

and Cellular IP. In all three extensions, the different Mobility Management Protocols (MMP) frameworks are utilised, for supporting triggering and transferring context during the handover operation. Entities like the Foreign Agent (FA) in the case of Mobile IP, the Mobility Anchor Point (MAP) in the case of Hierarchical Mobile IP and the Cellular IP Gateway (CIP-GW) were also used to store context locally. For evaluation, the work focused mainly on supporting security services since it was anticipated that this is one of the most important context transfer candidate services causing a heavy impact and possible interruption to a user's ongoing sessions. This led to a thorough analysis of the impact of AAA on the user's sessions. In this evaluation, different possible schemes were investigated namely: Reactive Context Transfer, Predictive Context Transfer and Gateway Supported Context Transfer. These were evaluated using both simulation analysis, testbed implementation, as well as mathematical analysis.

The second main outcome was a proposal of a new application for context transfer to support middleboxes and more specifically firewalls. Here, how context transfer can support multimedia session continuation during mobility was proposed and also support security by remotely configuring firewalls and by minimizing the risk of network attacks.

The third major outcome of this work was a context transfer module which is arranged to forward context transfer information related to a mobile host's sessions from a Previous Access Router (PAR) to a New Access Router (NAR) when handover takes place in a mobile communication network. This context transfer module is arranged to reside at one of the access routers and to provide a message framework for the interworking between itself, a module related the mobility management protocol and a module related to the context transfer candidate service. The module provides a message framework for the forwarding of context information between itself and a context transfer module residing at the other of the access routers. The module was integrated and evaluated as part of the Hybrid Mobility Management architecture in the EU IST EVOLUTE project [25] and in the Ambient Control Space of the Ambient Networks project [91].

1.4 Structure of Thesis

The thesis is structured as follows: Chapter 2 gives the state of the art of mobility management protocols for next generation all-IP networks. The chapter emphasises the importance of mobility management support protocols for providing improvement on the overall handover operation. Chapter 3 describes the motivation behind the research on context transfer and give an overview of possible research issues related to context transfer. Chapter 4 proposes context transfer extensions to mobility protocols providing a high level description on how these could be extended with context transfer capabilites. The chapter gives a more detailed specification for one of the protocols. Chapter 5 illustrates the impact of security protocols on handover performance and the benefits of the proposed schemes are demonstrated using both real and non-real time services. Four possible different schemes are compared here according to handover performance: Predictive context transfer, Reactive context transfer, Gateway supported context transfer and context transfer disabled. Chapter 6 evaluates context transfer for its support on middleboxes and firewalls. Two benefits of context transfer are demonstrated here and how context transfer could support middleboxes to provide multimedia session continuation; How context transfer can minimize the risk for network attacks by e.g. closing unused ports in firewalls. Chapter 7 describes how the proposed context transfer module approach can be integrated in the Evolute and Ambient Networks architectures. The final chapter concludes the research by giving a summary of the achievements and research contributions, the main outcomes and conclusions as well as possible future research issues. A list of the research contributions and publications can be found at the end of the thesis.

Chapter 2

Mobility Management in IP-based Networks

2.1 Scope

This chapter gives an overview of IP mobility. It aims to be self contained and gives a literature review of mobility management in IP-based networks. Primarily the main mobility management protocols proposed to handle terminal mobility are explained. It considers the differences between terminal, personal and session/service mobility. For terminal mobility the distinction between macro- and micro-mobility and the differences between handover and location management are also explained. The last part of the chapter emphasises the importance of supportive protocols like candidate access router discovery, paging, security considerations, and especially context transfer.

2.1.1 Introduction to IP mobility

If routing was based on host-specific routing then designing mobility protocols for an All-IP network would be straightforward. This is because in host-specific routing each node will have routing-table entries acting for every host in the network that will be updated while a mobile node roams across different networks and sub-networks. However using host-specific routing for a network the size of the Internet will clearly be unworkable. This is because routing tables will need to hold millions of entries of hosts and routers would spend most of their time and resources on executing table look-ups instead of actual traffic. Also frequent exchanges of routing updates would be required between neighboring routers that would consume most of the network's resources. Therefore IP routing is based on network prefixes mainly to satisfy this scalability problem. IPv4 and IPv6 are based on network prefix addressing which although satisfying scalability issues, minimizing use of resources and making routing more efficient, has its drawbacks when dealing with mobility [26].

One of the problems with traditional IP is that if a mobile node moves from one network to another and retains its IP address, under general IP circumstances it will lose communication with any other node. The reason for this is because IP packets destined to a specific address will be routed towards the router(s) that advertise reachability to the network-prefix of that address. If a node is not located on the link where its network-prefix says it's supposed to be located, then packets sent to that node will be undeliverable. The mobile node will not be able to communicate with any other node. A node may not move from one link to another if it wishes to communicate without at least changing the network prefix portion of its address to reflect its new point-of-attachment to the network.

2.1.2 Types of Mobility

IP mobility management can be categorized in to four different types, namely: Terminal, Session, Personal or Service Mobility (see Figure 2.1). Terminal Mobility allows a device to move between IP subnets, while continuing to be reachable for incoming requests and maintaining sessions across subnet changes. Session Mobility allows a user to maintain a media session even while changing terminals. For example, a caller may want to continue a session begun on a mobile device on the desktop PC when entering their office. Personal Mobility allows to address a single user located at different terminals by the same logical address. Service Mobility allows users to maintain access to their services even while moving or changing devices and network service providers.

2.1.3 Location vs Handover Management

The aim of location management is to allow a system to keep track of the locations of users between consecutive communications. This can be divided into two main tasks, namely location registration (or location update) and call delivery. Location registration is realised by a mobile node by periodically informing the system to update relevant location databases with an up to date location information. For call delivery the location of the mobile node is determined by the information available in the system databases during communication initiation. Call delivery can be further split into two steps: determining the serving database of the targeted mobile user and locating the visiting cell/subnet of the targeted mobile user (also known as paging). There are a number of challenges and research issues when it comes to location management techniques. In the inter-system roaming case evaluating the following is of importance: Reduction of signaling overheads, latency of service delivery and QoS guarantees in different systems. In the case of fully overlapped heterogeneous wireless networks the research challenges include the following: through which networks a mobile terminal

should perform location registrations; in which networks and how the up-to-date user location information should be stored; how the exact location of a mobile terminal would be determined within a specific time constraint. Research on location management is out of the scope this research but may be considered for future work.

Handover management is the process by which a Mobile Node maintains its connections active when it changes point of attachment. This is sometimes divided into intrasystem vs intersystem handover procedure. Intrasystem (also known as horizontal handover) is the handover taking place in homogeneous access networks (e.g. WLAN to WLAN). Intersystem handover (also know as vertical handover) is the handover taking place between heterogeneous networks WLAN to UMTS. Intrasystem handover is required when the signal strength of a serving base station becomes lower that a certain threshold value. Intersystem handover takes place either when a user is moving out of a coverage area and enters another overlaying network, when a user decides to change to another overlaid network for his/her future service needs or when distributing the overall network load among different systems to optimize for example the performance of each individual network. There are a number of challenges and therefore research issues for handover management in next generation all-IP based wireless systems: minimizing signaling overhead, QoS guarantees, handover latency, resources and routes setup delay, limiting disruption to user traffic, minimize handover failure, minimize packet loss, efficient use of network resources, scalability, reliability and robustness issues.

2.1.4 Macromobility vs Micromobility

In IP terminal mobility there is usually a distinction made between macromobility and micromobility (see Figure 2.1). Informally this terms are used to mean *mobility over a large area* (global mobility) and *mobility over a small area* (local mobility). Actually what is important to distinguish is the difference of mobility when roaming within and

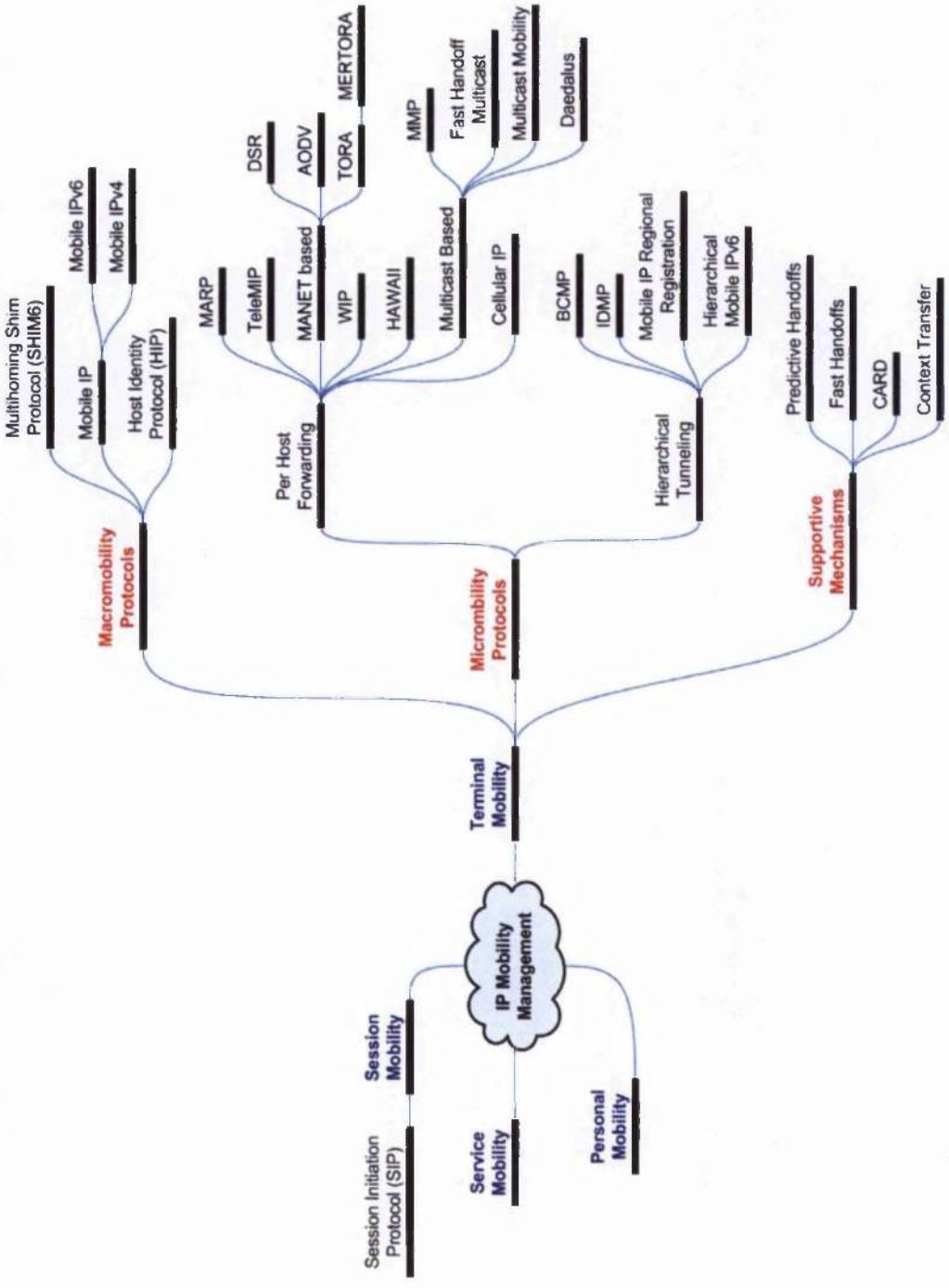

Figure 2.1: Types of Mobility Management Protocols

when roaming between administrative domains. These handover cases are significantly different because in the case of an inter-domain handover (change in administrative domains), the mobile host must change IP address, be re-authenticated and the user's QoS, charging scheme, policies may change. Furthermore issues such as speed and performance of the handover are less relevant simply because such handovers will be much rarer and also there is no guarantee of mobility support in the new administrative.

2.2 Macromobility

2.2.1 Mobile IP

The most popular protocol for IP mobility is Mobile IP [9] developed within IETF. Mobile IP makes mobility transparent to layers above IP and also enables the maintenance of active TCP connections. It does though face a number of limitations especially when it comes to real-time services. Mobile IP provides a good solution for mobility across the global Internet allowing nodes to maintain ongoing communications while changing links (see Figure 2.2).

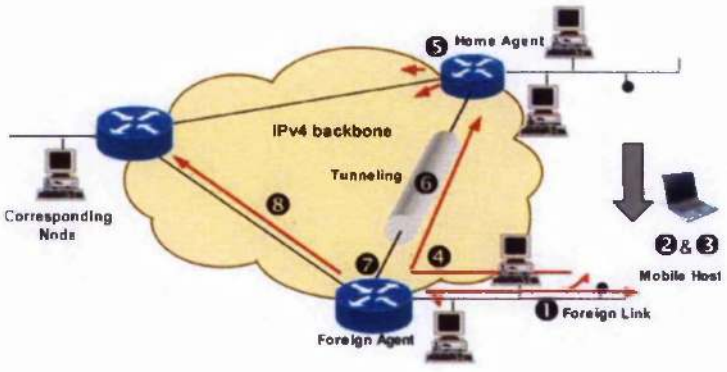

Figure 2.2: Steps involved in Mobile IPv4 when a mobile user changes link

1. Home Agents and Foreign Agents broadcast *Agent Advertisement* (Mobile IP messages) periodically to their attached links to identify home users and foreign users.

2. On reception of these *Agent Advertisements* the mobile nodes 'read' the advertisement to identify whether they resite on a home link or a foreign link.

   The following steps take place only if the node is in fact mobile and it is attached to a foreign link:

3. A care of address is given to the Mobile Node on the Foreign Link.

4. The mobile node registers the care-of-address obtained from the Foreign Agent with its Home Agent.

5. The Home Agent informs neighbor-routers of its reachability to the mobile node's home address.

6. On reception of packets destined to the mobile-node the Home Agent tunnels the packets to the Foreign Agent.

7. At the foreign agent the original packet is extracted from the encapsulated packet used for tunneling and forwarded to the Mobile Node.

8. Packets send by the Mobile Node are routed directly to their destination via the Foreign Agent.

In the Mobile IPv4 case all packets from the correspondent node are routed via the Home Agent to the mobile causing what is known as triangular routing that can be very inefficient. Also the IP-in-IP encapsulation used in Mobile IP adds a significant overhead. Another problem faced by Mobile IP is that firewalls become obstacles, which has been resolved with an extension known as reverse tunneling. A similar problem is

caused due to the shortage of publicly routable IPv4 addresses. These problems may have a serious impact on the quality of real-time sessions. Nevertheless Mobile IP is the most popular macro-mobility solution and is the de facto standard in this area. Mobile IPv6 [27] includes many functions also used by IPv4, with additions integrated in the IP protocol and improvement in routing usage allowing better support for mobility (i.e. network-layer mobility). Below is a list of the main advantages of Mobile IPv6 over Mobile IPv4.

- No need for Foreign Agents. Extra IPv6 features are introduced to allow a mobile node to operate in any location without the need of a local router.
- Route Optimization is integrated as a fundamental part of IPv6 protocol rather than just an option.
- IPv6 Neighbor Discovery is used instead of ARP (Address Resolution Protocol) improving robustness of the protocol.
- Ingress Filtering is performed normally by routers since packets contain both care-of address and home address.

2.2.2 Session Initiation Protocol (SIP)

In recent years, SIP was proposed as an alternative solution to the terminal mobility problem [10], [29]. SIP was originally proposed by the IETF as a general multimedia session initiation protocol. This enables two or more users to set up multimedia session between them. In SIP, the end-users are identified by SIP URIs (Unified Resource Identifier) which are very much like e-mail addresses (e.g. georgiades@surrey.ac.uk). Support for personal mobility is inherent in the SIP specification, since a user is addressed by a URI that is independent of the user's location and choice of terminal. Additionally, SIP could be used to support terminal mobility. In this case, when a

mobile host moves from one network to another, it can get a new IP address via DHCP and the SIP user agent (UA) can then resume any ongoing sessions by sending a Re-INVITE message to the corresponding hosts. The advantage of this approach is the introduction of mobility awareness at application layer. Furthermore, no changes are required to the IP stack on the end systems. SIP-based mobility support can improve the performance of realtime applications but any ongoing TCP connections will break during handover. Moreover, the host needs to acquire a new topologically correct IP address, while the delay incurred during address acquisition can be significantly large.

SIP can be used to support personal mobility. Personal mobility is the ability of end users to establish communication and also be identified regardless of location and equipment. For a user to roam while a session is active, IP mobility protocols need to be used. SIP can be used to complement IP mobility protocols like Mobile IP. The exchange of messages between a mobile user, a SIP server and a corresponding user in SIP, is analogous to the exchange of messages between a mobile user, a Home Agent and the corresponding user in Mobile IP. Since the aim was to overlay protocols with similar functionality it would mean that some of the actions performed would be duplicated. This needs to be considered if SIP mobility is to be used to support Mobile IP. Figure 2.3 shows a corresponding host which wants to invite a mobile host for communication. In this scenario the mobile host happens to recite in a foreign network and hence the redirect server informs the corresponding node of the mobile node's new SIP address. Using the new SIP address the corresponding sends an INVITE message at the mobile node's current location and direct communication is established.

The following was proposed in [25]. If a mobile host changes networks during a session it must inform the corresponding host of its new location (see Figure 2.4). It does this by sending a new INVITE containing the same Call-ID as in the original call setup. Within this INVITE SIP message the new IP address is included in the Contact field informing the corresponding host of its new IP location. The new address is also

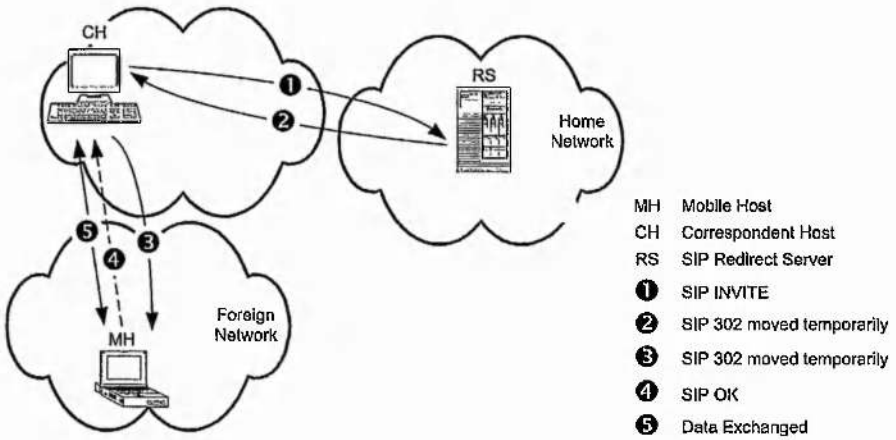

Figure 2.3: SIP for Mobility (1)

included in the "c" field of the Session Description Part (SDP) part of the message.

Besides informing the corresponding host of the change in network the mobile host should also register its new location with the location server as shown in Figure 2.5. Both SIP and Mobile-IP can be used to provide support for mobility management within an all-IP architecture. It is important to emphasise that SIP is an application layer protocol whereas Mobile-IP is a network layer protocol. Although belonging to different layers their operation can be compared especially from the IP traffic point of view. A SIP proxy server is analogous to a Home Agent in Mobile-IP. For example if both protocols are used, when a Mobile Host enters a foreign network it will need to register its new location address both with a SIP server at its home network and also with its home agent.

Therefore if SIP and Mobile-IP are going to collaborate what should be avoided is the waste of resources due to duplicate information and servers performing similar redirection services. What has been proposed in [25] is to use SIP mobility for real-time communication over UDP and Mobile IP for TCP communication. This will mean that when a host is about to transmit RTP streams it will use the care-of address and

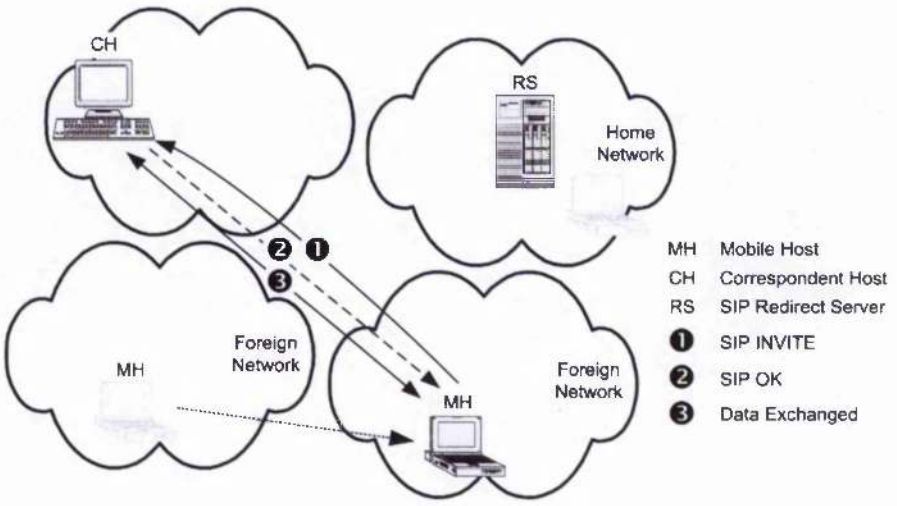

Figure 2.4: SIP for Mobility (2)

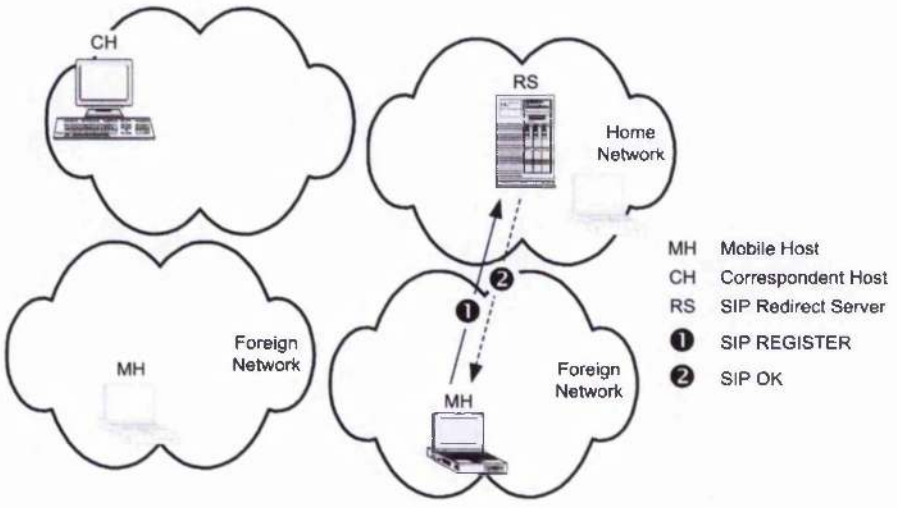

Figure 2.5: SIP for Mobility (3)

when it wants to establish a TCP connection it will use the home address thus routing traffic via the Home Agent. Whether to use Mobile-IP or SIP will be decided by a Mobile Policy Table (MPC). This is simply a look-up table indicating which services should use Mobile-IP and therefore deciding whether to use home or care-off address.

2.2.3 Host Identity Protocol (HIP)

IP addresses serve the dual role of being both end-point identifiers (Names of network interfaces on hosts) as well as Locators (Names of naming topological locations). The Host Identity Protocol (HIP) [32] supports an architecture that decouples the transport layer (TCP, UDP, etc.) from the networking layer (IPv4 and IPv6) by using public/private key pairs, instead of IP addresses, as host identities (see Figure 2.6). When a host uses HIP, the overlaying protocol sublayers (e.g., transport layer sockets and Extensible Security Protocol (ESP) Security Associations) are instead bound to representations of these host identities, and the IP addresses are only used for packet forwarding. However, each host must also know at least one IP address at which its peers are reachable. Initially, these IP addresses are the ones used during the HIP base exchange [33]. A further number of drafts have been proposed on how the HIP protocol could be extended to handle mobility and required extensions: mobility and multihoming [34], HIP domain name system descriptions [35], HIP rendezvous extension [36], HIP registration extension [37].

When a host moves to another address, it notifies its peer of the new address by sending a HIP UPDATE packet containing a LOCATOR parameter. This UPDATE packet is acknowledged by the peer, and is protected by retransmission. The peer can authenticate the contents of the UPDATE packet based on the signature and keyed hash of the packet. When using ESP Transport Format, the host may at the same time decide to rekey its security association and possibly generate a new Diffie-Hellman key [38];

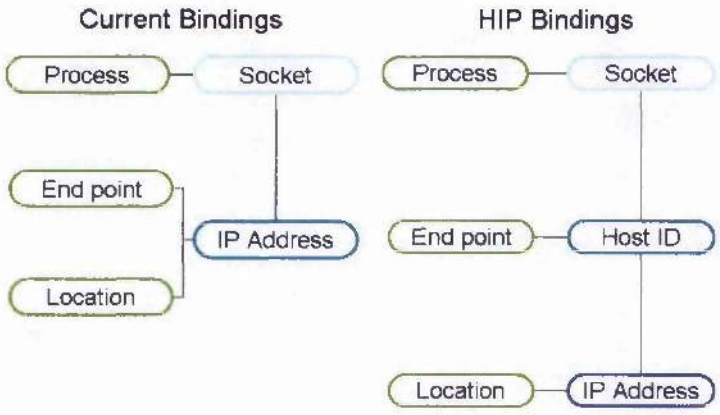

Figure 2.6: HIP: Separation of host identity from IP address

all of these actions are triggered by including additional parameters in the UPDATE packet, as defined in the base protocol specification and ESP extension. When using ESP (and possibly other transport modes in the future), the host is able to receive packets that are protected using a HIP created ESP SA from any address. Thus, a host can change its IP address and continue to send packets to its peers without necessarily rekeying. However, the peers are not able to send packets to these new addresses before they can reliably and securely update the set of addresses that they associate with the sending host. Furthermore, mobility may change the path characteristics in such a manner that reordering occurs and packets fall outside the ESP anti-replay window for the security association, thereby requiring rekeying.

2.2.4 Site Multihoming by IPv6 Intermediation (SHIM6)

The SHIM6 protocol is a layer 3 protocol for providing locator agility below the transparent protocols with failover and load sharing properties [39]. The host in a site which has multiple provider allocated IPv6 address prefixes, will use the SHIM6 protocol specified in this document to setup state with peer hosts, so that the state can later

be used to change to a different locator pair, should the original one stop working.

2.2.5 Discussion

Table 2.1 gives a comparisson of the different mobility management protocols which could be used for macro-mobility (global mobility) based on different attributes and capabilities. The first point regards the preservation of established connections which is a key cabability of terminal mobility in general. SIP was introduced to handle session initiation but there were also proposals to handle session mobility [29].

All four protocols are able to preserve established connections across a locator change. Mobile IP, HIP and SIP are able to preserve established connections even if both ends move simultaneously. SHIM6 can only achieve this only if the new locator has been previously communicated to the other prior to the move. A path outage refers to the situation where the path between the locator pair is broken somewhere in the middle of the network. Mobile IP cannot preserve connections between node and the home agent during such outages. SIP and HIP are capable of preserving connections across such outages but have no mechanisms for detecting such failures and decide on alternative paths. Mobile IP, HIP and SIP are able to accept new connections without waiting for name resolution, since DNS records do not need to be upgrated when moving. SHIM6 does not address this problem. Support referrals is the case where one device can redirect another device through a third device. Mobile IP and SHIM6 support by either name or upper layer identifier. HIP and SIP support referrals by name but not upper-layer identifier. Mobile IP provides a stable home address. SIP and HIP provide a stable upper layer identifier, SHIM6 does not attempt to address this problem. For many application it is desirable to have a stable address as they cache addresses for a significant length of time. Mobile IP only advertizes a single locator. When many locators are advertised to another device the other device can do load

Table 2.1: Comparison of Macromobility Protocols [40]

	Mobile IP (based on v6)	HIP	SIP	SHIM6
Preserve established connections	YES	YES	YES	YES
Support both ends moving simultaneously	YES	YES	NO	Only within know set
Span path outages	NO	NO	NO	YES
Resolve name to locators immediately after move	YES	YES	YES	NO
Support referrals	YES	only by name	YES	YES
Stable addresses	YES	Non-routable	YES	Assumed
Support load spreading	YES	YES	YES	YES
Multicast support	YES	YES	YES	YES
per- packet overhead (bytes)	0 if both have 20/40 in src away + 24 if dest away	0 (beyond IPsec transprt mode)	0	0 normally 8 if moved
connect overhead (messages)	0	0	4 for IPsec key negotiation	0
Locator change overhead (messages)	2 to update HA + 6/4 (cga)/0 if local (hmipv6) to update peer	4 to update peer	4 to update peer	4 to update peer
One end benefit	YES	NO	NO	NO
Typical deployment dependencies	HA if hmip used: MAP + config routers	Rendezvous svr, New RR, IPsec	SIP Location Server, SIP porxy servers	None
Control message auth check				
Minimum	On-path	On path + same node	On-path	Crypto
Maximum	Crypto	Crypto	Crypto	Crypto

spreading of different connections to the first device by using different locators. SHIM6 has the ability to advertise multiple locators. HIP also supports this option but is not as mature as SHIM6. Mobile IP supports sourcing multicast in the home address by tunneling it through the home agent. For SHIM6 and SIP multicast can be sourced from any source but there is no support for moving sessions. HIP does not support sourcing multicast. MIPv6 uses 20 bytes for the destination option header. When packets are reverse tunneled to a home agent, this becomes 40 bytes (the equivalent of an IPv6 header). If it uses a Type 2 routing header in packets sent to a home address then 24 bytes are required additionally. SHIM6 uses an 8-byte payload extension header with data packets. HIP uses the IP encapsulating security payload (ESP) within data packets i.e. the size is equal to an ESP header but is only an overhead if IPsec transport mode will be used anyway. Connect and locator change overhead are to do with the number of message exchanges required during the first exchange between a mobile node and a correspondent node and when the MN changes location. At the first exchange between MN and CN, Mobile IP does not generate any additional messages. At the time a mobile node moves away from home and decides to use route optimization, it generates 6 additional messages (Binding Update, Binding Acknowledgement, Home Test Init, Home Test, Care-of Test Init and Care-of Test).

SHIM6 assumes that the node is always at home and generates no message exchanges. HIP uses a 4-way handshake to negotiate IPsec state prior to being able to send data. During ongoing communication if there is a locator change, MIPv6 required 2 messages to update the Home Agent and 6 to update any correspondent node. SHIM6 generates 4 messages to update the peer. HIP generates 3 messages to update rendezvous server and a 3 message handshake to update each peer. MIPv6 requires a HA and if HMIP is used, a MAP is required plus configurable routers. HIP requires a rendezvous server, a new RR and IPsec. SHIM6 does not have any deployment dependencies. One end benefit refers to the fact that some protocols are more beneficial when only one end of

a connection supports the protocol. This allows a new device to gain immediate affect. MIPv6 provides benefit for a mobile node even without support for the correspondent nodes. Both SHIM6 and HIP require support in both ends before their benefits can be realised. Security aspects also differ for each protocol. MIPv6 at a minimum only verifies that control messages were originated by someone on the path between the two ends using a refer routability test but has the option of using cryptographically generated addresses (CGA) for more security. SHIM6 also uses a return routability test, plus at least a verification that the new locator is a locator of the same node. It can also use CGAs for more security. HIP however requires strong cryptographic checks on all control messages. IPsec is used in HIP for data security although for MIPv6 and SHIM6 is optional.

2.3 Micromobility

The idea of micro-mobility was introduced after a number of shortcomings were identified with Mobile IP [9] such as non-localized location management, triangular routing impact on packet delivery delay and in-flight packets being lost during handover [28]. In-flight packets refer to any packets destined to the MN during handover. Since then several micromobility schemes have been proposed to augment Mobile IP and provide a faster and smoother handoverthan what is achievable by Mobile IP alone. Hierarchical Mobile IP (HMIP) [11], Cellular IP [42], HAWAII [43], and Intra-Domain Mobility Management Protocol (IDMP) [44] are some examples of micro-mobility protocols. Moreover IETF has recently established another working group for dealing with Network-based Localized Mobility Management (NetLMM) [13].

The majority of these proposals agree that Mobile IP is suitable for handling macro-mobility (inter-domain mobility) but not the micro-mobility (intra-domain mobility) [49]. Network-prefix-based protocols like HMIP usually require the Mobile Node (MN)

to change its IP address as it changes its point of attachment. The change of IP address causes disruption to ongoing sessions not only due to delays in address acquisition, but also of its impact on other IP dependant protocols. Alternatevily micro-mobility protocols that use host-specific forwarding allow MNs to maintain the same IP address but this pose scalability problems as the number of routing entries becomes very high. In addition, these protocols, such as Cellular IP and HAWAII, assume a hierarchical network topology which defeats the robustness and flexibility of an IP routing protocol.

2.3.1 Network-prefix-based schemes

Mobile IP Regional Registration (MIP-RR) - With Mobile IP, a mobile node registers with its home agent each time it changes its care-of address. Here a Gateway Foreign Agent (GFA) is introduced to provide regional registrations in the visited domain. This is an optional extension to Mobile IPv4 aiming to reduce the number the signaling messages to the home network, and reduce the signaling delay when a mobile node moves from one Foreign Agent to another within the same visited domain.

Hierarchical Mobile IPv6 [11] introduces the Mobility Anchor Point and extends the Mobile Node's and Home Agent's operations. A Mobility Anchor Point is a router located in the Mobile Node's visited domain. It behaves like a temporary Home Agent for the Mobile Node thus reducing mobility signaling. By reducing the amount of signaling outside the local domain it can support Fast Mobile IP handovers to help mobile nodes in achieving seamless mobility. The introduction of the MAP concept minimizes the latency due to handovers between access routers since it will take less time to bind-update a local MAP than a distant HA.

2.3.2 Per-host forwarding schemes

Cellular IP [42] is a per-host forwarding protocol, where the next hop is determined by finding an exact match of the destination IP address in the routing table. This allows the mobile host to maintain the same IP address even when it moves from one subnet to another within a micro-mobility domain. Per-host schemes such as cellular-IP are very good candidates for micro-mobility as the signaling load will be reduced compared with basic mobile IP and also the handover will be faster. Cellular IP is a lightweight and robust protocol that is optimized to support local mobility but efficiently interworks with Mobile IP to provide wide area mobility support. It resolves the challenges by being optimized for wireless access networks and highly mobile users. Cellular IP offers a number of benefits. It offers small and cheap access points. Its distributed location management allows for the same protocol to be used across heterogeneous networks and therefore seamless migration between different environments. Using Cellular IP the local level of service can always be obtained. It is fully compatible with IP, it does not require new packet format or encapsulation and does not require extra address space.

HAWAII (handoff-aware wireless access Internet infrastructure) was another proposal for a domain-based approach for supporting mobility, introduces in 1999 in IETF [43]. HAWAII installs host-based forwarding entries in specific routers using specialized path setup schemes to support intra-domain micro-mobility and uses Mobile-IP as a default for supporting inter-domain macromobility. The architecture of HAWAII provides the following:

- straight forward QoS support by assigning a co-located care-of address to the mobile host.
- maintains end-to-end connectivity with little disruption as the mobile host moves by establishing special paths to the MH.

- uses soft-state mechanisms to maintain forwarding state to provide a degree of tolerance to router or lack failures within the network.

IDMP (Intra-Domain Mobility Management Protocol) [44] uses two dynamically auto-configured care-of addresses (CoAs) for routing the packets destined to mobile users. A global care-of-address (GCoA) is used to identify the mobile node in the current domain whereas a local care-of-address (LCoA) changes every time the mobile changes subnets and identifies the mobile node in the local subnet. However unlike HAWAII, MIP-RR or HMIPv6, IDMP is designed as a stand-alone solution for intra-domain mobility and doesn't assume the use of MIP for global mobility management. It uses a Mobility Agent (MA) similar to a MIP-RR GFA and a Subnet Agent (SA) similar to MIP FA in CoA mode to provide domain-wide and subnet-specific mobility services.

MER-TORA (Mobility Enhanced Routing Temporally-Ordered Routing Algorithm) [67], [68] is based on TORA ad hoc routing protocol [70][71] which was designed to decouple the generation of far-reaching control message propagation from the dynamics of the ad hoc network topology. To handle the handover operation a mobile node, MER-TORA exploits TORA's fast route restoration mechanism to establish new routing paths for the mobile node by changing the size of the routing table entry of the mobile node. In addition, tunneling is used between old and new access routers for diverting packets from the old to new location of MN. It is assumed that there is a virtual link, based on IP routing, for signaling between the access routers to manage handover and to exchange capabilities of access routers. MER-TORA modifies the TORA to run proactively instead of reactively as originally designed. In addition, the implementation of MER-TORA maintains the conventional network-ID based routing whenever possible, and resorts to tunneling or host-specific forwarding when mobile node moves to a new point of attachment. This has an advantage over other scheme in term of scalability. It also retains the robustness of the ad-hoc IP routing protocol, such as topological de-

sign freedom, reduced configuration and greater resilience although it does not usually provide optimal routes for communication after MN executed a handoff. MER-TORA assumes a brand new and complex routing protocol. However, although it is an effective protocol to handle mobility, it comes with several shortcomings. It is a complete IP routing solution for both the fixed network node and moving mobile nodes. Its implementation suggests the replacement of the existing intra-domain routing protocol. This changes the way Internet routing should be handled in the access network and hence would have serious deployment issue. In addition, this scheme is significantly more complex than Hierarchical Tunneling and Host-specific Forwarding. Implementing this protocol will need to gain confidence of IP community that it works properly in all circumstances and that they understand how to deploy, upgrade, and manage network with this protocol. Because of the shortcomings described above, MER-TORA has not received much attention in the IP-community.

TeleMIP (Telecommunications-Enhanced Mobile IP) [45], is based on the observation that current IP mobility schemes have a subnet (and finer granularity of location resolution) and mostly no scoping for the transmission of location updates. Cellular IP, for example, proposes a base-station-level (layer 2) granularity similar to cellular networks. The current subnet-based FA scheme in Mobile IP, on the other hand, leads to a change in care-of addresses at every subnet transition. A generalization of the FA concept was proposed by introducing a new node, the Mobility Agent (MA), at network layer (layer 3) granularity, higher than that of a subnet, thus reducing the generation of global location updates. By limiting intradomain location updates to the MA, the latency associated with intradomain mobility was further reduced without resorting to source-specific routes. Finally, our two-level mobility management scheme allows the use of private addressing (and, if necessary, non-IP mobility management) within the provider's own domain.

Another recently proposed protocol is known as NetLMM (Network-based Localized

Mobility Management) protocol. The requirements for localized mobility protocol have been analyzed in Kempf et al. [46] [47] and it is shown that none of the existing protocols completely fulfill them. One of these requirements is that the mobile node is not involved in mobility management. This is extremely attractive from the point of view of Ambient Networks, since it would allow to support legacy nodes supporting only plain IP to benefit from attachment to an Ambient Network. If such a node is acting as a router for a subnetwork attached to it (e.g. a Personal Area Network) this would also allow for composition of a legacy network to an Ambient Network, with the legacy network maintaining connectivity while moving. Because it is network based, the mobile node is not required to implement new mechanism in its IP stack, neither to change its IP address when it attaches to a new access router. NetLMM extends the MIPv6 protocol to allow the access router to send proxy local binding updates to the mobility anchor point (MAP) on behalf of the mobile node. Because this proxying introduces security risks such as IP spoofing and connection hijacking, a secure interface between the MN and the AR [48] has been developed.

MARP (Mobility Aware Routing Protocol), a micromobility protocol proposed in [72], eliminates some common deficiencies of micro-mobility protocols but retains the salient features. It makes use of both network-prefix-based routing and host specific forwarding but HSE are limited to a small set of routers thereby reducing the size of the forwarding table. This makes MARP scalable while effectively tackling intradomain mobility of MN on per-host basis. The routers with MARP capability can be deployed in MIP network in a seamless way as it interoparates with MIP as well as with the conventional prefix-based IP routing .

2.3.3 Discussion

Comparisons of the majority of the mentioned micro-mobility protocols can be found in [43], [57], [49] and [64]. Table 2.2 shows a summary of the characteristics of different local-mobility (micro-mobility) protocols. In some cases of the handover operation these are criteria that could be considered for selecting which local mobility protocol to use if any e.g. scalability, reliability. Topology and with which global mobility protocol it can interwork with should also be considered. Host specific can be considered as a standalone mobile routing protocol which does not rely on the conventional network-ID based IP routing protocols. Is there a preference for the MN to maintain the same IP address e.g. during ongoing sessions. The major difference between Hierarchical Tunneling and Host-specific For-warding schemes is that in tunnel-based protocols, MN needs to acquire a new care-of-address each time it moves on to a new access router, whereas in host-based protocol MN keeps its CoA. Hierarchical Tunneling scheme is an add-on built on top of the standard intra-domain routing protocol. This effectively hides the nodes' mobility from the routers, with mobility support confined to a few specialized nodes (i.e the mobiles themselves and the mobility agents). On the other hand, Host-specific forwarding scheme is tightly integrated with the topology of the mobile network and expose host mobility to routers. Tunnel-based scheme is simpler and more scalable as only the tunnel starting point and end-point are involved for handling mobility of MN. However they require that many routers store information about many mobile nodes resulting to scalability issues. This is because as the size of the network and number of mobiles grow, the forwarding table will grow, and eventually it will be too large to retrieve the information sufficiently quickly. The problem is likely to be most acute for the gateway. Nevertheless, host-specific scheme is more scalable in terms of address allocation and management, as MN is not required to up-date its address as it changes access router. For idle nodes paging support is useful. Handover delay and control packet require for address allocation and address returning is lower

with these schemes. Address allocation is a particular critical issue for IPv4 networks where the number of addresses is limited. Other difference between the different micro-mobility schemes are mainly on the technical implementation, such as usage of soft state expiry or explicit signaling to delete mobility management state, how paging areas are defined, use of packet snooping or explicit signaling to create and update routing information, whether the endpoint of signaling is at a cross-over router or at a gateway.

2.4 Supportive Mechanisms

2.4.1 Fast Mobile IPv6 (FMIPv6)

Mobile IPv6 enables a Mobile Node to maintain its connectivity to the Internet when moving from one Access Router to another, a process referred to as handover. During handover, there is a period during which the Mobile Node is unable to send or receive packets because of link switching delay and IP protocol operations. This *handover latency* resulting from standard Mobile IPv6 procedures, namely movement detection, new Care of Address configuration, and Binding Update, is often unacceptable to real-time traffic such as Voice over IP. Reducing the handover latency could be beneficial to non-real-time, throughput-sensitive applications as well.

FMIPv6 (Fast Mobile IPv6) [50] provides fast IP connectivity to a new point of attachment and therefore reduces packet loss and generally improves handover performance. During link configuration and binding update FMIPv6 configures the routing so that packets delivered to the old care of address are forwarded to the new. Moreover FMIPv6 provides support for configuring link information prior to handover in the new subnet while the mobile node is still attached to the old subnet. This reduces the amount of pre-configuration time in the new subnet. RFC4260 describes how a

Table 2.2: Comparison of Micromobility Protocols [72]

	Hierarchical Tunneling	Local Domain Tunneling	Host-Specific Forwarding	Mobile-Enhanced Routing
Protocol Example	Mobile Regional Tunneling HMIPv6 TeleMIP	NetLMM BCMP	Cellular IP HAWAII	MER-TORA
Mobile IP as Macromobility Protocol	YES	YES	YES	Optional*
Address Management	Varying co-located CoA	Static co-located CoA or home address	Static co-located C0A or home address	Static co-located CoA
Change of CoA during handover	YES	NO	NO	NO
Support prefixed-based routing	YES	NO	NO	YES
Packet Routing	Sequential Tunneling	Sequential Tunneling	Host-based Forwarding	Network-ID based routing
Packet redirection during handover	Tunneling	Tunneling	Host-based Forwarding	Tunneling
Type of updating message	Explicit signaling (based on Mobile IP)	Explicit Signaling	Explicit signaling or implicit data packet snooping	Explicit signaling
Topology Required	Hierarchical Mesh	Hierarchical Mesh	Hierarchical Tree	Hierarchical Mesh
Scalability	Good	Good/Average	Poor	Good
Reliability	Average	Average	Average/Poor	Average
Discovery mechanism	MIP Agent Advertisement	MIP Agent Advertisment	Layer 3 beaconing	Layer 3 beaconing
Paging Support	Yes (with extension)	No	Yes (built in)	No
Paging cache placement	Absolutely centralized	N/A	Located in selected paging nodes	N/A
Paging cache update mechanism	By regional binding update	N/A	By all uplink update packets and data packets	By regional binding update

Mobile IPv6 Fast Handover could be implemented on link layers conforming to the IEEE 802.11 suite of specifications.

2.4.2 Context Transfer

Context Transfer aims to minimize the impact of certain transport/routing/security-related services on the handover performance [13], [16]. When a mobile node (MN) moves to a new subnet it needs to continue such services that have already been established at the previous subnet. Such services are known as 'context transfer candidate services', and examples include AAA profile and IPsec state, header compression, QoS policy etc. Re-establishing these services at the new subnet will require a considerable amount of time for the protocol exchanges and as a result time-sensitive real-time traffic will suffer during this time. Alternatively, context transfer candidate services state information can be transferred, for example, from the previous access router to the new access router so that the services can be quickly re-established. A context transfer protocol will result in a quick re-establishment of context transfer candidate services at the new domain. It would also contribute to the seamless operation of application streams and could reduce susceptibility to errors. Furthermore, re-initiation to and from the mobile node will be avoided hence wireless bandwidth efficiency will be conserved.

TEXT (Time efficient Context Transfer) [58] protocol was also proposed to forward context and support the handover operation but in this case it aimed at forwarding actual traffic. It is developed based on the same philosophy as that in fast Mobile IP post-registration process to be specific, namely: to continue forwarding MNs traffic through a bi-directional tunnel between PAR and NAR as soon as the MN establishes an L2 connection with the NAR. The tunnel stays in place until the MN can use NAR as its default router to forward MNs traffic. In the same manner TEXT was proposed to start and complete transfer of critical context while MN is receiving its

data traffic through the tunnel via NAR. During that time, the NAR simply sends the MNs traffic without having looked into the details of the features associated with the MN. The PAR handles MNs feature processing and their associated context, i.e. acts as a context anchor for as long as the tunnel between NAR and PAR is in place. This way, the MN not only can receive its data, but also has its feature services processed at the PAR without disruption.

The context transfer protocol proposed in RFC 4067 is probably the Recently there have been many active discussions in the now closed IETF's SEAMOBY working group, aiming towards a protocol which would allow state information to be transferred between edge mobility devices. RFC 4067 is probably the most discussed describes the context transfer protocol, defining a framework of control structures that enable authorised context transfers and has been accepted as an experimental RFC [16].

2.4.3 Candidate Access Router Discovery (CARD)

To enable seamless IP-layer handover of a mobile node (MN) from one access router (AR) to another, the MN is required to discover the identities and capabilities of candidate ARs (CARs) for handover prior to the initiation of the handover. The act of discovery of CARs has two aspects: identifying the IP addresses of the CARs and finding their capabilities. This process is called *candidate access router discovery* (CARD). At the time of IP-layer handover, the CAR, whose capabilities are a good match to the preferences of the MN, is chosen as the target AR for handover [50]. IP mobility protocols, such as Mobile IP, enable mobile nodes to execute IP-level handover among access routers. Seamless IP mobility protocols will require knowledge of candidate access routers (CARs) to which a mobile node can be transferred. The CAR discovery protocol enables the acquisition of information about the access routers that are candidates for the mobile node's next handover. CAR discovery involves identifying a

CAR's IP address and the capabilities that the mobile node might use for a handover decision. There are cases in which a mobile node has a choice of CARs. The mobile node chooses one according to a match between the mobile node's requirements for a handover candidate and the CAR's capabilities. However, the decision algorithm itself is for further research.

2.4.4 Discussion

The primary objective of all of the mentioned mobility management protocols is to provide an improvement on the handover performance and thus maintain the handover delay to a minimum. For this reason previous research related to enabling seamless mobility over IP networks focused mainly on enhancing the handover procedure between access routers/base stations. However even with fast handover, packets will still be lost during change of attachment. To solve such a problem, several techniques have been proposed like Bi-Casting, HMIPv6, CARD and Context Transfer. Candidate AR discovery and context transfers form a very promising architecture to support handovers in IP networks. Nearly all work before has been based on setting up protocol state after handover by signaling new state information. With context transfers, it would be possible to keep practically all handover-related signaling within the wired links of the access network. The next chapter gives a more thorough description of the motivation behind Context Transfer as well as related research issues.

Chapter 3

Context Transfer Research Issues

In this chapter the motivation and research directions for context transfer are described. With the tremendous growth of mobile nodes in IP-based networks the routing paths through the network must be changed at every handover in order to deliver the host's IP traffic to the new point of attachment. To accommodate for this, protocols like Mobile IP have been proposed (see chapter 2). Because of the introduction of real time services such as VoIP, video etc. minimization of the impact of traffic redirection on the service becomes important. When establishing the new routing path (at the new access) the nodes must be configured to provide similar routing treatment to the IP packets as was provided along the old routing path.

Services like AAA, header compression, QoS, policies, PPP, multicasting, etc. could have a major impact on and when to establish the new routing path. These are referred to as context transfer candidate services in RFC 4067 and will be used in the rest of the thesis. From this stems a need to quickly re-establish context transfer candidate services without the mobile node performing all these protocol flows from scratch after the handover operation. Based on this motivation a number of research issues were considered and are described in this chapter.

3.1 Identifying Context Transfer Candidate Services

One of the research problems is to identify context transfer candidate services and also identify their impact on the handover performance. Below is a description of some network-related services, which are possible examples of context transfer candidate services.

AAA - Authentication, Authorization, and Accounting is a framework for controlling the access to computer resources, enforcing policies, inspecting usage, and providing the information required to bill for services. The time consumed by AAA transaction may affect the handover latency and consequently affect the ongoing sessions. During the handover, the interactions between mobile node and AAA servers need to be avoided. Context transfer could facilitate this by forwarding the AAA related information from the previous to the new access router [13].

IPsec state - where the AR may act as an IPsec gateway, in which case a security association between the MN and AR enables packets to be encrypted and decrypted between the two. IPsec [20] provides interoperable, high quality, cryptographically-based security for IPv4/IPv6. The security services offered by IPsec include access control, connectionless integrity, data origin authentication, protection against replays, encryption, and limited traffic flow confidentiality. These services are provided at the IP layer, offering protection for IP and/or upper layer protocols. These objectives are met through the use of two traffic security protocols, the Authentication Header (AH) and the ESP, and through the use of cryptographic key management procedures and protocols. The time consumed by these procedure may affect the handover latency and consequently affect the ongoing sessions. Context transfer could facilitate the IPsec key management during handover.

Multicast group membership - where the Access Router (AR) must know which multicast groups the mobile has already joined. A group key management protocol

supports protected communication between members of a secure group. A secure group is a collection of both senders and receivers communicating with other members of the group. A group key management protocol helps to ensure that only members of a secure group gain access to group data by gaining access to group keys. Context Transfer could facilitate the group key management during handover [21].

Quality of Service (QoS) - Quality of Service is of particular concern for the continuous transmission of high-bandwidth video and multimedia information. Establishing the initial QoS between a mobile node and routers in the network would require a significant number of message exchanges. Judging from existing QoS mechanisms such as DiffServ [22] and IntServ [23], re-establishing the initial QoS between the mobile node and the new access router could be very time consuming. This is undesirable and a protocol like context transfer could greatly facilitate such a service. The mobile node's QoS context could be forwarded from the previous access router to the new access router in the new subnet thus avoiding the message exchanges between mobile node and router for reinitiating the QoS at the new delivery path.

Header Compression - Real time applications in wireless environments face the problem of large packet overhead, especially for IPv6, thus header compression is required. A number of header compression schemes have been proposed by IETF. These compression schemes in general require from 1 to 4 exchanges between the last hop router and the mobile node before full compression takes place. Before this procedure completes, the header information sent over the radio network link still remains uncompressed. Context Transfer could be used to supply the new access router with the compression context used at the previous router [24].

3.2 Investigating different Context Transfer Schemes

Another research problem is to identify and compare different possible options of transferring the desired context information. Table 3.1 shows 6 different possible options or schemes for handling context transfer. The different schemes are divided into the one's that will require exchanges over the radio interface (thus involving the Mobile Node) and those that will only take place on the network side. Taking no context transfer support as a reference, one possible way to establish states at the new path is for the mobile to simply restart all protocol negotiations from scratch after the handover. With regards to context transfer one option will be to assign the MN responsibility for storing and sending context therefore updating any new access routers after the handover operation. However for some types of state it may be necessary for an access router to periodically inform the mobile node, so that it obtains an updated set of state information. A third option is to make the context transfer candidate protocols responsible for transferring their own states. This will however require modifications to all candidate protocols which would imply giving a set of requirements to each corresponding IETF WG. This is possible but it would require substantial effort for any new updates. A forth option will be to enhance the mobility management proposals which are responsible for the handover operation to transfer the state. This will demand modifications to the handover protocol used. A fifth option will be to introduce a central entity within the network to store a copy of the state and which could be downloaded and installed upon request before or after the handover operation. A sixth option is to have the previous access router informing the new access router of its state when the handover takes place. A dedicated protocol will need to be introduced for this. These are summarised here in the table below:

Which option is more suitable may depend on the particular protocol it tries to support and also the handover type i.e. predictive or reactive. Having no context transfer will

Table 3.1: Context Transfer Schemes

Signaling Over Air-Interface	No Signaling Over Air-Interface
Option 1: Restart protocol negotiations from scratch.	Option 4: Extend the Mobility Management Protocols to forward the state.
Option 2: The Mobile Node is responsible for transferring context to the new point of attachment.	Option 5: Utilise or introduce a central entity within the network to store a copy of the state which could be downloaded on demand.
Option 3: The Context Transfer Candidate Protocol is responsible for transferring its own state.	Option 6: Obtain the state from the previous access router when the handover takes place.

be the simplest option in the case where seamless handover is not necessary. Option 2, although it may be feasible for some state e.g. multicast group memberhship number it may not be the most appropriate option for forwarding security states that partly involve the network only. Option 3 could also be possible by e.g. triggering RSVP to deliver an RSVP soft state refresh at the new access router. Option 4 is a good way of utilizing an existing protocol framework which also aims to support the handover operation avoiding the need for a standalone. Utilization of the mobility management protocol also means good synchronization with the handover procedure and no need of new triggers or the need to listen to any handover event advertisements. Option 5 is probably the best choice if there is a possible central entity that can be utilised to store the context e.g. the MAP of hierarchical Mobile IPv6. Option 6 proposes a standalone protocol offering more flexibility and is the option discussed in the IETF under the SEAMOBY Working Group. Unlike options 4, 5 and 6 options 1, 2 and 3 have the disadvantage that the protocol signaling takes place across the radio interface which may be slower and more prone to errors. For these reasons the options 4,5 and

6 have been the focus of these work.

3.3 Summary and other research issues

A number of layer 2 context transfer solutions were also proposed. [75] describes recommended practices for implementation of an Inter-Access Point Protocol (IAPP) on a Distribution System (DS) supporting ISO/IEC 8802-11:1999 and IEEE 802.11 wireless local access network (WLAN) links. It describes how APs can interoperate on a common distribution system, using IAPP packets over TCP/IP or UDP/IP, as well as how RADIUS could be used to obtain information about one another. Regarding support for context transfer there were no requirements from the existing mechanisms of IEEE Std 802.11-1999 for the IAPP to carry context information between APs. However, the Context Block defined in IAPP MOVE packets could be utilised for this purpose. The actual information content and cryptographic protection of the context block will be the responsibility of the proposed standard. Layer 2 solutions and interoperability with these have not been covered in this research.

The use of context transfer is also seen in [76], more specifically PDP context transfer. In 3GPP networks a PDP context is a logical association between a Mobile Station and a public data network running across a GPRS network. The context delivers aspects such as routing, QoS (Quality of Service), Security, Billing etc. Context transfer solutions for 3GPP networks have not been considered in this work but as with IP mobility protocols providing a Layer 3 Context Transfer can provide a common solution that could be used by all access technologies in future all-IP networks. Knowing when to initiate context transfer is very important in order to get the timing right and forward the context seamlessly with the handover. Trigger signals are thus crucial in achieving exactly this. As mentioned in [13] the context transfer solution must define the characteristics of these trigger mechanisms used to initiate context transfer.

In general one of the main benefits provided by Context Transfer is the fact of avoiding re-establishment of services over radio link. This implies that the transfer of the desired state information is less prone to errors as renegotiations take place over a potentially error-prone link on the fixed network. It also means that the measures used to secure the transport of information between peers in an IP network could be sufficient for context transfer e.g. if for example IPsec is used between PAP and NAR. Therefore security issues will need to be considered on a CTCS by CTCS basis. In this chapter the reasons for the motivation behind the need for having context transfer were described as well as investigated different research issues. In the subsequent chapters the solutions proposed are described which are based on options 4, 5 and 6) and how they can support different Context Transfer Candidate Services and in particular security.

Chapter 4

Context Transfer Extension to Mobility Protocols

When a mobile host moves to a new base station it also needs to establish certain context transfer candidate services that have already been established at the previous base station and left behind. Such services include header compression, multicast group membership number, QoS policy, AAA profile and IPsec state. Re-establishing these services at the new base station will require a considerable amount of time for the protocol exchanges and as a result time-sensitive real-time traffic will suffer during this time. On the contrary preserving the context of the IP flows can contribute towards the seamless operation of the handover. As mentioned in Chapter 3 one of the options of transferring context information is to enhance the mobility protocols to provide context transfer. This option requires that the mobility protocol responsible for the handover is modified. So far the following mobility protocols have been considered for extension with context transfer capabilities:

- Mobile IP (Macromobility protocol) [9]

- Hierarchical Mobile IP (Micromobility tunnel-based protocol) [11]
- Cellular-IP (Micromobility host-specific forwarding protocol) [12]

The following sections describe different mobility protocol extensions to these protocols for forwarding the desired state information to the new access router and hence supporting context transfer.

4.1 Context Transfer extension to Mobile-IP

As a MN moves from one (sub)network to another, after establishing a link-layer connection at its new network, it sends a binding update (BU) packet to a foreign agent (FA) and its home agent (HA). This is in fact a handover process at the IP layer. Since its HA may be located far away from the MN's current point of attachment, making use of the HA while designing the context transfer operation was not considered. The closest entities involved are the MN, the previous FA (PFA), and the new FA (NFA). However, there is no message exchange taking place between the PFA and the NFA, as specified in Mobile IP [9]. Hence, explicit signaling is required for the NFA to request feature contexts from the MN's PFA. Figure 4.1 (Left) depicts how the context transfer operation can be triggered upon reception of a Mobile IP BU packet. Upon receiving a BU packet from the MN, the NFA sends a Context Update Trigger (CU-Trig) message to the PFA of the MN. The MN's PFA responds with Context Update Data (CU-Data) message, in which requested feature contexts of the MN are provided.

The context transfer procedure described above may not be sufficient, because if Mobile IP is applied for handling the mobility of MN across networks (macro-mobility), which is usually not as frequent as mobility across access routers or base stations, the handover performance may be relaxed. However, for fast handover and truly seamless handover performance, a proactive approach is preferred. Figure 4.1 (Right) shows the signaling

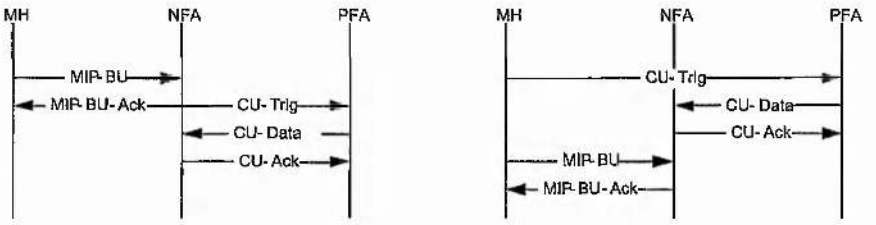

Figure 4.1: Context Transfer extension to Mobile IP (Left: Reactive Context Transfer, Right: Predictive Context Transfer)

sequence of a fast handover version of context transfer operation in Mobile IP. Here, it is assumed that MN is able to anticipate a change in network, and hence a handover, as well as to acquire information on the new foreign agent. MN sends a Context Update Trigger (CU-Trig) to its current FA (which would become PFA) prior to sending Mobile IP binding update to the NFA. Such CU-Trig message provides necessary information to the MN's current FA, and activates the FA to forward context information to the MN's NFA. The context transfer operation is likely to be carried out at the same time as the BU operation, since the MN may send the BU packet at any time after CU-Trig is sent to its previous FA. Performing context transfer prior to handover operation certainly promises better handover performance. If it is carried timely, the services used by MN at its previous network may be continued without any interruption.

4.2 Context Transfer extension to Hierarchical Mobile-IP

Hierarchical Mobile IP introduces the Mobility Anchor Point (MAP) as a local entity to assist with Mobile IP handovers. The MAP reduces the amount of signalling required outside the local domain and also supports Fast Mobile IP handovers to assist the mobile nodes in achieving seamless mobility [50], [51]. When a mobile node changes access points within a MAP domain only a single local Binding Update (BU) is required

with the MAP. This minimises latency in comparisson to Mobile IP where two BUs are send to MN's correspondent node and to the Home Agent. As with Mobile IP, the mobility solution is independent of the underlying access technology. Thus the interoperability issue, which is required for context transfer, between the different types of access networks is already taken care of by the mobility protocols.

When the mobile node changes access point within a local MAP domain it only registers its new local care of address with the MAP. The global care-of address which is already registered with the corresponding node(s) and the Home Agent does not change. Therefore when the mobile node does a local handover, it sends a BU to inform the MAP of its new local care-of address. What is proposed here is to use the BU packet as a trigger to initiate authorised context transfer from the MAP to the new access router (see Figure 4.2). MAP could be used as a central entity to store the context information and would download this to the new access router, using a context update (CU) packet, on reception of a BU packet.

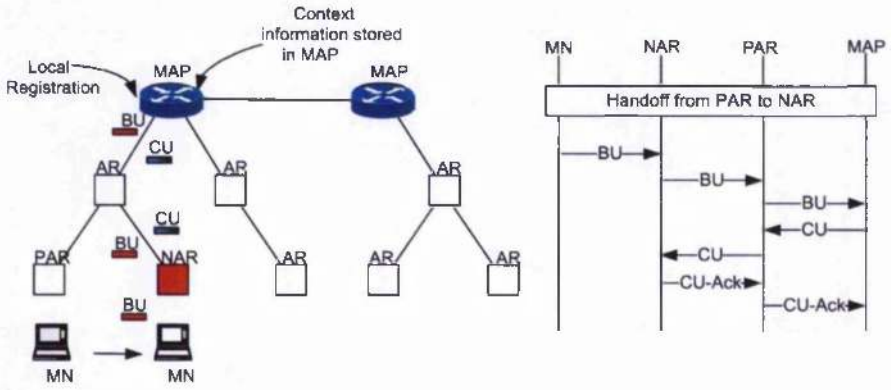

Figure 4.2: Context Transfer extension to Hierarchical Mobile IP

What is proposed in [11] is that the MAP should also be able to handle smooth handovers. When a MN handovers to a new MAP domain the MN may send a BU to the previous MAP requesting to forward packets addressed to the MNs new CoA. In

this scenario, the BU packet could be utilised to initiate context transfer between the previous and the new MAP.

4.3 Context Transfer extension to Cellular IP

The cellular-IP protocol has been designed to provide local mobility and handover support for frequently moving mobile hosts. Cellular IP can interwork with other mobility protocols like Mobile IP [9] and SIP [10] to support wide area mobility. During or immediately after handover, packet losses may occur due to delayed propagation of the new location information. The aim of cellular-IP is to minimize these packet losses in order to avoid a degradation of service quality as handovers become more frequent. The extensions to cellular-IP proposed in [15] are to offer extra functionality for forwarding the desired state information at the new base station. This context transfer mechanism will result in quick re-establishment of context transfer-candidate services at the new base station and interoperability with any layer 2 radio access technology. It would contribute to the seamless operation of application streams and would reduce susceptibility to errors. Re-initiation of services to and from the mobile node will be avoided and hence latency will be reduced.

4.3.1 Cellular-IP protocol extensions

Within a cellular-IP domain, during a handover from one Base Station (BS) to another, cellular-IP control packets could be used to initiate and transfer authorised context from the CIP-GW to the New Base Station (NBS). The context information will be stored at the CIP-GW and a copy of this context (state information) will be forwarded to the NBS. One of the main advantages of using cellular-IP is the distinction it makes between idle and active users. This separation allows the network to follow a mobile

node in active state from BS to BS and deliver packets without searching for the mobile host. By separating the caches for active and idle mobile hosts only a smaller cache needs to be searched for most of the packets which results in faster lookups and better scalability. This CIP advantage of separating active hosts from idle mobile hosts is also a benefit to the context transfer mechanism since it also targets active mobile hosts.

In order to incorporate this context transfer mechanism in the cellular-IP protocol the following enhancements are required:

- Introduction of a Context-Update (CU) packet
- Introduction of Context cache at each cellular-IP leaf node (Leaf node refers to any node that provides radio access to the mobile node).
- Re-configure the cellular-IP Route-Update packet.
- Introduction of a Context-Update request (CU-Req) packet
- Introduction of a Context-Update reply (CU-Rep) packet

In what follows, a description of each of these extensions is explained: Similarly to the Route update and paging update packets defined in [12] the context update packet will also be an ICMP packet. The basic format of an ICMP packet is shown in Figure 4.3.

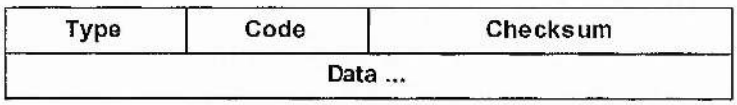

Figure 4.3: ICMP Packet Format

For the context update packet the source address will be the address of the CIP-GW and the destination address will be the NBS address. The type is a Cellular IP

control packet and the code is context-update. The payload of the context update packet carries authentication information in the same format as the route and paging update packets (see Figure 4.4) but carries control information in a different format (see Figure 4.5). The payload of the context-update packet carries authentication and control information in the following format [12].

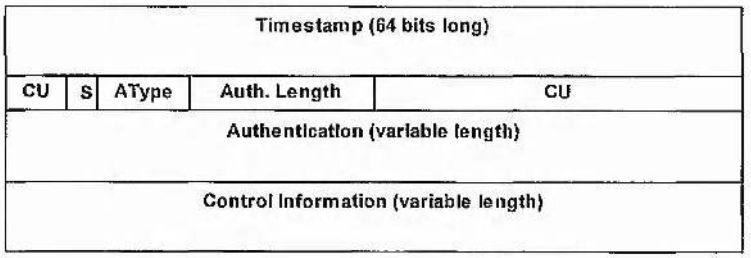

Figure 4.4: Payload of context-update packet

Timestamp - Contains a timestamp used to determine the order in which update packets are sent. The timestamp field is formatted as specified by the Network Time Protocol [9].

CU - Currently Unused. Must be set to 0.

S flag - Set to 1 to indicate semi-soft handover. Default value is 0. Any Cellular IP node that does not support semi-soft handovers may ignore this bit.

AType - Denotes the authentication method used. The default authentication is described in [73]. All authentication methods must utilize the timestamp field.

Auth. Length - Denotes the length of the authentication information in bytes.

Authentication - Contains the authentication information. Alternatively the Authentication Header [73] could be used for authenticating control packets. This is for further study.

Control information is encoded in the a Type-Length-Value format (see Figure 4.5).

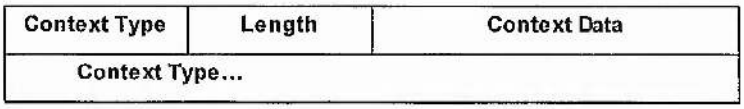

Figure 4.5: Control Information in Context-Update Packet

Context Type - Indicates the type of context information.

Length - Indicates the length (in bytes) of the following data field within. The length does not include the Type and Length bytes.

Context Data - Contains the context information of a single context type.

Similarly to the context-update packet the CU-Req will also be an ICMP packet. The source address will be the address of the new CIP-GW and the destination address will be the address of the previous CIP-GW. The type is a Cellular IP control packet and the code is CU-Req. The payload of the CU-Req packet carries a list of the desired context information. CU-Rep is also an ICMP packet. The source address will be the address of the previous CIP-GW and the destination address will be the address of the new CIP-GW. The type is a Cellular IP control packet and the code is CU-Rep. The payload of the context update packet carries the context information. Cellular IP nodes will need to be upgraded to maintain a Context Cache. Context Cache will maintain context information relating to each of the mobile hosts attached to that BS. The operation of Context Cache is summarised in Table 4.1.

One of the currently unused (CU) bits could be used as a flag which when set will indicate that the route-update packet was spawned due to a handover. The payload of the ICMP packet will be changed to the one shown in Figure 4.6.

H flag Set to 1 to indicate handover. Default value is 0.

When the route-update packet is received by the CIP-GW, if the H flag is set to 1, the CIP-GW will send a context-update packet towards the Mobile Host.

Table 4.1: Context Cache at the CIP-GW and at leaf nodes

	context cache
refreshed by	context-update packets or candidate protocol(s) packets
updated by	context-update packets or candidate protocol(s) packets
updated when	mobile host handovers to a NBS or when candidate protocol(s) renegotiate(s)
scope	active mobile hosts
purpose	maintain context information relating to the mobile host
location	CIP-GW and leaf nodes

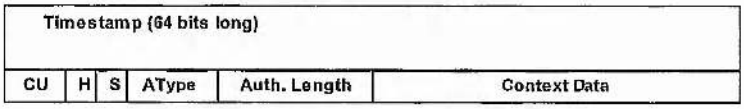

Figure 4.6: Route Update re-configuration

4.3.2 Routing

Route-update packet transmitted by the mobile host reaches the CIP-GW using shortest path hop-by-hop routing. Cellular IP nodes monitor these passing data packets and use them to create and update Route Cache mappings. These map mobile host IP addresses to downlink neighbours of the Cellular IP node. Packets addressed to the mobile host are routed along the reverse path, on a hop-by-hop basis, according to these Route mappings [12]. When the route-update packet reaches the CIP-GW, if the H flag of the route-update packet is set to 1, the CIP-GW will send a context-update packet towards the mobile host. The context-update packets will be routed along the reverse path on a hop-by-hop basis towards the mobile host. When the context-update arrives at the NBS, the NBS stores the context data in its context cache and it discards the packet. When the route-update packet reaches the new CIP-GW, if the H flag is

enabled and the GW identifies the MH as a newcomer to its domain, it requests the context information from the previous CIP-GW by sending a CU-Req packet. On reception of the CU-Req the previous CIP-GW forwards the desired context information to the new CIP-GW using a CU-Rep packet. The new CIP-GW 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 towards the mobile node. When the context update arrives at the NBS the NBS stores the context data in its context cache and discards the packet. handover is initiated from the mobile host by sending a route-update packet towards the cellular-IP gateway. When an active host approaches a new BS, it transmits a route-update packet and redirects its packets from the PBS to the NBS. The route-update packet will configure Route Caches along the way from the NBS to the CIP-GW. In most cases the paths leading to the PBS and NBS may overlap. In nodes where the two paths coincide, the route-update packet simply refreshes the old mapping and the handover remains unnoticed.

Whether the context transfer procedure takes place during or after the handover procedure, will depend on whether the cellular-IP handover used, was *semi-soft* or not. One of the extensions proposed in [12] aims to improve the performance of loss sensitive applications by introducing another type of handover called *semi-soft* handover. The handover procedure described in the previous section is known as *hard* handover and is where the mobile host switches from the PBS to the NBS all at once. With *semi-soft* handover the mobile host maintains communication with the PBS while establishing connection with the NBS. Packets intended to the mobile host are sent to both Base Stations, so when the mobile host eventually handovers it continues to receive packets without interruption [12]. The mobile host initiates the semi-soft handover by sending a route-update packet with the S flag set to 1 towards the CIP-GW via the NBS while continuing to listen to the PBS. This handover procedure will not only result in a

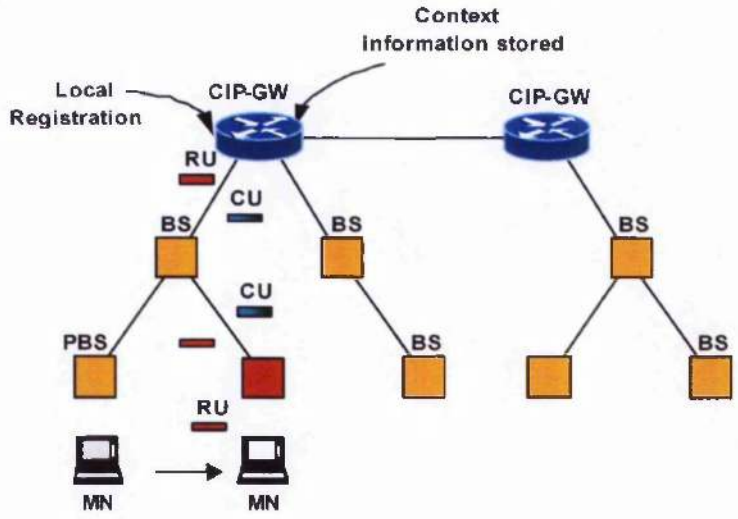

Figure 4.7: Context Transfer extension to Cellular-IP

smoother change over between base stations but it is also favoured by the context transfer extension since it provides us with a context transfer trigger (route-update packet) prior to handover. If the context transfer procedure completes before the mobile node attaches to the NBS, the NBS will have a copy of the desired state information prior to handover and consequently this will be the ideal case.

Knowing when to initiate context transfer is very important in order to get the timing right and forward the context seamlessly with the handover. Trigger signals are thus crucial in achieving exactly this. As mentioned in [13] the context transfer solution must define the characteristics of these trigger mechanisms used to initiate context transfer. The re-configured Route-Update message will be the trigger used at the Cellular-IP gateway to initiate Context Transfer from the Cellular-IP gateway to the new base station. When the route-update packet is received by the CIP-GW, if the H flag is set to 1, the CIP-GW will initiate context transfer. In the case of an intra-domain handover the CIP-GW will send a context update packet to the NBS. In the case of an

inter-domain handover (i.e. change of local domains and thus CIP-GWs) the CIP-GW will request for a copy of the context from the previous CIP-GW prior to sending a context-update packet to the NBS. The addition of the context transfer mechanism to the cellular-IP protocol should not add any disruption to the loss prone services. Here an extra packet to the cellular-IP protocol was introduced, the context-update packet, which will be used as a carrier to forward a copy of the context information from the PBS via the CIP-GW to the NBS.

4.3.3 Conclusion and Discussion

Since the context transfer mechanism proposed in this work is an extension to the cellular-IP protocol the zone of operation will depend entirely on the coverage of cellular-IP. Although cellular-IP was intended to provide mobility and handover support locally to the context transfer extensions proposed in this work provide both intra-domain and inter-domain handover support. As with the rest of the cellular-IP control packets the context-update packets will carry mandatory authentication information. In general since the context transfer extension proposed in this work is an extension to the cellular-IP protocol the security proposed for cellular-IP [12] covers the security requirements for a context transfer mechanism. This will avoid the need of using security mechanism such as IPsec [74] and TLS [73] which will create additional overhead on the header.

In this chapter context transfer extensions were proposed for Mobile IP, Hierarchical Mobile IP and Cellular IP. It was illustrated that the mobility management signaling can be initialized for triggering as well as for carrying context and that the protocol entities e.g. MAP, CIP-GW can be used for storing context. In the case of cellular IP the proposed enhancements were described in a detailed design specification. In the next chapter this proposal is evaluated.

Chapter 5

Performance Evaluation of Context Transfer enhanced Micromobility

5.1 Security Provisioning in IP Networks

In all-IP networks, the AAA infrastructure is used to authenticate and authorize the end hosts for access to network resources and for accounting and billing purposes. RFC 2865 describes the RADIUS protocol [19], a protocol for carrying authentication, authorization, and configuration information between a Network Access Server (NAS), which desires to authenticate its links and a shared Authentication Server. In a WLAN access network, the wireless Access Point (AP) acts as the NAS while a RADIUS server may act as an Authentication Server. The IEEE802.1x standard [23] has been proposed for port-based network access control for WLANs.

Figure 5.1 shows the different entities involved during IEEE 802.1x based authentication

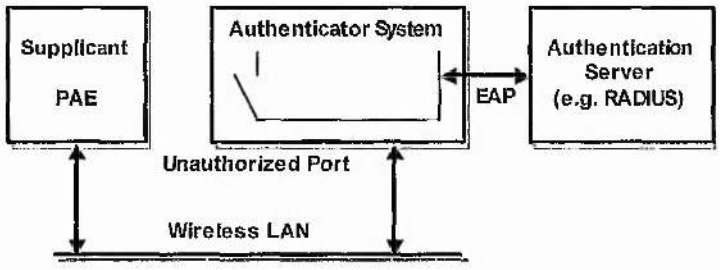

Figure 5.1: IEEE 802.1x Authentication

of a mobile host in a WLAN.

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

- Supplicant (Mobile User)
- Authenticator (Access Point)
- Authentication Server (RADIUS Server [19])

The message exchange between the mobile host (also referred to as supplicant) and the NAS takes place using the Extensible Authentication Protocol (EAP) [20]. The Extensible Authentication Protocol (EAP) allows arbitrary authentication methods using credential and information exchanges of arbitrary lengths. By using EAP, support for a number of specific authentication schemes known as EAP types may be added, including token cards, one-time passwords, and public key authentication using smart cards, certificates, and others. Strong EAP types such as those based on certificates offer better security against brute-force or dictionary attacks and password guessing than password-based authentication protocols. An Access Point (AP) that supports EAP is not required to have an understanding of the specific EAP type used in the EAP authentication process. It is aware only of when the EAP authentication process starts and ends.

EAP-TLS is a mutual authentication method, which means that both the client and the server prove their identities to each other [18]. During the EAP-TLS exchange, the supplicant sends its user certificate and the RADIUS server sends its computer certificate. If either certificate is not sent or is invalid, the connection is terminated. During the EAP-TLS authentication process, shared secret encryption keys are generated.

The authenticator must support 802.1x/EAP authentication and the supplicant and authentication server must support EAP/TLS authentication. As mentioned earlier, in this chapter, a context transfer solution for transferring AAA state information stored at the micro-mobility domain gateway to the mobile host's new base station once handover takes place is proposed. The new base station maybe within the same domain or a new domain, depending on whether the handover was inter-domain or intra-domain.

Figure 5.2 shows a signaling flow diagram of the EAP-TLS message exchanges between the mobile host, the New Access Point (NAP) and the RADIUS server before introducing the context transfer solution. The Point-to-Point Protocol (PPP) provides a standard method for transporting multi-protocol datagrams over point-to-point links. PPP also defines an extensible Link Control Protocol (LCP), which can be used to negotiate authentication methods. The Extensible Authentication Protocol (EAP) is a PPP extension that provides support for additional authentication methods within PPP.

The EAP-TLS conversation will typically begin with the New Access Point (NAP) acting as the authenticator sending an EAP-Request/Identity packet to the MN, and the MN responding with an EAP-Response/Identity packet containing the peer's userId. From this point onwards the EAP conversation take place between the Mobile Node and the RADIUS server with the NAP encapsulating and decapsulating the packets. Once having received the MN's Identity, the RADIUS responds with an EAP-TLS/Start packet. The EAP-TLS conversation will then begin, with the MN sending an EAP-

Response packet containing a TLS client hello handshake message. The EAP server will then respond with an EAP-Request packet containing a TLS server hello handshake. The data field of this packet will encapsulate one or more TLS records. This may include some or all of the following messages: TLS server hello handshake message, a TLS certificate, a Server Key Exchange, a Certificate Request, a Server Hello Done, TLS Finished and a TLS change cipher spec messages. The MN then responds to the EAP-Request with an EAP-Response packet containing the necessary keys and certificates. The message exchange continues until all the required authentication credential are exchanged and if successful the RADIUS server sends an EAP success message to the MN. As can be seen, multiple message exchanges are required between these entities before the network authenticates the mobile host. This delay could be very large especially if the RADIUS server resides far away from the new base station. Hence, it would be desirable to avoid this message exchange and find a faster to re-authenticate the mobile host.

5.2 Testbed evaluation

5.2.1 Preliminary Testbed Evaluation

Besides the theoretical reasons for the need for context transfer we have taken a pragmatic approach and evaluated a possible use case scenario. For this we investigated the impact of AAA on the handover performance of a mobile user. The handover performance was evaluated in the Wireless Network Testbed (WNT) at the Centre for Communications Systems Research (CCSR), University of Surrey. Figure 5.3 shows the network configuration used for this investigation. As we can see from the figure, the access network in this case is Wireless LAN (WLAN) and the Cellular IP (CIP) base stations are co-located with the WLAN access points (AP). At the edge of the network, there are CIP GWs co-located with Mobile IP foreign agents (FA) and SIP

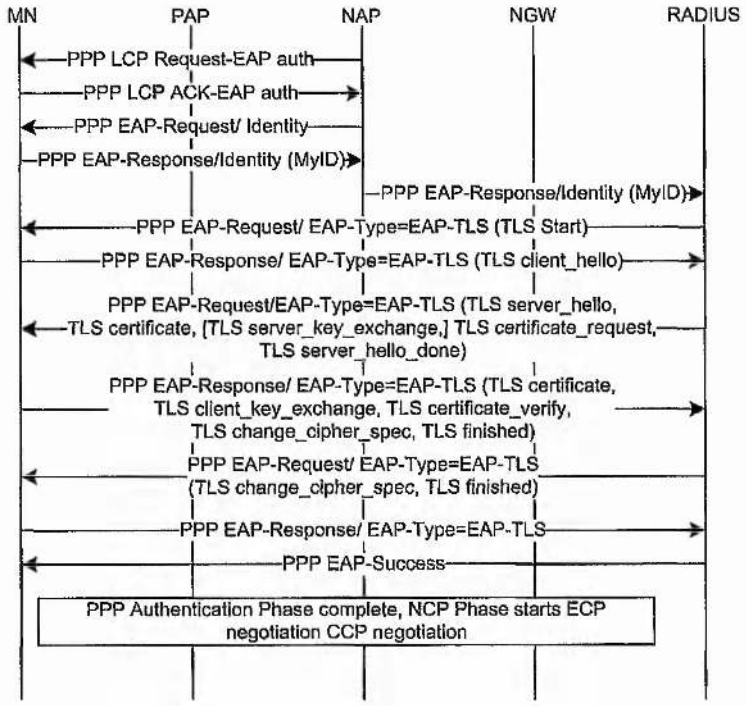

Figure 5.2: EAP/TLS procedure between Mobile Node and RADIUS server

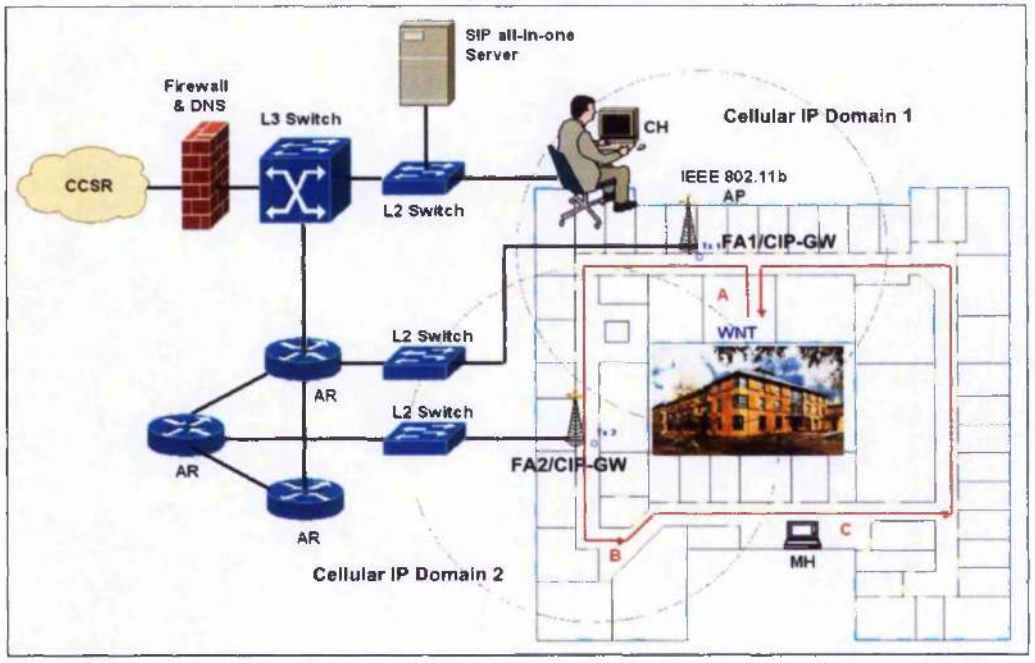

Figure 5.3: WNT Testbed Configuration

Proxies. When the MN hands off between the CIP domains, it acquires a new CoA and updates the HA. Furthermore, there is a RADIUS server (authentication server) that is used for authenticating mobile nodes.

The gateways are running on Linux PCs. The cellular IP base stations are set up on laptops that have wired interfaces to connect to the CIP-GW and wireless interfaces that act as IEEE802.11 access points. The AP functionality is realised by using wireless card enabled laptops supported by the open source hostAP driver [78]. The mobile node is a laptop with wireless connectivity. The HA service is provided by a Cisco router while the RADIUS server is also a Linux machine running the freeRADIUS software [80]. Finally the corresponding node is an FTP server.

In this test scenario, the MN first detects an access point, associates with it and is then authenticated by the RADIUS server using the EAP-TLS protocol. Mobility is

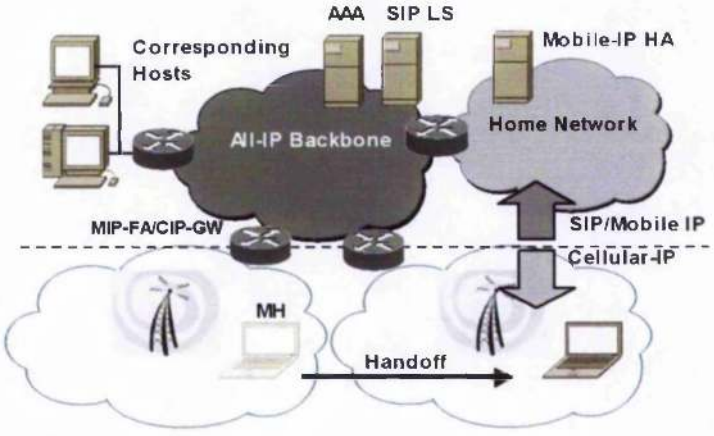

Figure 5.4: Handover between Cellular IP domains

handled as follows. Mobile IP is used to provide network macromobility and Cellular IP for network micromobility. The Session Initiation Protocol is used for providing session macromobility. Figure 5.4 shows a handover between two Cellular-IP domains. Similarly to the SIP/Cellular scheme the mobile host maintains its home address and thus during a SIP re-INVITE procedure the 'Call-ID' and 'c' fields in the SIP and SDP headers remain the same. This end system solution fits well with the Internet's design principle to obey the layer model since both have been designed to handle network layer mobility management. Cellular IP is used to support fast handover and paging as it is a highly efficient protocol for micromobility domains. Hence, we can optimise the handover performance with the interworking of SIP and Cellular-IP. When a handover takes place, the mobile host does not need to acquire a new IP address and it can keep using its home address. If the new point of attachment is within the same CIP domain, the host only has to send a Route Update message to the CIP Gateway. On the other hand, when the handover is between two CIP domains, first a Route Update is sent to the new gateway (GW). Then, the SIP UA on the mobile host will send a re-INVITE message to each of its corresponding hosts. This message will contain the address of

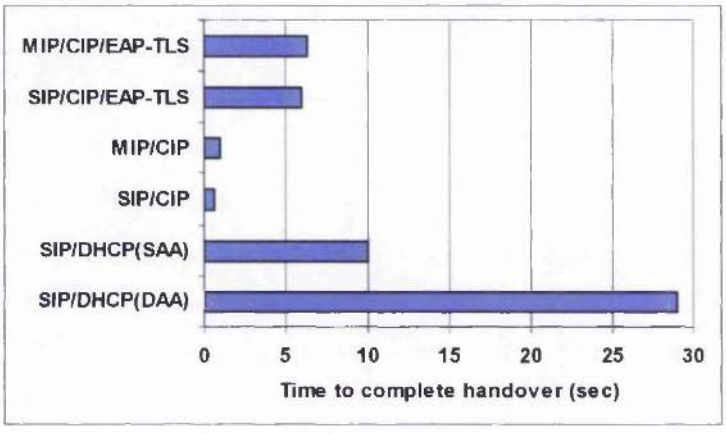

Figure 5.5: Comparison of various combinations

Cellular-IP Gateway in the Contact field and in the c field of the SDP header. Once again, the same Call-ID is maintained. Thus we assume that each domain GW will also have a SIP proxy server, which will act as the outbound proxy for the mobile hosts attached to the domain. When the INVITE is received by the corresponding host(s), it will reply back with a 200 OK message and the session(s) can be then resumed. The handover is completed by sending a REGISTER message to the home SIP server to inform it about the current location of the UA. One implication of using the Cellular-IP Gateway address is that all packets originated from the correspondent host will be encapsulated. The gateway receives and decapsulates these packets and performs a local binding table look-up to route the packets to the mobile host. On the uplink, the mobile host sends packets without any encapsulation. The handover is completed once the mobile host informs the home register of its current location.

Using the testbed configuration shown in Figure 5.3 the time to complete a handover under different integrated mobility schemes were obtained (see Figure 5.5). The figure shows the different possible combiantions used to established handover. Mobile IP together with Cellular IP, SIP together with Cellular IP and finally SIP with DHCP

(both Static Address Allocation (SAA) and Dynamic Address Allocation (DAA)). The main point to note here is the additional time required when EAP-TLS (Extensible Authentication Protocol - Transport Layer Security) is introduced [18]. As it can be seen from the graph an additional delay of 6 to 7sec is introduced in all schemes. How this delay can be minimised using CT can be found in the chapters to follow.

5.2.2 Context Transfer Solution

Figure 5.6 shows the resulting message flow when the AAA context transfer solution is used.

- Reactive Context Transfer: In this case, after the handover the MN sends a CXT-Trigger towards the new RAN. Upon the reception of the CXT-Trigger packet, the NAP (New Access Point) of the new RAN sends a CXT-Request message to the PAP (Previous Access Point) of the old RAN, which in turn forwards the requested AAA context in the CXT-Reply packet. The new RAN stores the context in its cache and forwards the context to the NAP in a CXT-Update packet. The NAP installs the context and then re-authenticates the client on the basis of the received information. This clearly demonstrates how the number of messages exchanged is reduced, thus avoiding communication with the RADIUS server but at the same time the client is authenticated by the network on the basis of the received context information.

- Predictive Context Transfer: Same as in the reactive case but the procedure up until the installation of the context at the NAP takes place before the handover.

It clearly demonstrates how the number of messages exchanged is reduced, thus avoiding communication with the RADIUS server but at the same time the client is authenticated by the network on the basis of the received context information. In this case,

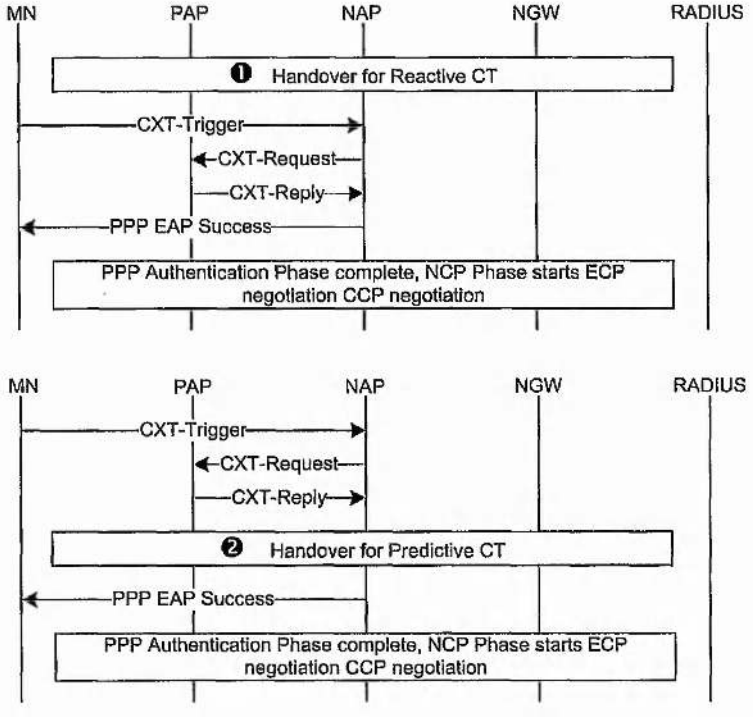

Figure 5.6: Reactive and Predictive Context Transfer

once the MN signals the handover, the new base station requests for AAA context from the previous gateway. Upon receiving the desired context, the new BS is able to authenticate the mobile host straightaway on the basis of the context information.

In order to incorporate this context transfer mechanism in the cellular-IP protocol the following enhancements are required:

- Introduction of a Context-Update (CU) packet
- Introduction of Context cache at each cellular-IP leaf node.
- Re-configure the cellular-IP Route-Update packet to indicate handover when it occurs and in such a case, to inform the new base station/gateway about the previous gateway.

- Introduction of a Context-Update request (CU-Req) packet
- Introduction of a Context-Update reply (CU-Rep) packet

Table 5.1 shows the EAP/TLS packets captured at the mobile host during the authentication procedure when an inter-domain handover takes place. For this set of observations, the context transfer has been disabled and therefore a full re-authentication is required. The handover is initiated by the Cellular IP 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 MN to the new access point (AP2) while successful authentication is indicated by the EAPOL Success message. Using the timestamps associated with these two messages, the time taken for a successful authentication can be calculated. The time difference between the Cellular IP Route Update packet and the EAP Success packet is used to determine the time taken for the handover from one BS to another and the subsequent re-authentication which in this case is:

$$\text{handover delay} = 56.523\text{-}48.304 = 8.219 \text{ sec}$$

All together the handover delay is about 8 seconds and this demonstrates that the EAP/TLS exchange is a significant delay component in this scenario. In contrast Table 5.2 shows the handover delay resulting when the Context Transfer mechanism is enabled. For this scenario the mobile host moves from AP2 back to AP1. As can be seen from the table, the handover delay has been significantly reduced from about 8 seconds to approximately 0.4 seconds. In this case, again the Route Update (with handover flag set) indicates the handover and then the context transfer takes place on between the new and previous acess points, followed by the 'reduced' re-authentication procedure based on the received context. Finally, AP1 informs the mobile host that it has been successfully authenticated by sending the EAP Success message as indicated in Table 5.2. It is important to note that the node is authenticated almost 20 times

Table 5.1: EAP/TLS signaling exchange (AAA Context Transfer Disabled)

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

Table 5.2: EAP/TLS signaling exchange (AAA Context Transfer Enabled)

Msg	Time (sec)	Source	Destination	Protocol	Info
1	59.786	MN	CIP-GW	CIP	Route Update
2	60.167	AP1	MN	EAP	Success

faster. The test was repeated a number of times and it has been observed that though the actual times vary the context transfer enabled handover is much faster than the one without context transfer scheme.

$$\text{handover delay} = 60.167\text{-}59.786 = 0.381 \text{ sec}$$

5.2.3 Effect on Real-Time Services

The two scenarios, Cellular-IP with Context Transfer (1) enabled and (2) disabled, were tested for the case where SIP/Cellular IP scheme is deployed for mobility management as a possible solution for handling mobility for real time services in all IP networks. For this test, the network is configured as shown in Figure 5.7 with the addition of SIP clients on the mobile host and the corresponding host.

A modified version of Linphone [79] was used as the SIP-based test application for evaluating the impact of the proposed context transfer solution on real time multi media services. A multimedia session is set up between the MN and the CH using the application. While the session is underway, the mobile host handovers to a new base station and the session is disrupted. To re-establish the session, the application on the MN sends a new session set up request (a SIP re-INVITE message) to the CH. The resulting SIP signaling exchanges between the mobile host (MN) and corresponding host (CH) are shown in Table 5.3 and Table 5.4 respectively. The handover delay is also shown, which in this case represents the time taken to re-establish the session after

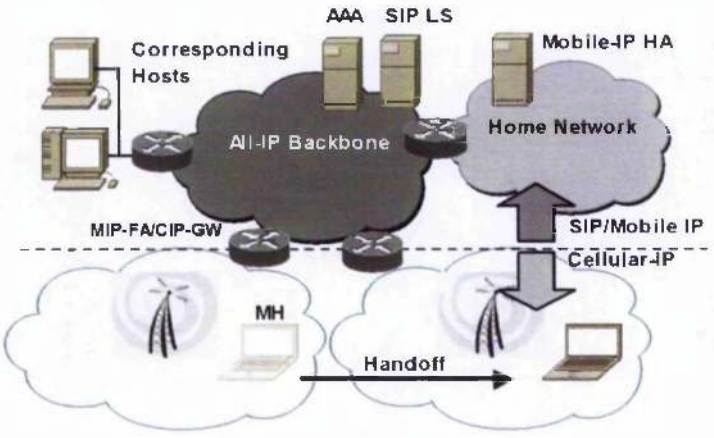

Figure 5.7: SIP/Cellular scheme

handover.

$$\text{handover delay} = 42.412\text{-}34.177 = 8.235 \text{ sec}$$

The results in Table 5.3 depict that the SIP client (MN) attempts to send a re-INVITE message towards the corresponding host (CH) and a REGISTER message towards the SIP location server (SIP LS) several times before reaching them successfully. This was due to the fact that the MN was not authenticated during the initial two attempts and so the packets could not go through to the CH via the new BS.

$$\text{handover delay} = 68.404\text{-}65.240 = 3.164 \text{ sec}$$

From Table 5.3 it is clear that it takes more than 8 seconds to re-establish the session. Hence, the multimedia session remains disrupted for this period of time. In contrast Table 5.4 shows the SIP signaling exchange when AAA context transfer is enabled. For this case the MN is authenticated significantly faster allowing the first re-INVITE and REGISTER messages to reach the CH and SIP Location Server, respectively. This

Table 5.3: SIP Signaling exchange (AAA Context Transfer Disabled)

Msg	Time	Source	Destination	Protocol	Info
1	34.177	MN	CH	SIP/SDP	INVITE
2	34.178	MN	SIP LS	SIP	REGISTER
3	34.178	MN	SIP LS	SIP	REGISTER
4	36.177	MN	CH	SIP/SDP	INVITE
5	36.178	MN	SIP LS	SIP	REGISTER
6	36.178	MN	SIP LS	SIP	REGISTER
7	38.977	MN	CH	SIP/SDP	INVITE
8	38.978	MN	SIP LS	SIP	REGISTER
9	38.978	MN	SIP LS	SIP	REGISTER
10	38.984	CH	MN	SIP	100 Trying
11	38.999	CH	MN	SIP	200 OK
12	39.001	CH	MN	SIP	200 OK
13	41.516	CH	MN	SIP	180 Ringing
14	42.407	CH	MN	SIP/SDP	200 OK
15	42.412	MN	CH	SIP	ACK

Table 5.4: SIP Signaling exchange (AAA Context Transfer Enabled)

Msg	Time	Source	Destination	Protocol	Info
1	65.240	MN	CH	SIP/SDP	INVITE
2	65.241	MN	SIP LS	SIP	REGISTER
3	65.242	MN	SIP LS	SIP	REGISTER
4	65.589	CH	MN	SIP	100 Trying
5	65.593	CH	MN	SIP	200 OK
6	65.594	CH	MN	SIP	200 OK
7	67.614	CH	MN	SIP	180 Ringing
8	68.403	CH	MN	SIP/SDP	200 OK
9	68.404	MN	CH	SIP	ACK

minimises the delay in re-establishing the session to about 3 seconds, which is mainly caused by the SIP signalling exchange and not by the authentication signaling exchange as in Table 5.4.

For both scenarios the interdomain handover was repeated several times and the results are shown in Figure 5.8.

These results indicate the improvement caused by the addition of a Context Transfer mechanism to Cellular IP. It is clear from Figure 5.8 that the handover delay when context transfer option was disabled was 8 seconds on average. The variation between the different handover attempts was due to processing time at the different access points, variation in the background network traffic, network attachment time and mobility protocol response time. On the contrary when the context transfer option was enabled the handover delay was significantly reduced to about 3.5 seconds on average. Figure 5.9 shows a breakdown of the handover delay into the major components. The total handover delay is mainly due to the Cellular IP (CIP), EAP/TLS and SIP message ex-

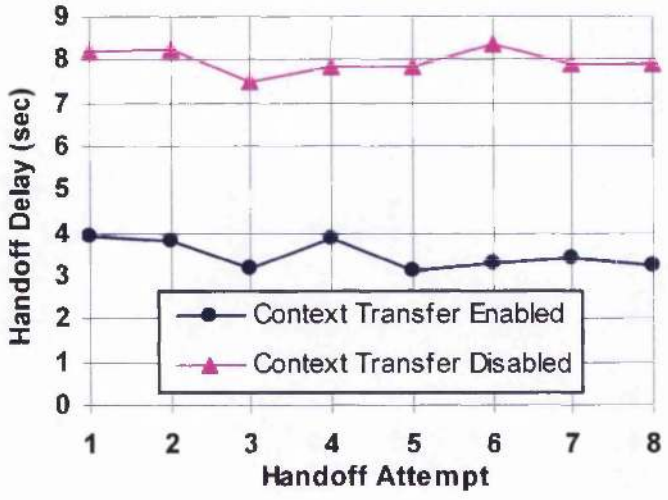

Figure 5.8: SIP signaling exchange: CT Enabled v CT Disabled

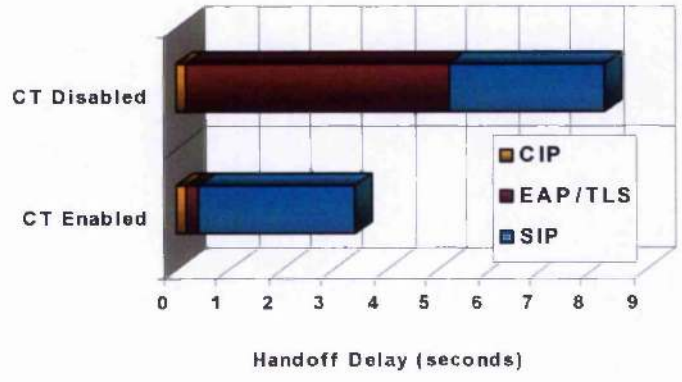

Figure 5.9: Handover delay reduction with Context Transfer

changes. Notice how when context transfer is disabled the full authentication procedure takes place introducing an undesired delay of about 5 seconds. The delay component caused by EAP/TLS is minimised to a couple of milliseconds when context transfer is introduced, reducing the overall delay from 8 to about 3.5 seconds (see Figure 5.9). The results presented here demonstrate the tremendous effect of deploying context transfer mechanism and how it aids in realizing a seamless and secure handover. For this solution existing messages of cellular IP 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, it can be claimed that the proposed AAA context transfer solution will reduce the overall handover delay by a factor of twenty. 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 mobile host without contacting the AAA server. Furthermore additional results presented here demonstrate the effect of AAA Context Transfer on SIP multimedia services when the scheme was integrated in the interworking mobility solution of SIP/Cellular IP. Due to the fast re-authentication process, the handover performance of the multimedia application was greatly enhanced and the SIP session was re-established with much reduced delay. This work demonstrates how the context transfer mechanism improves the overall handover performance and hence aids in realizing seamless and secure mobility management in all IP infrastructures.

5.2.4 Impact of AAA Context Transfer on TCP Performance

In this next test scenario, the MN first detects a base station and associates with it, and is then authenticated by the RADIUS server using the EAP-TLS protocol. It acquires a CoA from the corresponding MIP FA and registers with the HA. Furthermore, a cache entry for this node is created at the base station and the gateway and the AAA context is stored therein along with routing information. The MN then starts

communication with a corresponding host, which in this case is an FTP server. While the MN is downloading a big file from the server, it roams to another CIP domain and the authentication/registration procedure has to be repeated again before it can resume the FTP session. When context transfer is used, an improved handover performance is expected because the feature context is transferred between the access routers and the need for re-establishing such context from scratch is eliminated. The effect of context transfer on non real-time services like FTP, TELNET etc. is also investigated. Such services are as important for the mobile as are the real-time services.

In the absence of context transfer, the MN will follow exactly the same authentication procedure as was used when it first connected to the network, involving the RADIUS server. In contrast, when context transfer is used, the security context of the MN is transferred from the previous GW to the new GW and through to the new BS. The new BS can then use this security context to authenticate the MN straight away without involving the RADIUS server. The process involving registration with Home Agent remains the same in both cases. For both the cases, the TCP throughput for the FTP session is measured.

In Figure 5.10 the TCP throughput for both cases (with and without context transfer, marked by NCT and CT respectively in the graph) is shown. When context transfer scheme is deployed in the network, TCP throughput performance improves significantly. This trend is repeated every time a handover takes place, though the actual values are different. This was due to processing time at the different entities, variation in the background network traffic, packet send rate from the corresponding node, network attachment time, mobility protocol response time and the fact that the handover was executed at different times during downloading time.

It is clear from the TCP sequence number traces that context transfer helps reduce the impact of handover on the application (FTP, in this case). Note that in the absence of any context transfer mechanism, the file transfer stops for almost 7 seconds, during

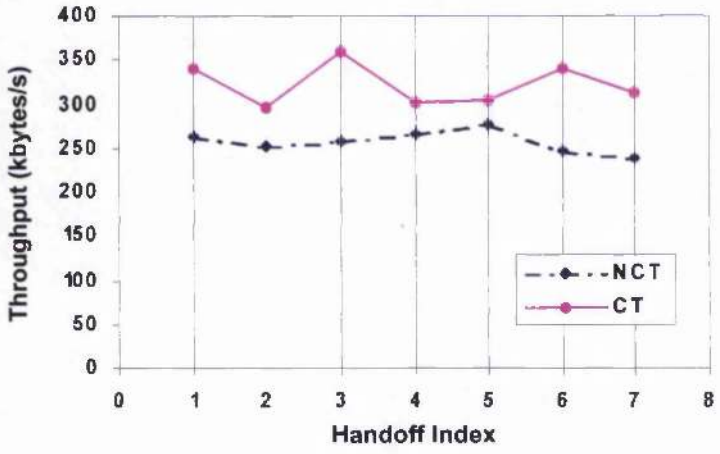

Figure 5.10: TCP throughput for the FTP session

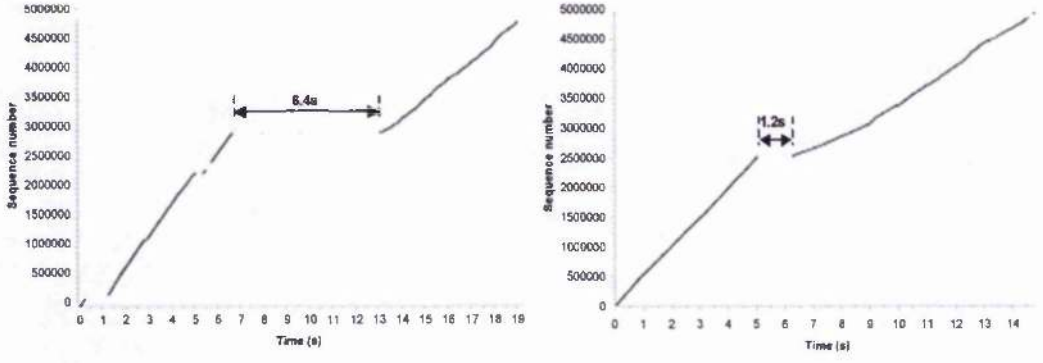

Figure 5.11: TCP Sequence Number Traces (Left: CT disabled, Right: CT enabled)

which the MN authenticates again with the new BS and then registers with the HA after acquiring a new CoA from the FA. On the other hand, when context transfer is activated, the file transfer is disrupted only for just over a second. In this case, the delay caused by re-authentication is avoided and the disruption is mainly due to Mobile IP procedures. Based on the Figure 5.10, the increase in throughout is almost 40% for some cases. Furthemore, the TCP sequence numbers for both cases were plotted, as shown in Figure 5.11. One can see that by using context transfer the downloading time was reduced from about 19 to 14 seconds. The results presented here demonstrate the tremendous effect of deploying context transfer mechanism and how it can improve the performance of services and applications during handovers.

5.3 Simulation Analysis

In this section the simulation model used for analysis is described. Unlike the testbed analysis this gave a more realistic scenario by including multiple number of users, multiple number of access points, the ability to see the impact of global mobility, authentication across the internet, impact of communication with the home domain and a more suitable network architecture for investigating different mobility management schemes and context transfer schemes. Figure 5.12 shows the network level view of the model. This model was based on the mobility simulation model used in [72]. It consists of a Home Agent (HA), an internet cloud, 1 gateway (GW), 16 ARs, and a variable number of MNs. The traffic-source and traffic-sink nodes represent correspondent nodes (CNs) of MNs as packet sender and receiver, respectively. ARs are grouped into 4 paging areas (PA), i.e. PA1-PA4. The MNs move around within the network (coverage area of 800m x 800m) and can attach to any of the 16 ARs (each covering 100m x 100m). A bursty ON/OFF traffic source is used. The traffic ON/OFF times have values of 30 and 90 seconds, respectively. During an ON period, CN has been set to create

packets at a rate of 50 packets per second. All traffic is delivered from the CNs to the MNs. Packets are being sent over the global Internet to reach the gateway (GW) of the micro-mobility domain, which then delivery packets to the point of attachment of the MN using a delivery mechanism specified by the micro-mobility protocols depending on the selected scenario.

Unless otherwise stated, the mobility model used in the simulations is the random waypoint model [83]. When a simulation begins, MNs are first placed randomly in the simulation area. Then, each node selects a destination position in a random fashion and moves towards it with a velocity selected from a predefined range. Once the destination point is reached, the node stops there for a pause time of exponentially distributed value with a mean of 60 seconds. This procedure is repeated throughout the simulation. An ideal wireless model that assumes perfect coverage, no propagation delay and no transmission errors is used. Hence, packets transmitted over the wireless interface encounter no transmission error or loss. All routers are assumed to have an unlimited buffer size. As such, the only reason for packet loss is merely due to interruption during handover. Furthermore, handovers at layer two and below are instantaneous i.e. hard handovers. All fixed links are of 10Mbps, with delay of 5ms, whereas the effective data rate of wireless link is 1.5Mbps [84]. For evaluation purposes values of timer associated to MN mobility states are the same as the configuration for Cellular IP, i.e. ready time (30 seconds), route update time (3 seconds), route timeout (9 seconds), paging update time (180 seconds), paging timeout (540 seconds). Paging functions incorporated in HMIP are based on description in [11].

5.3.1 AAA Enhancements

The base model was extended to incorporate AAA functionality included for authenticating and authorizing a mobile client (see Figure 5.13). The AAA server implemen-

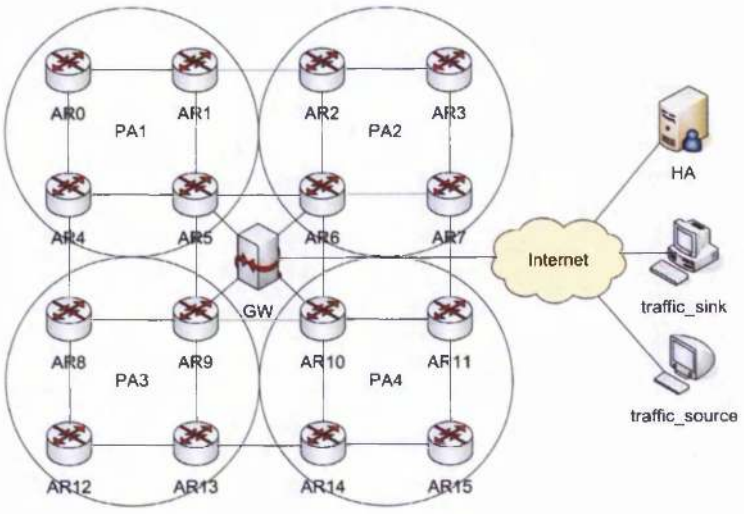

Figure 5.12: OPNET Base Simulation Model

tation is based on RADIUS and RFC 2865, the access points and mobile client were enhanced to support the Extensible Authentication Protocol (EAP) based on RFC 3748, and both the client and the server were implemented to support the EAP-TLS protocol according to [18]. The purpose of each these protocols is briefly summarized here:

- RADIUS [19] describes a protocol for carrying authentication, authorization, and configuration information between a network access server and the nodes which it desires to authenticate.

- EAP [20] the Extensible Authentication Protocol (EAP) is an authentication framework designed to support multiple authentication methods. It typically runs directly over data link layers such as Point-to-Point Protocol (PPP) or IEEE 802, without requiring IP.

- TLS [21] Transport Level Security (TLS) provides for mutual authentication,

integrity-protected ciphersuite negotiation and key exchange between two endpoints.

- EAP-TLS [18] allows a PPP peer to take advantage of the protected ciphersuite negotiation, mutual authentication and key management capabilities of the TLS protocol, described in RFC2246 [21].

In this extended scenario a roaming mobile user will need to be authenticated before being able to continue any ongoing sessions with the correspondent node (traffic source and traffic sink in Figure 5.13). The main aim of this setup is to investigate the impact possible authentication and authorization requirements will have on the user's active sessions during handover. When a mobile node changes point of attachment (i.e. access point) it will need to be authenticated based on the EAP-TLS signaling (see Figure 5.13 in previous section).

5.3.2 Context Transfer Extensions

Figure 5.14 shows a number of possible ways proposed to extend cellular IP to support context transfer. Note that these proposed solutions are all extensions to the micromobility protocol in this case cellular IP. Figure 5.14 (a) and (b) on the diagram simply show the cases for no authentication and with authentication mainly to highlight the context information stored at the leaf access point. Figure 5.14 (c) and (d) show the cases of predictive context transfer and reactive context transfer similarly to how they are defined in RFC4067. Notice however that in all cases the context passes via the gateway due to the nature of cellular IP. Figure 5.14 (e) shows the proposal where a copy of the AAA context is kept at the gateway and downloaded during handover at the new access point. Figure 5.14 (f) and (g) and two further proposals which will allow a whole branch or tree under the gateway to receive the context and allow service to the user but these i.e. 5.14 (f) and (g) are for further study.

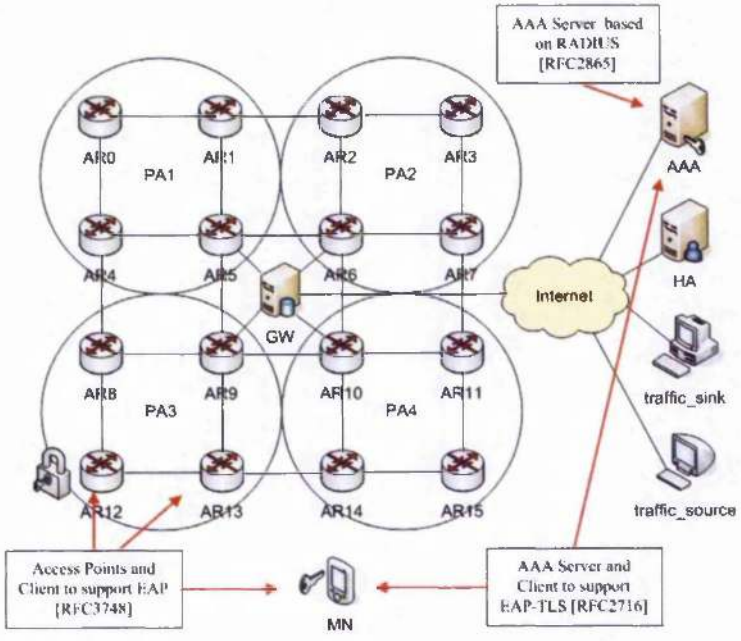

Figure 5.13: OPNET Simulation Model with AAA enhancements

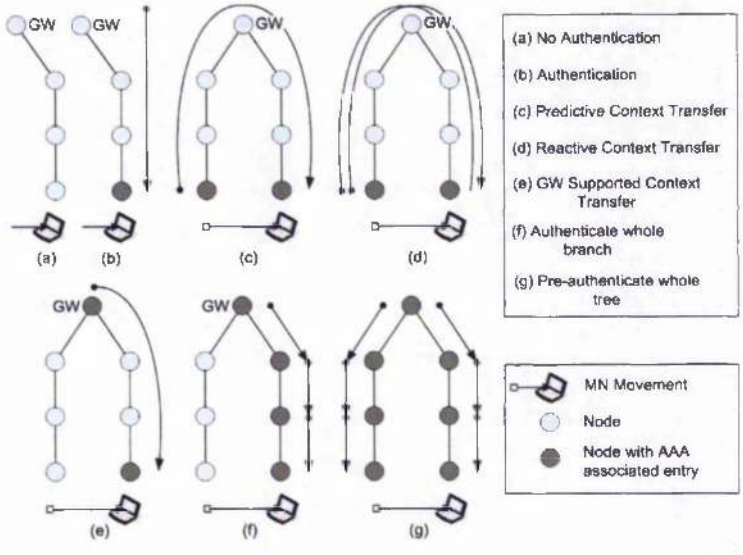

Figure 5.14: Context Transfer Schemes

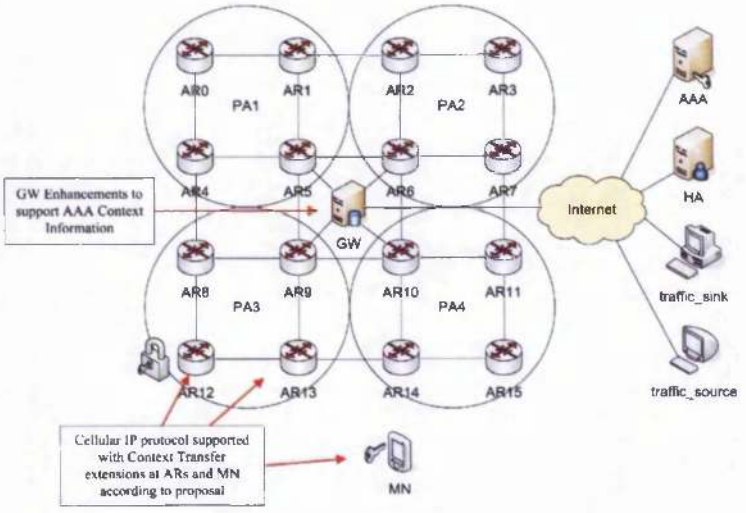

Figure 5.15: OPNET AAA Simulation Model enhanced with Context Transfer Schemes

As opposed to Figure 5.15 the following enhancements were made to the model in order to support context transfer at the mobile node, access routers and gateway (see Figure 5.15) for evaluating the proposed schemes.

Figure 5.16 is a plot of the Number of Packets lost per handover at the mobile node (MN) for different values of Constant Bit Rate (CBR) for the data traffic send from the Corresponding Node (CN).

As can be seen all context transfer extension proposals give a significant improvement in the number of packets lost during handover (see Figure 5.16). The main reason for this is that under the normal AAA scheme the signaling procedure takes place on a global scale between the home and the visited network of the mobile user. On the other hand the proposed schemes allow for the avoidance of the full AAA procedure thus minimizing the overall handover delay and hence the resulting packet loss. In the simulation model the delay caused by the normal AAA scheme procedure depends mainly upon the average internet delay. In our model this is set to 100ms. It can be however expected to vary between 30 to 300ms in general [94].

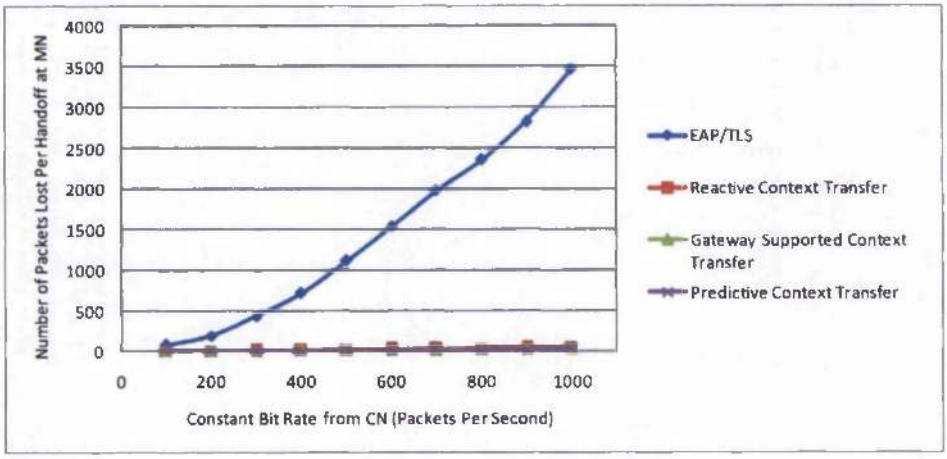

Figure 5.16: Packets Lost per handover at MN for different CBR values of data traffic send from the CN

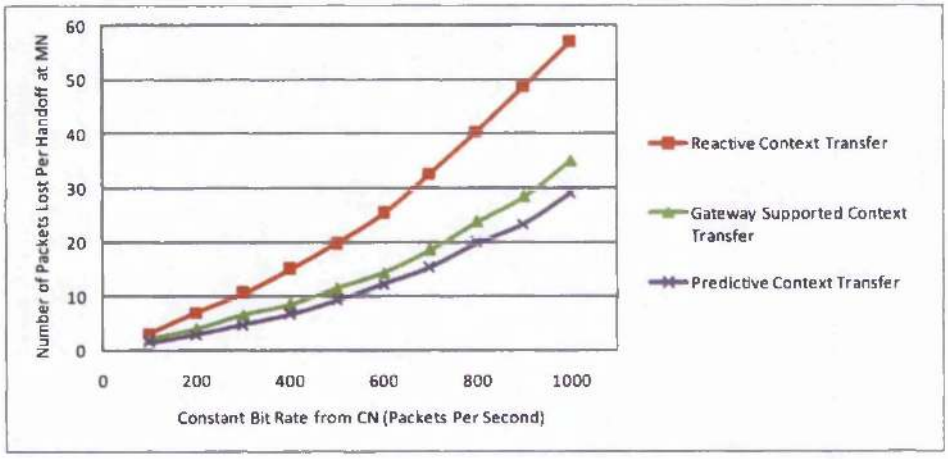

Figure 5.17: Packets Lost per handover at MN for different CBR values of data traffic send from the CN

Figure 5.17 shows a more detailed version of Figure 5.16 for the results of the proposed schemes. As expected the predictive context transfer scheme offers the best solution by keeping the AAA impact on the handover delay to a minimum. Due to the fact that the AAA context is transferred to the new access point prior to handover the access point is prepared to accept and give access to the mobile user. On the contrary in the reactive case context transfer is initiated by the mobile user once the handover is accomplished. Therefore by the time the context is obtained from the previous access point an additional delay is introduced (in this scenario approximately 2 seconds). Storing context at the gateway and downloading this on the access point on demand is an alternative solution which due to the nature of cellular IP it saves time as compared to the reactive context transfer scheme (see Figure 5.14 (d) and (e)).

The simulator results show that the number of packets lost are not actually proportional to the CBR of the data traffic send from the CN. This is due to the fact that the more the packets in the network the greater the network congestion resulting to some packets also being lost due to playout delay (i.e. packets arrive to late to be acceptable).

5.4 Analytical Modeling

In this section, the analytical model used to evaluate the performance of context transfer is described. Figure 18 shows the model used for analysis. The aim is to investigate the additional delay component faced by the MN during the handover operation introduced by the context transfer protocol procedure as compared to the AAA procedure. Furthermore to obtain a quantitative measure of the performance the packet loss during this period for different scenarios is calculated. In Table 5.5 the parameters that will be used for our analysis are defined.

The time required for the EAP-TLS signaling exchanges to be carried out can be regarded as the time from when the LCP Request-EAP auth message is send by the NAP

Table 5.5: Parameters used for quantitative analysis

Parameter	Meaning
$\lambda_{\alpha-\beta}$	Transmission rate from α to β
ϵ_{α}	Processing delay + Routing table lookup at α
δ_{α}	Latency across link α = Propagation delay + Link Layer delay
$\delta_{\alpha-\beta}$	One way delay between α and β
$\delta_{(scheme)}$	Total time required by the protocol to complete it's procedure
$\sigma_{(scheme)}$	Total number of packets lost during the protocol's procedure
w	One way delay across wireless link
δ_{nar-gw}	One way delay between new access router (nar) and gateway (gw)
δ_{par-gw}	One way delay between previous access router (par) and gateway (gw)
δ_{gw-aaa}	One way delay between gateway (gw) and AAA Server (aaa) - Internet Delay
ϵ_{α}	Processing delay + Routing table lookup at α
λ_{CN-MN}	Transmission rate from CN to MN

to the time the MN receives the EAP success message (see Figure 5.1). Theoretically one can say that end-to-end delay from α to β can be defined as the total sum of the delays across each link as well as the processing delay at each entity in between:

$$\delta_{\alpha-\beta} = \sum_{i=1}^{n} \delta_i + \sum_{j=1}^{m} \epsilon_j \tag{5.1}$$

Where n is the total number of links between α to β, and m is the total number of entities (e.g. access routers) when a packet is processed. The total time taken from the signaling exchanges of a scheme can be determined by:

$$\delta_{scheme} = \sum_{i=1}^{n} x_i \delta_i + \sum_{j=1}^{m} y_j \epsilon_j \tag{5.2}$$

Where x_i is the number of times a packet traverses link i and y_j the number of times a packet traverses, is processed by entity j.

Four schemes are evaluated in total namely predictive context transfer, reactive context transfer, gateway supported context transfer and full authentication (see Figure 5.14 for a description of these).

In Figure 5.18 a MN hands-off from the previous AP (PAP) to the new one (NAP) and requires re-authentication. A number of delay components between the main entities involved are defined. These include: (1) the delay a packet experiences across the wireless link, (2) between the APs (NAP or PAP) and the gateway (GW), and (3) between the GW and the AAA server where it is assumed equivalent to the internet delay. The processing delay at each entity is small enough to be ignored. Table 5.5 explains the parameters used for our analysis. Table 5.6 shows the different values for each parameter. For the parameter values used please refer to [85], [86] and [87].

From the model shown in Figure 5.18 and the signaling flow illustrated in Figure 5.2 the total time required for the AAA procedure can be determined:

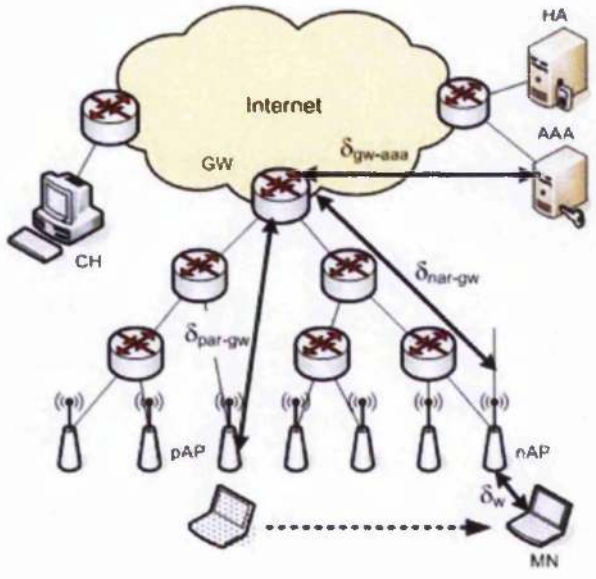

Figure 5.18: Model used for quantitive analysis

Table 5.6: Typical values for used parameters [86], [87], [94]

Parameter	Value Range	Used Value
δ_w	10-50ms	variable
δ_{nar-gw}	10ms	10ms
δ_{par-gw}	10ms	10ms
δ_{par-gw}	10ms	10ms
δ_{gw-aaa}	30-100ms	50ms
ϵ_α	0.001ms	ignore
λ_{CN-MN}	100-1000 pps	variable

$$\delta_{AAA} = 6\epsilon_{mn} + 3\epsilon_{ar} + 4\epsilon_{aaa} + 11\delta_{w} + 8(\delta_{nar-gw} + \delta_{gw-aaa}) \tag{5.3}$$

Similarly based on Figure 5.2 the time required for the context transfer signaling exchanges to be carried out can be regarded as the time from when the CXT Trigger is sent from the MN to the NAP to the time the MN receives the EAP success message from the NAP. From Figure 5.18 the procedure time required by the different schemes are:

$$\delta_{RCT} = \epsilon_{mn} + \delta_{w} + 2(\delta_{nar-gw} + \delta_{par-gw}) + 3\epsilon_{ar} \tag{5.4}$$

$$\delta_{PCT} = \epsilon_{mn} + \delta_{w} + \delta_{nar-gw} + \delta_{par-gw} + 2\epsilon_{ar} \tag{5.5}$$

$$\delta_{GCT} = \epsilon_{mn} + \delta_{w} + 2\delta_{nar-gw} + 2\epsilon_{ar} + \epsilon_{gw} \tag{5.6}$$

As can be seen, in the case of δ_{AAA} a lot more message exchanges take place across the wireless link. Using the time required by the four schemes δ_{AAA}, δ_{RCT}, δ_{PCT} and δ_{GCT}, and the rate at which the MN is receiving packets from the CN_{CN-MN} the number of packets lost can be determined.

$$\sigma_{AAA} = \lambda_{CN-MN} \times \delta_{AAA} \tag{5.7}$$

$$\sigma_{RCT} = \lambda_{CN-MN} \times \delta_{RCT} \tag{5.8}$$

$$\sigma_{PCT} = \lambda_{CN-MN} \times \delta_{PCT} \tag{5.9}$$

$$\sigma_{GCT} = \lambda_{CN-MN} \times \delta_{GCT} \tag{5.10}$$

Figure 5.19 illustrates how CT can minimize the number of packets lost for different number of bit rates. The three context transfer schemes also have minor differences between them with the predictive context transfer performing better due to the fact that the context is forwarded before the handover operation thus it is not accounted for when it comes to delay. The GW supported context transfer also minimized the

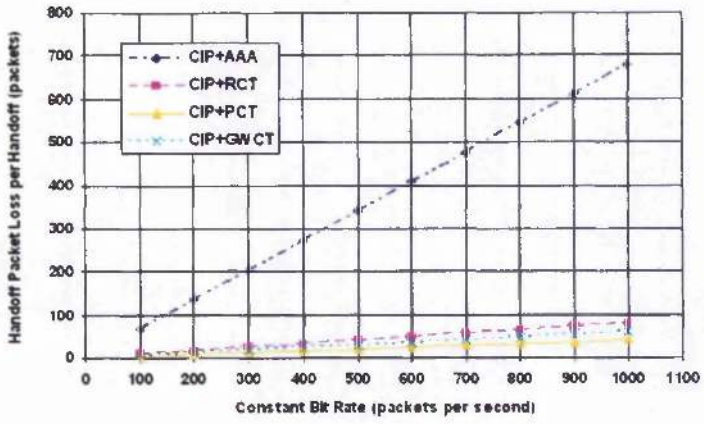

Figure 5.19: Additional packet lost per handover for each scheme for different transmission rate values

delay and packet loss as in the cellular IP scenario the GW is topologically closer to the new AR as compared to the previous access router. Figure Figure 19 also illustrates the advantage of avoiding signaling exchanges across the internet as in the case of the AAA scheme.

Comparing Figure 5.19 with Figures 5.16 and 5.17 one can see that the analytical results are similar to the simulation results. The main difference is that the analytical model does not take into account any packet buffering at the routers, playout delay, and network congestion which are the cause for more packet loss in the simulation case as traffic is increased. Also the AAA scheme gives a better performance under the analytical evaluation whereas the Context Transfer Schemes perform better under the simulation evaluation. However the results are comparable thus improving the confidence of the simulation results.

5.5 Conclusion

In this chapter a performance evaluation of a micromobility protocol enhanced with context transfer capabilities was carried out. At first a testbed evaluation was achieved where cellular IP was extended with context transfer capabilities for supporting AAA and was evaluated in three different scenarios namely: real time services, non real time services, and moving network scenarios. The results showed that in all three cases the handover delay was reduced significantly and therefore the throughput was improved. It was also realised that there was very little addition delay because of context transfer compared to the scenario without authentication.

Using a simulation model it was possible to investigate different possible schemes including predictive context transfer, reactive context transfer and gateway context transfer. All three schemes showed significant improvement in the handover delay and thus the number packets lost was reduced as opposed to the case where the full AAA had to take place. The schemes do also vary between them in terms of performance which will become apparent depending on the size of the network which the micromobility protocol covers. The larger the network the more hops the context transfer protocol messages have to cover either from the PAR to the NAR or from the GW to the NAR. These results were also verified analytically where the handover performance was evaluated for different values different values of transmission rate.

Chapter 6

Middlebox Context Transfer

This chapter proposes to use context transfer as a means for supporting middleboxes during mobility. A middlebox is defined as any intermediary device performing functions other than the normal, standard functions of an IP router on the datagram path between a source node and destination node [88]. Middleboxes enforce application specific policy-based functions such as packet filtering (firewall operation), Network Address Translation (NAT), Virtual Private Network (VPN) tunneling, Intrusion detection, Load balancing (to balance load across servers, or even to split applications across servers by IP routing based on the destination port number) etc. A MN may roam among heterogeneous wireless networks which may be protected by separate middleboxes such as Firewalls/NATs, any ongoing sessions in the old RAN may be interfered with by the firewall in the new RAN. Furthermore when the MN leaves the previous network, any open ports used for this MN's sessions will only close upon timeout leaving the firewall susceptible to numerous attacks [89], [90]. As shown in Figure 6.1, a MN may be connected to WLAN1 and communicate with a corresponding node using a certain application.

As an example while FW1 allows the MN's session to traverse when the MN's handovers

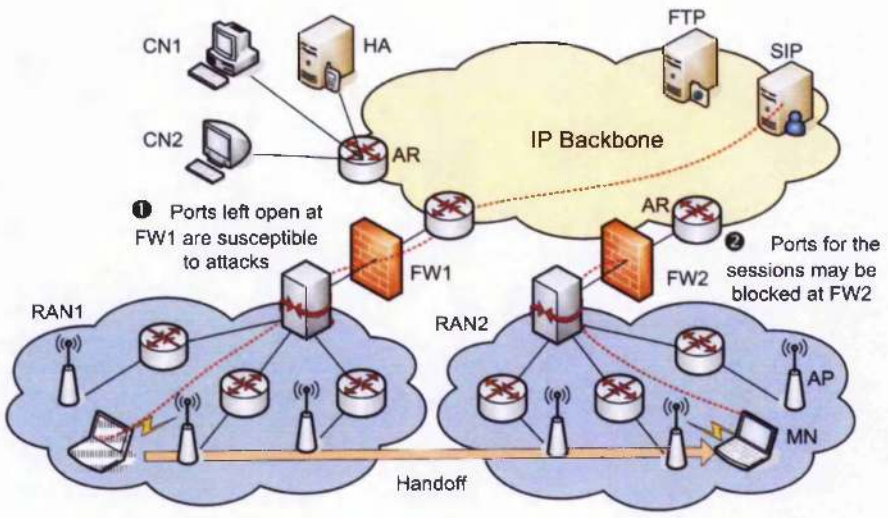

Figure 6.1: Loss of session for a MN which hands off between different radio access networks

to WLAN2, even if authorized to access the network, FW2 may block any of the users' ongoing sessions. Furthermore, the delay introduced in order to re-configure the Firewall in the new RAN i.e. 'pinholing' (the term firewall pinhole is used to describe a port that is opened through a firewall to allow a particular application to gain controlled access to the protected network), adds a significant delay to the handover latency and consequently may deteriorate the performance of the multimedia sessions.

Here the impact this may have on the multimedia sessions of a mobile user is further evaluated. Several protocols (e.g. H.323, SIP, RTSP etc) and mechanisms have been developed to support multimedia mobile applications in a future all-IP networking architecture, meeting the demands of mobile end users. However, there are certain issues associated with the handling of multimedia sessions in such a mobile environment:

Dynamic IP address and port: While the user hands off in a new RAN, he obtains a new IP address from entities such as FA (Mobile IP entity) or DHCP. This means

that a new association must be established at the middlebox for the new obtainable IP address. Furthermore multimedia signaling protocols like H.323, SIP, RTSP etc. use dynamic port to establish communication between the involved entities. These two restrictions prevent the use of static rules for middlebox devices such as Firewalls and NATs. As an example, in SIP protocol the pinholes are created according to the SDP information that is conveyed at SIP messages.

IP address fields: Headers in the multimedia signaling protocols (for example in SIP protocol the headers- contact, record-route, via, from, to) contain fields that use IP addresses instead of domain names. As an effect, these addresses are private IP addresses and need to be translated to public routable IP addresses.

Media Transport: Multimedia payload is usually conveyed from protocols such as RTP that are blocked by middlebox devices such as Firewalls/NATs. Each application uses specific RTP ports to convey the media information.

Lifetime issues: The binding between public and private IP addresses (NAT) and pinholing from incoming and outgoing traffic (Firewall) must be associated to the lifetime of each connection. These bindings will timeout on inactivity. Typical value of this inactivity is in the range of 60 seconds [19]. In case this occurs, the end-user does not receive any incoming traffic. As the number of mobile nodes within a RAN increases, the number of pinholes in the middlebox (Firewall/NAT) is increased and as an effect there is an increase for possible security compromise of the middlebox.

Session Re-Establishment: Suppose that a MN is establishing a multimedia session (e.g. SIP session, RTSP session) in a RAN and in the same time the user is experiencing a handover towards a new RAN while the session is still active. It is important to maintain the multimedia communications/session in the new RAN. This means that the session state characteristics (e.g. session id, RTP incoming/outgoing ports) must be transferred in the new RAN. The method of accomplishing session transfer depends

on the media signaling protocol. For example, in SIP MNs send a 'Re-Invite' message towards the CN. This also necessitates SIP signaling traversal from the new RAN. After the new session is re-established in the new RAN, the real-traffic communication (i.e. RTP traffic) of the media path is established. This also necessitates the dynamic potholing of the appropriate RTP ports in the NAT/Firewall at the new RAN. The above procedure can be repeated for each active multimedia session that the MN has established with the corresponding CNs. While the MN moves to a new RAN, the bindings in the old RAN remain open until there is a timeout (typically 60 seconds). This is a security vulnerability that can be spoofed by a legitimate user and as an effect the NAT/Firewall may be compromised.

Latency and Jitter: Middleboxes can degrade QoS by introducing latency and jitter. An issue is not only how fast the firewall can interact with the network traffic, but how fast it can process multimedia packets. First, the call setup process has to be done using H.323 or SIP. The presence of a NAT necessitates extra processing of each packet associated with port number.

6.1 Solution Description

During the handover, the interactions between MN, the multimedia servers and the middleboxes must be minimized. Context transfer could facilitate the above procedure by forwarding the pre-established bindings of active sessions from the middlebox (Firewall/NAT) of the old RAN to that in the new RAN. Middlebox Context Transfer solution is proposed which can be used to forward middlebox associated bindings of the MN from the old RAN to new RAN. Figure 6.2 shows a MN in RAN1 communicating with two corresponding nodes (CN1 and CN2) and then handing off to RAN2. What is proposed in this research is that upon handover, context transfer exchange can be used between the involved middleboxes to update each other's traffic control status

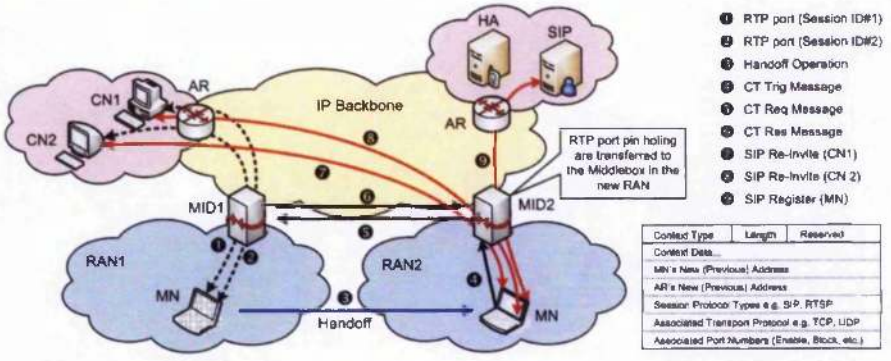

Figure 6.2: Session re-establishment using middlebox context transfer

dynamically based on these new changes. This involves a mutual communication between the middleboxes of the involved RANs. Context Transfer Protocol is used for the handshaking of this communication as shown in Figure 6.2. When the MN hands off in the new RAN2, it communicates with the middlebox by initiating a CU-Trig message. This message is sent from the MN to MID2. MID2 then requests from MID1 any bindings associated to the MN using a CU-Req message. MID1 in turn replies with a CU-Rep containing associated bindings which can then be used by MID2 to update its session or traffic control configuration.

The middleboxes in the old and new RAN may exchange the following information for the active sessions of each MN:

- Media Signalling Protocol Type (e.g. SIP, MGCP, RTSP) for which the relevant middlebox states must be transferred from the old RAN in the new RAN.

- For each active multimedia session, the information regarding the relevant open RTP ports is transferred within the context transfer protocol, so that a fast establishment 'pinholing' procedure is accomplished in the Firewall at the new RAN.

- Information associated with the session ID, traffic type, port numbers, whether should be enabled, blocked or treated according to the local policies.

- Mobile Node's new and previous IP addresses (Upon handover the MN may move to a new RAN belonging to a different administration domain and thus a new IP address is assigned).

Furthermore the handover of the MN in the new RAN can leave states (such as firewall pinholes) in place for some time in the old RAN. Such open holes may be subject to security vulnerabilities leading to middlebox compromise and Denial of Service (DoS) attacks. Context transfer can alleviate these drawbacks by deleting states along the old path and help limit any security vulnerabilities that middleboxes may face. This can be achieved using the CU-Req packet which can inform FW1 to update its traffic control table. A number of security threads are possible especially from a malicious MN. A MN which has not been authenticated and authorized before moving on the network can potentially request for context to be transferred to specific firewalls causing network disruptions. Multiple context transfer requests can also cause DoS attacks. Also a rogue firewall may transfer undesired context to neighbor-firewalls causing again network/service disruptions as well as possible DoS attacks. To avoid such attacks it is assumed that there is some kind of security (trust) relationship between the Firewall in the initial RAN and the MN which initiates the context transfer. A security association is also assumed between the involved firewalls. As proposed in [7], IPsec should be supported between the involved Firewalls. It is preferable that such a secure channel should be set up prior to context transfer to avoid additional latency and any impact on the handover performance of the MN. How these security associations are established, will depend on the security architecture and principles defined in the Ambient Networks project [23].

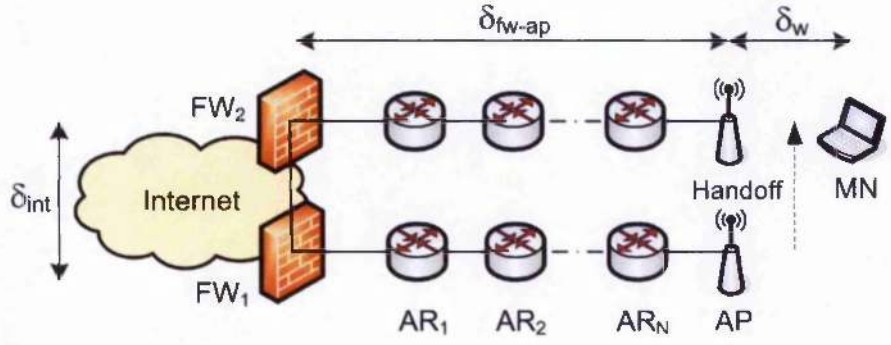

Figure 6.3: Middlebox Context Transfer model used for analysis

6.2 Performance Evaluation

6.2.1 Analytical Modeling

Figure 6.3 shows the model used for quantitative analysis. In this scenario the MN moves from one domain to another which is protected by a separate FW. For our analysis a number of delay components between the involved entities are defined.

Table 6.3 shows the components together with the selected values for analysis.

Based on Figure 6.2 the time required for the context transfer signaling procedure to be carried out can be regarded as the time from when the MN sends the CU Trig to the time MID2 opens the necessary ports. From the model on Figure 6.3 the total delay for the context transfer scheme δ_{CT} can be determined as follows:

$$\delta_{RCT}(openports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + 2\delta_{int} + 2\rho_{fw} \tag{6.1}$$

$$\delta_{PCT}(openports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + \delta_{int} + 2\rho_{fw} \tag{6.2}$$

Similarly we are interested in the time it takes for context transfer to close any ports left open at the old MID. This can be regarded as the time from when the MN sends

Table 6.1: Parameter Description

Parameter	Meaning
δ_w	One way delay across wireless link
δ_{int}	One way delay between old (FW1) and new Firewall (FW2) Internet Delay
δ_{fw-ap}	One way delay between access point (ap) and Firewall (fw)
δ_{FW}	Time it takes for a Firewall to close an unused open port
δ_h	Time it takes for the handover operation to complete
λ_{CN-MN}	Transmission rate from CN to MN
ϵ_α	Processing delay + Routing table lookup
σ_a	Packets Lost at node a
σ_h	Packets Lost during the handover operation
σ_T	Total number of packets lost

Table 6.2: Selected values for Analysis [86], [87], [94]

Parameter	Value Range	Used Value
δ_w	10-50ms	20ms
δ_{int}	30-100ms	30ms
δ_{fw-ap}	10-20ms	10ms
δ_h	40ms-2sec	200ms
ϵ_{fw}	30-60s	50s
ρ_{fw}	0.001ms	ignore
λ_{CN-MN}	100-1000 pps	100-1000 pps
ϵ_α	0.001ms	ignore

the CU Trig to the time MID1 receives the CU-Req packet. Using Figure 6.3 this can be determined as follows:

$$\delta_{RCT}(closeports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + \delta_{int} + \rho_{fw} \tag{6.3}$$

$$\delta_{PCT}(closeports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + \rho_{fw} \tag{6.4}$$

Using equations 6.1 to 6.4 and the selected values from Table 6.2 the total handover delay for each of the schemes could be calculated:

$$\delta_{RCT}(openports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + 2\delta_{int} + 2\rho_{fw} \tag{6.5}$$

$$\delta_{RCT}(openports) = 0.001ms + 20ms + 10ms + (2 \times 30ms) + (2 \times 0.001ms) \tag{6.6}$$

$$\delta_{RCT}(openports) \approx 90ms \tag{6.7}$$

$$\delta_{PCT}(openports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + \delta_{int} + 2\rho_{fw} \tag{6.8}$$

$$\delta_{PCT}(closeports) = 0.001ms + 20ms + 10ms + 30ms + (2 \times 0.001ms) \tag{6.9}$$

$$\delta_{PCT}(closeports) \approx 60ms \tag{6.10}$$

$$\delta_{RCT}(closeports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + \delta_{int} + \rho_{fw} \tag{6.11}$$

$$\delta_{RCT}(closeports) = 0.001ms + 20ms + 10ms + 30ms + 0.001ms \tag{6.12}$$

$$\delta_{RCT}(closeports) \approx 60ms \tag{6.13}$$

$$\delta_{PCT}(closeports) = \epsilon_{mn} + \delta_w + \delta_{fw-ap} + \rho_{fw} \tag{6.14}$$

$$\delta_{PCT}(closeports) = 0.001ms + 20ms + 10ms + 0.001ms \tag{6.15}$$

$$\delta_{PCT}(closeports) \approx 30ms \tag{6.16}$$

It is clear that the time required for the reactive schemes to complete is more than the predictive schemes as the signaling have to travel across a longer path to complete operation. The reactive scheme for opening the ports will add an addition delay of 30ms to the total handover delay in this case due to the additional request message between the two firewalls. In the predictive case for opening the ports however the delay of 60ms should not necessarily be added to the total handover delay as it may occur before or in the worst case during the handover operation which may in fact take longer. The scheme delays for closing the ports i.e. 60ms and 30 ms for reactive and predictive respectively are relatively small compared to default firewall configuration settings which close ports after 60 seconds or more if left open.

In the case of reactive context transfer for opening the ports at FW2, by using the time required by the context transfer scheme and the rate of which MN is receiving packets from CN, λ_{CN-MN}, the additional number of packets lost, σ_{RCT}, after a handover can be determined.

$$\sigma_{RCT} = \lambda_{CN-MN} \times \delta_{RCT} \tag{6.17}$$

In the case of the RCT scheme for opening ports and a corresponding node transmitting 500pps this will imply:

$$\sigma_{RCT} = \lambda_{CN-MN} \times \delta_{RCT} \tag{6.18}$$

$$\sigma_{RCT} = 500pps \times 90ms \tag{6.19}$$

$$\sigma_{RCT} = 45 \texttt{ packets lost per handover} \tag{6.20}$$

The number of packets lost during handover can be calculated can be calculated as follows:

$$\sigma_h = \lambda_{CN-MN} \times \delta_h \tag{6.21}$$

$$\sigma_h = 500pps \times 200ms \tag{6.22}$$

$$\sigma_h = 100 \texttt{ packets lost per handover} \tag{6.23}$$

The total number of packets lost during handover including the context transfer operation can be calculated as follows:

$$\sigma_T = \lambda_{CN-MN} \times (\delta_h + \delta_{RCT}) \tag{6.24}$$

$$\sigma_T = 500pps \times (200ms + 90ms) \tag{6.25}$$

$$\sigma_T = 145 \texttt{ packets lost per handover} \tag{6.26}$$

The number of packets lost depends on the number of packets per second sent to the Mobile Host and the handover delay. Comparing the number of packets lost during the handover operation σ_h i.e. 100 packets, as compared to the number of packets lost in the case which includes the context transfer operation σ_T i.e. 145 packets, the increase is fairly significant but the advantage of this is that the mobile node will obtain session continuation which otherwise will be lost. As will be seen in the next section, in a real scenario the handover delay is much higher than what was estimated in the analytical model.

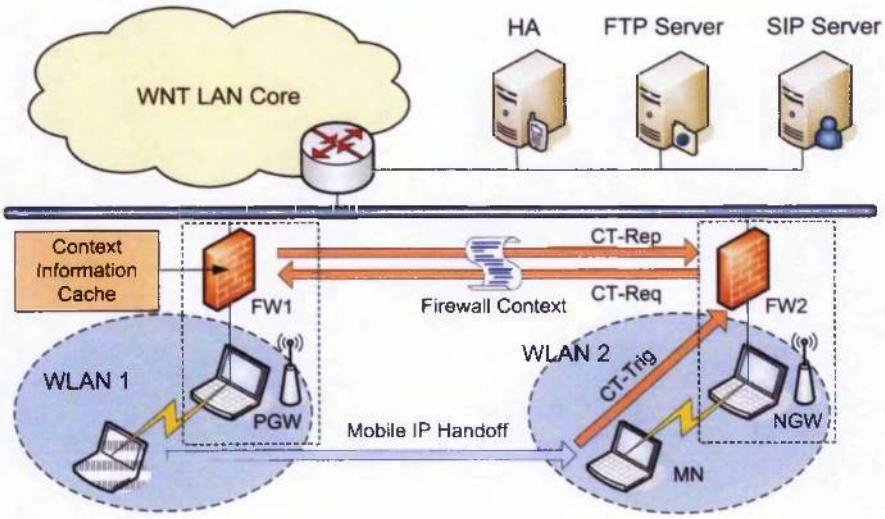

Figure 6.4: Context Transfer in both directions

6.2.2 Testbed Evaluation

The proposed solution was also evaluated using a testbed implementation. Figure 6.4 shows the setup configuration. In this scenario the mobile node has an ongoing SIP session which wants to maintain upon handover. It is assumed that FW1 has been pre-configured to permit SIP calls (destination TCP or UDP port number set to 5060) whereas FW2 was not. After the handover the mobile node sends a CT-Trig to FW2. Upon reception of the CT-Trig, FW2 sends a CT-Req message to FW1, which in turn triggers a CT-Rep packet back to FW2. The CT-Req packet is used both as a trigger for the CT-Rep packet requesting firewall port status regarding the sessions of the specific mobile client but also informs FW1 about the sessions that the mobile client was using in order to close any unused open ports dynamically without depending on the timeouts.

The CT-Rep packet also contains port and protocol information associated with the mobile client's sessions. When FW2 receives this packet it has sufficient information

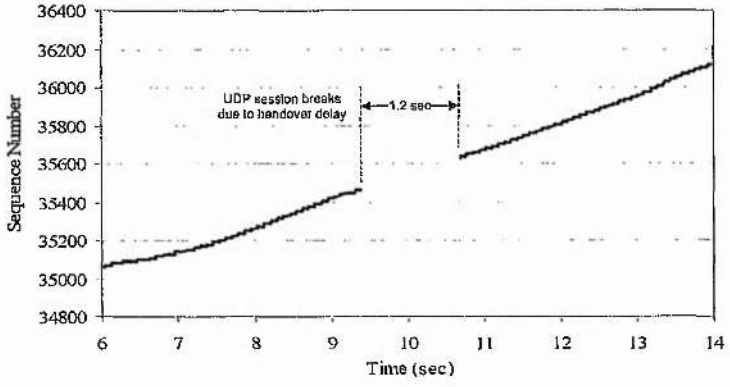

Figure 6.5: The impact of handover delay on a single UDP communication stream

to either enable or block any specific port, session number, traffic type etc. related to the mobile clients communication sessions and thus update its settings dynamically.

The performance of the Middlebox Context Transfer scheme was evaluated in the WNT for: (a) Informing the Firewall in the new RAN to make certain ports available (pinholing the Firewall) and (b) Informing the Firewall in the previous RANs to close certain unused ports. Figure 6.5 shows a UDP trace from traffic received at the mobile node. It must be noted that no Firewall was used for this measurement. The handover delay was measured from the time the last packet of the first part of the stream. This delay was caused by the combination of Mobile IP, Cellular IP and processing time at the client. The same procedure was repeated 10 times and the handover delay was ranging between 0.9-1.38 seconds with average of about 1.2 seconds. Figure 6.6 shows a UDP trace for the case when Middleboxes are set in the two domains configured with different policies: the first is set to allow UDP traffic through port 6970 whereas the second does not. Therefore once the mobile client handovers to the new domain the stream he expects to receive is blocked by the new Middlebox and the traffic is lost. Figure 6.7 shows a UDP trace for the case where Middlebox Context Transfer is enabled between the two GWs. Context transfer is used to forward information associated to

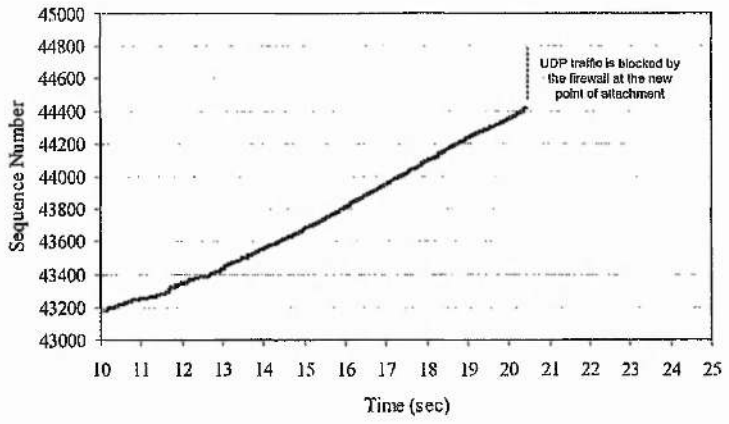

Figure 6.6: Middlebox blocks communication stream after handover

the mobile client's streams e.g. protocol type, port numbers, so that the new FW can update its policies allowing the required streams to pass through. The handover delay was again measured from the time the last packet of the first part of the stream was received until the time of the first packet of the second part of the stream was received. In this case the delay was approximately 1.7 seconds. For establishing confidence in the results the same procedure as in Figure 6.7 were repeated 6 times and the results are shown in Table 6.3. The handover delay was ranging from 1.39 to 1.83 seconds, with average of 1.633 seconds. This was an additional delay of about 0.5 seconds as compared to the case with Context Transfer i.e. 1.185 however this approach ensures session continuation.

Another set of measurements has been established. The aim of this setup was to measure the time required from the trigger sent by the mobile node to FW2, to the time it took for FW1 to close the specific ports associated with the mobile node's sessions. The results are shown in Table 6.4. It has been observed that using context transfer to inform FW1 it took on average approximately 1.36 seconds. It has to be noted that 30 and 60 seconds are common default timeouts configured at the Firewalls

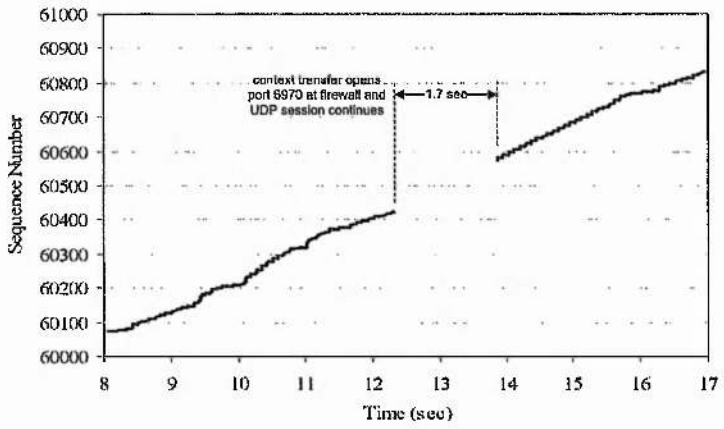

Figure 6.7: Firewall Context Transfer allows communication stream to continue

to close an unused port [20].

These results illustrate the fact that knowledge of the mobility of a user can allow the Middlebox in the old RAN to dynamically close any open ports which are related to the active sessions of the mobile user. The time of 1.36 seconds is significantly smaller to the static timeouts of e.g. 60 seconds which the firewalls are commonly configured to do. This time difference can give a much smaller opportunity for attacks such as port scanning. Therefore the proposed solution not only maintains session re-establishment in the new RAN, but minimizes any middlebox compromise in the old RAN due to user's mobility.

6.3 Conclusion

This chapter described and evaluated a mechanism of forwarding secure state information between middleboxes belonging to different Radio Access Networks. Two applications for context transfer were proposed:

Table 6.3: The impact of handover on the session re-establishment

Handover Index	Time (sec)	FWCT*
1	1.31	No
2	1.35	No
3	1.14	No
4	1.21	No
5	0.97	No
6	1.13	No
Mean Value	1.185	No
1	1.78	Yes
2	1.83	Yes
3	1.55	Yes
4	1.61	Yes
5	1.39	Yes
6	1.64	Yes
Mean Value	1.633	Yes

*FWCT - Firewall Context Transfer Enabled

Table 6.4: Time required by the context transfer scheme to close unused ports at FW1

handover Index	Time for closing ports (sec)
1	1.42
2	1.35
3	1.29
4	1.41
5	1.26
6	1.40
7	1.39
8	1.38
9	1.39
10	1.31
Mean Value	1.36

- How context transfer could support middleboxes to provide multimedia session continuation.
- How context transfer can minimize the risk for network attacks by closing unused ports in the middleboxes (Firewalls).

In the first case it was demonstrated that the context transfer protocol could be employed for the purpose of forwarding certain security information between the firewall in the old path and the firewall in the new path of the sessions of a mobile user. It is generally possible that a new firewall may be blocking certain session ports, communication protocols, security protocols which will result in session discontinuation. An even more critical scenario is if the new firewall is blocking any mobility management used for handling mobility. For example if the host identity protocol is used for host mobility but the new firewall interference with this end-to-end protocol then the

handover will fail. Similarly if SIP is used for session mobility or SHIMv6 is used for handling mobility and multihoming any interference of the Firewall may cause session handover failure or session termination. The second proposal was the use of context transfer for supporting the firewall in the old path for updating its configuration and minimizing security attacks. Defualt values for firewalls to close any unused ports are of the order of 60 seconds. With the use of context transfer it was possible to remotely reconfigure the firewall in a time of about 1 second.

Chapter 7

Context Transfer Module

The previous chapters investigated possible ways of how mobility protocols extended with context transfer capabilities, evaluated their performance for supporting a AAA scenario and proposed and evaluated a middlebox support application for context transfer. In this chapter, a module-based approach for context transfer is proposed as a way for next generation networks. If context transfer is considered as a localized mobility management support protocol due to its *orthogonal to the handover operation* nature it usually requires the involvement of new and previous access routers as well as the involvement of a mobile node before and after a handover operation. In a scenario where context transfer is used to support a standardized end-to-end mobility management protocol both context as well as mobility management operations have to be supported on the mobile node thus adding complexity on the terminal. Furthermore to support both predictive as well as reactive handover we require the involvement of the terminal before and after handover which implies shifting the protocol's procedure in the overall network attachment and handover management operation. For the protocol to be adopted as part of future architectures the specification and signaling requirements have to be concise (as opposed to a generalised form) simplified as much as possible

and the right applications to be identified. In this chapter context transfer is considered from an architectural point of view which resulted in the proposal of a standalone context transfer module.

7.1 A module-based approach

When considering the integration of RFC4067 in different architecture scenarios the problem space and scope of context transfer widened. The majority of scenarios that context transfer could be utilised for during a handover operation may go beyond the simple case of forwarding information from previous to new access router. Although the MN may handover to a geographically localised router, routing wise it may be across different domains. Moreover it could involve heterogeneous access technologies. Also scenarios of only forwarding context information from specific nodes e.g. GWs or nodes with a co-located FW may be considered. Furthermore context transfer could support different services that may be available on different nodes in the network e.g. in the case of AAA we would like to support the leaf nodes of the network acting as authenticators whereas in the GW scenario we may just want to support the GWs.

This has led the research in investigating a module-based approach. Implementing RFC4067 will require a context transfer state machines at the MN, PAR and NAR, a protocol implementation context transfer cache for storing information at PAR and NAR, and an interface to candidate transfer service protocol or software module. A context transfer module will minimize complexity and simplify integration. A context transfer module will be required at the entities which store the context and it will require interfaces to the mobility management scheme used as well as the candidate context transfer service it aims to support. A standalone context transfer module was therefore proposed which is arranged to forward context information related to a mobile node's sessions from a previous access router to a new access router when handover takes place

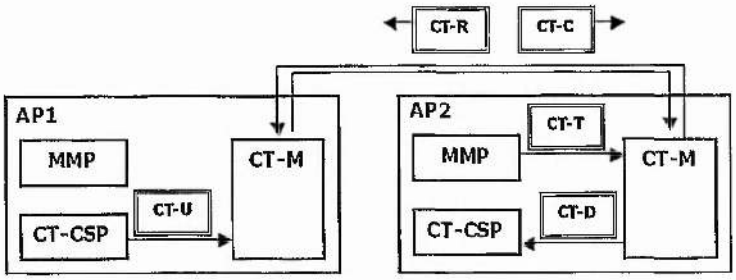

Figure 7.1: Reactive Context Transfer

in a mobile communication network.

This context transfer module (see CT-M in Figure 7.1) is arranged to reside at one of the access routers and to provide a message framework for the interworking between itself (CTM), the mobility management protocol (MMP) and the context transfer candidate service protocol CTCSP. The messages exhanged between these modules have the following roles: CT-U to update the context, CT-R to request for context, CT-C to transfer context, CT-T to trigger the context operation, CT-D to download context. This context transfer module has been designed as a plug in module configurable to cooperate with a plurality of different mobility management protocols and utilise a trigger used by the mobility management protocol to trigger context transfer. The main advantage of this is that only the access routers are involved in the context transfer exchange and the mobile node remains unaware of this context transfer is taking place. Figure 7.1 and Figure 7.2 shows how the context transfer module can handle both reactive and predictive cases.

This module-based approach was evaluated in both IST Evolute [25] as well as the IST Ambient Networks [91] proposed architectures and these will be described in the following sections of this chapter.

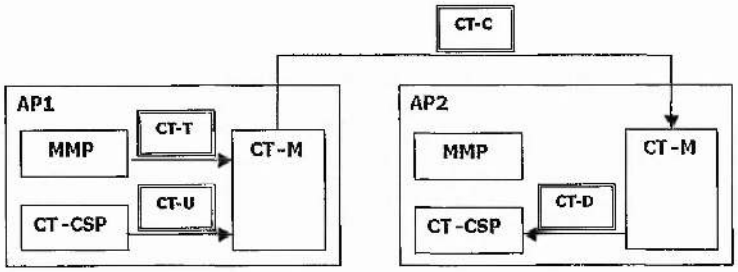

Figure 7.2: Predictive Context Transfer

7.2 Context Transfer Module as part of the Evolute Architecture

Evolute was an EU IST project which aimed at designing, specifying and developing an all-IP-based network infrastructure that will offer seamless multimedia services to users who access the network via a variety of different wireless technologies. This included a multilayer mobility management scheme to efficiently handle mobility for different types of services (real time and non real time multimedia traffic) using either network layer solutions (Mobile IP), or application layer ones (SIP) along with various IETF micromobility approaches (e.g. cellular IP, HAWAII, IDMP etc.) Other objectives included specification and development of an intelligent service provisioning environment for mobile users based on SIP and to provide fast and secure access to mobile multimedia services using a scalable and robust AAA architecture. Particularly for this research the project also aimed at designing, specifying and developing an efficient scheme for transferring context information from a mobile's old access to the new access network in order to enhance the performance of horizontal and vertical handovers. One of the main outcomes of the Evolute project was the proposal of the Evolute Mobility Gateway (EMG) which was created to support the hybrid SIP/Mobile IP multilayer mobility management architecture, described in [17]. The main aim of EMG was to

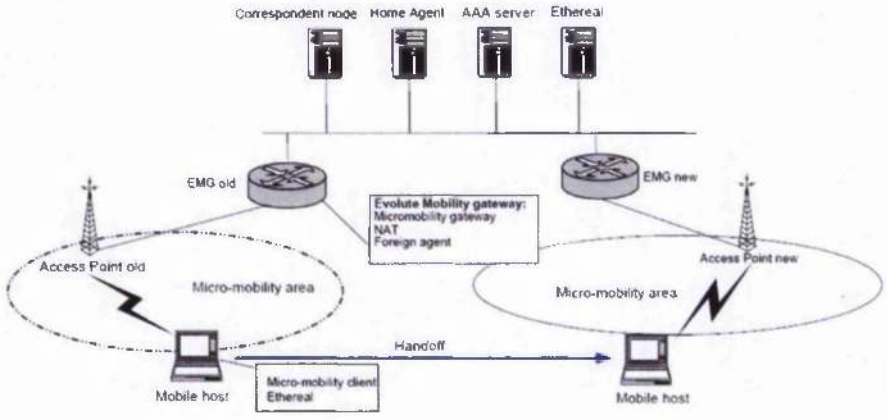

Figure 7.3: EMG testbed configuration

provide the means for supporting interworking between macromobility and micromobility protocols and was therefore tested using SIP/Mobile IP and HMIP/Cellular IP accordingly. The last step in the integration was the inclusion of the context transfer module.

7.2.1 Quantitative Evaluation

7.2.2 Description of Measurements

For the evaluation the main purpose was to collect and analyze measurements regarding the WLAN infrastructure. For the collection of the measurements the network architecture is configured as shown in 7.3. Here we show the set of measurements aimed at estimating the handover delay and the way this delay affects the throughput at the link between a mobile host and a correspondent node. A TCP session is set up between the MH and the CN using TTCP (Test TCP). While the session is underway, the mobile host handovers to a new access point under a new EMG and the session is disrupted. To re-establish the session, Cellular IP, EAP/TLS and Mobile IP messages must be

exchanged. TCP trace is used for the analysis of the measurements and ethereal as a packet sniffer.

7.2.3 Hardware and Software Characteristics

This section presents the hardware and software characteristics of the entities used to perform all tests.

The Evolute Mobility Gateway (EMG) is based on the hybrid SIP/Mobile IP/Cellular IP multilayer mobility management architecture, described in [17]. It includes a Mobile IP Foreign Agent co-located with the Cellular IP Gateway functionality. For Cellular IP the open source software from the University of Columbia was used [77]. The integration of the context transfer implementation with EMG involves the incorporation of the context module into the EMG source code.

The Access Points were based on the hostAP open source software. This software has been enhanced in order to allow the APs to send context information to the Cellular IP leaf nodes so that the local cache and the cache at the EMG can be updated. In this way the APs are able to handle the context that is sent by the EMG to the new AP when there is a handover. Access Points support IEEE 802.1x port-based authentication framework. The server part (Authenticator) that is included in the APs is provided by the Host AP driver [78] along with an Intersil Prism2/2.5/3 chipset, on the WLAN base station.

The IEEE 802.1x framework applies to the wireless hosts (Supplicants) as well. This is provided by an open source module that implements the client part of IEEE 802.1x port-based authentication framework [78]. Linphone and Kphone [79] is used as a SIP user agent, a STUN client has been integrated with Linphone and Kphone making it capable of bypassing the NAT/Firewall problems and establish real-time connections with any host. For Mobile IP, Dynamics HUT Mobile IP [81] is used.

For the Home Agent the Dynamics HUT Mobile IP [81] has been used in the HA as well. For the correspondent node a PC with one network card. For the AAA Server a PC with one network card, acting as a Radius server. The open source FreeRadius Server is used for this.

All PCs used in this experiment had the following specifications: Processor: Intel Pentium 4, CPU: 2.4 Ghz, Memory: 256 Mb, OS: Linux version 2.4.22.

7.2.4 Results

The results included here illustrated the effect of the context transfer scheme. To achieve the handover Mobile IP was used as the macro mobility protocol. The handover was repeated several times and the results appear in Figure 7.4. At Figure 7.5 the mean delay of the handover with context transfer enabled or disabled is presented. Also Figure 7.5 shows a breakdown of the handover delay into the major components:

- Cellular IP
- EAP/TLS
- MIP message exchanges

From Figure 7.4 one can see that the handover delay when context transfer option was disabled, was 4.3 seconds on average. On the contrary when the context transfer option was enabled the handover delay was significantly reduced to about 2.5 seconds on average. Figure 7.5 shows a breakdown of the handover delay into the major components. The total handover delay is mainly due to the Cellular IP (CIP), EAP/TLS and MIP message exchanges. It is worth noticing also that when context transfer is disabled the full authentication procedure takes place introducing an undesired delay of about 2 seconds. The delay component caused by EAP/TLS is minimised to a couple

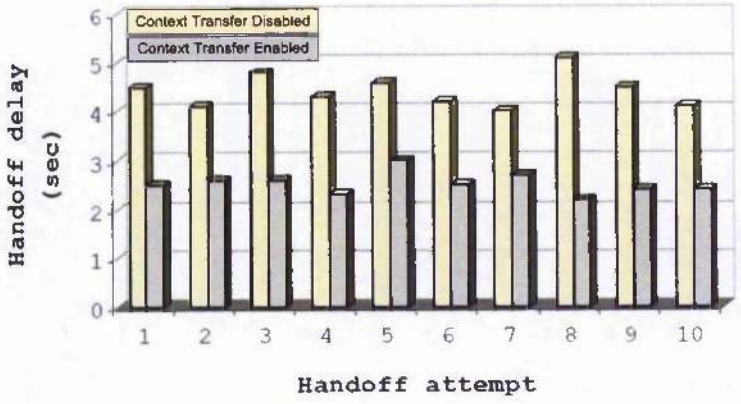

Figure 7.4: handover delay when Context transfer is enabled and when Context transfer is disabled

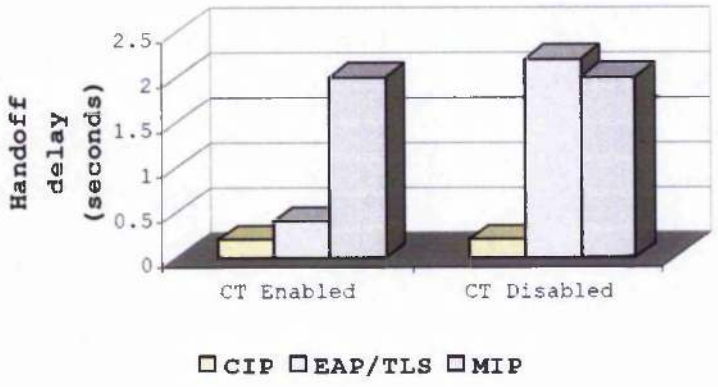

Figure 7.5: Mean delay of the handover: CT enabled v CT disabled

of milliseconds when context transfer is introduced, reducing the overall delay from 4.3 to about 2.5 seconds (see Figure 7.5).

From these results it is deduced that the performance of the EMG is improved by having context transfer module extensions. The handover delay is minimised and the throughput achieved, even in the case of multiple mobile hosts, is satisfactory.

7.3 State Transfer in Ambient Networks

Ambient Networks (AN) aim to embrace the heterogeneity arising from the different network control technologies such that they appear homogeneous to the potential users of network services [91]. One of the biggest challenges is to support the provisioning of seamless and secure mobility in such a heterogeneous environment. Until now, mobility management solutions dealt mainly with user terminal handovers between two wireless access points in an operator-controlled infrastructure; these handovers were predominantly initiated by physical relocation. However, in the emerging network scenarios considered within AN, the term "mobility" has a wider sense and involves system responses to any changes in the user and network environments, including changes in radio and network/application resources as well as commercial conditions. Furthermore, mobility solutions need to support a larger variety of mobile entities. Accordingly, it is no longer possible to envisage a single mobility paradigm that can address this diverse set of requirements. Instead, the concept of a set of solutions is introduced that can be flexibly combined and integrated on demand.

Within the context of AN, a state transfer module (STM) is proposed in order to support mobility management and retain multimedia session continuity upon handover. The idea of state transfer was introduced in [13], and [16] as a solution to minimize the impact of certain transport/routing/security-related services on the handover performance. When a mobile node (MN) moves to a new subnet it needs to maintain

such services that have already been established at the previous Radio Access Network (RAN). In [13] such services were referred to as 'context transfer candidate services', and examples of these services include AAA profile, IPsec state, header compression, QoS policy, multicast membership number, and session maintenance etc [17] and [65]. Re-establishing these services at the new subnet will require a considerable amount of time for the protocol exchanges and as a result time-sensitive real-time traffic will suffer during this time. Alternatively, context transfer candidate services state information can be transferred, for example, from the previous RAN to the new RAN so that the services can be quickly re-established.

As opposed to [16] the state transfer solution proposed in this work utilizes the AN concept, module-based approach and framework by providing a standalone state transfer module in a well defined ambient control space. The advantages of doing this include but are not limited to: synchronization with a plurality of mobility management protocols; utilization of triggers and signaling received by a common handover selection tool with the mobility management protocols, the possibility to trigger state transfer not only from the mobile terminal, use of well defined AN signaling and transport layer protocol for such a peer-to-peer communication, utilization of the module-based approach and well defined interfaces.

7.3.1 Ambient Network Architecture

The AN architecture is described in [91] and aims to support existing network services of heterogeneous networks. A core requirement in ANs is the ability for networks to compose - that is support mechanisms that achieve on-the-fly negotiations and agreements across different administrative domains; and provide the ability to reconfigure in a self managed way. There are three main components of the architecture, as shown in Figure 7.7.

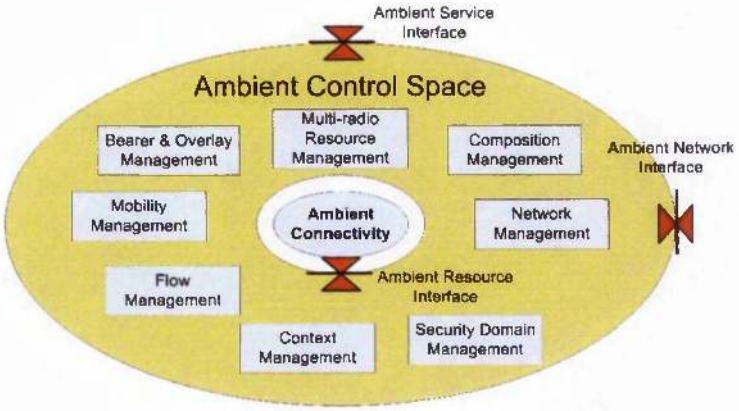

Figure 7.6: Ambient Control Space

1. The Ambient Control Space (ACS) consists of cooperating control functions. It is designed such that, although a small number of control functions is required, additional functions can easily be added or removed. The control functions can be broken down into functional entities (FEs).

2. The Ambient Connectivity abstraction layer provides the ACS with a generic, technology independent view of the underlying connectivity.

3. The Ambient network interfaces:

   - The Ambient Network Interface (ANI) connects the components of the ACS belonging to different AN; composition takes place across this interface.
   - The interface between the ACS and the connectivity is the Ambient Resource Interface (ARI); providing a homogeneous way to deal with radio access technologies and internetworking procedures.
   - The Ambient Service Interface (ASI) provides the interface to the applications and services, so that they can use the functionality provided by the ACS.

The Ambient Control Space (ACS) is the environment within which a set of modular control functions can co-exist and cooperate. The environment includes plug and play concepts that allow the ACS to bootstrap and discover the set of present functions dynamically. Further, a naming structure and registration mechanisms are defined to ensure that new functionality can be developed and integrated without impacting the overall system design and implementation. More information on the Ambient Networks Architecture can be found in [92].

7.4 Handover and Location Management

Within the ACS, the main functional entity (FE) supporting the mobility toolbox is the HandOver and Locator Management (HOLM). HOLM manages the IP layer connectivity during handover events by supporting mobility protocols and mechanisms. Figure 7.4 depicts the system architecture of HOLM and the other FEs that interact with it. It should be stressed that HOLM is not a monolithic set of protocols or modules that are available at every node, but instead the appropriate modules are used based on the specific node requirements. For example, a user terminal may require a different set of modules or protocols than an access router needs. At the core of HOLM lies the Handover Selection and Execution Control (HOSEC) module, which performs the following tasks:

- Mobility tool selection: The HOSEC module contains the decision engine for tool selection. The decision engine is a simple rule-based system that can be executed on a small mobile device.

- Protocol initiation: Depending on the specific implementation of the protocol, daemons may have to be started if this has not yet been done during system

startup. The protocol state machine may need to be initiated so that the protocol can change into operational mode.

- Coordination of composition control: Mobility control functions may need to participate in network composition.
- State monitoring and control runtime behavior: HOSEC can monitor the current state for each state machine and use this information to perform appropriate actions (e.g., coordinate the sequence of protocol steps and transitions between states).
- In the following subsections the FEs that support HOLM Triggering, Multiple Radio Resource Management (MRRM) and Policy are described.

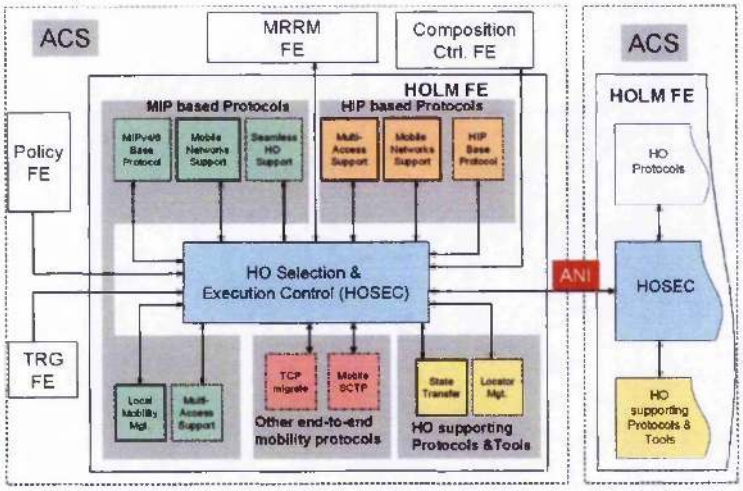

Figure 7.7: Handover and Location Management

By embracing the Ambient Networks architecture described above a State Transfer Module (STM) was proposed as part of the ACS. The modular concept of the Ambient Networks architecture provides an ideal placeholder for a standalone state transfer module. STM has been designed as a plug-in module configurable to cooperate with a

plurality of different mobility management protocols and utilise handover triggers and events [93]. The main advantage of this is that only the access routers are involved in the context transfer exchange and the mobile host remains unaware of the context transfer is taking place. Furthermore in AN we consider state transfer as module to be used in a plug-n-play fashion in a well defined architecture as opposed to a protocol which is bound by the entities involved [16] considers mobility management solutions which dealt mainly with user terminal handovers between two wireless access points in an operator-controlled infrastructure). AN involves system responses to any changes in the network environments as well as the user, including changes in radio and network/application resources as well as commercial conditions. The idea of subscribing to handover triggering events in AN is to avoid the need to synchronize with the mobility management protocol used as both STM can subscribe to the same triggering events consumed by the mobility management protocols. Moreover STM can be benefited from the well defined AN signaling and transport layer protocol and well defined interfaces for such a peer-to-peer communication in a heterogeneous network environment.

7.5 STM interfaces to other AN modules

This module belongs to the Handover and Location Management FE [95]. It is designed to complement mobility management tools during the handover operation by providing state forwarding of services like AAA, QoS, Header Compression which are established after the handover operation at the new point of attached. In the context of this work we demonstrated how STM could be utilised to remotely configure a Firewall which may interfere with any ongoing sessions at the new router or point of attachment, as part of the Ambient Networks architecture. The same protocol could be utilised in the future to forwards other protocol states like AAA, QoS or even states of FEs to avoid re-establishment of this at the new point of attachment. The motivation of

demonstrating the support of STM for Firewalls comes down to the importance and the impact of security during mobility and the user's ongoing sessions. Within ACS, STM interacts mainly with HOSEC (Handover Selection and Execution Module) which is responsible for handover slection and event notifications via the Triggering FE. STM subscribes to the Triggering FE as shown in Figure 7.8 and waits for a trigger from HOSEC during the handover operation.

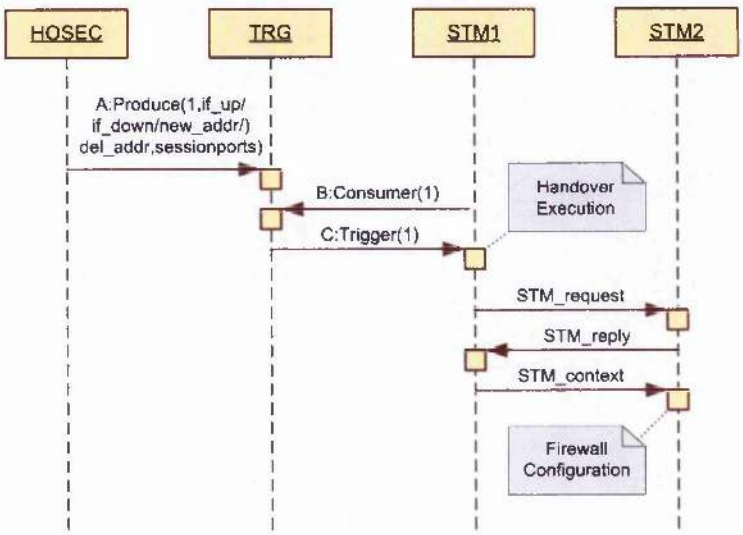

Figure 7.8: STM interactions with other ACS modules

The main interface required for communication within HOSEC is the Mobility Tool Interface (MTI). The communication between STM and HOSEC is implemented via Triggering FE [93]. STM registers to listen to MTI-common trigger ID (channel), using MTI-STM. The MTI-common channel is used to receive broadcast triggers from MTI where the MTI channel will be used for private communication between STM and HOSEC (see Figure 7.9).

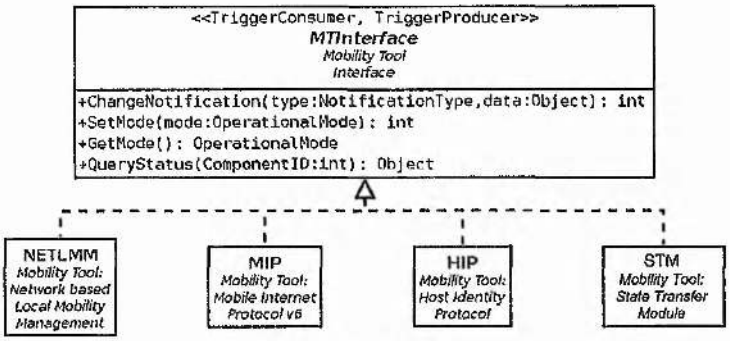

Figure 7.9: Mobility Tool Interface (MTI) used by STM

7.5.1 STM-GSLP Message Structure

The STM protocol is designed to signal bilateral operations between the ST Modules of two ACSs which are involved during the handover operation. The protocol makes use of an XML-based description to exchange information about the required state information for the different candidate services.

Each STM-GSLP XML message contains a message element, which forms the root element of each STM-GSLP message. The message element contains a body element and an optional header element.

Figure 7.10 shows the tree XML elements for an STM-GSLP message. The information present in the header may be omitted in normal protocol operation when provided to/from the underlying transport protocol.

7.5.2 STM-GSLP Message Types

STM REQUEST - Send from source STM to target STM requesting initiation of state transfer. This message includes a list of the possible STM types that could be transferred in the STM CXT message.

```
<message xsi:GenericStateCarrier="gsc.xsd">
        <header>
                address information of ACS#1
                address information of ACS#2
        </header>
        <body>
                message type
                authentication
                error code etc.
        </body>
        <payload>
                cxt type 1
                cxt type 2
                cxt type 3.
        </payload>
</message>
```

Figure 7.10: XML message structure

STM RESPONSE - Send from target STM to source STM to acknowledge acceptance of STM operation as well as an indication on which context could or could not be sent and a response to STM REQ.

STM CANCEL - Send from the target STM to reject the STM as a response to the STM INIT or at any time during the STM procedure for terminating the process.

STM CONTEXT - Used for carrying the actual context information. The payload of this packet will be subdivided to carry multiple STCS types. It will also contain a field whether certain context should be installed or removed. Having separate messages for installing or removing context could be considered.

STM ACK - Used for acknowledging a successful reception of the sent context. This message will be sent immediately after the content of STM CXT have been received.

STM SUCCESS - An optional message after completion of installation or removal of the desired context to indicate successful transfer.

STM ERROR - An error message will be useful in case either the source or target

STMs would like to request retransmission or terminate the process due to an unexpected failure.

7.5.3 Prototyping

In the AN train scenario prototype (see Figure 7.11) we utilize STM to remotely configure a middlebox at the new point of attachment (router providing access) based on a user's ongoing sessions. A middlebox is defined as any intermediary device performing functions other than the normal, standard functions of an IP router on the datagram path between a source node and destination node. Middleboxes enforce application specific policy-based functions such as packet filtering (firewall operation), Network Address Translation (NAT), Virtual Private Network (VPN) tunneling, Intrusion detection, Load balancing (to balance load across servers, or even to split applications across servers by IP routing based on the destination port number) etc. Here we con-

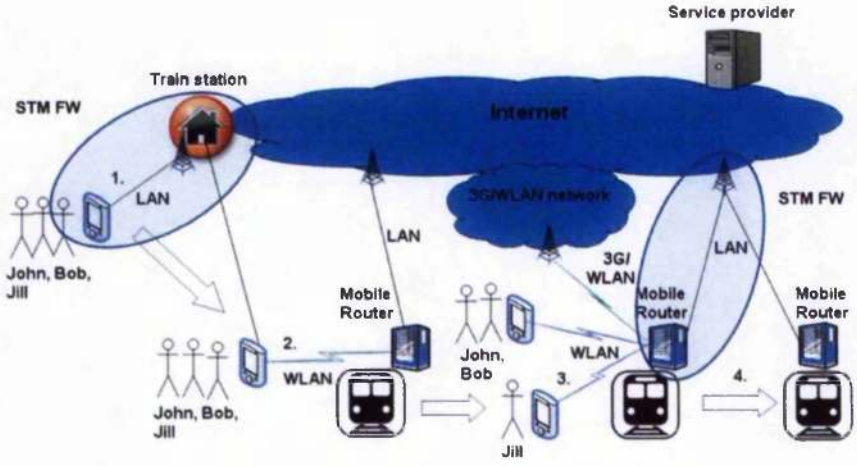

Figure 7.11: State Transfer Module in the Ambient Networks' Train Scenario

sider the scenario where a Mobile Node (MN) may roam among heterogeneous wireless

networks which may be protected by separate middleboxes such as Firewalls/NATs and any ongoing sessions in the old RAN may be interfered with, by the firewall in the new RAN. When the moving PAN or node moves to a new domain STM can support multimedia session continuation during mobility by remotely configuring the on-path Firewalls from the mobile node or through the previous access router. The prototype illustrated that the transfer of secure state information among RANs during mobility of the train could support security provisioning, minimize firewall security vulnerabilities and support in maintaining the mobile users' sessions which may otherwise be dropped.

7.6 Conclusion

This chapter described a module-based approach to state transfer and how it was integrated in the Evolute project as well as the Ambient Networks proposed architectures for next generations networks. In the Evolute the module was integrated as part of the EMG and evaluated by supporting in a hybrid multilayer mobility management architecture. In the Ambient Networks architecture the module was integrated as part of the Ambient Control Space framework and evaluated by supporting remote Firewall configuration in a moving train scenario. In both scenarios the context transfer module supported the handover operation and minimized impact on any ongoing sessions. A context transfer module will require interfaces to a handover triggering event or mobility management protocols as well as interfaces to the context transfer candidate service it supports.

Chapter 8

Conclusion

8.1 Summary and Contributions

In the context of this work three main research contributions have been proposed related to context transfer, namely: extensions to IP mobility management protocols, a new application for the protocol for supporting firewall/middlebox configuration and a standalone module for supporting integration in next generation architectures. Context transfer was evaluated in the Wireless Network Testbed at the University of Surrey as well as by OPNET simulations and analytical modeling.

At first, the idea of extending existing mobility protocols to support context transfer was considered without the need to introduce a standalone protocol. How key entities like FA in MIP, MAP in HMIP, and CIP-GW in CIP could be used as central entities to store information locally was also proposed and specified. Enhancement on how the binding update packet of Mobile IP (and HMIP) and RU in Cellular IP could be used as triggers for context transfer between the new leaf router and the gateway were also made. A more detailed specification was developed for the Cellular IP protocol and the solution was evaluated for its support on avoiding any additional delays introduced

by the AAA operation. For this solution existing messages of the Cellular-IP protocol were used as triggers and additional messages where 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 additional EAP/TLS delay by a factor of 20. This is because the full EAP/TLS procedure is avoided by transferring the AAA context to the NBS, thus enabling it to re-authenticate the MH without contacting the AAA server. Due to the fast re-authentication process, the handover performance of the multimedia application was greatly enhanced and the SIP session was re-established faster than before. Furthermore, we have evaluated the impact of the Context Transfer solution on non-real-time services. Based on the results shown here, the AAA Context Transfer solution has improved the performance of TCP-based applications significantly and the throughput increase can be as high as 40%. Besides the classic cases of predictive and reactive schemes as proposed in RFC 4067 we also proposed a scheme which included the support of the local gateway for storing context information. All schemes demonstrated the Context Transfer contribution on improving the overall handover performance in all-IP infrastructures, while supporting the provisioning of multimedia services in a seamless manner.

Two solutions for supporting Firewalls/Middleboxes at domain edges were also proposed: (a) supporting middleboxes for session continuation during mobility and (b) supporting security by remotely configuring firewalls and by minimizing the risk of network attacks. We demonstrated how the context transfer protocol could be employed for the purpose to forward certain security information from the old to the new middlebox to support session maintenance during mobility and also in the same time notify the previous middlebox to close unnecessary open ports for improved security and resolve vulnerability. Preliminary results have been included for the latter case illustrating how knowledge of the mobile movements could facilitate in closing unnecessary open ports without depending on the long timeouts. This technique could be highly consid-

ered in environments like Ambient Networks where security between involved entities is intrinsic to the security architecture. The transfer of secure state information among RANs during the mobility of users could support their security provisioning, minimize security vulnerabilities which they face and also assist in maintaining a mobile users sessions which may otherwise be dropped.

The final research contribution focused on a modular approach for context transfer as opposed to a protocol. In next generation networks the mobility management procedure may become so complex that together with network attachment, authorization, authentication etc. may result in a number of protocol and module components in next generation networks. Thus a module was proposed that could reside in the nodes sending and receiving context with the option of leaving the mobile node unaware of any context interactions thus minimizing protocol interactions between Mobile Node and Network. This module was evaluated as part of the Evolute's Hybrid Multilayer Mobility Management Architecture as well as the Ambient Networks Architectures. In both architectures the conclusion was that a modular approach can be beneficial by providing a plug-n-play standalone module, utilisation of other handover event triggers, synchronization with a plurality of mobility management protocols, easy integration in next generation networks, while assisting and improving the overall mobility management.

All three research contributions offered significant improvements to the handover performance, support for mobility management and integration in next generation networks.

8.2 Future Work

For future work it will be interesting to investigate other context transfer candidate services besides security e.g. transferring the multicast group membership number, QoS or header compression attributes could provide performance benefits. An important

issue here is that a candidate service for context transfer should only be considered in cases where it is possible to obtain a correct context transfer for the service in a given implementation and deployment, that is, one which will result in the same context at the new access router as would have resulted had the mobile host undergone a protocol exchange with the access router from scratch. Also any future context transfer solutions must inter-work with existing and emerging IP protocols, in particular, those protocols supporting mobility in an IP network. Also they must provide a performance that is equal to or better than re-initializing the context transfer-candidate service between the mobile host and the network from scratch. Otherwise, context transfer is of no benefit.

Besides handover delay and packet loss, future work may investigate other performance criteria like signaling overhead, scalability, complexity, synchronisation and compatibility with mobility protocols. It must be noted that the context research achieved in this work aims to forward certain context to assist the handover management operation. However context awareness does have a wider meaning in next generation network and can be met in other fields like mobile middleware, sensor networks, ambient networks etc. Ideas investigated in this thesis could be considered in these fields as well.

New security considerations also arise if context transfer is to be employed in the next generation networks. New security requirements, the transfer of security state information and security issues of context transfer need to be further investigates for both homogeneous and heterogeneous environments. What security state information can be transferred e.g. state of authentication, state of authorization, cryptographic keys, key lengths need further investigation. This requires analysis of current authentication protocols including RADIUS, Diameter and COPS for identifying context transfer candidate parameters and examine the feasibility. Whether IPsec is sufficient for establishing a security association and a trust relationship between the different domain/entities that are involved during security context transfer also needs further research.

Finally the area of mobility management may become more complex as it tries to satisfy global scenarios across heterogeneous radio access technologies. This will have an impact on the way mobility management protocols, schemes, solutions and support protocols may be adopted. Next generation architectural proposals will also have an influence as to what protocols should be used or how they could be integrated.

Bibliography

[1] TS 23.002, *Network Architecture (Release 6) TS 23.002 v6.3.0*, Technical Specification Group Services and Systems Aspects, 3rd Generation Partnership Project, December 2003.

[2] TS 22.258, *Service Requirements for All-IP Networks (AIPN)*; Stage 1, Technical Specification Group Services and System Aspects, 3rd Generation Partnership Project, December 2005.

[3] D. Wisely, P. Eardly, L. Burness, *IP for 3G-Networking Technologies for Mobile Communications*, Wiley, 2002.

[4] www.3gpp.org, *3rd Generation Partnership Project* homepage, [*Last Accessed: December 2007*].

[5] www.3gpp2.org, *3rd Generation Partnership Project 2* homepage, [*Last Accessed: December 2007*].

[6] G. Patel, S. Dennet, The 3GPP and 3GPP2 Movements Towards an All-IP Mobile Network, IEEE Personal Communications, vol. 7, no. 4, pp. 62-64, August 2000.

[7] TR 23.882, 3GPP System Architecture Evolution; Report on Technical Options and Conclusions (Release 7), Technical Specification Group Services and System

Aspects, 3rd Generation Partnership Project, April 2007, [*Last Accessed: December 2007*].

[8] TR 23.882, 3GPP System Architecture Evolution; Report on Technical Options and Conclusions (Release 8), Technical Specification Group Services and System Aspects, 3rd Generation Partnership Project, December 2007, [*Last Accessed: December 2007*].

[9] C.E. Perkins, *IP Mobility Support for IPv4*, RFC 3344, IETF, August 2002.

[10] J. Rosenburg, H. Schulzrinne, G. Camarillo, A. Johuston, J. Peterson, R. Sparks, M. Handley, E. Schooler, *SIP: Session Initiation Protocol*, RFC3261, IETF, June 2002.

[11] H. Soliman, C. Castelluccia, K.El. Malki, L. Bellier, *Hierarchical Mobile IPv6 Mobility Management (HMIPv6)*, RFC4140, IETF, August 2005.

[12] Z.D. Shelby, D. Gatzounas, A. Campbell, C.Y. Wan, *Cellular IPv6*, IETF, July 2001.

[13] J. Kempf, Ed., Problem Description: Reasons for Performing Context Transfers Between Nodes in an IP Access Network, RFC 3374, IETF, September 2002.

[14] C.J. Mitchel, *Security for Mobility*, IEE Telecommunications Series 51, Wiley, ISBN: 0-86341-337-4, December 2003.

[15] M. Georgiades, C. Politis, N. Akhtar, R. Tafazolli, *Context Transfer Extension to Cellular-IP*, draft-georgiades-seamoby-ctecip-01.txt, IETF, December 2003.

[16] J. Loughney, M. Nakhjiri, C. Perkins, R. Koodli, *Context Transfer Protocol (CXTP)*, RFC4067, IETF, July 2005.

[17] C. Politis, K.A. Chew, N. Akhtar, M. Georgiades, R. Tafazolli, T. Dagiouklas, *Hybrid Multilayer Mobility Management with AAA Context Transfer Capabilities*

for All-IP Networks, IEEE Wireless Communications Magazine, vol .11, no. 4, pp. 76-88, August 2004.

[18] B. Aboba, D. Simon, *PPP EAP TLS Authentication Protocol*, RFC 2716, IETF, October 1999.

[19] C. Rigney, S. Willens, et.al, *Remote Authentication Dial In User Service (RADIUS)*, RFC 2865, IETF, June 2000.

[20] B. Aboba, L. Blunk, et.al, *Extensible Authentication Protocol*, RFC 3748, IETF, June 2004.

[21] T. Dierks, C. Allen, *The TLS Protocol (Version 1.0)*, RFC 2246, IETF, January 1999.

[22] C. de Laat, G. Gross, L. Gommans, J. Vollbrecht, D. Spence, *Generic AAA Architecture*, RFC 2903, IETF, August 2000.

[23] K. Ahmavaara, H. Haverinen, R. Pichna, *Interworking Architecture between 3GPP and WLAN Systems*, IEEE Communications Magazine, November 2003.

[24] IST-BRAIN, project website, http://www.ist.brain.org/ [*Last Accessed: December 2004*].

[25] IST-EVOLUTE, project website, intranet.evolute.gr [*Last Accessed: December 2004*].

[26] C. Perkins, Mobile IP Design Principles and Practices, Mobile IP Design Principles and Practices, Prentice Hall, November 1997.

[27] D. Johnson, C. Perkins, and J.Arkko, *Mobility Support in IPv6*, draft-ietf-mobileip-ipv6-24, IETF, June 2003.

[28] I.F. Akyildiz, J. Xie and S. Mohandy, *A Survey of Mobility Management in Next-Generation All-IP-Based Wireless Systems*, IEEE Wireless Communications, special issue on Mobility and Resource Management, vol. 11, no. 4, pp. 16-28, August 2004.

[29] E. Wedlund, H. Schulzrinne, *Mobility Support using SIP*, Proceedings of ACM/IEEE International Workshop on Wireless and Mobile Multimedia (WoWMoM), pp. 76-82, August 1999.

[30] H. Schulzrinne, E. Wedlund, *Application-layer mobility using SIP*, ACM SIGMOBILE Mobile Computing and Communications Review archive, vol. 4, no. 3, pp. 47-57, July 2000.

[31] N. Banerjee, S.K. Das, A. Acharya, *SIP-based Mobility Architecture for Next Generation Wireless Networks*, Proceedings of the 3rd IEEE International Conference on Pervasive Computing and Communications (PerCom 2005), 0-7695-2299-8/05, March 2005.

[32] R. Moskowitz, and P. Nikander, *Host Identity Protocol Architecture*, RFC 4423, August 2005.

[33] R. Moskowitz, *Host Identity Protocol*, draft-ietf-hip-base-05, March 2006.

[34] T. Henerson, *End-Host Mobility and Multihoming with the Host Identity Protocol*, draft-ietf-hip-mm-04, IETF, June 2006.

[35] P. Nikander, *Host Identity Protocol (HIP) Domain Name System (DNS) Extensions)*, draft-ietf-hip-dns-08, October 2006.

[36] J. Laganier, *Host Identity Protocol (HIP) Rendezvous Extension*, draft-ietf-hip-rvs-05, June 2006.

[37] J. Laganier, *Host Identity Protocol (HIP) Registration Extension*, draft-ietf-hip-registration-02, June 2006.

[38] E. Rescorla, *Diffie Hellman Key Agreement Method*, RFC 2631, IETF, June 1999.

[39] E. Nordmark, M. Bagnulo, *Level 3 multihoming shim protocol*, draft-ietf-shim6-proto-05.txt, IETF, May 2006.

[40] D. Thaler, *A Comparison of Mobility-Related Protocols*, draft-thaler-mobility-comparison-00.txt, IETF, December 2006.

[41] G. Tsirtsis, H. Soliman, V. Park:, *Dual Stack Mobile IPv4*, Internet Draft, draft-ietf-mipv4-dsmipv4-00.txt, July 2006.

[42] A.T. Campbell, J. Gomez, S. Kim, A.G. Valko, C.Y. Wan, Z.R. Turanyi, *Design, Implementation and Evaluation of Cellular IP*, IEEE Personal Commuications, vol. 8, no. 4, pp. 42-49, August 2000.

[43] R. Ramjee, T.L. Porta, S. Thuel, K. Varadhan, S.Y. Wang, *HAWAII: A Domain-Based Approach for Supporting Mobility in Wide-Area Wireless Networks*, IEEE/ACM Transaction Networks, vol. 10, no. 3, pp. 396-410, June 2002.

[44] A. Misra, S. Das, S.K. Das, *IDMP-Based Fast Handoffs and Paging in IP-Based 4G Mobile Networks*, IEEE Communications Magazine, vol. 40, no. 3, pp. 138-145, March 2002.

[45] S. Das, A.Misra, P.Agrawal, S.K.Das, *TeleMIP: Telecommunication Enhanced Mobile IP Architecture for Fast Intra-Domain Mobility*, IEEE Personal Communications Systems Magazine, vol. 7, no. 4, pp. 50-58, August 2000.

[46] J. Kempf, Goals for Network-based Localized Mobility Management (NETLMM), draft-ietf-netlmm-nohost-req-05.txt, IETF, October 2006.

[47] J. Kemph, Problem Statement for Network-based Localized Mobility Management, draft-ietf-netlmm-nohost-ps-05.txt, IETF, September 2006.

[48] C. Vogt, *Security Threats to Network-Based Localized Mobility Management*, draft-ietf-netlmm-threats-04.txt, IETF, September 2006.

[49] A.T. Campbell et.al, *Comparison of IP Micromobility Protocols*, IEEE Wireless Communications Magazine, vol. 9, no. 1, pp. 72-82, February 2002.

[50] R. Koodli, *Fast Handovers for Mobile IPv6*, RFC 4068, IETF, July 2005.

[51] P. McCann, *Mobile IPv6 Fast Handovers for 802.11 Networks*, RFC 4260, IETF, November 2005.

[52] E.K. Paik, Y. Choi, *Prediction-based fast handoff for mobile WLANs*, 10th International Conference on Telecommunications(ICT), ISBN 0-7803-7661-7, vol. 1, pp. 748-753, February 2003.

[53] P. McCann, Mobile IPv6 Fast Handovers for 802.11 Networks, RFC 4260, IETF, November 2005.

[54] C. Tan, S. Pink and K. Lye, *A fast handoff sheme for Wireless Networks*, Proc. of Second ACM International Workshop on Wireless Mobile Multimedia (WOWMOM), ACM SIGMOBILE, pp. 83-90, August 1999.

[55] K.El. Malki, *Low Latency Handoffs in Mobile IPv4*, draft-ietf-mobileip-lowlatency-handoffs-v4-0.3.txt, November 2001.

[56] K.El. Malki and H.Soliman, *Hierarchical Mobile IPv4/v6 and fast handoffs*, draft-elmaki-soliman-hmipv4v6-00.txt, March 2000.

[57] F.M. Chiussi, D. A. Khotimsky, and S. Krishnan, *Mobility Management in Third-Generation All-IP Networks*, IEEE Communications Magazine, vol. 40, no. 9, pp. 124-135, September 2002.

[58] M. Nakhjiri, A. Singh, *Time Efficient conteXt Transfer (TEXT)*, draft-nakhjiri-seamoby-text-ct-01.txt, IETF, September 2002.

[59] M. Liebsch, A. Singh, H. Chaskar, D. Funato, E. Shim, *Candidate Access Router Discovery (CARD)*, RFC 4066, IETF, July 2005.

[60] http://www.opnet.com, *OPNET Modeler Wireless Suite*, OPNET Technologies, [*Last Accessed: December 2007*].

[61] A. Stephanie, A. Mihailovic, A.H.Aghvami, *Mechanims And Topology for Fast Handover in Wireless IP Networks*, IEEE Communications Magazine, vol. 38, no. 11, pp. 112-115, November 2000.

[62] J. Kemph, *Instructions for Seamoby and Experimental Mobility Protocol IANA Allocations*, RFC 4065, July 2005.

[63] T. Haitao, J. Eisl, M. Georgiades, *Mobility Support: Design and Specification*, FP6-CALL4-027662-AN P2/D9, December 2006.

[64] I.F. Akyildiz, J. Xie, S. Mohanty, *A survey of Mobility Management in Next Generation All-IP-based Wireless Communications*, vol. 11, no. 4, pp. 16-28, August 2004.

[65] S. Leggio, J. Manner, K. Raatikainen, *Achieving Seamless Mobility in IP-based Radio Access Networks*, IEEE Wireless Communications Magazine, vol. 12, no. 1, pp. 54-59, February 2005.

[66] E. Fogelstroem, A. Jonsson, C. Perkins, *Mobile IPv4 Regional Registration*, draft-ietf-mipv4-reg-tunnel-04, IETF, October 2006.

[67] A.O. Neill, G. Tsirtsis, and S. Corson, *Edge Mobility Architecture*, draft-oneil-ema-02.txt, IETF, March 2000.

[68] A.O. Neill, S. Corson, *An Approach to Fixed/Mobile Converged Routing*, Technical Report TR-2000-5, University of Maryland, Institude for Systems Research, March 2000.

[69] M. Nakhjiri, A. Singh, *Time Efficient conteXt Transfer (TEXT)*, draft-nakhjiri-seamoby-text-ct-01.txt, IETF, September 2002.

[70] V. Park, M.S. Corson, *Temporally-Ordered Routing Algorithm (TORA), Version 1 Functional Specification*, IETF, November 2000.

[71] V. Park, M.S. Corson, *A highly adaptive distributed algorithm for mobile wireless networks*, in proceedings of INFOCOM'97, Kobe, Japan, April 1997.

[72] K.A. Chew, *Mobility Management in All-IP Mobile Networks*, PhD Thesis, Communication Centre for Systems Research, University of Surrey, January 2004.

[73] Transport Layer Security (tls) WG, IETF, http://www.ietf.org/html.charters/msec-charter.html, [*Last Accessed: December 2007*].

[74] IP Security Protocol (ipsec) WG, IETF, http://www.ietf.org/html.charters/ipsec-charter.html, [*Last Accessed: December 2007*].

[75] IEEE Std. 802.11f/D2.0, 'Draft Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation, July 2003.

[76] 3GPP TR 25.936 v 4.0.0, *Handovers for Real Time Services from PS Domain*, 3GPP, March 2001.

[77] Cellular IP source code, http://www.comet.columbia.edu/cellularip/, [*Last Accessed: December 2007*].

[78] HostAP source code, http://hostap.epitest.fi, [*Last Accessed: December 2007*].

[79] Linphone, http://www.linphone.org, [*Last Accessed: December 2007*].

[80] FreeRADIUS, http://www.freeradius.org,[*Last Accessed: December 2007*].

[81] Dynamics Mobile IP, http://dynamics.sourceforge.net, [*Last Accessed: December 2007*].

[82] Ethereal, http://www.ethereal.com/, [*Last Accessed: December 2007*].

[83] E. Hyytia, J. Virtamo, *Random waypoint mobility model in Cellular Networks*, IEEE Wireless Networks, vol.13, no.2, p177-88, April 2007.

[84] A. Kamerman, G. Aben, *Net Throughput With IEEE 802.11 Wireless LANs*, Proceedings of the IEEE Wireless Communications and Networking Conference, September 2000.

[85] K. Kong, M. Song, K. Park, C. Hwang, *A comparative analysis on the signaling load of Mobile IPv6 and Hierarchical Mobile IPv6: Analytical Approach*, IEICE Transactions on Information and Systems, vol. 89, no. 1, pp. 139-149, January 2006.

[86] W. Lai, J. Chiu, *Improving handoff performance in wireless overlay networks by switching between two-layer IPv6 and one-layer IPv6 addressing*, IEEE Journal on Selected Areas in Communications, vol. 23, no. 11, pp. 2129-2137, November 2005.

[87] S. Lo, G. Lee, W. Chen, J. Liu, *Architecture for Mobility and QoS support in all-IP Wireless Networks*, IEEE Journal on Selected Areas in Communications, vol. 22, no. 4, pp. 691-705, May 2004.

[88] B. Carpenter, *Middleboxes: Taxonomy and Issues*, RFC3234, IETF, February 2002.

[89] M. Frantzen, F. Kerschbaum, E. Schultz, and S. Fahmy, *A framework for understanding vulnerabilities in firewalls using a dataflow model of firewall internals*,

Computers and Security, Science Direct, Elsevier, vol. 20, no. 3, pp. 263-270, May 2001.

[90] E. Schultz, *When firewalls fail: lessons learned from firewall testing*, Network Security, Elsevier, vol. 1997, no. 2, pp. 8-12(5), February 1997.

[91] N. Niebert, A. Schieder, H. Abramowicz, G. Malmgren, J. Sachs, U. Horn, C. Prehofer, and H. Karl, *Ambient Networks: An Architecture for Communication Networks Beyond 3G*, IEEE Wireless Communications, vol. 11, no. 2, pp. 14-22, April 2004.

[92] N. Niebert, A. Schieder, J. Zander, R. Hancock, *Ambient Networks: Co-operative Mobile Networking for the Wireless World*, Wiley, ISBN: 978-0-470-51092-6, April 2007.

[93] A. Surtees, R. Aguero, J. Eisl, M. Georgiades, *Mobility Management in Ambient Networks*, Vehicular Technology Conference, Dublin, Ireland, April 2007.

[94] A. Fei, G. Pei, R. Liu, and L. Zhang, *Measurements on delay and hop-count of the Internet*, IEEE GLOBECOM, November 1998.

[95] E. Perera, B. Roksana, S. Herborn, M. Georgiades, J. Eisl, E. Hepworth, *A Mobility Toolbox Architecture for All-IP Networks: An Ambient Networks Approach*, IEEE Wireless Communications, April 2008.

Publications

Journal Publications

- M. Georgiades, T. Dagiuklas, K. Moessner, "On the enhancement of Mobility and Multimedia Communications in Heterogeneous RANs", Int. J. Internet Protocol, to be published in 2008.

- E. Perera, R. Boreli, S. Herborn, M. Georgiades, J. Eisl, E. Hepworth, "A mobility toolbox architecture for All-IP networks: An Ambient Networks Approach", IEEE Wireless Communications Magazine, vol. 15, no. 2, pp. 8-16, April 2008.

- M. Georgiades, K. A. Chew, R. Tafazolli, "Advances in Micro-Mobility Management using a Mobility-Aware Routing Protocol", Research Letters in Communications Journal, Hindawi Publications, vol. 2007, no. 2, December 2007.

- M. Georgiades, N. Akhtar, C. Politis, R. Tafazolli, "Enhancing mobility management protocols to minimize AAA impact on handoff performance", Elsevier International Journal for the Computer and Telecommunications Industry, Computer Communications, vol. 30, no. 3, pp. 608-618, February 2007.

- C. Politis, T. Diagiuklas, N. Akhtar, K. Chew, M. Georgiades, R. Tafazolli, "Mobility Management and AAA Support Characterisation for Providing Seamless

Multimedia Services over IP-based Infrastructures", IEEE Wireless Communications Magazine, vol. 11, no. 4, pp. 76-88, August 2004.

Conference Papers

- N. Akhtar, M. Georgiades, C. Politis, R. Tafazolli, *SIP-based End System Mobility Solution for All-IP Infrastructures* , IST Mobile & Wireless Communications Summit 2003, Aveiro, Portugal, June 2003.
- M. Georgiades, K. Chew, C. Politis, R. Tafazolli, *Context Transfer Extension to Mobility Protocols for All-IP Based Infrastructures*, Wireless World Research Forum (WWRF), 9th meeting, Switzerland, July 2003.
- N. Akhtar, M. Georgiades, C. Politis, R. Tafazolli, *Real-time Evaluation of Mobility Management Schemes for IP-based WLAN Infrastructures*, International Evolute Workshop, Surrey, UK, November 2003.
- M. Georgiades, C. Politis, N. Akhtar, R. Tafazolli, *Context Transfer Extension to Cellular-IP*, draft-georgiades-seamoby-ctecip-01.txt, December 2003.
- M. Georgiades, N. Akhtar, C. Politis, R. Tafazolli, *AAA Context Transfer for seamless and secure multimedia services over All-IP networks*, 5th European Wireless Conference (EW'04), February 24-27, 2004, Barcelona, Spain.
- M. Georgiades, N. Akhtar, R. Tafazolli, *Security Context Transfer to complement IP-based Mobility Management schemes*, Networking 2004, Athens, Greece, May 2004.
- M. Georgiades, N. Akhtar, C. Politis, R. Tafazolli, *Impact of AAA Context Transfer on TCP Performance in all-IP Networks*, IST Mobile & Wireless Communications Summit 2004, Lyon, France, June 2004.

- M. Georgiades, H. Wang, R. Tafazolli, *Security of Context Transfer for Future Mobile Communications*, submitted to Wireless World Research Forum (WWRF), 12th meeting, Toronto, Canada, November 2004.

- M. Georgiades, N. Akhtar, M. Ghader, Z. Li, S. Gultchev, R. Tafazolli, *Surrey's Next Generation Wireless Network Testbed*, IEEE Mobile and Wireless Communication Networks, MWCN 2005, Marrakech, Morocco, September 2005.

- M. Wang, M. Georgiades, R. Tafazolli, *Security provisioning in an integrated WLAN/WPAN infrastructure and its impact on the handoff performance*, Wireless World Research Forum (WWRF), 15th meeting, Paris, France, December 2005.

- M. Georgiades, T. Dagiuklas, R. Tafazolli, *Middlebox Context Transfer support for multimedia sessions in All-IP network*, International Wireless Communications and Mobile Computing Conference (IWCMC 2006) Sheraton Wall Centre, Vancouver, Canada, July 2006.

- M. Georgiades, T. Rinta-aho, F. Meago, *Handover Management Process in Ambient Networks*, Wireless World Research Forum (WWRF), 17th meeting, Heidelberg, Germany, November 2006.

- R. Aguero, A. Surtees, J. Eisl, M. Georgiades, *Mobility Management in Ambient Networks*, IEEE 65th Vehicular Technology Conference VTC2007, Dublin, Ireland, April 2007.

- J. Eisl, M. Georgiades, T. Jokikyyny, R. Boreli, E. Perera, K. Pentikousis, *Management of multiple mobility protocols and tools in dynamically configurable networks*, First Ambient Networks Workshop on Mobility, Multiaccess and Network Management (M2NM 2007), October 2007.

Book Contributions

- A. Zugenmaier, M. Georgiades, P. Schoo, et.al. *Chapter 5: Security in Ambient Networks*, Ambient Networks: Co-operative Mobile Networking for the Wireless World, Wiley, edited by N. Niebert, A. Schieder, J. Zander, R. Hancock, ISBN: 978-0-470-51092-6, April 2007.

- J. Eisl, J. Makela, R. Aguero, S. Uno, M. Georgiades, et.al. *Chapter 9: Ambient Networks Mobility Management*, Ambient Networks: Co-operative Mobile Networking for the Wireless World, Wiley, edited by N. Niebert, A. Schieder, J. Zander, R. Hancock, ISBN: 978-0-470-51092-6, April 2007.

Patent

- M. Georgiades, C. Politis, N. Akhtar, R. Tafazolli, *Context Transfer Module in Wireless Communication Networks*, UK Patent Application No. GB 04 09085.8, April 2004.

Award

- M. Georgiades, *Context Transfer support for IP-based mobility management*, was awarded the Nokia Prize at the CCSR Awards for Research Excellence 2004, Surrey, UK, April 2004.