Distributed Home Agent Mobility Management for IP Based Cellular Network

Chi Wah Yung

A dissertation submitted for the degree of

MASTER OF PHILOSOPHY (MPhil)

to

School of Electrical and Electronic Engineering
Faculty of Engineering, Computer and Mathematical Sciences

THE UNIVERSITY OF ADELAIDE, South Australia

March, 2013

MAJOR SUBJECT: WIRELESS COMMUNICATIONS & NETWORKING

Supervisors:

Prof Reginald Coutts, School of Electrical & Electronic Engineering

Prof Derek Abbott, School of Electrical & Electronic Engineering

Table of contents

List of figures

List of tables

List of acronyms

2.5G	Intermediate Generation of Cellular Network between 2G and 3G
2G	Second Generation Cellular Network
3G	Third Generation Cellular Network
4G	Forth Generation Cellular Network
AAA	Authentication, Authorization and Administration
AN	Access Network
BSC	Base Station Controller
CDMA	Code Division Multiple Access
CN	Correspondent Node
CoA	Care Of Address
CSD	Circuit Switched Data
DHAA	Dynamic Home Agent Anchoring Scheme
D-MIP	Distributed Agent Mobility Management Platform
FA	Foreign Agent
FDMA	Frequency Division Multiple Access
HA	Home Agent
HSPA	High Speed Packet Access
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPv6	Internet Protocol Version 6
ISO	International Organization for Standardization
IWF	Inter-working Function
MIP	Mobile IP Protocol
MIPv6	Mobile IP Protocol for Internet Protocol Version 6
MN	Mobile Node
MS	Mobile Station
NS2	Network Simulator 2
OPNET®	OPNET® is either registered trademarks or trademarks of OPNET Technologies, Inc. in the United States and/or other countries.
OSI	Open System Interconnection
PDA	Personal Digital Assistant
PDSN	Packet Data Serving Node
RAN	Radio Access Network
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UMTS	Universal Mobile Telephone System

Abstract

The convergence of wireless networks both fixed and mobile with the Internet is creating a revolution in the way wireless networked resources interact with each other. This thesis is concerned with mobile networks and proposes to deal with the mobility management problems for the mobile computing devices in the next generation of multi-technologies integrated IP based mobile networks. In order to do this, a new distributed home agent approach for mobility management has been developed that harmonizes the concept of micro-mobility and macro-mobility management in order to enable seamless mobility management on different kinds of wireless network environment especially interaction with the legacy cellular network in which resources are limited and expensive. The major contribution of this thesis is three-fold.

Firstly, this thesis proposes network access architecture and a distributed mobility management scheme, which enables the mobility of a mobile device in a cellular packet data network in order to reduce the latency and network traffic required to handle the mobility management functionality. A detailed design of the distributed mobility management scheme is presented for the implementation and the conceptual model is analysed.

Secondly, simulation of the mobility management schemes using two different network simulation packages to enable a comparison of the simulator functionalities is presented.

Finally, the results of the simulation and suggested future work are presented.

Declaration of originality

This work contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution to C. W. Yung and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of the thesis, when deposited in the University Library, being available for loan, photocopying, and dissemination through the library digital thesis collection, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University‘s digital research repository, the Library catalogue, the Australasian Digital Thesis Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: _____ (C. W. Yung) **Date:** _____

Publications

The following are some of the publications of the candidate which are related to the theme of this thesis.

1. C. W. Yung, R. P. Coutts, and D. Abbott, ―Design of distributed agent mobility management platform (D-MIP) for IP-based wireless networks," *Proceedings of 4G Mobile Forum*, San Diego, USA, July 2005.
2. C. W. Yung, R. P. Coutts, and D. Abbott, ―Modeling and simulation of dynamic home agent anchoring scheme for mobility management of IP based wireless networks*,"* *Proceedings of 2006 Global Mobile Congress*, Beijing, China, Oct 2006.

Notable seminars and workshops

3. University of Adelaide Seminar, ―Design of distributed agent mobility management platform (D-MIP) for IP-based wireless networks," University of Adelaide, 2005.

Acknowledgments

It is my great pleasure to express my gratitude to all the people who have supported me greatly during the pursuit of my study. Without their encouragement and help, it would not have been possible for me to complete this dissertation.

Firstly, I would like to give my best appreciation to my two supervisors, Professor Reginald P. Coutts and Professor Derek Abbott, for their support, inspiration and encouragement to me throughout the course of the study. They have helped to provide me with required motivation and confidence in myself, and the courage to take this challenging task and finish this thesis so smoothly and quickly.

Secondly, I would like to thank all the staff and postgraduate associates in the School of Electrical and Electronic Engineering at the University of Adelaide. Their friendship, help and encouragement deserve my sincere appreciation. Also, their timely support, discussion and sharing made my study journey a lot easier and constantly inspired me on new knowledge and research topics.

Lastly but not the least, I want to give my special thankfulness to my family and friends. Their strong support and encouragement made this study possible.

Chapter 1: Introduction

1.1 Background and problems of mobility management

Over the past two decades, the technology of cellular networks has evolved over four generations and currently a fourth generation (4G) [1] is already deployed commercially and is the expected platform for further mobile technology development. First generation (1G) Frequency Division Multiple Access (FDMA) communication technology in the 1980‘s only supports voice mobile calls. Second generation (2G), 2.5th generation (2.5G) and commercial deployed 3rd generation (3G) and 3.5G mobile technologies are migrated from a circuit based backhaul to an IP packet based backhaul network. The packet data transfer speed from 2G network to 3G or 3.5G networks has greatly improved. For example, a 3G network of CDMA2000 EVDO[1] provides access to mobile devices with forward link air interface speeds of up to 2.4 Mbit/s with Rev. 0 and up to 3.1 Mbit/s with Rev. A. A 3G to 3.5G network of Universal Mobile Telephone System (UMTS) supports up to 21 Mbit/s data transfer rates in theory with High Speed Downlink Packet Access (HSDPA). A more advanced air interface and IP Core has been developed which is the basis of 4th generation (4G) mobile technology standard development termed Long Term Evolution (LTE). Note that, LTE is now being

[1] The North American 3G technology (3GPP2)

deployed right across the world and its ongoing development (e.g. to LTE Advanced) is being undertaken in 3GPP (the Third Generation Partnership Project). For example the 3GPP Long Term Evolution (LTE) has a downlink peak rate of at least 100 Mbit/s. Over the past decade, the data transfer rate between Mobile Station (MS) and Base Station (BS) has improved from 9.6kbps in 2G network to 21 Mbit/s in 3G or 3.5G network. The first mobile standard to support[2] mobile data transfer was the 2G GSM technology. It implemented Circuit Switched Data (CSD) that enabled the Mobile Station (MS) to transmit data over a voice channel up to 9.6kbit/s. A 2.5G, a packet switching data solution called GPRS was introduced in the mid 1990s, which brought in a milestone in supporting wireless mobile computing in using IP packets switching for mobile data call. During this period the demand for greater data transport started to overtake circuit transport. Thus due to the better channel utilization and lower operating costs of packet switching, most of the Radio Access Network (RAN) is migrating from circuit switching to packet switching in order to provide a seamless integration for packet data solution. In addition, 2.5G and 3G networks were able to accommodate more concurrent mobile subscribers and higher network capacity per base station. With the commercial deployment of the 3G and 3.5G networks in most countries around the world, high-speed packet data communication became the major trend of mobile communications that continues to grow unabated. In addition, a high-speed cellular packet data solution became an important medium to enable the initial convergence of wired Internet and wireless network. The data transfer rates will continue to increase to 100s of MBit/s with the introduction of multi-antenna technologies such as MIMO in 4G networks which are now being incorporated in LTE[3]. With the high speed of data transfer, mobile broadband has

[2] Other than within a voice circuit as on the old fixed PSTN network.
[3] MIMO has been incorporated in WiMAX earlier but WiMAX will not be discussed to any degree in this thesis.

become a reality. The ‗Internet Generation' has become accustomed to having broadband access wherever they go and not just at home or in the office.

As 3G, 3.5G and 4G cellular networks are commercially deployed around the world, more and more services that require high-speed packet data are being further developed and deployed. Some of the most popular services include video communications, social networking, Internet browsing, and multimedia information services enabling new forms of news delivery. As different devices are developed to utilize the 2.5G, 3G and 4G networks for more sophisticated services, the need for data connection of portable devices to the cellular network is expected to increase as well. Portable devices such as Laptops, PDAs, Tablets and other intelligent devices use the internet protocol as the transport protocol for communication with other wired or wireless devices or services. In this emerging context, IP Mobility is one of the key schemes for maintaining connectivity between networks between portable devices to servers or other devices. Also, frequent handoff including between distinct networks is expected in the next generation of cellular networks because of the decreased radio coverage area per base-station in order to provide a high speed radio link [2] and multiple networks.

However, as mobile station (MS) accesses data using any packet switching method, mobility management becomes an increasing problem. This thesis examines micro-mobility effects on using a new distributed home agent anchoring approach, analysing the performance of the new approach with different types of traffic patterns and compares this with using a Mobile IP protocol.

The problem of mobility management was initially discovered when cellular network infrastructure migrated from the 2G circuit-based network to 2.5G/3G packet data networks. Inside a 2G circuit-based network, data is sent over the air by setting up a full voice channel and data from a mobile station transmits over the voice channel in a modulated format. The voice channel is connected at all times during the mobile call, which leads to poor utilization of the radio channel and subscribers are also charged for the connection at all times even where there is not any data being sent over the channel. When the mobile station hands off to another base station, a new radio channel is established but the mobile station still maintains a path to the current voice channel through switching to the Inter-Working Function (IWF) until the MS moves out of the serving zone of the IWF. Therefore, from the mobile station perspective, the point of attachment to the IP network is always fixed. Mobility is only an issue if the mobile station moves out of the IWF serving zone and the call needs to be re-established. This operation is called hard handoff. In a 3G network, most or a majority portion of the Access Network (AN) is packet switched instead of circuit switched. Packet switching in the radio access network provides advantages on channel utilization and additional subscriber features such as _always on' and packet based charging. However, this approach introduces some issues in mobility management. The following paragraphs summarise the mobility management problems in the current IP based packet switching cellular wireless network.

1) Inside a cellular network, continuous radio coverage is provided through dividing the geographical areas into many overlapped cells. Each cell has at least one base station to transmit and receive the radio signal from the mobile station (MS). When a MS is moving from one serving cell to another serving cell, the point of attachment of the MS to the radio access network is changed. For a packet data Radio Access Network (RAN), if the MS

moves to another cell during the data call, the assigned IP address will become invalid in the new point of attachment. This causes permanent connection loss because the MS is no longer inside the original network that the IP address used. Therefore, the MS either needs to re-negotiate a new IP address during handoff or implement a mobility management protocol such as Mobile IP [3] protocol.

2) When the MS moves to another serving cell, the location of the MS needs to be updated to Base Station Controller (BSC), which allows the BSC to keep track of the location of the MS for incoming call paging. In most cellular networks, location updates are managed by the network infrastructure such as by using a registration message sent from each base station to the MS. The MS location is already kept through network layer messaging mechanism of the radio access network that is called link layer messages. However, it is not available for usage in higher layer applications.

There are various methods that have been proposed in the past few years to solve the mobility issues such as the Mobile IP Protocol [3] and other micro-mobility solution including uMIP [4], Cellular IP [5][6][7], Hawaii [8][9] and Hierarchical MIP [10] with paging capability [11]. Micro-mobility management protocols including using distributed approaches are also developed [12][13]. A detail survey of micro mobility management protocol can also be found in [14][15]. Mobile IP has been in use for over a decade and able to solve the mobility problems by introducing Home and Foreign agents, which are mainly responsible to redirect the packet from the MS‘s home network to the foreign network care of address. Mobile IP was well adopted in the sector community after being standardized by the IETF and became a popular solution to manage mobility in IP based network. However, there are two major disadvantage of the Mobile IP protocol when implemented inside a radio access network.

1) To implement the Mobile IP protocol, each IP network attached to every base station requires a constant broadcast –Agent Advertisement" message. When the MS moves from one network to another, the MS does not immediately know that it has been handed off to another network until receiving the Agent Advertisement message only to find out the network is a foreign network. The MS also can request an Agent Advertisement message in an ad hoc manner by sending an Agent Solicitor broadcast or multicast message. Mobile IP relies on the IP protocol for determination of the home or foreign network and to enable the packet forwarding mechanism. However, this causes considerable amount of delay and signaling overheads during handoff for a cellular type network. In addition, the size of a cell and handoff continues to decrease in next generation networks in order to provide higher data transfer rate, especially inside an urban area, and therefore packet delay and signaling overheads will also increase.
2) Mobile IP uses a tunneling method to forward the IP packets from the home network to the care of an address in the foreign network. The extra routing causes delay of the packet delivery and in fact if the MS is moved to a large distance (e.g. a different country), considerable packet delay can be expected, which will degrade the data transfer rate constantly and could cause constant packet loss in any real time application. These problems will affect most services in a 3G or 4G network such as Video and Voice over IP.

In addition, most of the current micro-mobility management protocols are mainly addressing mobility in a single domain than using a RAN technology and do not take into account the impact on Quality of Service (QoS) factors. Note that, QoS [16] is a key element on next generation networks in order to enable seamless operation between different access

technologies such as Wireless LAN (WLAN), ad hoc wireless networks, future cellular network including mobile WiMAX [17][18].

On the other hand, the trend for future cellular networks is for of an all-IP based architecture providing ubiquitous coverage through a variety of radio access network technologies such as WiMAX, 3G, 3.5G and 4G technologies. These networks will provide high-speed, packet-mode operation in which different radio technologies are implemented for different geographical needs. For example, WiMAX technology will be more suitable for wide area with a good line of sight and without large number of buildings. The current 3G and 3.5G technology provides coverage in an urban area that may have many buildings and other obstructions. Within smaller areas such as the café or home, Wi-Fi provides a cost effective way for the connection. Mobile devices will be capable of roaming, make and break connection with all those different access technologies and be capable of dynamically configuring their transceivers accordingly. This is the vision that provides the motivation for this thesis. Indeed this is the motivation behind a wide area of research being conducted into future cellular networks and supporting protocols. The paradigm shift from circuit to packet mode is well advanced in today‘s networks. The efficiencies, both in terms of packet-mode operation and operational costs of IP networks will necessitate the change. Although the direction of future cellular networks may be evident, the pathway to reach such a state is not. The IETF is playing a major role in developing standards to support mobility in IP-based networks. The most advanced to date comes in the form of the proposed standard, Mobile IP, which provides an IP based solution to host mobility. Several vendors and research groups have implemented Mobile IP. Its support continues to grow. However, in realizing the vision, several more advancements need to be made. Some of these include:

- Inter-working of IP core networks and the various RAN technologies
- Quality of Service on voice, video and data transmission.

- Seamless handoff between different radio technologies.
- Robust, IP based protocols capable of providing mobility management functionality equivalent to that of the circuit-switched protocols such as ANSI-41, GSM-MAP operating over Signaling System Number 7.

Using a distributed approach to manage the problem of IP mobility management is a fairly new area and developing in the past few years in order to address the issue of mobility management. Recently, a few different approaches have been proposed which employed distributed approach for mobility management are listed from [19] to [24]. Also, there are various proposals on addressing mobility management using link-layer fast handover schemes in IEEE 802.11-based wireless network [25][26]. In addition, the IEEE standard 802.21-2008 [27] provides a framework for handovers between heterogeneous wireless networks such that the same framework is applicable to different network types which differ at the link-layer and below. For the IP version 6 space, there are various proposal on handling mobility management in the IP version 6 network such as the hierarchical Mobile IPv6 [28], with fast handover mechanism for Mobile IPv6 [29] and dual stack host and router support [30]. However, this thesis is mainly focused on IP version 4 and IP version 6 and its related mobility management techniques and protocols are not included in the scope of this thesis. Some others approaches were proposed before which used different network layers as an approach to address the IP mobility management function such as the TIMIP [31], end to end support [32], SCTP [33], transport layer [34], and SIP protocol [35][36][37]. Some other approaches are using cross layer methods to solve the mobility issue such as the WiSwitch [38] and the Pre-application Mobility Management Platform [39][40]. Due to the wide spread landscape and approaches of IP mobility management research, the following objectives and scope is fixed in order to develop a manageable research scope based on the resource, timeframe and tools available.

1.3 Research objectives and scope

The objectives of this thesis are listed below:

- Provide a survey of IP based networking protocols and IP based mobility management scheme and highlights how mobility management affects the cellular wireless network developments. This research has not considered SIP, IP version 6 or IPv6 related protocols development such as Mobile IPv6 protocol in order to reasonably confine the scope for this research.
- Design the requirement, scheme and protocol for an enhanced IP based mobility management which can be used for different wireless network technologies.
- Examine the new scheme and protocol using network simulation techniques by simulation using two different network simulators.
- Draw conclusions and discuss the merits and shortcomings of the proposal.

1.4 Roadmap overview of thesis

This section provides a brief overview of the structure of the thesis. It is intended to help readers to quickly search for the topic and contents of interest to enable an understanding of context for the new protocol developed for this thesis.

Chapter 2 provides an introduction to IP based data networks. The Internet Protocol is described in detail along with other protocols that are integral for the functioning of IP based mobility management.

Chapter 3 provides a detailed explanation of the Mobile IP and recent enhancement of Mobile IP.

Chapter 4 introduces the current cellular and wireless networks development. This chapter presents an introduction to 3G networks as well as the latest WiMAX and LTE Wireless networks. The mobility issues are described for these technologies.

Chapter 5 presents the requirements, design and simulation of a Distributed Home Agent Mobility Management Platform (D-MIP). A new distributed home agent platform is designed and simulation results using ns2 simulator is presented.

Chapter 6 presents a Dynamic Home Agent Anchoring Scheme (DHAA) which is based on the design of D-MIP from Chapter 5 and was extended with detail analysis and diagrams. Then, simulation of the DHAA scheme is carried out using OPNET® simulators. Simulation results is presented and discussed. As DHAA is an extension of the D-MIP in Chapter 5, the structure of Chapter 6 is similar to Chapter 5.

Chapter 7 provides a conclusion and discussion of future work relating to the work in this thesis.

1.5 Statement of original contributions

This thesis designed and simulated a new distributed mobility management protocol which aims to reduce communication overheads compared with the traditional Mobile IP protocol. A network simulation is carried out using two distinct simulation software packages in order to analyse the benefit and details of the newly designed protocol.

Chapter 2: Internetworking and internet protocols

This chapter provides a summary of the Internet Protocol (IP) based networking concept. A general introduction of the development of networking technologies is followed by a detailed description of the Internet Protocol (IP) that forms the foundation of most modern wired and wireless data networks.

2.1 Fundamentals of internetworking

The traditional telephone network is developed on the basis of a circuit-switched and connection oriented model. A dedicated connection is reserved and temporarily creates a connection between both the caller and the called parties for the duration of a call. Such systems are inefficient in supporting transmissions that consists of short burst of data and then followed by long idle duration when there is no data transmitted. Packet switching networks offer better efficiencies that enable greater network scalability than circuit-switched networks for such bursty data traffic. Packet switching replaces the centralized switches with distributed routers and each with multiple connections to adjacent routers. Data is divided into a series of small packets that are independently routed on a hop-by-hop basis from the

source point to its destination point. This approach allows messages to be multiplexed on the available paths on a statistically determined basis, gracefully adapting the transmission to traffic levels and optimizing the use of existing link capacity without pre-allocating link bandwidth [45]. Using a packet switching method, computers or mobile devices are always connected, and continually online. Also, it does not need a process to setup the connection just like the dial up process in the old telephone network. Hence, it does not require any dial up time for the connection. Packet switched networks have their origins in the early 1960‘s and the Internet in its early realization began around 1969 as U.S Department of Defense (DoD) funded experiment to interconnect research sites in the U.S. However, the full commercial realization of the internet did not arrive till the 1990s. The first generation of protocols used was superseded around 1981 by a protocol suite that is now synonymous with the Internet, namely TCP/IP (Transmission Control Protocol/Internet Protocol). Note that, TCP/IP has since become the most widely used open systems interconnection protocol suite ever developed. The Internet is the worldwide collection of interconnected computer networks that use IP as the means of interconnection. An IP‘s ability to allow computer networks using a variety of transmission technologies, comprising computers running a variety of operating systems to communicate with each other has been the factor behind the explosive growth of the global Internet. Knowledge of the principles of internetworking is an important pre-requisite for understanding the protocols described and analyzed in this thesis. This section aims to provide a general overview of internetworking concepts. In the following sections, an introduction of the concept of internetworking and protocols are presented.

2.2 The layered approach

To understand the concept of internetworking, the layered approach is the most important concept upon which most of the internetworking protocols and systems are built. The network communication space is sub-divided into different layers and each layer is responsible to handle a specific functionality of in order to provide a reliable and error free communication protocols to both ends of the computers. The internetworking protocols discussed in this document conform to a layered structure. The central protocol the Internet Protocol called IP is considered to be part of a five-layered system that follows the OSI[4] model. A detailed description of the OSI model is in the next section. IP is a network layer protocol which is in the layer 3 in the OSI model. This layer is responsible for transporting packets across a network and deciding which routes packets should take. All packet transfers are based on an addressing scheme called the IP address.

2.3 OSI reference model

The International Organization for Standardization (ISO) developed the OSI reference model to facilitate the open interconnection of computer systems [41]. The reference model identifies logically ordered layers for all the functions required for a communications session between two computers. The model defines mechanisms for passing data between two machines that share the same network such as LAN or WAN. The OSI model defined a network interconnection into seven different layers and each of the layer performs specific functionalities. The layers are organized based on the natural sequence of events that occur during a communications session as illustrated in Figure 2-1. They are described as below:

[4] OSI stands for Open Systems Interconnect model.

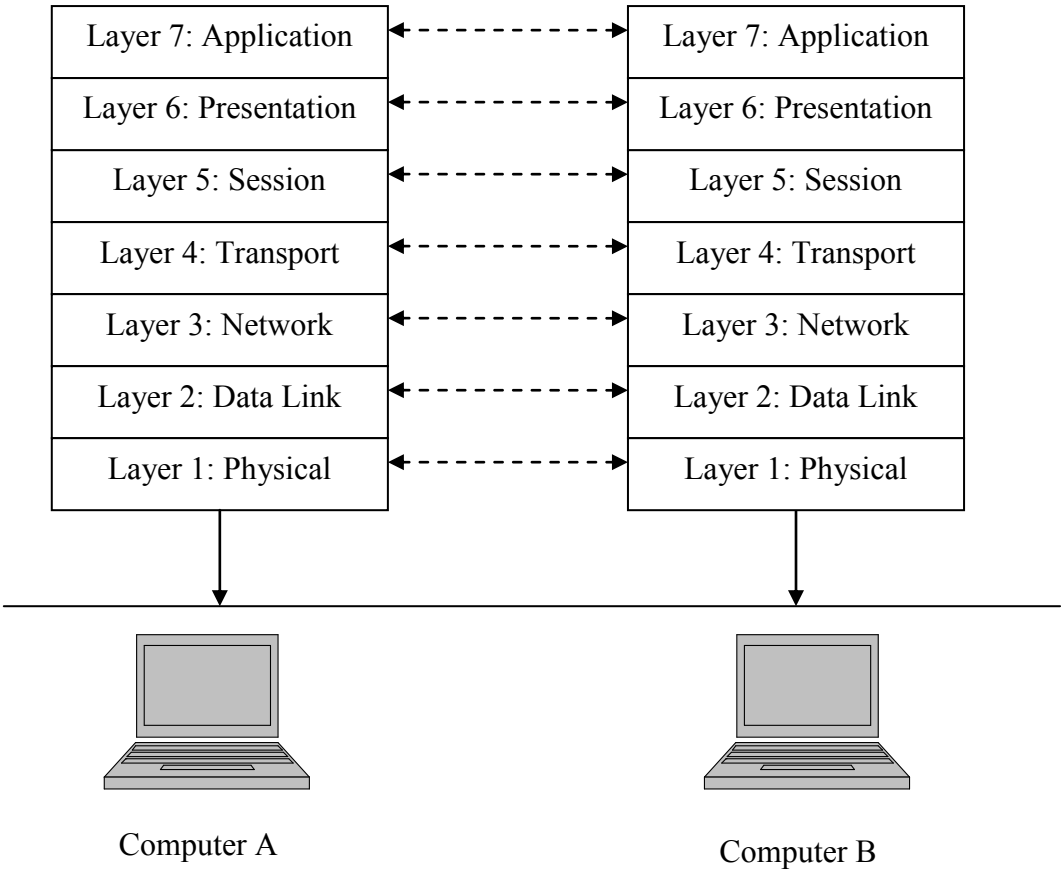

Figure 2-1: The OSI reference model

Layer 1: The Physical Layer

The bottom layer, or Layer 1, of the OSI reference model is called the physical layer. This layer is responsible for the transmission of the bit stream. It accepts frames of data from Layer 2, the data link layer, and transmits their structure and content.

Layer 1 is also responsible for the reception of incoming streams of data, one bit at a time. These streams are then passed on to the data link layer. The physical layer, quite literally, operates on only 1s and 0s. It has no mechanism for determining the significance of the bits it transmits or receives. It is solely concerned with the physical characteristics of electrical and/or optical signaling techniques. This includes the voltage of the electrical current used to

transport the signal, the media type and impedance characteristics, and even the physical shape of the connector used to terminate the media. Transmission media includes any means of actually transporting signals generated by the OSI's Layer 1 mechanisms. Some examples of transmission media are coaxial cabling, fiber-optic cabling twisted-pair wiring.

Layer 2: The Data Link Layer

Layer 2 of the OSI reference model is called the data link layer. As all the layers do, it has two sets of responsibilities: transmit and receive. It is responsible for providing end-to-end validity of the data being transmitted. On the transmit side, the data link layer is responsible for packing instructions and data into frames. A frame is a structure indigenous to the data link layer that contains enough information to make sure that the data can be successfully sent across to its destination. Implicit in this definition is that the data link layer contains its own address architecture. This addressing is only applicable to other networked devices that reside locally on the same data link layer domain. Successful delivery means that the frame reaches its intended destination intact. In some protocols, the frame also contains a mechanism to verify the integrity of its contents on delivery such as a CRC check. Numerous situations can result in transmitted frames either not reaching the destination or becoming damaged and unusable during transit. It is the data link layer's responsibility for detecting and correcting any and all such errors. The data link layer is also responsible for reassembling the binary streams that are received from the physical layer back into frames. The physical and data link layers (1 and 2) are required for each and every type of communication regardless of whether the network is a LAN or wide-area network (WAN). Together, these two layers provide all the mechanisms that software applications need to contact and communicate with other devices connected to the same LAN. In Figure 2-2, all the user machines can directly access

the local server. Consequently, they do not require the use of network layer protocols or addressing to communicate with each other.

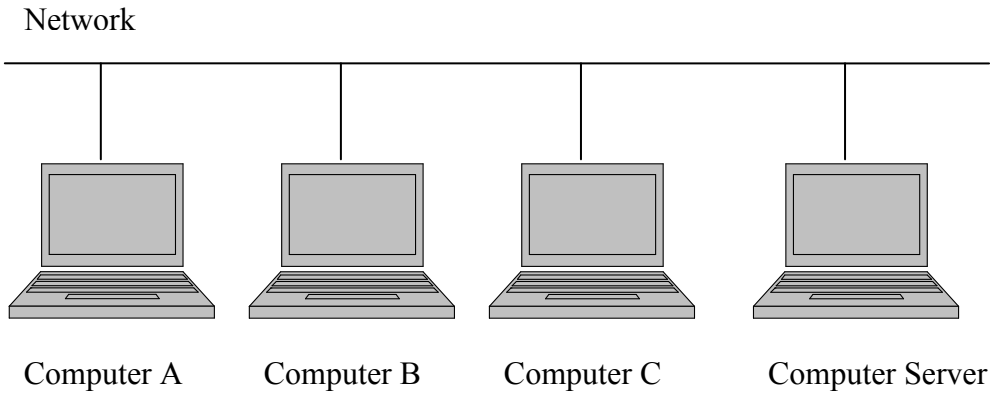

Figure 2-2: Network contains the physical and data link layers.

These two layers are also highly interrelated and, consequently, come bundled together in products. When you purchase LAN hardware (Ethernet, Token Ring, FDDI, and so on), for example, you have simultaneously selected both a physical layer and a data link layer specification.

Layer 3: The Network Layer

The network layer enables internetworking. The protocols at this layer are responsible for establishing the route to be used between the source and destination computers. This layer lacks any native transmission error detection/correction mechanisms and, consequently, is forced to rely on the end-to-end reliable transmission service of either the data link layer or the transport layer. Although some data link layer technologies support reliable delivery, many others do not. Therefore, Layer 3 protocols (such as IP) assume that Layer 4 protocols (such as TCP) will provide this functionality rather than assume Layer 2 will take care of it. Figure 2-3 illustrates the same network as Figure 2-2. The only difference is that a second network has been connected to it via a router. The router effectively isolates the two data link

layer domains. The only way to communicate between these two domains is through the use of network layer addressing.

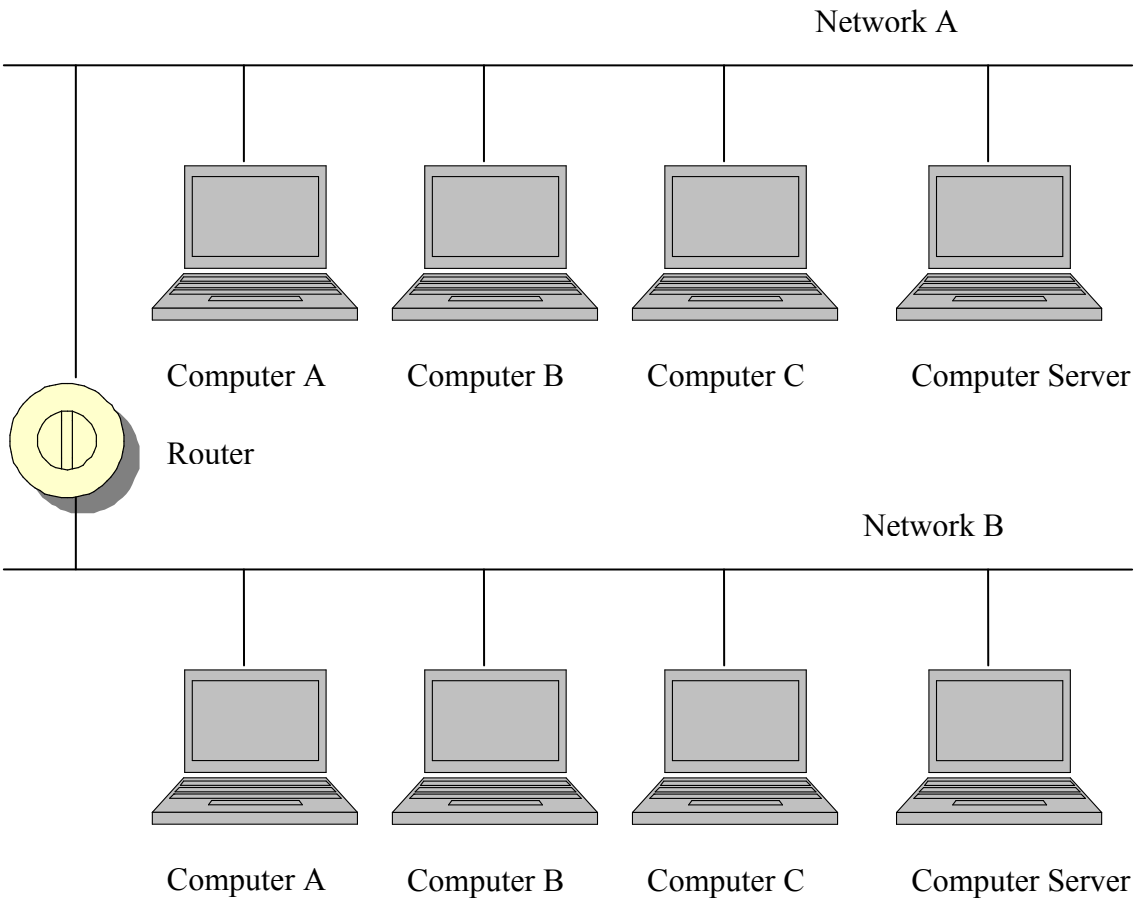

Figure 2-3: Interconnection of different network by the network layer

In this situation, if a user on Network 1 needed to access information stored on the server of Network 2, network layer addressing would be needed. The network layer can perform this intermediary function because it has its own addressing architecture, which is separate and distinct from the data link layer machine addressing. The network layer mechanisms have been implemented in a series of protocols that can transport application data across LAN segments, or even WANs. These protocols are called routable protocols because their datagrams can be forwarded by routers beyond the local network. Routable protocols include IP, Internetwork Packet Exchange (IPX), and AppleTalk. Each of these protocols, as well as

the other routable protocols, has its own Layer 3 addressing architecture. This addressing architecture is used to identify machines that are connected to different networks. Routers are needed to calculate the routes and forward the data contained within the routable protocol packets to machines that lie beyond the local link of the transmitting machine. Note that, IP has emerged as the dominant routable protocol. Consequently, this entire thesis reinforces the fundamentals of routing using only the IP protocol in the examples and illustrations. Unlike the first two layers, the use of the network layer is optional in data communications. The network layer is required only if the computer systems reside on different networks, or if the communicating applications require its services. In the first case, the different LAN domains would have to be interconnected somehow; otherwise, the communications could not occur. Alternatively, application software could require the use of either network or transport layer mechanisms, regardless of how the communicating devices are interconnected.

Layer 4: The Transport Layer

Layer 4, the transport layer, provides a similar service to the data link layer, in that it is responsible for the process to process communication. Unlike the data link layer, the transport layer can provide this function beyond the local LAN segment. It can detect packets that were either damaged or lost in transit and can automatically generate a retransmit request. Another function of the transport layer is the re-sequencing of packets that, for a variety of reasons, may have arrived out of order. The packets may have taken different paths through the network, for example, or some may have been damaged in transit. In any case, the transport layer can identify the original sequence of packets and put them back into that sequence before passing their contents up to the session layer. Much like the interrelationship between the first and second layers, the third layer of the OSI reference model is usually tightly interrelated with the fourth layer. Two specific examples of routable protocol suites

that tightly integrate these two layers are open standard TCP/IP and Novell's IPX/SPX (Internetwork Packet Exchange, Sequenced Packet Exchange). Together, these layers provide the mechanisms that enable the transfer of information between source and destination machines across a communications network that spans beyond a Layer 2 domain. These layers also provide other functions such as re-sequencing packets received out of order and retransmitting packets not received or received damaged.

Layer 5: The Session Layer

Layer 5 of the OSI model is the session layer. Many protocols bundle this layer's functionality into their application layers. Some specific examples of session layer services are Remote Procedure Calls (RPCs) and quality of service protocols such as RSVP the bandwidth reservation protocol.

Layer 6: The Presentation Layer

Layer 6, the presentation layer, is responsible for managing the way that data is encoded. Not every computer system uses the same data encoding scheme, and the presentation layer is responsible for providing the translation between otherwise incompatible data encoding schemes, such as American Standard Code for Information Interchange (ASCII) and Extended Binary Coded Decimal Interchange Code (EBCDIC). The presentation layer can be used to mediate differences in floating-point formats, as well as to provide encryption and decryption services.

Layer 7: The Application Layer

The top, or seventh, layer in the OSI reference model is the application layer. Despite its name, this layer does not include user applications. Instead, it provides the interface between

those applications and the network's services. This layer can be thought of as the reason for initiating the communications session. For example, an email client might generate a request to retrieve new messages from the email server. This client application automatically generates a request to the appropriate Layer 7 protocols and launches a communications session to get the needed files. Note that most of today's networking protocols use their own layered models. These models vary in the degree to which they adhere to the separation of functions demonstrated by the OSI reference model. It is quite common for these models to collapse the seven OSI layers into five or fewer layers. It is also common for higher layers to not correspond perfectly to their OSI-equivalent layers. Additionally, models may not even describe the full spectrum of the OSI's layered functions! The IEEE's layered functional model, for example, is solely for LANs and MAN—it does not extend above the data link layer.

Ethernet, Token Ring, and even FDDI are compliant with this model.

2.4 Client / server model

In an Internet environment, the method of communication between two nodes is based on the client-server paradigm. This model forms the basis of most network communications, and allows us to conceptualize the functioning of the layered approach described above. Clients formulate requests, send them to servers and await replies. Servers execute and manage processes that await requests and perform actions. The paradigm is more than just software, it comprises various separate technologies working together to provide a service. In a mobile device, it always works as a client and accesses the services from a server at the other end. Therefore, if a client is moving from one network to another network, the mobility management function is critical to allow the client continuous access the server for the required services. Some services, if they do not required a continuous session, can be re-

connected or be re-established after a client moves from one network to another. One example of this kind of service is the Internet web browsing on a HTTP service. The web browser on the client side can refresh the HTTP pages and in most circumstances it will not affect the user experience. However, some services which require a continuous session will more likely be affected by the movement of the mobile device from one network to another. Therefore, mobility management is critical for these applications.

2.5 Internetworking architecture

An interconnected network consists of hosts / computers which are connected by a series of intermediate devices called routers. A Host is used as a means for any end-user‘s (client or server) computer to connect to a network. The general Internet architecture consists of network segments interconnected by devices called routers. Routers are devices that use IP addresses to make decisions on where to send IP packets. From a layering perspective routers are composed of OSI layers 1-3. They have two or more network interface layer connections. Figure 2-4 shows an internetworking architecture which consists of two hosts interconnected by a router [42].

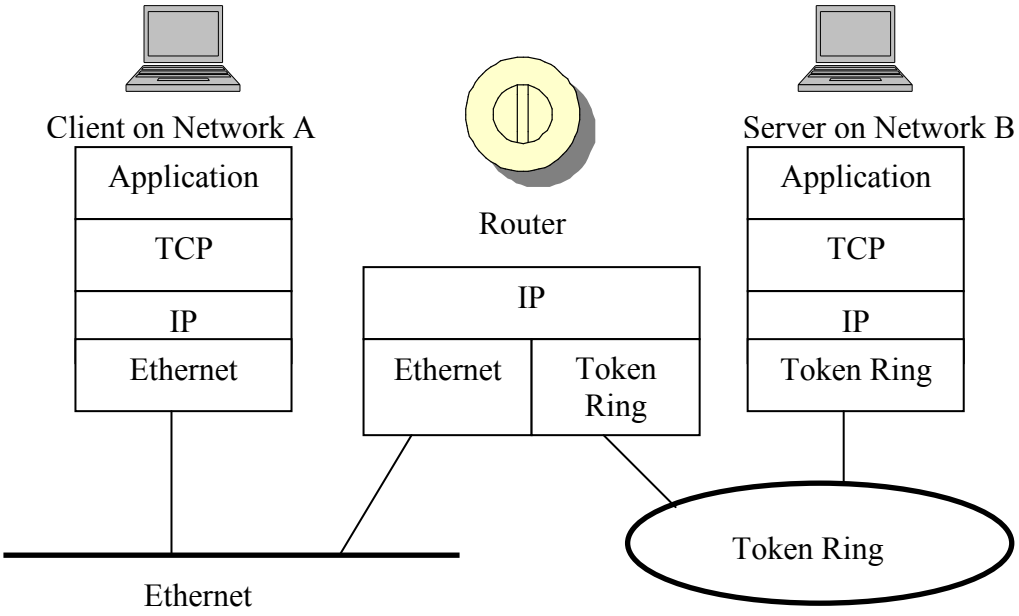

Figure 2-4: Internetworking architecture

2.6 Information encapsulation and decapsulation

It is important to understand how hosts send each other packets of information. This flow of information occurs vertically along the protocol stack. A sending host application will generate data that flows down the protocol stack with each subsequent layer with a header appended into the information. The header contains control information for the management and processing of the specific function in that layer. This process is called encapsulation. At the destination host, it proceeds up the protocol stack towards the destination application, with each layer removing its header and performing error checking and / or error correction in a reverse process. This is referred as the decapsulation process. Figure 2-5 [43] shows the concept of encapsulation and decapsulation.

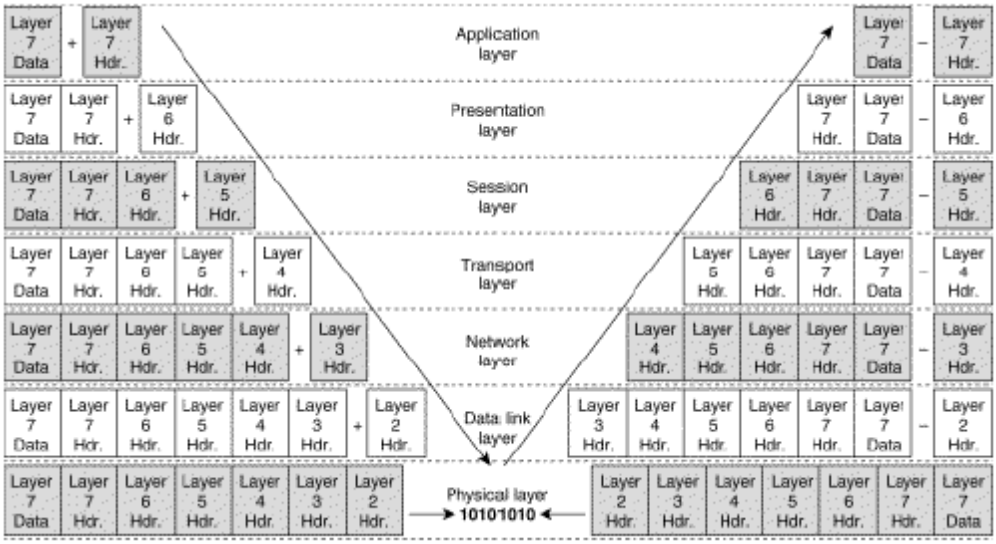

Figure 2-5: The concept of encapsulation and decapsulation.

2.7 The Internet Protocol

The Internet Protocol (IP) [44] is the layer 3 protocol in the OSI 7 layer concept. It provides an unreliable datagram delivery service across networks. It is connectionless because each datagram is independently transmitted between one and others and no session is formed to manage a group of packets. It also is an unreliable protocol because the network does not guarantee delivery of the datagram but a best effort approach is often used. A simple error control protocol (ICMP[5]) is called upon to notify the source should network problems occur. This notification being carried within IP is also a best-effort basis and the delivery is not guaranteed. Any reliability constraints, including flow-control, are the responsibility of the higher layer protocols (e.g. TCP). Figure 2-6 [45] shows the IP datagram format in details.

[5] ICMP stands for Internet Control Message Protocol

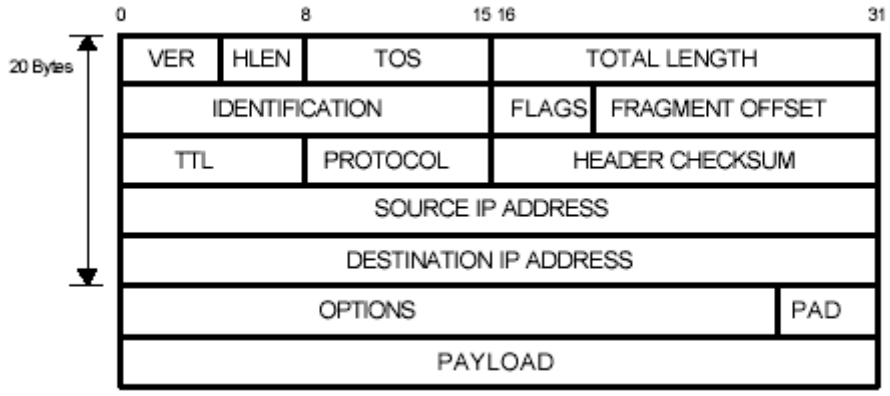

Figure 2-6: IP datagram

The details definition of each fields are as follows.

VERSION (VER): This is a 4-bit field which is used to indicate to the IP software module the version used for this datagram.

HEADER LENGTH (HLEN): This is a 4-bit field which contains the length of the datagram header measured in 32-bit words.

TYPE OF SERVICE (TOS): This is an 8-bit field comprising 5 sub fields as illustrated below:

- Precedence: 3-bit field which is ignored today
- D, T, R, M: Specify type of transport , where D=minimize delay,
- T=maximize throughput, R=maximize reliability, M=minimize monetary cost

TOTAL LENGTH: This is a 16-bit field which measures the total length of the IP datagram header and payload) measured in octets. Maximum size of datagram is therefore 2^{16} or 65,535 octets, PAYLOAD LEN=HLEN-TOT. LENGTH.

IDENTIFICATION: This is a 16-bit field used to assign a unique (within the bounds of wraparound) number to each datagram generated by a host. This field is important in the reconstruction of fragmented datagrams.

FLAGS (FL): This is a 3-bit field which controls and aids in the fragmentation/reassembly process. The fields are defined as **Bit 0:** Reserved; **Bit 1:** DF 0=May fragment, 1=don‘t fragment, **Bit 2:** MF 0=Last fragment, 1=more fragments.

FRAGMENT OFFSET: This is a 13-bit field which specifies where in a fragmented byte stream this fragment belongs. It is measured in units of 8 octets and is relative to the first fragment which has an offset of zero.

TIME TO LIVE (TTL): This is an 8-bit field which specifies an upper bound on the life time of a datagram in a network. The value is decremented by 1 by each router that processes the datagram. When this field reaches zero the datagram is destroyed and an ICMP message is generated.

PROTOCOL: This is an 8-bit field which indicates the higher layer protocol that created the data being carried in the datagram. TCP has protocol value=00000001.

HEADER CHECKSUM: This is 16-bit field used to compute a checksum for the IP header portion of a datagram. This provides a low level of error detection for IP datagrams. As some fields of the IP header change as routers process datagrams, this value is recomputed at each processing point.

SOURCE IP ADDRESS: This is the 32-bit IP address of the sender.

DESTINATION IP ADDRESS: This is the 32-bit IP address of the receiver.

OPTIONS: This is a variable length field. There may be zero or more options in a datagram. The current options defined are:

1. Security and handling restrictions
2. Record route (each router records its IP address)
3. Timestamp (each router records IP address + time)
4. Loose source routing (specify list of IP addresses that must be traversed by datagram)
5. Strict source routing (specify list of the *only* IP addresses datagram can traverse)

PADDING: In order to ensure that the IP header ends on a 32-bit boundary, padding is added using zero bits.

All hosts that are part of an IP network must have an IP address assigned which is used to identify the host in the network. IP addresses are the key identifiers in moving information around a network which is used to deliver datagrams. An IP address is a 32-bit number that is made up of network prefix portion and host portion. Routing is based on the network prefix portion. Routers also advertise reachability to the network prefixes which they serve. This information is exchanged with other routers using routing protocols that have the effect of creating forwarding entries that allow routers to deliver a packet to the destination on hop-by-hop basis. Each hop is treated independently of the previous hop. IP addresses are written in dotted decimal notation and convey information on the prefix length of the network prefix portion, e.g. 192.168.22.1/24. This implies the first 24 bits make-up the network prefix and the remaining 8 bits are the host portion. Figure 2-7 illustrates the IP address format.

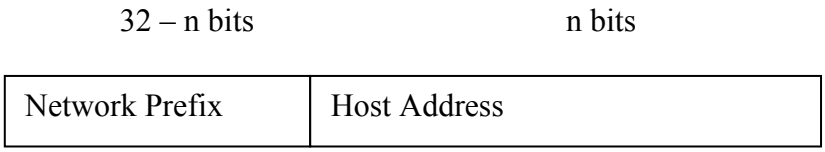

Figure 2-7: IP address format

All hosts and routers in an IP network maintain a routing table. The table is checked for every forwarding decision when handling IP packets by a host or a router. A host has two entries in its routing table; one for a default router used for packets destined for external nodes and one for internal nodes. Entries in a routing table are created using routing protocols, manually or a combination of both. As routes within the Internet can change, routing tables contain

dynamic information. Routing table entries are classified into three broad categories. These are listed below:

- Host specific route: This entry has a prefix length of 32 bits. Completely specifies the destination node.
- Network prefix route: This entry identifies a separate network and provides a next hop when the prefix length is between 1-31 bits.
- Default route: This entry provides a match for all destination IP address, having a prefix length of 0 bits. When determining the next hop, the routing algorithm must select the route with the largest matching prefix length. Therefore the order of precedence in selecting a route is host-specific, failing that network prefix, and failing that default route. A simplified (logical) sample routing table is shown below in Table 2-1 [50]. Each route type is included.

Table 2-1: Example of a routing table.

Destination	Next Hop	Interface
192.168.0.0/32	Host1	Ethernet1
255.254.21.0/8	Host2	Ethernet2
0.0.0.0/0	Host2	Ethernet2

2.7.1 Internet Control Message Protocol

The Internet Control Message Protocol (ICMP) is an error control protocol, which is designed as part of the IP implementation. It provides basic feedback on errors based on abnormal network conditions or responds to basic queries such as echo request such as the PING protocol. It is a subset of IP as is defined in RFC 792.

2.8 Transport protocols

This section provides an introduction of the OSI layer 4 Transport protocol. Two of the most common transport layer protocols in use over IP based networks are introduced. These are the Transport Control Protocol (TCP) and User Datagram Protocol (UDP). Here, TCP provides a reliable data transport between two hosts and UDP provides a lightweight but non-reliable service for fast delivery and low overheads. The TCP protocol is used for loss sensitive applications such as file transfer and UDP is used for real-time or delay sensitive applications such as a video call or voice call.

2.8.1 Transmission Control Protocol

It is noted that TCP provides a connection-oriented, reliable datagram delivery service for higher layer applications. It also provides the following functionality [46]:

Virtual circuit connection: TCP establishes a point-to-point link between two hosts by defining two end-points. It is a virtual connection. A virtual connection is maintained in software only. There is no physical end-to-end link. Each end-point consists of the pair [IP address, port number].

Buffered transfer: TCP provides buffers in both receive and transmit sides. Also, the flow of data is under control by a flow control mechanism.

Reliable data stream transfer: TCP guarantees a complete, ordered and error free data delivery from the sender to the receiver application. TCP divides a data stream into octets and delivers the exact stream sequence from the sender to the receiver.

Full duplex operation: TCP provides two independent connections between two hosts. It allows data to flow in either direction independent of the other. TCP adds a minimum of 20 bytes of header to each application layer segment.

2.8.2 User Datagram Protocol

UDP provides a connection-less, unreliable datagram delivery service for higher layer applications. UDP is a lightweight transport protocol that adds only 8 bytes of header overhead to the application data. No acknowledgement or ordering functionality is provided as with TCP. Note that, UDP provides the following functionality [46]:

Virtual circuit connection: UDP provides protocol ports in support of multiple communication sessions between two hosts. The ports consist of the pair [IP address, port number].

Optional checksum field: UDP provides the facility to calculate a checksum that covers the UDP header and part of the IP header. Application programs using UDP must be able to provide reliability mechanisms to handle duplication, out-of-sequence delivery, link layer failures.

Chapter 3: Mobile IP mobility management protocol

To initially overcome the mobility management problem, Mobile[6] IP was and still is one of the most popular protocols, which was standardized by IETF and it was subsequently enhanced to the Mobile IPv6 protocol. Due to the signaling overheads and global mobility characteristics, Mobile IP is normally classified as a Macro-mobility protocol. This thesis will not investigate Mobile IPv6 but it summarizes some improvements of the Mobile IPv6 protocol, which could be used for future work.

Mobile IP provides an IP based mobility solution, which allows a Mobile Node (MN) to maintain network connectivity when the MN moves from one network to another network. A permanent IP address is assigned to every MN called a Home Address. When an MN handoff occurs to another subnet, it acquires a temporary IP address called Care of Address (CoA), which is a topology suitable for the visiting network. When an MN is attached in the home

[6] To understand the origin of the term _mobile' it is useful to distinguish be between _static mobility' as is the case of a user terminal accessing the Internet at different static locations and for which the Mobile IP protocol was framed and _continuous mobility' in the case of a user terminal moving in space (e.g. in a moving vehicle) and accessing the Internet via a mobile device.

network, MN uses its permanent home IP address for communication. A location register and router on the home subnet called the Home Agent (HA) keeps track of an MN‘s binding that maps a home address into the corresponding CoA. When the MN moves into a new subnet, a MN acquires the CoA from the Foreign Agent (FA) and sends registration requests to inform the HA of the new CoA. HA also acts as the MN to attract all packets delivered to the MN and using an IP tunnel to deliver the packets to the MN via the FA. This approach allows IP packets, originally sent to the MN‘s home IP address, to be forwarded to the MN‘s CoA at the foreign network. Recently, Mobile IPv6 provides a new feature called route optimization, which allows dynamic binding update to the Corresponding Node (CN) in the IPv6 network. After the binding update, the CN can send packets to the MN directly through the CoA of the MN and minimize further IP packets forwarding using the tunneling methods. The following sections describe the detailed mechanism of Mobile IP.

3.1 Mobile IP

Mobile IP [47] provides a network layer solution to address the mobility problem in an IP based network. Mobile IP allows a mobile node to move between networks while retaining the same IP address and its connections. Mobile IP is the IETF standard described in RFC's 2002 - 2006. The base Mobile IP protocol as defined in RFC2002 [48] allows a mobile node to move between networks without changing its IP address. A permanent IP address (termed the home address), that identifies the mobile nodes *home network*, is used in all transport layer connections and for routing when at home. A second, topologically address, termed the *care-of address,* is obtained and registered by the mobile node when roaming in a *foreign network* and used for routing in such situations. This care-of address reflects the nodes current point of attachment. Mobile IP introduces the following functional entities [49]:

Home Agent (HA) – is a router on a mobile node‘s home network that contains a database holding the permanent IP address to current care-of address mapping of all mobile nodes that provides service. The HA *tunnels* datagrams for delivery to the mobile node when it is away from home and maintains current location information for the mobile node.

Foreign Agent (FA) – is a router on a visited network that provides routing services to the mobile node while registered within its domain. The FA decapsulates and delivers datagrams to the mobile node that were tunneled by the HA.

Mobile Node (MN) – is a host or router that changes its point of attachment from one network or subnet to another. A mobile node may change its location without changing its IP address. It may continue to communicate with other Internet nodes at any location using its (static) IP address.

Figure 3-1 [50] shows the basic operation of Mobile IP when a mobile has roamed into a foreign network and registered its care-of address with its HA, while continuing communication with a correspondent node, who is unaware of the location change by the mobile node. The operations are summarized as below:

- Mobility agents, such as the HA and FA, broadcast their capabilities on their respective subnets.
- Mobile nodes listen to these advertisements and determine if they are on a home or foreign network.
- Mobile nodes register their care-of address (obtained through advertisements by FA or dynamically via DHCP server) with their HA if roaming. If on their home network a

mobile node works without the support of mobility agents. Data packets destined to MN's home address are intercepted by HA and then tunneled to FA, who delivers the packets to the MN in the foreign network.

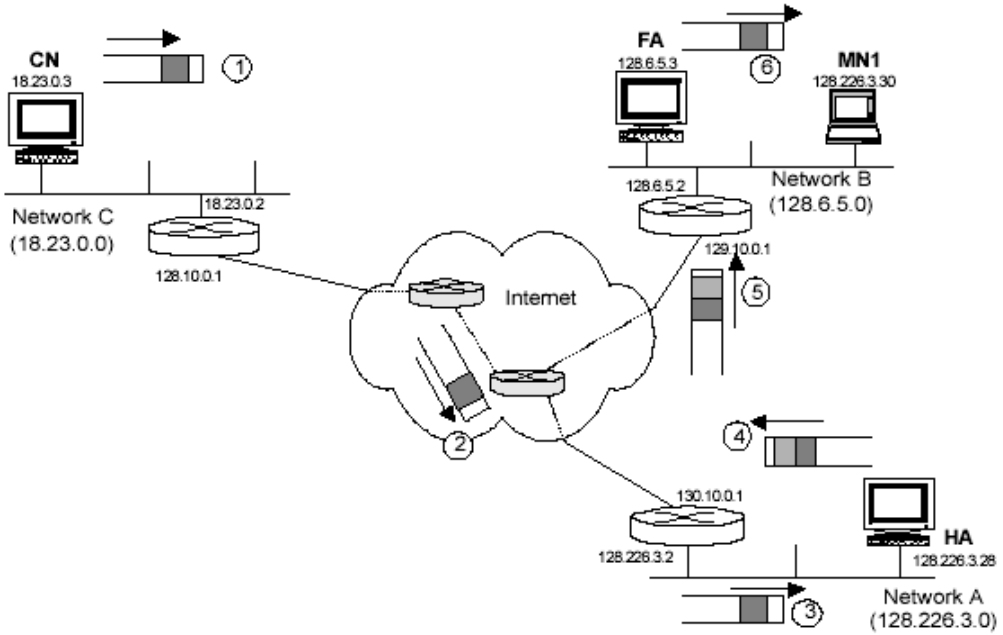

Figure 3-1: Basic operation of Mobile IP

In Figure 3-1, a correspondent node (CN) in network C sends a datagram to mobile node 1 (MN1). This datagram is addressed to MN1 home address, which results in delivery to network A (steps 1, 2 and 3). On MN1s home network, the HA acting as a proxy for MN1 intercepts this datagram, consults a routing table which contains MN1s current care-of address, and encapsulates the datagram in another IP header which has the current care-of address as the destination address. The datagram is then routed through the Internet and arrives at the FA identified by the care-of address (steps 4 and 5). The FA decapsulates the datagram and recognizes the destination IP address as that of MN1 who has registered within its domain. The FA then employs link layer resolution to deliver the datagram to MN1 (step 6). Although not shown, the CN may also be a mobile node. The routing of datagrams between CN-HA-FA is referred to as *triangle routing*. Datagrams originated by MN1 are delivered via normal

IP routing mechanisms and do not need to be directed through the home network or HA. It is well known that the triangle routing is an inefficiency and causes packet delivery delay while the MN is in the foreign network. Mobile IPv6 and Mobile IP route optimization [64] scheme address this issue.

The Mobile IP protocol can be broken down into three main operations, namely:

- **Agent discovery:** The process whereby mobility agents advertise their presence so the mobile node can detect it is visiting a foreign network connection.
- **Registration:** The process whereby a mobile node informs its HA of the current IP address acquired in the foreign network.
- **Tunneling:** The method by which datagrams are redirected to the mobile node which the mobile node is in the foreign network and using a different network address.

Each of the three procedures is discussed in detail in the following sections.

3.1.1 Agent discovery

In order for mobile nodes to be able to use Mobile IP the services of the HA and FA must be available. These two mobility agents advertise their presence on a network by using extensions to ICMP router discovery messages (RFC1256). The ICMP router discovery provides a method by which a host can automatically determine a default router through which datagrams can be sent. This protocol is defined for use on broadcast/multicast capable networks. Router discovery consists of two components: *router advertisements* and *router solicitation*. Router advertisements are broadcast by certain routers and contain a list of one or more IP addresses for the network interface on which the message was sent, including a

preference level for each advertised address. Router solicitation messages are broadcast by hosts to solicit the immediate sending of a router advertisement message.

3.1.1.1 Router advertisement

Mobile IP defines an extension to the ICMP router advertisement message termed a *mobility agent advertisement.* This extension carries a list of care-of addresses supported by the FA. The HA also uses this extension, not to advertise care-of addresses, but for other administrative matters. A further extension termed ‗prefix length extension' may be included by a mobility agent to enable the mobile node to determine the network prefix of each address supplied in the router advertisement portion of the message. If present, this extension follows the mobility agent advertisement extension. This extension can be used by a MN to determine if it has moved to a new subnet. Figure 3-2 shows the structure of ICMP router advertisement message, the mobility agent advertisement extension, and the prefix length extension. Figure 3-3 shows the encapsulation of the messages [45].

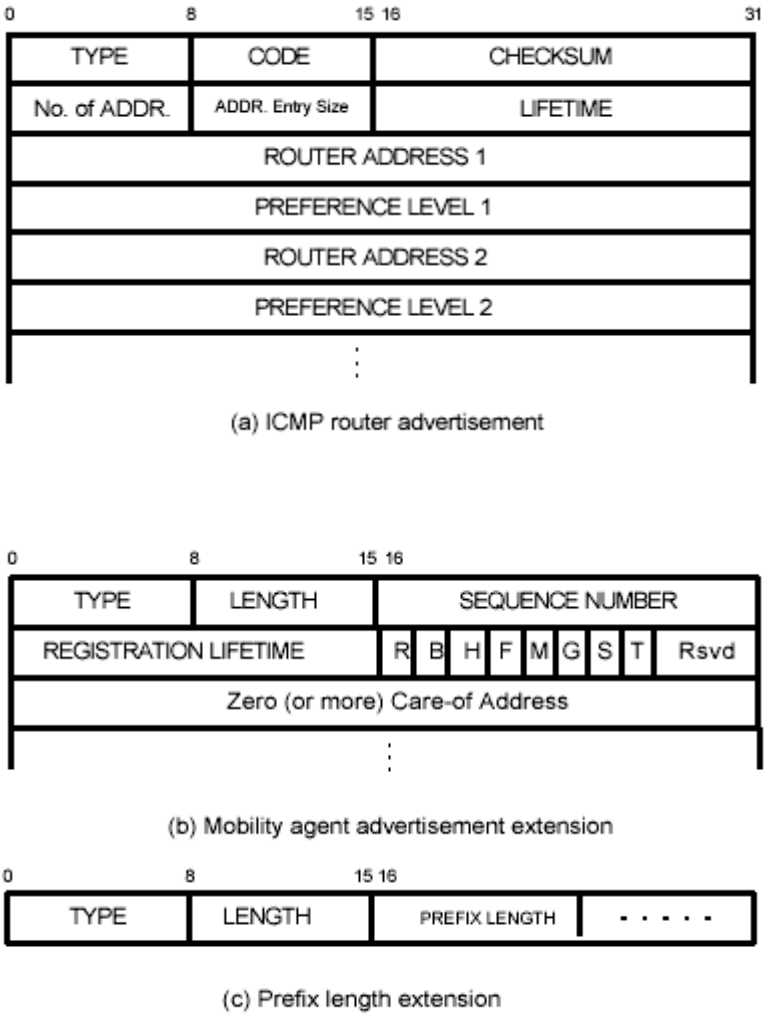

Figure 3-2: Router advertisement methods

IP HEADER	ICMP Router Advertisement	Mobility Agent Advertisement Ext	Prefix Length Extension

Figure 3-3: Agent advertisement message

The fields are defined as follows:

(a) ICMP Router Advertisement

TYPE: This is a 8-bit field with value 9 for ICMP router advert.

CODE: This is a 8-bit field with value 0

CHECKSUM: This is a 16-bit field that uses the same method of calculation as for an IP header checksum.

No. of Addresses: The number of router addresses (32-bit words) advertised in this message address entry size The number of 32-bit words of information per each router address. In this version the value is 2, i.e. one for the address and one for the preference

LIFETIME: The maximum number of seconds that the router addresses may be considered valid.

Router Address[i]: The sending router‘s IP address(es) on the interface from which the message is sent. (i=1…No. of Addresses)

Router Preference[i]: The preferability of each router address as a default router, relative to other routers on the same subnet. (i=1…No. of Addresses)

(b) Mobility Agent Advertisement Extension

TYPE: This is a 8-bit field with value 16 for this extension

LENGTH (6 + 4*N): where N is the number of care-of addresses advertised

SEQUENCE NO.: This is a 16-bit number that is incremented by the mobility agent each time an agent advertisement is sent

REGISTRATION LIFETIME: This value indicates to the mobile node the maximum lifetime (in seconds) that this mobility agent is willing to accept in a registration request. A value of 65,535 (18hrs) indicates infinity.

FLAGS

R: Registration required. Registration through a foreign agent is required rather than using a co-located care-of ADDRESS.

B: Busy. Indicates that the foreign agent cannot accept registrations from additional mobile nodes.

H: Home Agent. This bit indicates that the mobility agent offers home agent services.

F: Foreign Agent. This bit indicates that the mobility agent offers foreign agent services.

M: Minimal Encapsulation. This bit indicates the mobility agents supports tunneling of datagrams using encapsulation technique defined in RFC2004

G: Generic Route Encapsulation (GRE). This bit indicates the mobility agents supports tunneling of datagrams using encapsulation technique defined in RFC1701

S: Smooth Hand-off. This bit is not yet a part of RFC2002, but has an intended use in an internet draft (work-in-progress) to improve hand-off performance of Mobile IP. This is discussed in a later section.

T: Tunneling. This bit is specified in RFC2344 and indicates support by mobility agent for reverse tunneling. Support for RFC2344 is not mandatory for a mobility agent

CARE-OF ADDRESS: The advertised care-of address(es) provided by this foreign agent. If F bit is set then at least one care-of address is to be present.

(c) Prefix Length Extension

TYPE: Value of 19

LENGTH N: where N = value of ‗number of addresses' field in the ICMP router advertisement portion of the message

Prefix Length: The number of leading bits that define the network of the corresponding router address listed in the ICMP router advertisement portion of the message

3.1.1.2 Router solicitation

Any host can send a router solicitation message to elicit a router advertisement message. If mobility agents are present on the subnet then their replies will also contain a mobility agent

advertisement extension. Note that in the absence of a default router on the subnet a foreign agent must offer default router capability to a MN. Figure 3-4 shows the structure of ICMP router solicitation message and the way in which the agent advertisement messages are encapsulated in IP datagrams.

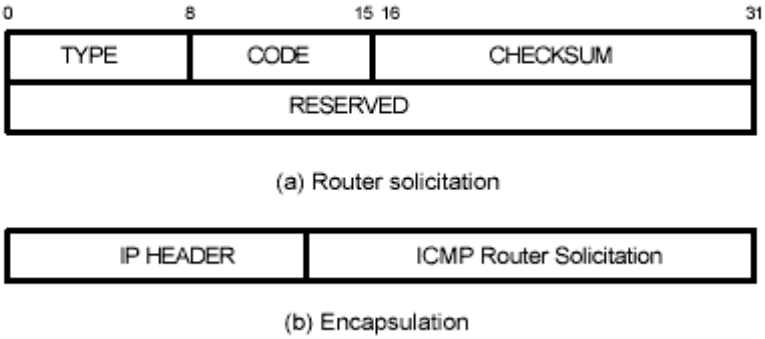

Figure 3-4: Router solicitation message

A mobile node should only send a solicitation message in the absence of an agent advertisement. When sending a solicitation message the destination IP address should be the all-routers multicast address (224.0.0.2) or the limited broadcast address (255.255.255.255). Because ICMP router discovery messages can be sent by non-mobility agent routers, a MN must be able to distinguish between two such uses. This is achieved by calculating the IP payload length (Total Length field – Header Length field) and comparing against the ICMP field length indicated by the _No. of Addresses' field. If a non-zero difference exists then a mobility agent advertisement extension exists and the advertising router has FA, HA or combined functionality.

3.1.2 Registration

Registration is the process that allows a mobile node to communicate current reachability information to its HA. A mobile node will invoke a registration process for the following reasons:

- Whenever there is a change in care-of address
- To renew a current registration that is about to expire
- To deregister a care-of address when the mobile node returns home

RFC2002 defines two registration messages, registration request and registration reply. These new messages are exclusive to Mobile IP and are defined using UDP well known port 434. The message format consists of the fixed length (24 octets) registration component followed by any extensions (variable length). Figure 3-5 shows the encapsulation of the message formats. The HA maintains a dynamic storage containing mobility bindings for the mobile nodes it provides service for. The mobility bindings map the permanent IP address (home address) to the current care-of address of the mobile node and hold the remaining registration lifetime for that binding. These bindings are updated by registration requests received from mobile nodes. Figure 3-6 illustrates the process of registration in which a mobile node obtains a care-of address through a FA and registers this with its HA [48]. If a mobile node uses a co-located care-of address then the registration request is sent directly to the HA. The HA sends the registration reply directly to the mobile node. Only a mobile node can originate a registration request. A FA or HA can only send a registration reply in response to a registration request. In the above scenario the FA acts as a relay point for the registration messages, however it does interpret and use the information in the registration message and may add or remove extensions as required. The FA cannot modify the MN assembled portion of the registration request as this would invalidate the authenticator value in the MN-HA authentication extension.

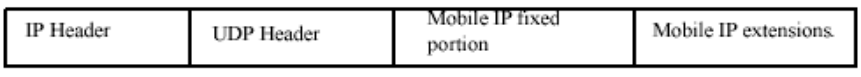

Figure 3-5: Mobile IP message

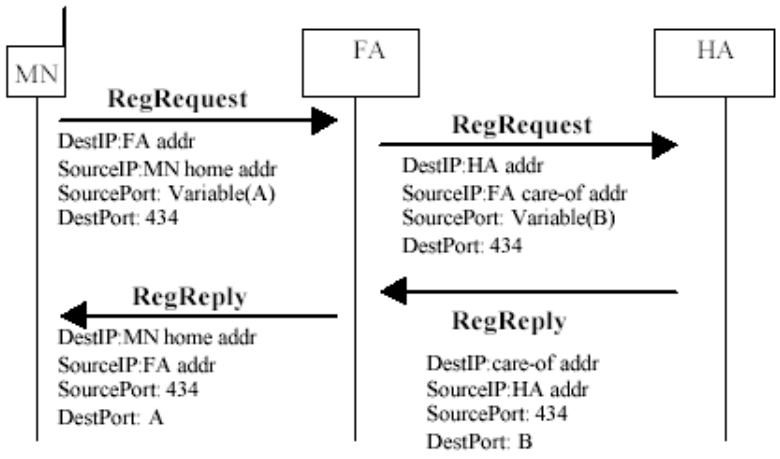

Figure 3-6: Registration message sequence

3.1.2.1 Registration request

A mobile node issues a registration request to modify the mobility binding held at the HA. The message format is given below in Figure 3-7 [47].

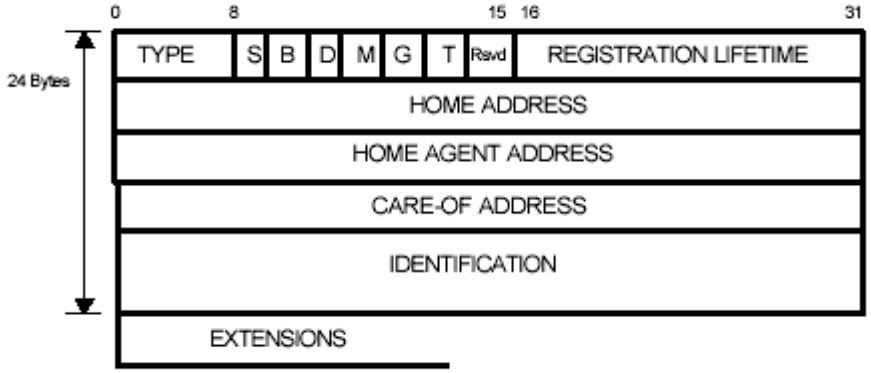

Figure 3-7: Registration request message

The fields are defined as:

TYPE: Value of 1 for registration request

FLAGS:

S: Simultaneous bindings. When set indicates to the HA that the MN wishes to retain any prior mobility bindings in addition to this one.

B: Broadcast datagrams. When set indicates to the HA that the MN wishes to receive any broadcast messages on the home network.

D: Decapsulation. When set indicates to the HA that the MN is capable of decapsulating any datagrams tunneled to it using an agreed tunneling protocol. This implies MN is using a co-located care-of address.

M: Minimal Encapsulation. This bit indicates the MN requests tunneling of datagrams to it using the encapsulation technique defined in RFC2004.

G: Generic Route Encapsulation (GRE). This bit indicates the MN requests tunneling of datagrams to it using the encapsulation technique defined in RFC1701

T: Tunneling. This bit is specified in RFC2344 and indicates MN desire to use reverse tunneling. Support for RFC2344 is not mandatory.

MN REGISTRATION LIFETIME: This is the requested duration of lifetime by the MN for this mobility binding

HOME ADDRESS: The permanent IP address of the mobile node

HOME AGENT: The IP address of the home agent

CARE-OF ADDRESS: The IP address acquired by the mobile node identifying the tunnel endpoint.

IDENTIFICATION: A 64-bit value used for matching registration requests with registration replies and for replay protection.

EXTENSIONS: Any extensions added by the mobile node or FA

The base Mobile IP specification, RCF2002, defines three extensions, all related to authentication functions. These are covered in detail in the security aspects section.

3.1.2.2 Registration reply

A registration reply is sent in response to a registration request by the HA and is relayed by the FA, if one exists in the path. Figure 3-8 shows the registration reply message structure.

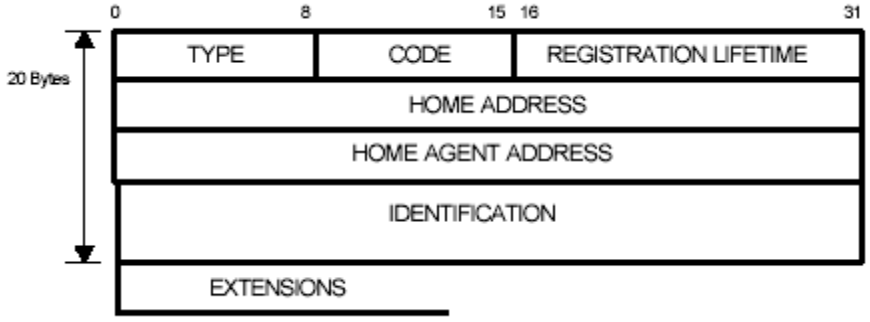

Figure 3-8: Registration reply message

The fields are defined as:

TYPE: Value of 3 for registration reply

CODE: Value indicating the result of the registration

REGISTRATION LIFETIME: The duration for which this mobility agent is willing to offer the registration

HOME ADDRESS: The permanent IP address of the mobile node

HOME AGENT: The IP address of the home agent

IDENTIFICATION: A 64-bit value used for matching registration requests with registration replies and for replay protection.

EXTENSIONS: Any extensions added by the HA or FA

The base Mobile IP specification, RCF2002, defines three extensions, all related to authentication functions. These are covered in detail in the security aspects section. The code field indicates the result of the registration. Three broad categories are defined for the reply.

- Registration successful

- Registration denied by foreign agent
- Registration denied by home agent

Code values of 0 or 1 indicate a successful registration. The reader is referred to RFC2002 for a list of code values indicating denial of registration.

3.1.3 Tunneling

Tunneling is the mechanism employed by the HA to deliver datagrams destined for the mobile node. Tunneling techniques are used in a variety of applications in the Internet, such as multicast, multi-protocol operation and VPN services. Many tunneling specifications have been released as RFC's such as, L2TP (RFC2661), IP-in-IP Encapsulation (RFC2003). Tunneling works by encapsulating the original datagram with a new header that defines an intermediate destination IP address. Routing delivers the datagram to the intermediate destination, where decapsulation occurs revealing the original destination IP address. Normal routing again delivers this datagram to its original intended destination. With Mobile IP tunneling occurs whenever a datagram destined for a mobile node arrives at the home network and is intercepted by the HA, who determines that the MN is roaming. In order to deliver the datagram a new IP header is appended to the original datagram containing the MN care-of address as the destination address. The care-of address may terminate at the FA or the MN (if MN is using a co-located care-of address). Mobile IP requires that all FA and HA support tunneling datagrams. Any mobile node intending to use co-located care-of address is also required to support tunneling. The following tunneling protocols are specified in RFC2002:

1. IP-in-IP encapsulation (RFC2003) (mandatory)
2. Minimal encapsulation(RFC2004) (optional)
3. Generic Route Encapsulation (GRE, RFC1701) (optional)

In order to use IP-in-IP encapsulation, it is a requisite to know in advance that the tunnel end-point can perform decapsulation. For Mobile IP, it is a protocol requirement that IP-in-IP encapsulation be supported. Figure 3-9 shows the how a datagram is encapsulated using RFC2003.

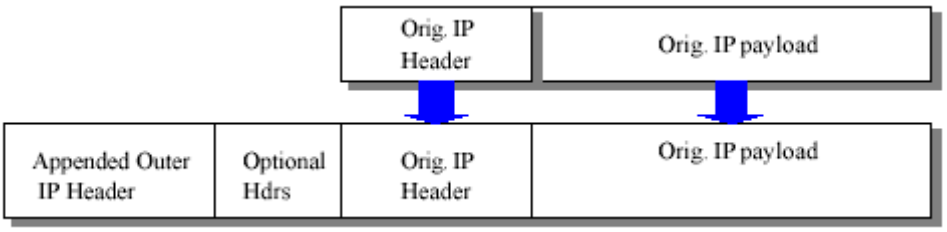

Figure 3-9: IP in IP encapsulation message

The outer IP header defines the end-points of the tunnel, i.e. the encapsulator (source IP address) and decapsulator (destination IP address). The inner IP header (including any options present) is copied from the original IP datagram unchanged apart from decrementing the TTL field as a result of processing. Encapsulation cannot be performed on a datagram with TTL=0. Similarly, if on decapsulation the inner IP header TTL is found to be 0 then the datagram is discarded. Optional headers may precede the outer IP header. Such an optional header may be the IP authentication header (RFC1826).

Chapter 4: Introduction to cellular and wireless networks

This chapter provides an overview of third generation and wireless mobile technologies. Third generation networks introduce packet-mode operation of the air interface and IP based core networks. Third generation networks deliver high speed data rates to users enabling applications such as video telephony to become a reality. This chapter introduces the third generation technologies leading to the latest 4G wireless technologies WiMax[7] and LTE. Integration with Mobile IP is also discussed in the later section.

4.1 Third generation cellular networks

3G is the third generation of mobile phone standards and technology, superseding 2G. It is based on the International Telecommunication Union (ITU) family of standards under the International Mobile Telecommunications program, IMT-2000. Note that 3G technologies enable network operators to offer users a wider range of more advanced services while achieving greater network capacity through improved spectral efficiency. Services include wide-area wireless voice telephony, video calls, social networking applications and broadband wireless data, all in a mobile environment. Enhancement of 3G data capability using include High Speed Packet Access (HSPA) data transmission capabilities has enabled deliver speeds

[7] With the rapid deployment of LTE and the take up of WiMax, LTE and LTE Advanced are becoming the Global 4G standard.

up to 14.4 Mbit/s on the downlink and 5.8 Mbit/s on the uplink. 3G networks are wide area cellular _voice centric' mobile networks which evolved to incorporate high-speed internet access and video telephony. Various countries such as China have developed variants of the 3G technology standards. The European 3G effort is termed Universal Mobile Telecommunications Service (UMTS), a collaborative effort by the RACE and ACTS programs [51]. IMT-2000 identified global spectrum bands for 3G systems. During the May 2002 meeting of the WARC, three alternative bands with a total of 519 MHz of spectrum, were allocated for IMT-2000 applications. Subsequently three further bands identified were 806-960 MHz, 1710-1885 MHz, and 2500-2690 MHz.

The major objectives for IMT-2000 3G systems are listed below:

1. Use of common global frequency bands.
2. Global roaming independent of radio access technology.
3. Providing a common operating band for the radio technologies in different environments, such as vehicular, pedestrian and indoor.
4. High transmission speed (e.g. for Internet access) relative to environment.
5. Compatibility of services within IMT-2000 and fixed networks.
6. Spectrum efficiency, quality, flexibility and overall cost improvement as a result of the utilization of advance technologies.
7. Virtual Home Environment (VHE) concept –service portability.

With second-generation air interface standards presenting the bottleneck for wireless Internet access, 3G systems were essentially designed from the bottom-up, i.e. bearer services before applications. The most featured part of IMT-2000 was the standardization of air interface

specifications, or radio transmission technologies (RTTs). The minimum requirements set forth in IMT-2000 for these RTTs are listed below:

1. 144 kb/s in vehicular environment
2. 384 kb/s in pedestrian environment
3. 2.048 Mb/s in indoor office environment
4. 9.6 kb/s in satellite environment.

Originally 10 proposals were put forward [51], for the deadline in 1998. Of those four were selected as IMT-2000 RTTs in January 2000. In order to harmonize these RTTs to allow for a common world standard, two groups were subsequently established;

- 3GPP-third generation partnership project looking at harmonizing the European Telecommunications Standards Institute (ETSI), Association of Radio Industries and Business (ARIB), Telecommunication Technology Committee (TTC), Telecommunication Technology Association of Korea (TTA), and T1 proposals from Wideband Code Division Multiple Access (W-CDMA).
- 3GPP2 –third generation partnership project2 looking at Telecommunication Industry Association (TIA) and TTA.

These groups were also looking at core network aspects including the transition to all-IP based networks. The family of IMT-2000 RTTs is briefly described below:

- W-CDMA Frequency Division Duplex (FDD) proposed jointly by ETSI and ARIB (Japan). Uses wide band carriers in new or existing spectrum.
- W-CDMA Time Division Duplex (TDD) proposed jointly by ETSI and ARIB (Japan). Employs the use of Time Division CDMA (TD-CDMA) in the unpaired spectrum bands.

- CDMA2000 proposed by TIA as an evolution path for IS-95 networks. Uses existing (IS-95) and new (wider) carriers in existing or new spectrum.
- EDGE proposed by both ETSI/UWC as an evolutionary path that provides convergence of the air interface for GSM and IS-136 standards. Allows an operator without new 3G spectrum to meet IMT-2000 requirements. Includes a wideband component.

Figure 4-1 [4] shows the roadmap for the evolution of cellular networks from first generation to the third generation systems, in terms of the RTTs.

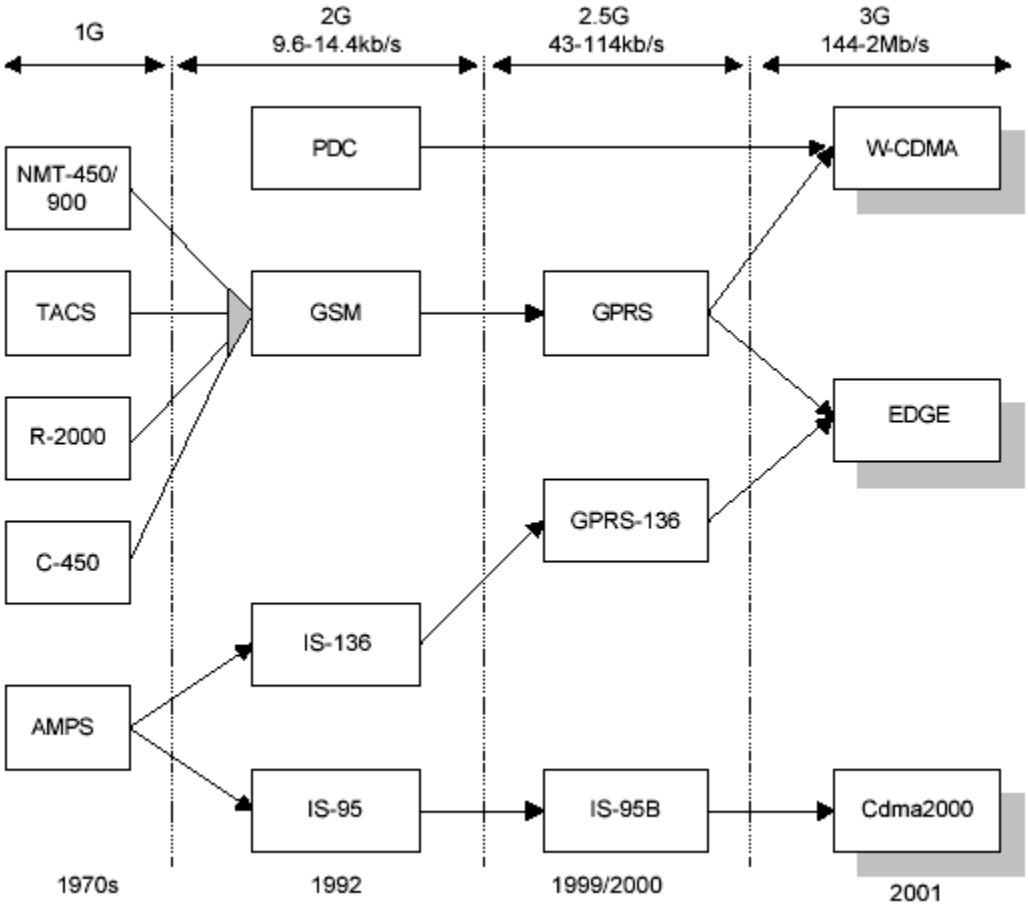

Figure 4-1: Evolution of the 3G standards

In addition to specifying RTTs, IMT-2000 also has a study group dedicated to specifying core network architectures in support of 3G services. This work is being carried out by ITU-T SG5 (the partnership projects are also concentrating on this issue). It is envisaged that IMT-2000 networks will evolve from existing wireless networks and in particular the interim network architectures being deployed at present. Two important aspects in IMT-2000 are the ability to allow seamless roaming among the different RTTs and the concept of a ‗virtual home environment' (VHE), in which a users subscribed services are available as they roam across networks [52][53]. To facilitate global roaming and VHE aspects, a necessary requirement is the interworking of mobility application protocols used in current second generation systems i.e. ANSI-41, GSM-MAP and PDC-MAP. Core network architecture standards work is a continuing part of the IMT-2000 standards. The general concept of network architecture evolution is expected to distinguish between a generic radio access network (GRAN) component and a core network component. Several research projects have been conducted to investigate the interoperation of RTTs attached to a common core network. One such project is Radio Access Independent Broadband on Wireless (RAINBOW) project, part of Europe's UMTS efforts [52]. The idea behind the separation of RAN and core network is partly to realize global roaming and partly to facilitate a paradigm shift in telecommunication network design. 2G systems are limited by their vertical architecture, with all system aspects defined in a rigid, top-to-bottom manner, from bearer services to applications. Any network enhancement requires end-to-end testing. In contrast, future networks will need to be modular and flexible with respect to network changes. Support for plug-and-play style operation will become a requirement. The goal is to develop IP based core networks that will evolve from being parallel networks supporting data services, to being the single core network with integrated IP-based signaling protocols. The present direction sees GSM adopt GPRS (driven by 3GPP) for a core network architecture and IS-95 networks adopt Mobile IP (driven by

3GPP2) based core network architectures. Figure 4-2 shows the representation of 3G network architecture [54].

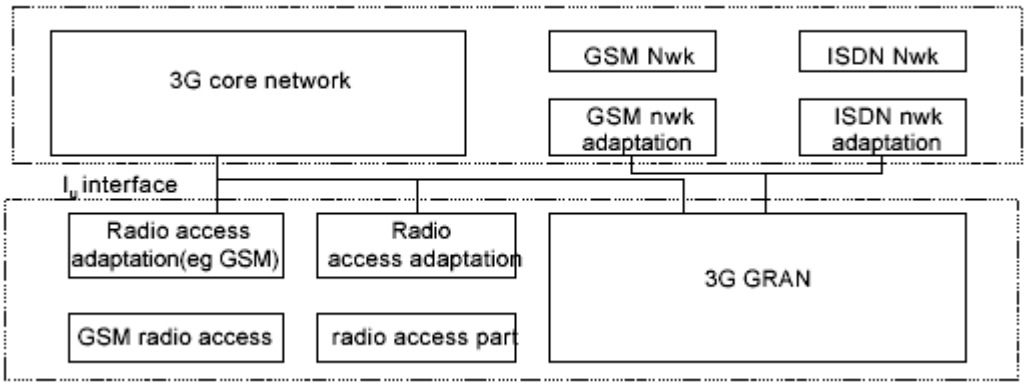

Figure 4-2: 3G system architecture

4.2 WiMAX wireless network

IEEE specifications of 802.16 commonly known as Worldwide Interoperability for Microwave Access is an established technology for fixed wireless access. The first standard for 802.16 appeared in October 2001, which defines the air interface and Media Access Control (MAC) protocol for a wireless metropolitan area [55]. The actual idea of 802.16 begun at a meeting called by US National Wireless Electronics Systems Testbed (N-WEST) of the U.S. National Institute of Standards and Technology in August 1998. This led to the formation of 802.16 Working Group. 802.16a was completed in January 2001 and the new 802.16d was approved in June 2004 which is now named 802.16-2005. Most recent version is the 802.16e [56] to support mobility up to speeds of 70-80 miles per hour. The issue of true mobility[8] is of particular importance in WiMAX because its promised potential of seamless broadband connectivity which can only be effective if it is available while on the move. The MS scans all the downlink signals from the BS in order to get an unoccupied downlink channel from the BS. Generally there exist many BSs existing with individual service area

[8] By true mobility it is meant continuous mobility at vehicular speeds

overlapping with each other. These BS works at different frequencies and the exact number of frequencies being used generally differs from country to country depending upon the frequency spectrum available for allocation by the government. The MS scans the wireless network to establish initial network access and periodically for selecting suitable Target BS (TBS) for a handover to maintain connectivity while moving from the range of one BS to that of other. The latest WiMAX standard [55] recommends network assisted handover as used in cellular mobile systems, where the BS currently serving the MS provides information about the other Base stations in the network. The serving BS periodically broadcasts information about other BSs to the MS. The IEEE 802.16e standard supports temporarily suspending the uplink and downlink communication between the MS and BS in order to allow the MS to perform scanning for neighboring BSs. While communications are suspended, the data streams must be buffered on either side. Any improvement on the time it takes for the MS to complete its scanning operation improves the performance of the communications, i.e. reduces delay. However large latency of handover will mean visible disruption for services like VoIP and other real-time applications. Hence reducing the latency resulting from network scans and handover will mean improved usability and acceptability for IEEE 802.16e networks. Details of Handover techniques in WiMAX are discussed, then a thorough analysis of how this it is implemented in existing wireless technologies is provided. Thereafter novel approaches for reducing handover latency are proposed. Finally, simulations demonstrating the performance of the proposed techniques are explained. New clause 6.3.22 in [56] defines the handover process in WiMAX.

4.3 Long Term Evolution

LTE (Long Term Evolution) represents the transition of UMTS variant of 3G also developed through the 3rd Generation Partnership Project (3GPP). 3GPP Release 8 defines the first stage

of 4G mobile communications technology, including an all-IP flat networking architecture. LTE provides downlink peak rates of at least 100 Mbit/s, 50 Mbit/s in the uplink and RAN (Radio Access Network) round-trip times of less than 10 ms. LTE supports flexible carrier bandwidths, from 1.4 MHz up to 20 MHz as well as both FDD (Frequency Division Duplexing) and TDD (Time Division Duplex). The goals for LTE include improving spectral efficiency, lowering costs, improving services, making use of new spectrum and reformed spectrum opportunities, and better integration with other open standards. The architecture that will result from this work is called EPS (Evolved Packet System) and comprises E-UTRAN (Evolved UTRAN) on the access side and EPC (Evolved Packet Core) on the core side. EPC is also known as SAE (System Architecture Evolution) and E-UTRAN is also known as LTE.

The main advantages with LTE are high throughput, low latency, plug and play, FDD and TDD in the same platform, improved end-user experience and simple architecture resulting in low operating expenditures. LTE will also support seamless connection to existing networks such as GSM, CDMA and WCDMA and while requiring partial equipment upgrade is becoming the defacto 4G standard although there are differences in frequency bands of operation in the different regions to date.

In parallel with the LTE radio access, packet core networks are also evolving to the flat SAE architecture. This new architecture is designed to optimize network performance, improve cost-efficiency and facilitate the uptake of mass-market IP based services. There are only two nodes in the SAE architecture user plane: the LTE base station (eNodeB) and the SAE Gateway, as shown in Figure 4. The LTE base stations are connected to the Core Network using the Core Network–RAN interface, S1. This flat architecture reduces the number of involved nodes in the connections. Existing 3GPP (GSM and WCDMA/HSPA) and 3GPP2

(CDMA2000 1xRTT, EV-DO) systems are integrated to the evolved system through standardized interfaces providing optimized mobility with LTE. For 3GPP systems this means a signaling interface between the SGSN and the evolved core network and for 3GPP2 a signaling interface between CDMA RAN and evolved core network. Such integration will support both dual and single radio handover, allowing for flexible migration to LTE. Control signaling for mobility is handled by the Mobility Management Entity (MME) node, separate from the Gateway. This facilitates optimized network deployments and enables fully flexible capacity scaling.

4.4 Mobile IP integration

Mobile IP support is implemented in the North American variant of 3G (i.e. CDMA2000) network by housing the FA function in the PDSN. HAs are likely to exist in various private networks. In addition to the FA entity, an AAA infrastructure is expected to be used for transport of all Mobile IP related signaling messages between a visited and home domain. While the ANSI-41 signaling network verifies the MS, the AAA infrastructure serves to verify the credentials of the user and reconciliation of charges for data services. Figure 4-3 illustrates the deployment of Mobile IP in a CDMA2000 network [57].

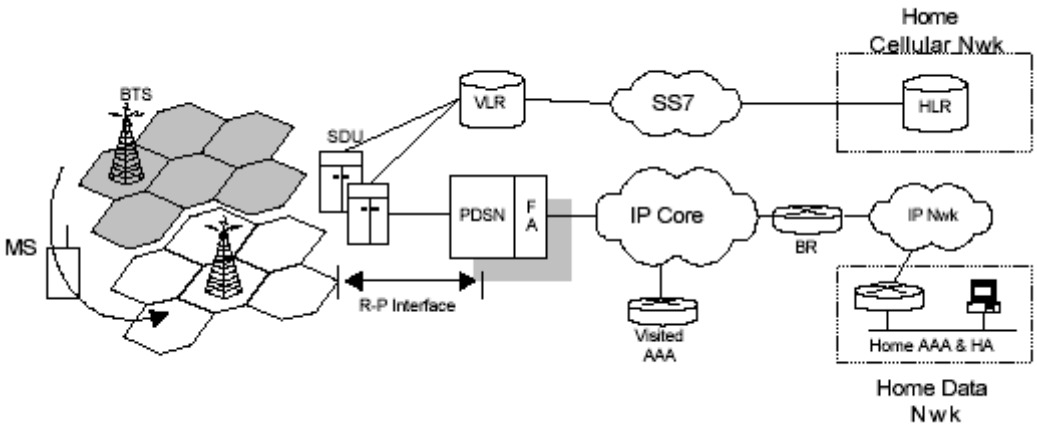

Figure 4-3: Mobility support in CDMA2000 network

The MS must establish link layer connectivity with the PDSN first. This is achieved by initializing a PPP session. During the PPP - IPCP phase, the MS must use the Mobile IPv4 option, in order to successfully propose its home address as the IP address to avoid being assigned a new one. Once L2 is configured, the PDSN will send an agent advertisement containing the care-of address that should be registered by the MS with its HA. On completion of the Mobile IP registration process, the MS receives packets by way of HA interception and tunneling to the PDSN/FA. As long as the MS continues to be served by the same PDSN, no further registration requests are required other than renewal requests. PDSN anchoring is a current topic of research within 3GPP2 as is mobility management across the R-P interface.

Chapter 5: Design, analysis and simulation of distributed agent mobility management platform and protocol (D-MIP)

5.1 Introduction

This chapter presents a new IP-based mobility management platform and protocol called Distributed Agent Mobile IP (D-MIP) which aims to provide a new mobility management platform with stable signalling overheads as well as minimizing packet delay while the mobile node is not in its home network. This platform will be more likely suitable for deployment in the current cellular infrastructure with moderate handoff frequency and mobile nodes may not have a fixed home network or a permanent IP address. The ideal application for D-MIP is for the client server applications based internet users who always initiate a connection from the mobile node but require a fast handoff and start-up time.

5.2 Architecture

D-MIP adopts a Distributed Agent approach to handle mobility management of an IP based device inside a cellular network. Also, D-MIP is designed to minimize the need for extra network equipment and leverage the concept and compatibility of MIP for macro mobility management. D-MIP introduces two network entities to handle mobility management. They are the D-MIP Name Server and Distributed Agent. The combination of these two network

entities can efficiently provide mobility management and location management of a MN inside an IP based network.

A D-MIP Name Server (NS) connected to the internet can act as a point of reference for locating a specific MN using a name based approach for this particular domain. Inside D-MIP, a MN does not have a permanent IP address, but each MN is assigned a unique name by the operator, which is maintained inside the D-MIP NS. When any CN needs to initiate communication to an MN, the CN needs to query the NS for the current IP address of the MN and then initiate the connection. The IP address of any MN registered for packet data service is updated to a NS when the IP address is allocated or changed.

A D-MIP Distributed Agent (DA) is located inside each subnet to handle mobility management for all MNs inside a cellular network. The DA is similar to a Home/Foreign Agent in MIP. However, the sequence and protocols of D-MIP are enhanced with the Distributed Agent approach. Figure 5-1 shows the architecture of D-MIP.

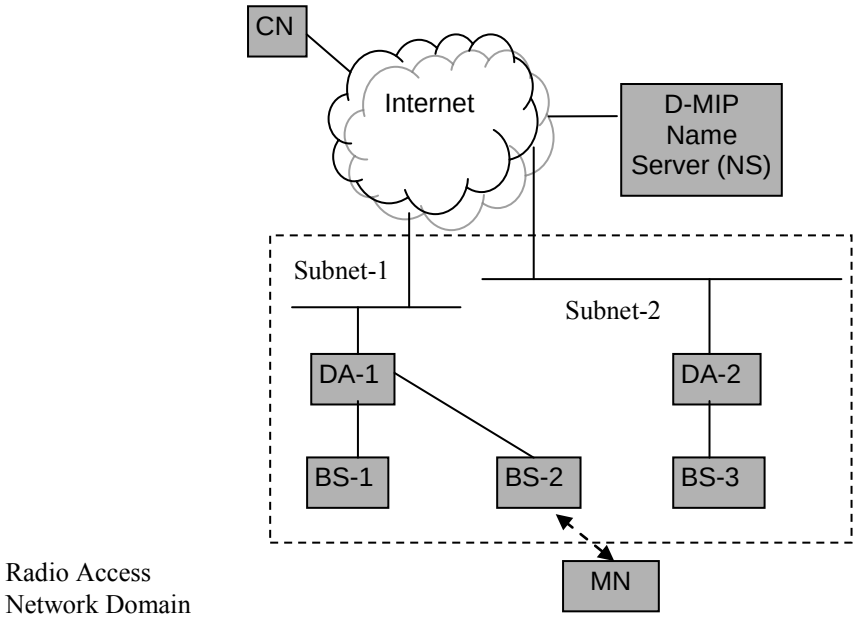

Figure 5-1: Architecture of D-MIP

A Distributed Agent (DA) is located inside each subnet of a radio access network and acts as a mobility agent for all the MNs inside the subnet. The DA performs the following functions:

- Controls all the packet routing and filtering for the subnet to which it belongs.
- Maintains IP address pool for usage from the MN for its particular subnet.
- Attracts and forwards IP packets to another DA.
- Performs registration and de-registration of MN. After registration, an IP address is allocated to the MN. This function is similar to the Dynamic Host Configuration Protocol (DHCP) protocol.

Coordinates and communicates with other DA‘s to keep track of movement of a MN in a subnet level.

5.3 D-MIP platform characteristics and design goals

The design goal of a D-MIP is to enable a simple, generic and cost effective way of deploying the mobility management mode for existing cellular networks. The design goals of D-MIP and its characteristics are described as following:

- To minimize extra network elements for mobility management.
- The design of D-MIP should enable deployment into different radio networks regardless of access technologies such as CDMA, UMTS, or WLAN.
- To minimize the cost of operating and ownership of D-MIP platform. D-MIP minimizes extra network elements. Also, D-MIP does not need a fixed permanent address for each MN, which allows for simple management of the MN without the need for the operator to permanently reserve IP addresses for each MN. This approach can better utilize the IP address allocation as needed.

5.4 D-MIP protocol operation

The D-MIP protocol operation description is divided into the following operations, which are:

1. MN initiates a new packet data session
2. CN initiates a new packet data connection to a MN
3. MN handoff to a different subnet
4. MN subsequent handoff to another subnet
5. MN closed the packet data session.

After generalizing the above scenarios, most possible conceptual MN movements inside the cellular network will be covered. The detailed protocol operations are described in the following sections with a simple configuration of D-MIP. Figure 5-2 shows a simple configuration of D-MIP in with three Subnet (Subnet-1 to Subnet-3) and three base stations (BS-1, BS-2 and BS-3) configuration.

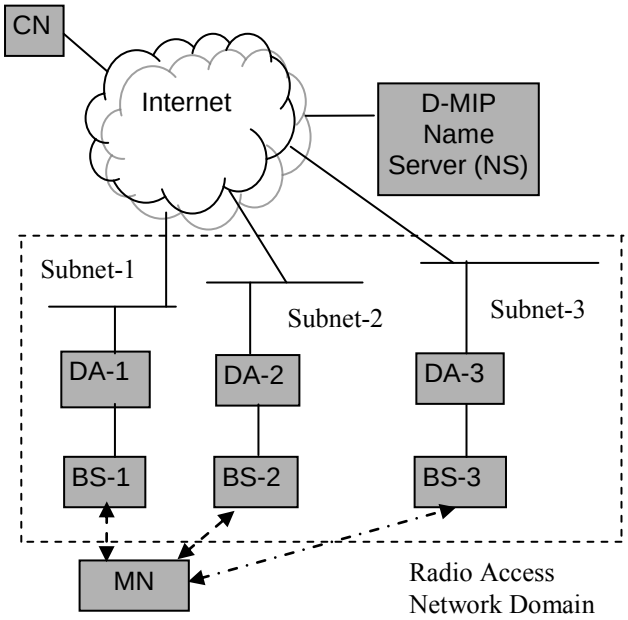

Figure 5-2: A simple configuration for a D-MIP

5.4.1 MN initiates a new packet data session

Inside D-MIP, when MN initiates a packet data session inside a subnet called Subnet-1, MN determines the need for a request for a new IP address. If MN did not have a topology correct IP address stored inside an IP address list or an MN simply did not have any IP address allocated, the MN requests a new IP address from the specific DA of the subnet to which the MN currently is attached. This is DA-1. The MN will need to send a Registration Message to DA-1 and specify whether a new IP address is requested. Then, DA returns a Registration Reply to MN. MN can use the allocated IP address to send and receive any IP packet from Subnet-1.

5.4.2 CN initiates a new packet data connection to a MN

When a CN initiates an IP connection to MN, CN needs to query the D-MIP NS using MN‘s domain based name, which is specified by the operator. If an MN already opened a packet data session, NS returns the current MN IP address to CN. If MN is not in a packet data session, which means the MN was not allocated an IP address, NS will send a layer 2 paging request to MN and request MN to initiate a packet data session. After MN successfully initiates the packet data session, NS sends CN the current MN IP address. Then, CN can initiate any IP connection to MN.

5.4.3 MN handoff to another subnet

MN detects handoff either using a lower network layer trigger or using the Agent Advertisement protocol from current MIP implementation. When MN detects handoff, the following steps are performed:

1. MN registers with the new subnet DA named as DA-2 to acquire a new IP address and added to the current list of IP addresses. Inside the registration message, a new address field contains MN‘s IP address at Subnet-1.
2. DA-2 sends an update message with MN‘s newly allocated IP address at Subnet-2 to DA-1 and NS in order to register the new IP address of MN. Then, DA-1 forwards IP packets to the new IP address of MN at Subnet-2.
3. MN now has a topology correct IP addresses inside Subnet-2 for all new connections.
4. Similar to MIP HA and FA, DA-1 attracts IP packets from Subnet-1 and using IP tunnel to redirect packets to DA-2 and deliver to MN subsequently.

5.4.4 MN subsequent handoff to another subnet

When MN handoff occurs to another new subnet named Subnet-3 using the same detection methods, MN uses the following handoff sequence:

1. MN sent a registration request to the DA-3 with the IP address at Subnet-2.
2. DA-3 sends an update message to DA-2 for the new address of the MN.
3. When DA-2 receives the update message, it will start routing packets to the new IP address of MN at Subnet-3. Also, DA-2 forwards the update message to its previous DA which is DA-1 since DA-2 already having a complete list of IP addresses of MN. This is a new mechanism compared with MIP. MIP always sends registration requests and updates location information back to the Home Agent, which could be very far away and hence introduces packet delay for all the data transfers until MN‘s moves back to its home network. Using this distributed approach, DA-2, which is the neighbour subnet of subnet-3 handle all the subsequent new IP address updates to all the previous Subnet Agents. Therefore, handoff signaling overheads between MN and DA-1 is minimized and packet delay is reduced since data packets are not always

forwarded by the agent in MN's home network. Data packets are only forwarded to MN from the last agent where MN initiates the connection which mostly are to a neighbour subnet. A brief overview of the sequence of events are shown in Figure 5-3.

4. These sequence of events are repeated when MN is handover to any subsequent DAs.

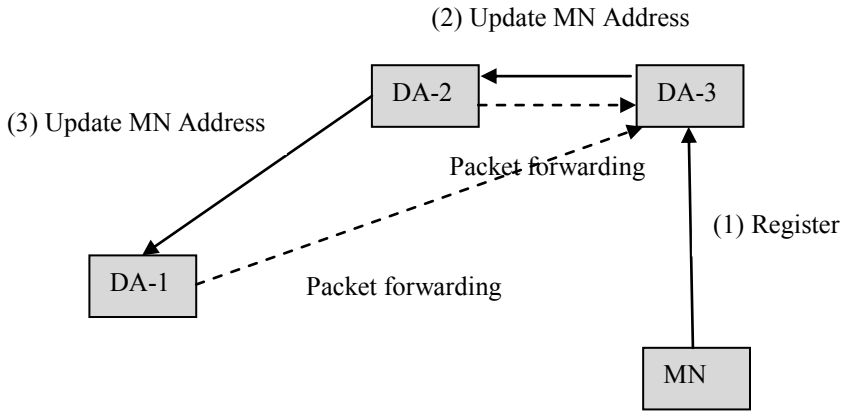

Figure 5-3: Handoff sequence of MN when register to DA-3

5.4.5 MN closes packet data session

MN can close the packet data session by sending a deregistration message to the corresponding DA, such as DA-3 if MN is inside Subnet-3. Then, DA updates other DAs for deregistration and subsequently releases all the IP addresses and packet tunnelling. DA also updates NS for release of all IP addresses.

5.5 Benefits of D-MIP

The benefits of D-MIP are summarized as follows:

1. IP addresses are dynamically assigned. MN does not need to have a permanent assigned address, which is expensive to maintain for network operators.

2. D-MIP can be implemented either inside a global network or a private network. Many cellular operators have a private IP based networks, which can implement D-MIP without major reconstruction of the IP network.
3. The concept of a Home Agent is eliminated. A Distributed Agent (DA) will coordinate all the mobility management needs using the distributed approach. Also, when a MN releases all sessions inside a subnet, the CoA for the MN are released. This can dynamically utilize the IP addresses required by the MN and release any unused IP addresses.
4. D-MIP reduces the path length of packet forwarding by dynamically assigning the functionality of the Home agent to any agent in the network called DA. This mechanism reduces the packet forwarding path compared with MIP, which always uses the Home Agent to forward all the packets that may be located very far away.
5. D-MIP can be deployed in different kinds of Radio Access Network, as D-MIP does not need any special router or equipment deployed inside the cellular network. Compared with HAWAII and Cellular IP, D-MIP provides a cost effective solution for implementing mobility management.

5.6 Simulation of D-MIP

Simulations are conducted to investigate two major performance issues on a mobile network; i.e. packet loss during handoffs and packet delay while the MN is roaming to a foreign network. A simulation model is constructed, which consists of a home network, two foreign networks and a wireline network to which CN is connected to. Figure 5-4 [9] shows the network model for the simulation. CN and W1 are defined as wireline nodes. BS1, BS2 and

[9] Measurement of delay in Layer 2 and Layer 3 was not available

BS3 are subnets that are with a wireless base-station. In order to investigate how the overall mobility management performance is affected by the home network's link delay time, the link delay time between the MN's home network link L1 is varied from 200 ms to 1000 ms. This simulates the overall packet delay from CN to MN and packet loss during handoff using the ns2 [58] simulator. The same sets of simulations are conducted with MIP for comparison.

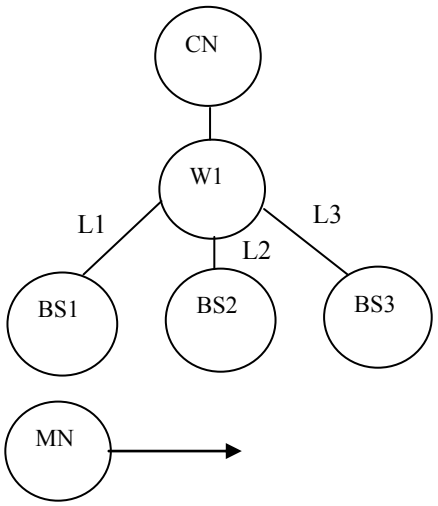

Figure 5-4: Simulation model

Figure 5-5 shows simulation results of packets lost during MN handoff from BS1 to BS2 and from BS2 to BS3. When MN handoff occurs from BS1 to BS2, the packet lost is similar to the MIP and was indicated at an identical line in Figure 5-5. This is due to the link delay introduced by MN's home network. However, when MN handoff occurs from BS2 to BS3, DMIP can reduce packet delay of the link by using the DA in BS2 for packet tunnelling. This in turn reduces the packet lost during handoff. This means MN can handoff to a different BS more seamlessly with reduced packet loss.

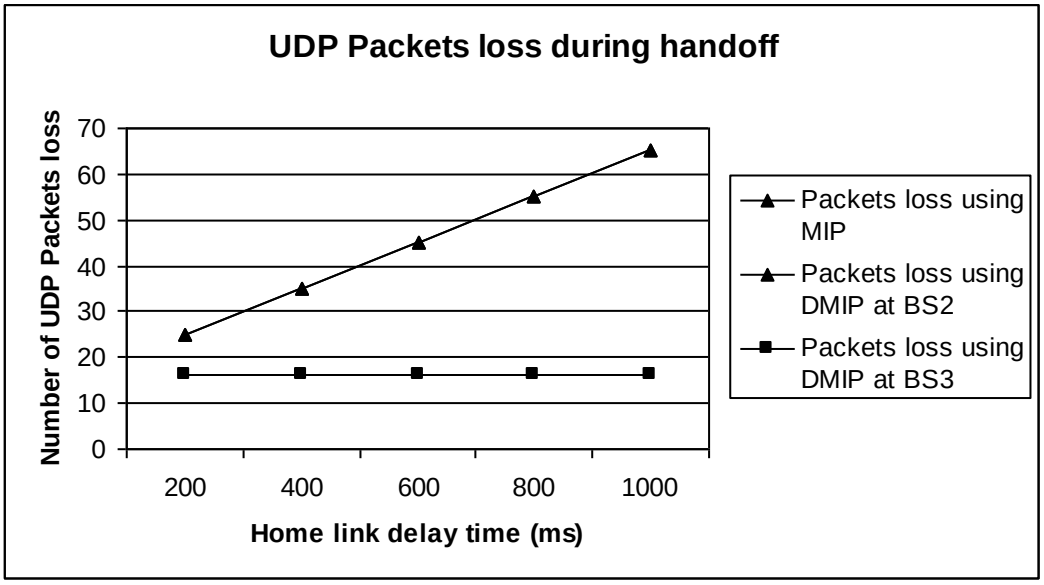

Figure 5-5: UDP Packet loss during handoff

Figure 5-6 shows simulation results of packet delay time from CN to MN while MN is roaming to a foreign network; i.e. BS2 and BS3. When compared with MIP, MN has similar packet delay time for the first handoff from BS1 to BS2 and was indicated at the same line in Figure 5-6. However, when MN handoff subsequently occurs to more distance base station, DMIP maintains the overall packet delay to a constant low value. However, MIP still maintains the long packet delay time due to MN always being anchored at the original home agent.

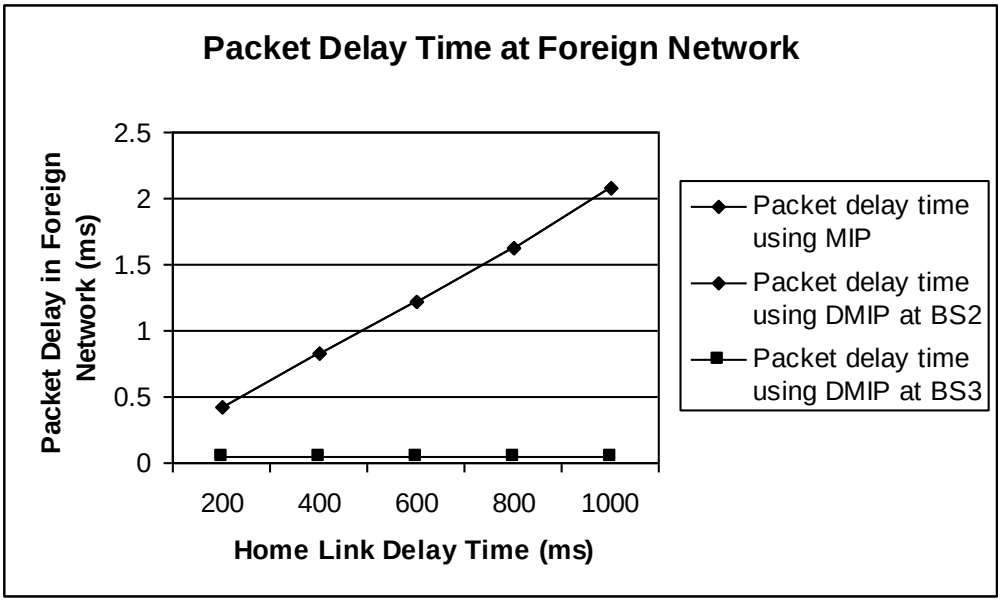

Figure 5-6: Packet delay time at foreign network

Chapter 6: Analysis and simulation of Dynamic Home Agent Anchoring (DHAA) scheme using OPNET®

6.1 Introduction

This chapter presents the design and simulation of an IP-based mobility management scheme called Dynamic Home Agent Anchoring (DHAA) for mobility management of mobile devices in 4G wireless network infrastructure that have a combination of IPv6 and IPv4 network. DHAA scheme provides an alternative mobility management solution for 4G cellular network infrastructure which needs to provide seamless mobility management support for mobile devices which communicate with heterogeneous types of cellular networks at the same time. DHAA is developed based on the D-MIP in Chapter 5 and extended into a more general conceptual view. A reference architecture is developed and a different simulator was used to carried out the simulation.

This rest of this Chapter is organized so that Section 6.2 describes the detail design of DHAA scheme, Section 6.3 presents the simulation results of DHAA using OPNET® Modeller [59] network simulator.

6.2 Dynamic Home Agent Anchoring (DHAA) scheme

6.2.1 Terminology

Mobile Node (MN) is an Internet Protocol enabled Mobile Node which capable to communicate using IP protocol. However, the lower layer protocols may various which depend on the Radio Access Network (RAN) that MN attached to. Examples of RAN technology are CDMA, UMTS and Wimax.

Corresponding Node (CN) is a fixed host resides on a fixed location and attached to the internet and communicate with MN using the IP protocol.

Base Station (BS) is a wireless Base Station which transmits and receive radio frequency using a specific radio technologies to a MN.

Subnet (SN) is an IP network that has unique network address as defined by the IP protocol. Mobile node (MN) is attached to different Subnet as it moves to different location.

Dynamic Agent (DA) resides at every Subnet which implements the DHAA scheme.

Temporary IP Address ($IP_{d\text{-}SN\#}$) is an IP address assigned by a DA which is topology correct inside a specific Subnet.

Permanent IP Address (IP_{perm}) is a permanent IP address which a MN assigned. This address will be a unique identifier for MN inside the internet. Other existing terminology and its functionalities are in [18].

6.2.2 Reference architecture

DHAA is a new approach which allows a Mobile Node (MN) can be anchored to any Home Agent (HA) at any time dynamically. To reduce packet delivery delay from the triangular routing in a Mobile IP network, HA which has the shortest distance to MN‘s anchoring network can be selected and anchored. Also, MN does not mandatory requires a permanent

IP address or a permanent HA anchored for mobility management functions such as packet interception and packet tunnelling.

DHAA scheme requires an enhancement of Mobile IP Home Agent (HA) and Foreign Agent (FA) by combining the HA and FA functionalities into single entity named Dynamic Agent (DA). DA will act as both HA and FA functionalities as defined in Mobile IP. A network which implemented DHAA scheme are named as DHAA enabled network. Figure 6-1 shows a simple architecture of DHAA enabled network in a cellular based network environment. There are 3 Base Stations defined which are BS-1, BS-2 and BS-3. Mobile Node (MN) is the normal cellular device which communicates with a BS using a specific radio technology such as CDMA or UMTS. DA-1 and DA-2 are Dynamic Agent which implements the Mobile IP HA and FA functionalities with the new DHAA extension. Corresponding Node (CN) is an IP based host which the MN communicated with using the IP network protocol. Subnet-1, Subnet-2 and Subnet-3 are different subnet inside an IP based network.

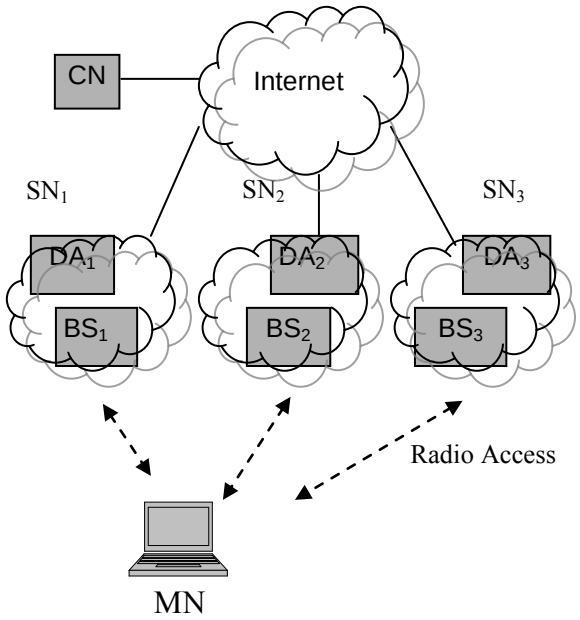

Figure 6-1: A DHAA enabled network

6.2.3 Design goals for a Dynamic Home Agent scheme

DHAA are designed with the following goals which allows a robust and generic scheme for implementation in various kinds of radio networks. They are:

- To enable deployment into different radio networks regardless of access technologies such as CDMA, UMTS, or Wimax.
- Reduce the packet delay of triangular routing by allowing Home Agent can be reassigned as a MN moved to a distance location.
- Compatible with Mobile IP protocol for macro mobility management.
- Reduce packet lost during handoff from one base station to another.

6.2.4 DHAA scheme overviews

DHAA scheme enhances the Mobile IP protocol by allowing reassignment of Home Agent on-the-fly during MN moving from one Subnet to another. As MN always handoff to its neighbours cells from the concept of cellular network, a dynamically assigned Home Agent can reduce packet forwarding delay compared with static Home Agent assignment in Mobile IP because the Home Agent could be long distance away from the current location of MN.

Dynamic Agent is assigned to MN when MN moved into a subnet which MN do not have a topology correct IP address. When MN moves into a new Subnet, MN acquires a temporary IP address from the Agent Advertisement information that broadcasts by every Dynamic Agent. The Temporary IP address can only be used to initiate IP based communication from MN. For IP communication initiate from CN, the permanent IP address of MN will be used.

When MN moves away from a subnet, the dynamic agent will release its temporary IP address once received a de-registration from MN. A detailed description of DHAA scheme in protocol level is in [60]. Details descriptions of DHAA scheme are illustrated by the following four scenarios.

6.2.4.1 MN permanent address routing for macro mobility

MN can be assigned a permanent IP address (IP_{perm}) if MN needs any IP communication initiates from CN. In this case, a DA which resides at the Home network of MN and topology correct to IP_{perm} will become MN's Home Agent. The operation of this case will be identical to Mobile IP protocol besides DA replaces the functionalities of Home Agent and Foreign Agent. MN registers with DA_0 and DA_0 will stop all packet forwarding to MN's other dynamics IP address. However, when MN moves from one subnet to another subnet, DHAA scheme can still be used to reduce packet delay through the usage of a Temporary IP address ($IP_{d\text{-}SN}$) that describes in the following sections.

6.2.4.2 MN start packet data session from a subnet

When MN starts a packet data session in SN_1, MN registers to DA1 which is the Mobility Management Agent inside the same Subnet. There are 2 outcomes from this registration process:

1. If MN assigned a permanent IP address and MN resides in its home network said S_1, MN does not need a new IP address because it already has a topology correct IP address which is IP_{perm}. DA_1 return registration successful to MN and stop all tunnelling for MN because MN is back to its home network.
2. If MN resides in a foreign network or MN do not have IP_{perm} assigned, MN needs a topology correct IP address in order to communicate inside the subnet. MN acquires a

new dynamics IP address (IP_{d-1}) from DA_1's Agent Advertisement information. Then, MN registers to DA_1. There are 2 outcomes from the registration process; i) If MN did not have IP_{perm} defined, DA_1 return registration successful only as shown in Figure 6-2 route S_1 and S_2. ii) If MN have IP_{perm} defined, DA_1 inform the DA in MN's Home network said DA_0 to update MN's location as shown in Figure 6-3 route S_1, S_3, S_4 and S_2.

After the registration process completed, MN can communicate with CN using a Temporary IP address IP_{s-1} as shown in Figure 6-3 route S_5 and S_6.

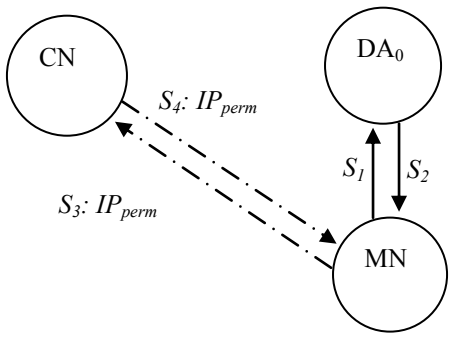

Figure 6-2: Permanent IP address routing of MN

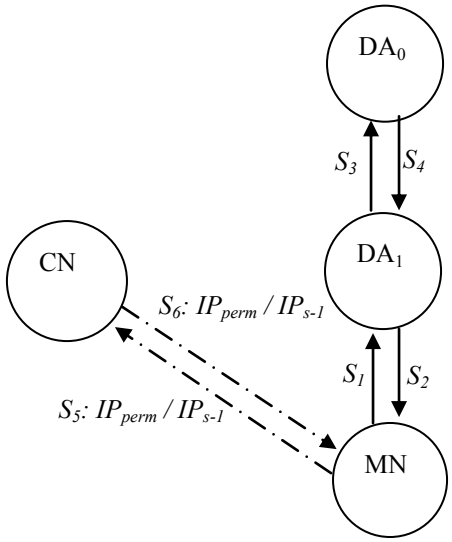

Figure 6-3: MN start packet data session inside a subnet

6.2.4.3 MN handoff to another subnet

When MN handoff from SN_1 to SN_2, MN detects handoff from the Agent Advertisement from DA_2 and identified its inside the coverage of another SN. Then, MN registers to DA_2 for its entry into the Subnet of DA_2. There are 2 outcomes from the register process:

1. If MN has a permanent IP address, DA_2 further inform MN's Home Agent which is DA_0 to start tunnelling packet to IP_{s-2}.
2. If MN does not have $IP_{perm,}$ DA_2 informs MN's previous Dynamic Agent which is DA_1 to start tunnelling packet to IP_{s-2} as shown in Figure 6-4 route S_1, S_3, S_4 and S_2.

After the registration process completed, the MN can communicate with the CN using a new Temporary IP address IP_{s-2} as shown in Figure 6-4 route S_5 and S_6. Also, packet data send to MN's IP_{perm} / IP_{s-1} will be forwarded by DA_1 to the MN as shown in Figure 6-4 route S_7 and S_8.

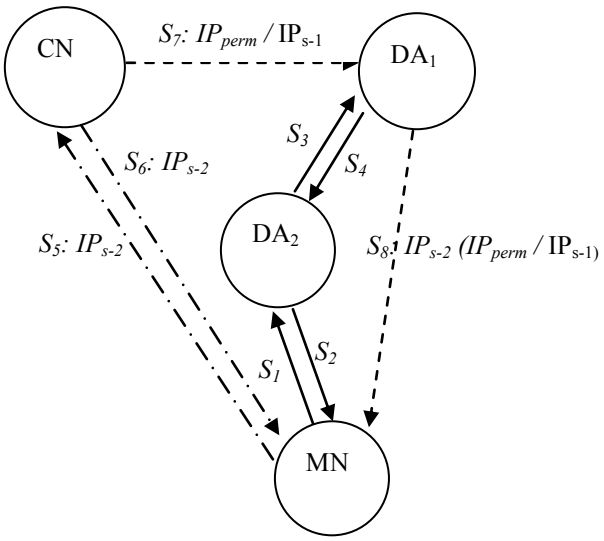

Figure 6-4: MN handoff to another subnet

6.2.4.4 MN further handoff to another subnet

When the MN further handoff from SN_2 to SN_3, the MN detects handoff from the Agent Advertisement information broadcasts from DA_3. Then, the MN registers to DA_3 for its present. There are 2 outcomes from the registration process:

8. If a MN has a permanent IP address, DA3 further registration information to MN's Home Agent which is DA1 and MN's previous Dynamic Agent which is DA2. Then, DA1 and DA2 can start tunnelling packet to IP_{S-3}.
8. If a MN does not have a permanent IP address (IP_{perm}) DA3 forwards registration information to the MN's previous Dynamic Agent only which is DA2. Then, DA2 can start tunnelling packet to IP_{S-3}. Also, DA2 will forwards registration information to DA1 and DA1 can start tunnelling packet to IP_{S-3} as shown in Figure 6-5 route S_{10}.

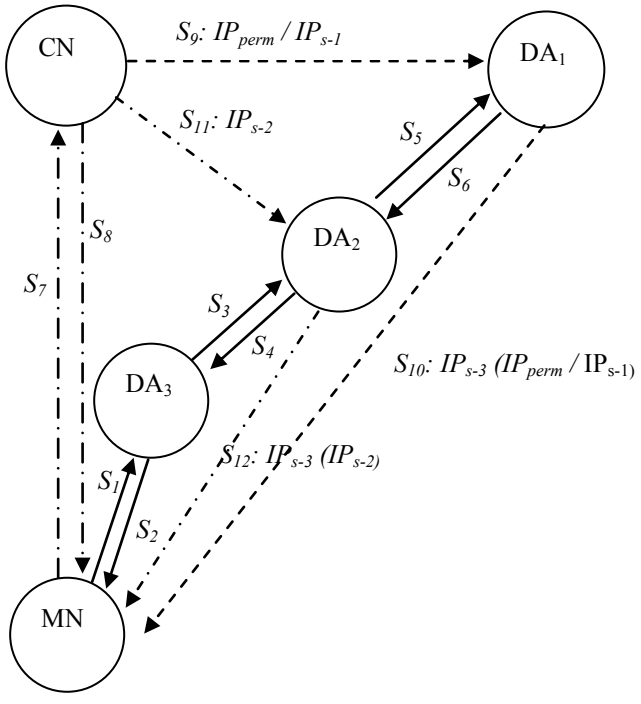

Figure 6-5: MN further handoff to another subnet

6.3 Simulation of DHAA scheme

Simulation is conducted using OPNET® to simulate the behaviour of the traffic for a mobile node moving from a home network to a far away foreign network. Two simulation models are constructed, which consists of MN resides in its home network and then moved to a foreign network. Figure 6-6 shows the network model for the simulation of MN resides in its home network. CN is Corresponding Node which is a wire line node sends IP traffic packets to MN continuously. HA and FA are home agent and foreign agents which locates on different subnets. When MN moves from its home network to foreign network, it registers with the FA and packets which destination are send to MN‘s permanent address are intercepted by HA and forwards to MN‘s Care Of Address via FA. This model shows in Figure 6-7. A constant stream of network traffic is configured which send from CN to MN for both network model. And the results are compared with using Mobile IP as well as using DHAA scheme.

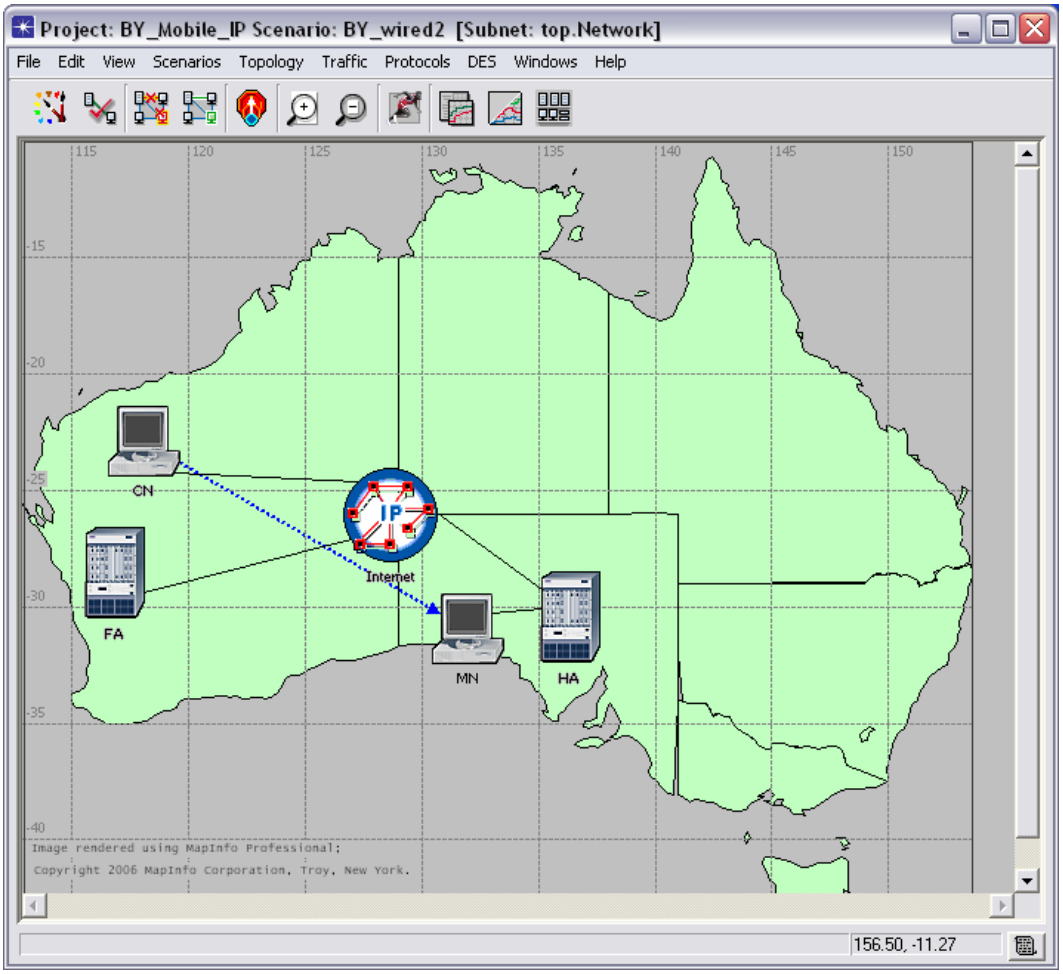

Figure 6-6: Simulation model of MN at home network

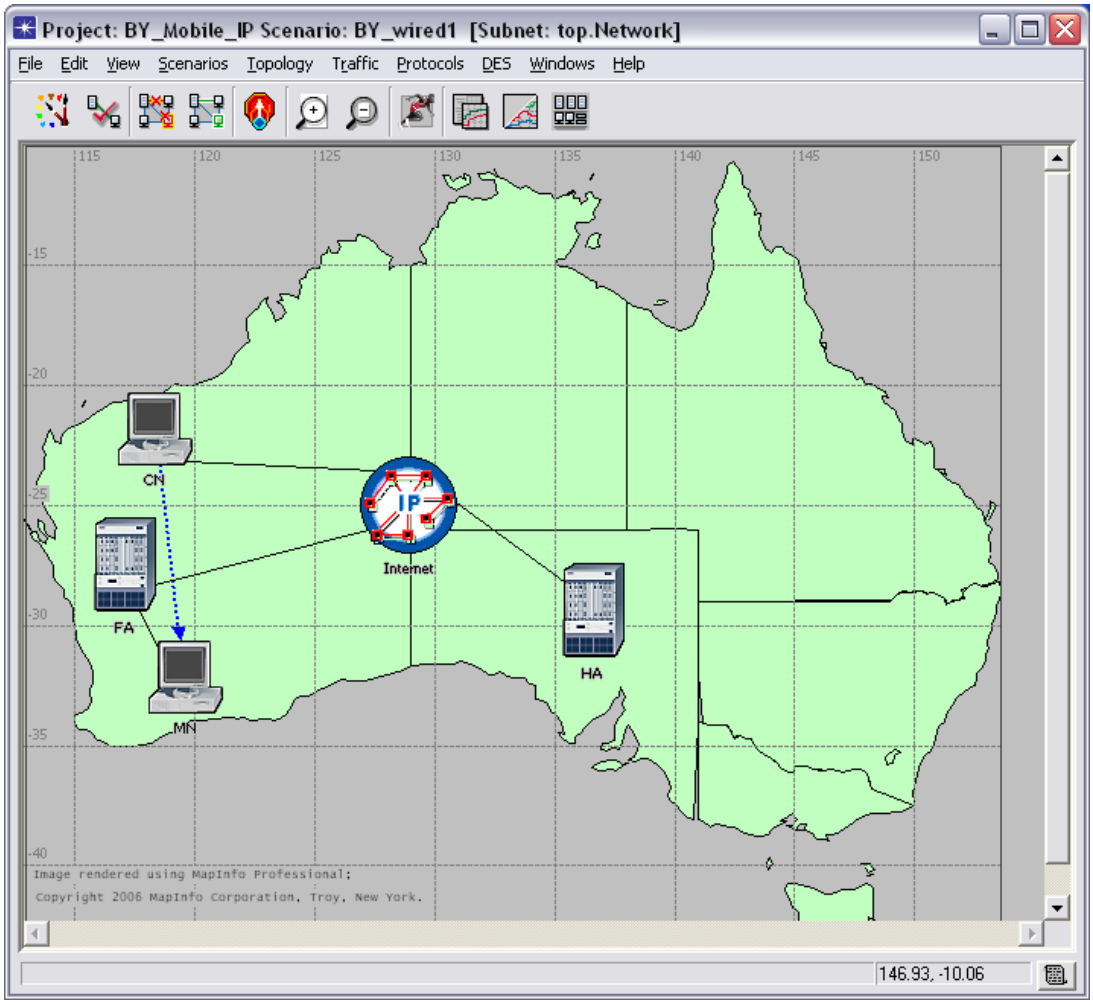

Figure 6-7: Simulation model of MN at foreign network

Figure 6-8 shows the IP traffic of the simulation for MN resides on its home network. The top graph shows traffic sent from CN to MN in bits per second in Y-axis. X-axis shows the simulation time elapsed. The traffic started at the middle of the simulation and its shows a shape rise of traffic once the source is activated. The second graph from top shows the traffic received by MN in bits per seconds. The X-axis also shows the simulation time elapsed. The received traffic is corresponded to the traffic sent with packet delay introduced from the link. The3rd graph shows Mobile IP tunnelled traffic sent from CN to MN and the 4th graphs shows the Mobile IP tunnelled traffic received at MN. Both graphs didn‘t show any tunnelled traffic

as MN is resided in its home network. Therefore, there is no tunnelled traffic involved and packet delay is same as a fixed node condition.

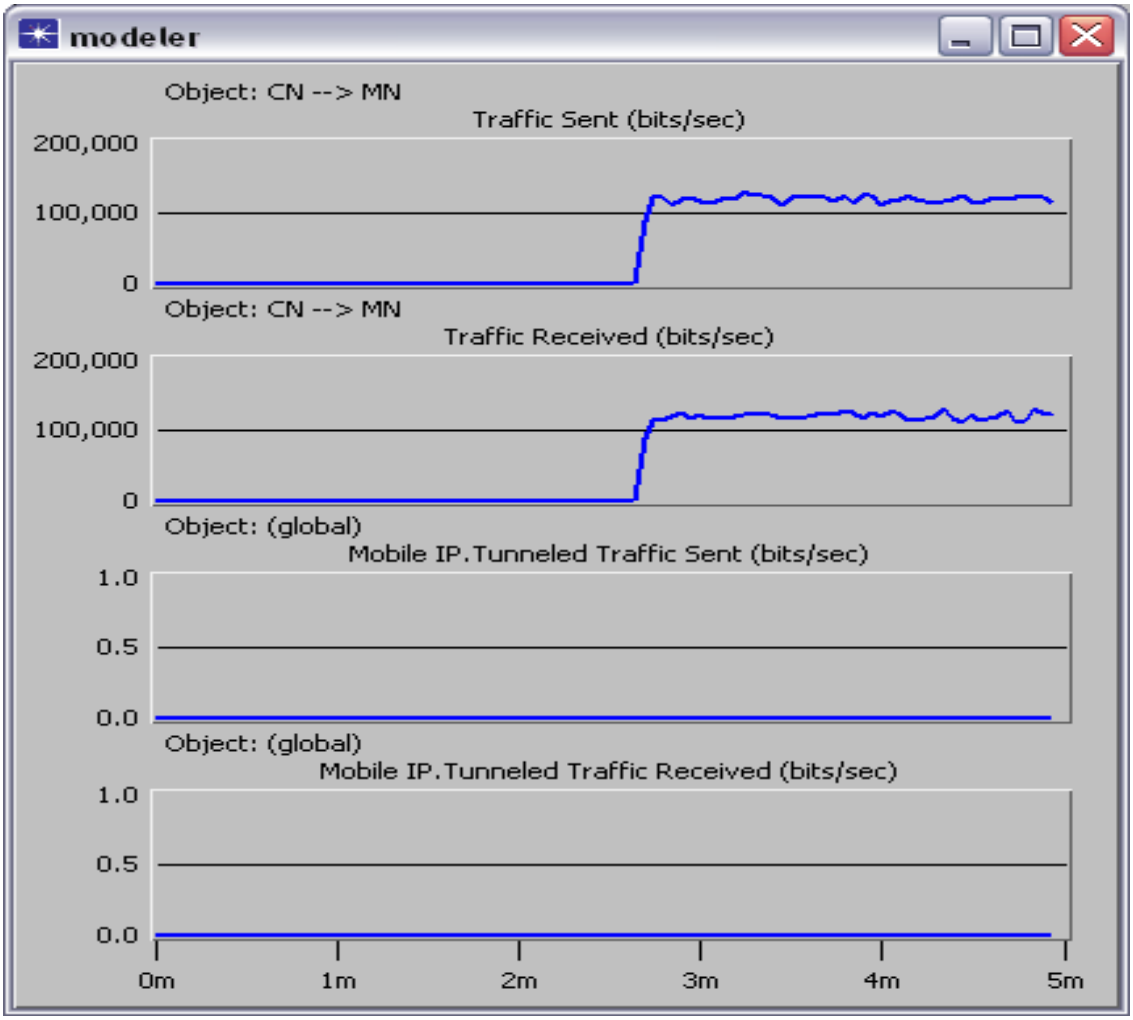

Figure 6-8: Network traffic of MN at home network

Figure 6-9 shows the IP traffic of the simulation for MN resides on the foreign network. The bottom two graphs show the amount of tunnelled traffic sent through HA. It can be seems that IP traffic forwarded from HA once MN moved to another subnet.

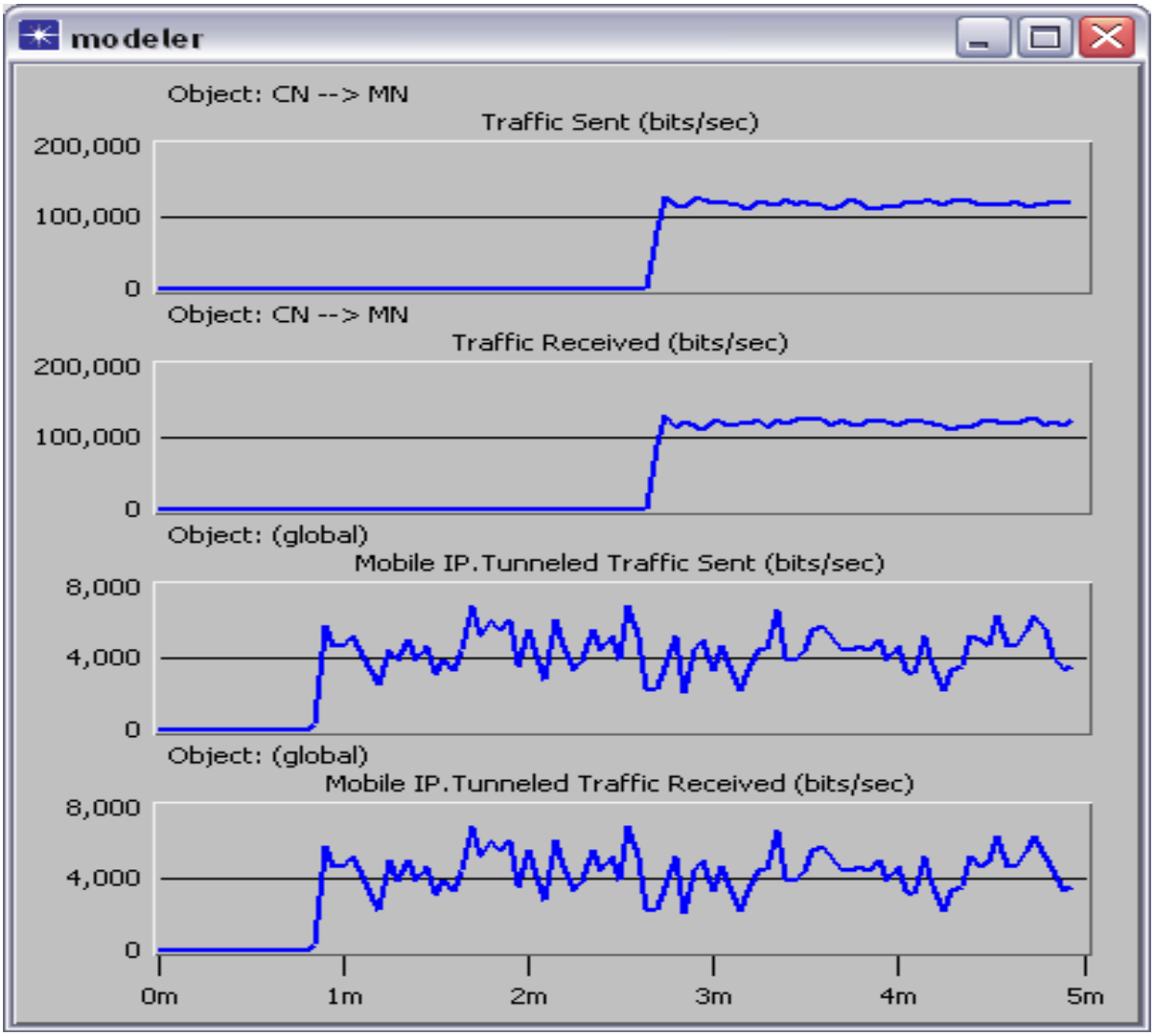

Figure 6-9: Network traffic of MN at foreign network

Figure 6-10 shows comparisons of the traffic tunnelled from the Home Agent using only Mobile IP and Mobile IP with DHAA scheme. The red and blue line shows the tunnelled traffic sent from the HA using MIP without DHAA. The green line shows the simulation results of using MIP with DHAA scheme. It is observed that, with the usage of dynamically assigned Home Agent, tunnelled traffic is reduced.

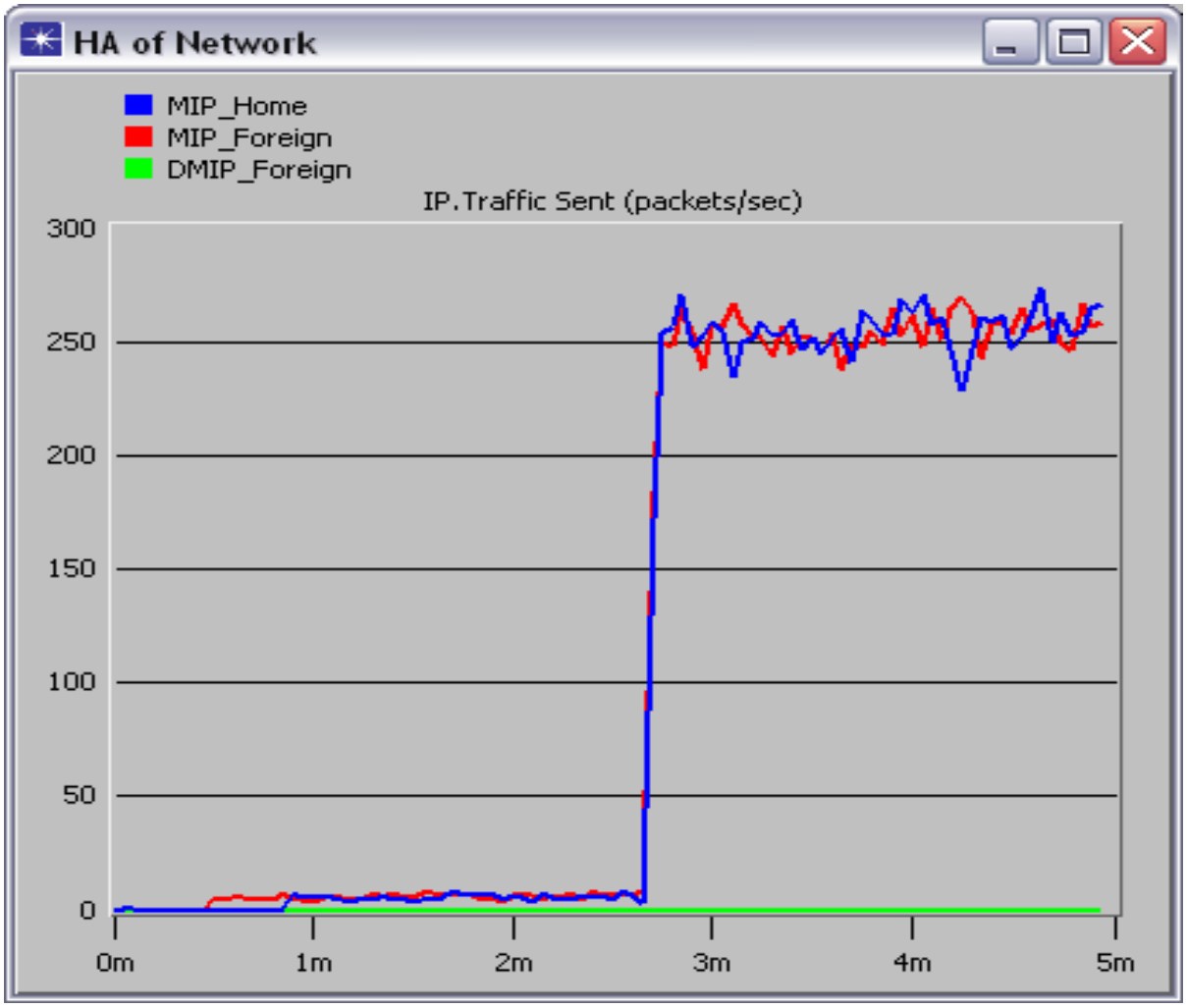

Figure 6-10: IP traffic forwards from Home Agent

Overall the simulation results support the benefits of the proposed DHAA scheme which enables the integration of 2G, 3G and 3.5G assets with targeted 4G investment to realise a better 4G network with better broadband data performance.

DHAA is an extension of the D-MIP described in Chapter 5 and a more detail reference model, conceptual design and simulation was conducted using a different network simulator.

Chapter 7: Conclusion and discussion

7.1 Conclusion

In this thesis, an overview of the evolution of mobile technology in conjunction with IP based technology is presented. The unique combination of mobile phone technology with the IP based protocols created an issue of mobility management and it is becoming more and more important as more services are moving from a fixed line internet connection to a mobile connection. The cellular technology did not provide a fixed point of connection to the mobile node and connection are constantly hands off from one mobile phone base station to another. The point of attachment of mobile node is constantly changing.

After an overview of the IP networking concept, cellular phone technology, an introduction of a key mobility management protocol Mobile IP and its recent enhancement is presented. Then, the design of a Distributed Agent Mobility Management platform and D-MIP protocol is presented. D-MIP is a generic platform that can be deployed in different kinds of radio access network such as WLAN, CDMA or UMTS. Also, D-MIP utilizes IP based technology and design, which minimizes change in network elements and data communication equipment. D-MIP allows easy deployment in current cellular networks without major equipment changes. D-MIP provides stable signalling overheads to the most Mobile Node. From the

simulation results using the ns2 simulator, it can be seen that D-MIP reduces the problems of link delay time from the network that the home agent located. As MN moves further away from its home network, the home network link delay time will increase and packet delay and packet lost are increased in MIP. D-MIP can reduce this problem and maintaining a stable handoff performance and packet delay due to the triangular routing of MIP.

Also, an extended design of a Dynamic Home Agent Anchoring (DHAA) Scheme is presented at a later chapter which is designed based on the D-MIP concepts. DHAA enhances the current Mobile IP approach by allowing Mobile Host (MH) attach to any Home Agent inside the IP network. It efficiently reduces signalling overheads and packet delay in the current IP network compare with using Mobile IP protocol and compatible with Mobile IPv6. Using DHAA, different networks infrastructure can be combined togethers and forms a 4G network. From the simulation results of DHAA using OPNET®, it can be seen that DHAA scheme reduces the problems of traffic sending from a distance Home Agent. As MN moves further away from its home network, the home network link delay time will increase and packet delay and packet lost are increased in MIP. DHAA can reduce this problem and maintaining a stable handoff performance and packet delay due to the triangular routing of MIP. The requirement definition, design and simulation of D-MIP and DHAA provided an enhancement to the current approach of mobility management. The aim is to develop an high level approach which can be implemented in different kinds of networks as well as study the approaches using simulation techniques. Two simulators were used to simulate the D-MIP and DHAA scheme and the results were presented.

7.2 Future work

Having defined the protocol in detail the following options exist to extend the work of this thesis.

- In order to gain real life data for the performance, the protocol can be implemented in an open source operating system such as FreeBSD. A laboratory test-bed could be constructed using WLAN type devices for BSs and Pentium PCs for Mobile nodes.
- Extend D-MIP functionality to Mobile IPv6. Similar deficiencies exist when using Mobile IPv6 in cellular network backbones. Handoff optimisation for Mobile IPv6 is an area of current research. Mobile IPv6 handoffs must contend with minimizing delays associated with Neighbour Discovery and Duplicate Address Detection (for both stateful and stateless auto-configuration) procedures. A hierarchical structure is still applicable in the Mobile IPv6 model. Therefore this work can be extended to Mobile IPv6. A brief example of applying the D-MIP concept is to apply the same concept at Chapter 6 Section 6.3.4 in a Mobile IPv6 situation for handling the mobility issue in MIPv6.
- Investigate the performance of D-MIP while operating in a more recent wireless technologies such as WiMax and 3.5G Cellular network.

References and bibliography

[1] E. Dahlman, S. Parkvall and J. Skold, "4G: LTE/LTE-Advanced for Mobile Broadband," Academic Press 2011.

[2] C Zhang, S. L. Ariyavisitakul and T. Meixia, –LTE-Advanced and 4G Wireless Communications," IEEE Communications Magazine, vol. 50, issue 2, Feb 2012, pp. 102-103.

[3] C. Perkins, Ed., –IP Mobility Support for IPv4, Revised," *Internet RFC 5944*, Nov 2010, The Internet Engineering Task Force (IETF).

[4] Prasan de Silva , Micro Mobile IP: Mobility Management in IP based Cellular Network, M.E. Thesis, University of Canterbury, Christchurch, New Zealand, March 2001.

[5] A. Valkó, —Cellular IP: A New Approach to Internet Host Mobility," ACM SIGCOMM Computer Communication Review, vol. 29, no. 1, Jan. 1999, pp. 50–65.

[6] A. Campbell, A. Valko, J. Gomez and Z. Turanyi, –Cellular IP," Old Internet draft, draftietf-mobileip-cellularip-00, January 2000, The Internet Engineering Task Force (IETF).

[7] A. T. Campbell, J. Gomez, S. Kim, A. G. Valko, C. Y. Wan and Z. R. Turanyi, –Design, implementation, and evaluation of Cellular IP," IEEE Personal Communication, vol 7, no. 4, Aug 2000, pp. 42-49.

[8] R. Ramjee, K. Varadhan, L. Salgarelli, S. R. Thuel, S.-Y. Wang and T. La Porta, HAWAII: A Domain-based Approach for Supporting Mobility in Wide-area Wireless Networks," IEEE/ACM Transactions on Networking, vol. 10, no. 3, pp. 396-410, June 2002.

[9] R. Ramjee, T. Porta, S. Thuel, K. Varadhan and L. Salgarelli, –IP micro-mobility Support using HAWAII," Old Internet draft, draft-ietf-mobileip-hawaii-00, expired, July 2000.

[10] E. Gustafsson, A. Jonsson, and C. Perkins, –Mobile IP Regional Registration," Old Internet draft, draft-ietfmobileip-reg-tunnel-03, July 2000, The Internet Engineering Task Force (IETF).

[11] H. Haverinen and J. Malinen, –Mobile IP Regional Paging," Old Internet Drafts, draft-haverinen-mobileip-reg-paging-00, June 2000, Mobile IP Working Group, The Internet Engineering Task Force (IETF).

[12] A. T. Campbell, J. Gomez, IP Micro-Mobility Protocols, *ACM SIGMOBILE, Mobile Computing and Communication, Review (MCCR)*, Vol. 4, No. 4, October 2001, pp. 45-54.

[13] X. Sun, S. Li, T. Liu and L. Zhang, –A distributed-agent scheme in Mobile IP networks," Proceedings of IEEE Conference on Electrical and Computer Engineering, Vol. 2, May 2004, pp-637-640.

[14] I. F. Akyildiz, J. Xie and S. Mohanty, –A Survey of Mobility Management in Next-Generation All-IP-Based Wireless Systems," IEEE Wireless Communications, vol. 11, no. 4, pp. 16-28, August 2004.

[15] Z. Zhu, R. Wakikawa and L. Zhang, –A Survey of Mobility Support in the Internet," RFC 6301, draft-zhu-mobility-survey.txt, July 2011, The Internet Engineering Task Force (IETF).

[16] A. Veres, A.T. Campbell, M. Barry and L. H. Sun, –Supporting Service Differentiation in Wireless Packet Networks Using Distributed Control," *IEEE JSAC, special issue on Mobility and Resource Management in Next-Generation Wireless Systems*, vol. 19, no. 10, Oct. 2001, pp. 2081–2093.

[17] IEEE Std 802.11r-2008: IEEE standard for local and metropolitan area networks – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 2: Fast Basic Service Set (BSS) Transition, IEEE Standards Association, July 2008.

[18] IEEE Std 802.16-2009: IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems, IEEE Standards Association, May 2009.

[19] P. Seite and P. Bertin, –Dynamic Mobility Anchoring,” Old Internet draft, draft-seite-mext-dma-00.txt, May 2012, The Internet Engineering Task Force (IETF).

[20] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury and B. Patil, –Proxy Mobile IPv6,” IETF RFC 5213, August 2008.

[21] P. Bertin, S. Bonjour and J-M Bonnin, –A Distributed Dynamic Mobility Management Scheme Designed for Flat IP Architectures,” Proceedings of 3rd International Conference on New Technologies, Mobility and Security, (NTMS 2008), Morocco, Nov 2008, pp. 1-5.

[22] H. Chan, –Proxy Mobile IP with Distributed Mobility Anchors,” GlobeCom 2010 Workshop on Seamless Wireless Mobility, Miami, USA, 6-10 December 2010, pp. 16-20.

[23] M. Fisher, F. U. Anderson, A. Kopsel, G. Schafer and M. Schlager, –A Distributed IP Mobility Approach for 3G SAE,” 19th International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC 2008), France, September 2008, pp. 1-6.

[24] P. Bertin, S. Bonjour and J-M Bonnin, –Distributed or Centralized Mobility?” Proceedings of IEEE Global Communications Conference (GLOBECOM 2009), Honolulu, Hawaii, Dec 2009, pp. 1-6.

[25] S. Pack, J. Choi, T. Kwon and Y. Choi, –Fast-Handoff Support in IEEE 802.11 Wireless Networks,” IEEE Communications Surveys & Tutorials, vol. 9, no. 1, pp. 2-12, 1st Quarter, 2007.

[26] H. Yokota, A. Idoue, T. Hasegawa and T. Kato, –Link Layer Assisted Mobile IP Fast Handoff Method over Wireless LAN Networks,” Proceedings or ACM MOBICOM 2002, pp. 131-139, September 2002.

[27] IEEE Std 802.21-2008: IEEE standard for local and metropolitan area networks – Part 21: Media Independent Handover Services, IEEE Standards Association, January 2009.

[28] H. Soliman, C. Castelluccia, K. ElMalki and L. Bellier, –Hierarchical Mobile IPv6 (HMIPv6) Mobility Management, IETF RFC 5380, October 2008.

[29] R. Koodli, "Fast Handover for Mobile IPv6," RFC 4068, July 2005, The Internet Engineering Task Force (IETF).

[30] H. Soliman, Ed., "Mobile IPv6 Support for Dual Stack Hosts and Routers," RFC 5555, June 2009, The Internet Engineering Task Force (IETF).

[31] A. Grilo, P. Estrela and M. Nunes, "Terminal Independent Mobility for IP (TIMIP)," IEEE Communications Magazine, vol. 39, no. 12, pp. 34-41, December 2001.

[32] A. C. Snoeren and H. Balakrishnan, "An End-to-end Approach to Host Mobility," Proceedings of ACM International Conference on Mobile Computing and Networking (MOBICOM 2000), August 2000, pp. 155-166.

[33] M. Riegel and M. Tuexen, "Mobile SCTP," Old Internet draft, draft-riegel-tuexen-mobile-sctp-09.txt, November 2007, The Internet Engineering Task Force (IETF).

[34] D. A. Maltz and P. Bhagwat, "MSOCKS: An Architecture for Transport Layer Mobility," Proceedings of IEEE INFOCOM'98, vol. 3, pp. 1037-1045, 1998.

[35] J. Rosemberg, H. Schulzrinne, G. Camarillo, A. Johnston, R. Sparks, A. Handley, and E. Schooler, "SIP: Session Initiation Protocol," IETF RFC 3261, June 2002.

[36] H. Schulzrinne and E. Wedlund, "Application-layer Mobility Using SIP," Mobile Computing and Communication Review, vol. 4, no. 3, pp. 47-57, July 2000.

[37] S. Salsano, C. Mingardi, S. Niccolini, A. Polidoro and L. Veltri "SIP-based Mobility Management in Next Generation Networks," IEEE Wireless Communication, vol. 15-2, April 2008.

[38] S. Giordano, D. Lenzarini, A. Puiatti and S. Vanini, "WiSwitch: Seamless Handover between Multi-provider Networks," Proceedings of Second Annual Conference on Wireless On-demand Network Systems and Services (WONS 2005), pp. 224-235, January, 2005.

[39] M. Chang, H. Lee and M. Lee: "A Per-application Mobility Management Platform for Application-specific Handover Decision in Overlay Networks," Computer Networks, vol 53-11, July 2009, pp. 1846-1858.

[40] M. Chang, M. Lee and H. Lee: "Per-Application Mobility Management with Cross-Layer Based Performance Enhancement," IEEE WCNC 2008., March 31 2008-April 3 2008, pp 2822–2827.

[41] National Academy Press, "The Evolution of Untethered Communications," Report presented to the US DoD and DARPA, 1998, ISBN: 978-0-309-05946-6.

[42] D. Corner, Internetworking with TCP/IP: Principles, protocols, and architecture, Prentice Hall, 2006.

[43] Internetworking Technology Handbook, Cisco Systems, Revision 17 April 2012.

[44] J. Postel, "Internet Protocol," RFC791, September 1981, The Internet Engineering Task Force (IETF).

[45] D. E. Comer, Internetworking with TCP/IP, Volume I:Principles, Protocols, and Architecture, Second Edition, Prentice-Hall, 1991

[46] C. L. Hedrick, "Introduction to the Internet Protocols," Rutgers University, 1987

[47] J. Solomon, Mobile IP: The Internet Unplugged, Prentice-Hall, 1998

[48] C. Perkins, "IP Mobility Support," RFC2002, October 1996, The Internet Engineering Task Force (IETF).

[49] C. E. Perkins, Mobile IP: Design principles and practices, Addison Wesley, 1998

[50] C. Perkins, "Mobile IP," *IEEE Communications* p84-99, May 1997

[51] M. Zeng, A. Annamalai and V. K. Bhargara, "Recent Advances in Wireless Communications," *IEEE Communications*, September 1999.

[52] I. Kriaras, A. W. Jarvis, V. E. Phillips and D. J. Richards, "Third Generation Mobile Network Architectures for UMTS," Bell Labs Technical Journal, Summer 1997.

[53] M. Torabi and R. Buhrke, ―Third-Generation Mobile Telecommunications and Virtual Home Environment:A Prioritization Analysis," Bell Labs Technical Journal, July-September 1998.

[54] A. Salkintzis, ―A Survey of Mobile Data Networks," IEEE Communications Surveys, Q3 1999.

[55] IEEE standard 802.16 IEEE standard for Local and Metropolitan Area Networks; Part 16: Air Interface for fixed broadband wireless networks.

[56] IEEE standard 802.16 IEEE standard for Local and Metropolitan Area Networks; Part 16: Air Interface for fixed and mobile broadband wireless networks; Amendment 2.

[57] I. Guardini, P. D'Urso and P. Fasano, ―The Role of internet Technology in Future Mobile Data systems," IEEE Communications, November 2000

[58] T. Issariyakul and E. Hossain, ―Introduction to Network Simulator Ns2", Second Edition, Springer 2012.

[59] OPNET® Modeler User Manual, OPNET.

[60] C. W. Yung, R. P. Coutts and D. Abbott, ―Design of distributed agent mobility management platform (D-MIP) for IP-based wireless networks," *Proceedings of 4G Mobile Forum Conference 2005*, San Diego, USA.