



TESIS DOCTORAL

FROM CENTRALIZED TO DISTRIBUTED
MOBILITY MANAGEMENT. A CONTRIBUTION TO
FUTURE MOBILE NETWORKS

GESTIÓN DE LA MOVILIDAD
CENTRALIZADA Y DISTRIBUIDA. UNA CONTRIBUCIÓN PARA
LAS REDES MÓVILES DE PRÓXIMA GENERACIÓN.

Autor: Javier Carmona Murillo

Conformidad del Director:

Fdo.:

Departamento de Ingeniería de Sistemas Informáticos y Telemáticos
Julio 2015



TESIS DOCTORAL

FROM CENTRALIZED TO DISTRIBUTED
MOBILITY MANAGEMENT. A CONTRIBUTION TO
FUTURE MOBILE NETWORKS

GESTIÓN DE LA MOVILIDAD
CENTRALIZADA Y DISTRIBUIDA. UNA CONTRIBUCIÓN PARA
LAS REDES MÓVILES DE PRÓXIMA GENERACIÓN.

Javier Carmona Murillo

Departamento de Ingeniería de Sistemas Informáticos y Telemáticos
Julio 2015

ACKNOWLEDGEMENTS

I would like first to express my gratitude to all those who helped me in the completion of this thesis. First of all, to my thesis supervisor, José Luis González Sánchez for his support and help throughout this period. Then, I would like to thank to all my colleagues at GITACA Research Group, with whom I enjoyed this journey.

To the Department of Telematic Engineering at University Carlos III of Madrid and the Centre for Telecommunications Research at King's College London, for giving me the opportunity to enjoy some of the most fulfilling professional and personal experiences in my life.

Finally, and above all, the most special thanks to my friends and family for their absolute and unconditional support.

Thank you.

¡Gracias!

Resumen

En los últimos años, la gestión de la movilidad en Internet ha sido un campo de investigación muy activo en el ámbito de las comunicaciones. Los protocolos de gestión de la movilidad se encargan de mantener activas las conexiones durante el movimiento del usuario entre distintas redes heterogéneas. Los actuales diseños de estos protocolos se basan en una gestión centralizada en la que algunos nodos de la red, denominados *anclas de movilidad*, gestionan el tráfico y la señalización de todos los usuarios. Sin embargo, estos mecanismos centralizados deben diseñarse de nuevo para hacer frente a los requerimientos de las redes de nueva generación y al enorme aumento del tráfico previsto para los próximos años.

Con este objetivo, las soluciones centralizadas se están viendo desplazadas por mecanismos que gestionan la movilidad de forma distribuida (DMM), más adecuados para estos requerimientos. La idea principal de DMM es ofrecer una solución en la que las funciones de movilidad estén distribuidas en distintos nodos que se encuentren, desde el punto de vista de la topología de la red, más cerca de los usuarios finales. Así, se consigue un encaminamiento más óptimo y una gestión de los recursos de la red más eficiente.

A pesar de que muchos protocolos, aún en fase de diseño, se están desarrollando para funcionar de forma distribuida, existen situaciones en las que DMM provoca mayores costes en la red y, por tanto, su rendimiento puede verse afectado. En estos casos, los protocolos centralizados resuelven la movilidad de una forma más eficiente y son, por tanto, preferibles. De esta forma, se espera que las arquitecturas de redes móviles de próxima generación puedan disponer de soluciones híbridas en el que la gestión de la movilidad de una parte del tráfico se mantenga centralizada mientras que el resto pueda ser gestionado de forma distribuida.

Partiendo de esta situación, esta tesis trata sobre el análisis, el diseño y la evaluación de los protocolos de gestión de la movilidad en IPv6. Específicamente, presentamos tres propuestas que se enmarcan dentro de cada uno de estas etapas en el desarrollo de la gestión de la movilidad. En primer lugar, se propone una solución centralizada denominada *LinkWork Mobile MPLS*. La segunda propuesta es una solución distribuida denominada DM3. Finalmente, la tercera propuesta desarrollada es un mecanismo híbrido, denominado *Hybrid DMM*.

Finalmente, el rendimiento de las soluciones ha sido evaluado mediante análisis y simulación, con el objetivo de estudiar el comportamiento de los protocolos en función de los principales costes considerados en la movilidad, como el coste de señalización, el coste del transporte de los paquetes y la latencia en el *handover*.

Palabras claves: Gestión de la movilidad centralizada y distribuida, protocolos de movilidad, redes inalámbricas y móviles, evaluación analítica, simulación de red, evaluación experimental.

Abstract

Over the last few years, IP-based mobility management in the Internet has been one of the most active research fields in communications. Mobility management protocols are responsible for maintaining the ongoing communications while the user roams among distinct networks and also to provide reachability to the mobile users in such heterogeneous environment in terms of access. Existing IP mobility support protocols are all based on a centralized mobility anchor that manages the traffic and signaling of the mobile nodes. However, centralized mobility management protocols need to be redesigned in order to cope with the recent trends in mobile Internet and current increasing mobile data traffic demand.

In order to address these limitations which inherently occur in Centralized Mobility Management (CMM) protocols, Distributed Mobility Management (DMM) solutions are being developed to efficiently handle the current mobile traffic explosion. In DMM, the core idea is that the mobility anchors are distributed within the network, topologically closer to the users, with the aim to provide an almost optimal routing support and an efficient use of network resources to improve the scalability required for next generation mobile networks.

However, and as already alluded above, despite the fact that a number of mobility management approaches are in-design phase towards a more distributed operation aiming to mitigate the problems related to centralized operation, there are instances where DMM incurs higher costs and the performance of the network might be compromised. In fact, in some of these cases, CMM seems to solve the mobility problem more efficiently and therefore should be preferred. In this context, future mobile network architectures might potentially exhibit a hybrid centralized-distributed behavior in which the mobility management of some traffic will be kept centralized, while mobility support for other applications can be distributed.

To cope with this evolution, the thesis concerns analysing, designing, and evaluating IPv6 mobility management protocols. Specifically, we propose three novel approaches which cover each of this evolutionary stage. Our first scheme, LinkWork Mobile MPLS is a centralized solution. The second proposal called DM3 is based on the distributed paradigm. Finally, the third proposal is a Hybrid DMM solution, suitable to tackle future mobile network architectures.

In order to evaluate the proposed schemes, we carry out analysis and simulations to measure the performance of the protocols in terms of mobility cost and handover latency.

Keywords: Centralized and Distributed Mobility Management, mobility protocols, wireless and mobile networks, analytic evaluation, network simulation, experimental evaluation.

LIST OF ABBREVIATIONS

3GPP	Third Generation Partnership Project
3G, 4G, 5G	3 rd , 4 th , 5 th Generation Mobile Networks
ABA	Access Binding Acknowledgement
ABU	Access Binding Update
AMA	Access Mobility Anchor
AP	Access Point
AR	Access Router
BA	Binding Acknowledgement
BCE	Binding Cache Entry
BS	Base Station
BU	Binding Update
CBR	Constant Bit Rate
CDN	Content Delivery Network
CMM	Centralized Mobility Management
CN	Correspondent Node
CoA	Care-of Address
DAD	Duplicate Address Detection
DiffServ	Differentiated Services
DMM	Distributed Mobility Management
DM3	Distributed Mobility Management MPLS
DSMIPv6	Dual-Stack Mobile IPv6
EPS	Evolved Packet Core
FA	Foreign Agent
FH	Fast Handover
FMIPv6	Fast Handovers for MIPv6
FPMIPv6	Fast Proxy Mobile IPv6
GPRS	General Packet Radio Service
GRE	Generic Routing Encapsulation
GSM	Global System for Mobile Communications
HA	Home Agent
HB-DMM	Host-based DMM
HMIPv6	Hierarchical MIPv6
HNP	Home Network Prefix
HoA	Home Address
IETF	The Internet Engineering Task Force

ILER	Ingress Label Edge Router
IntServ	Integrated Services
IP, IPv4, IPv6	Internet Protocol, IP version 4, IP version 6
IPsec	Internet Protocol Security
L2, L3	Layer 2, Layer 3
LER	Label Edge Router
LERG	Label Edge Router Gateway
LDP	Label Distribution Protocol
LMA	Local Mobility Anchor
LMD	Local Mobility Domain
LN	Linkage Nodes
LSP	Label Switched Path
LSR	Label Switched Router
LTE	Long Term Evolution
LW-MMPLS	LinkWork Mobile MPLS
MA	Mobility Anchor
MAC	Media Access Control
MAG	Mobile Access Gateway
MAR	Mobility capable Access Router
MDA	Mobility Distributed Anchor
MIH	Media Independent Handover
MIIS	Media Independent Information Service
MIPv4	Mobile IPv4
MIPv6	Mobile IPv6
MN	Mobile Node
MOS	Mean Opinion Score
MPLS	MultiProtocol Label Switching
MTU	Maximum, Transmission Unit
NA	Neighbor Advertisement
NB-DMM	Network-based DMM
ND	Neighbor Discovery
NELER	New Egress Label Edge Router
NS	Neighbor Solicitation
OSI	Open System Interconnection
OSPF	Open Shortest Path First
PBA	Proxy Binding Acknowledgement
PBU	Proxy Binding Update
PCoA	Proxy Care-of Address
PELER	Previous Egress Label Edge Router

PMIPv6	Proxy Mobile IPv6
PSNR	Peak Signal to Noise Ratio
QoS	Quality of Service
RA	Router Advertisement
RIP	Routing Information Protocol
RS	Router Solicitation
RSS	Received Signal Strength
RSVP	Resource ReserVation Protocol
RTT	Round Trip Time
RTP	Real Time Protocol
RWP	Random WayPoint (mobility model)
SDN	Software Defined Networking
TCP	Transmission Control Protocol
TE	Traffic Engineering
UDP	User Datagram Protocol
WiFi	IEEE 802.11 WLAN
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless LAN

CONTENTS

List of Figures	xv
List of Tables	xix
1 Introduction	1
1.1 Problem statement	3
1.2 Objectives and contributions	4
1.2.1 Objectives	4
1.2.2 Contributions	5
1.3 Thesis Outline	6
1.4 Publications	7
2 Related Work	11
2.1 IPv6 Mobility Protocols	13
2.1.1 IPv6 features to support Mobility	14
2.1.2 Centralized Mobility Management	17
2.1.3 Distributed mobility management	25
2.1.4 Distributed hybrid solutions	28
2.2 Quality of Service for IP Mobile Networks	29
2.2.1 Mobility and MPLS integration	31
2.3 Summary	35
3 Analysis and Modeling	37
3.1 Introduction and contributions	39
3.2 Background in the analysis of IPv6 mobility management	40
3.3 Analytical Model	44
3.3.1 Network model	44
3.3.2 Mobility models	44
3.3.3 Traffic model	45
3.4 Mobility cost	46
3.4.1 Signaling Cost	47
3.4.2 Data packet delivery cost	48
3.4.3 Tunnelling Cost	50
3.4.4 Handover latency	51
3.4.5 Packet loss	52
3.4.6 Buffer size	53
3.5 Concluding remarks	55

4	LinkWork Mobility Management MPLS: A centralized proposal	57
4.1	Introduction and contributions	59
4.2	Proposed architecture	61
4.2.1	LinkWork Mobile MPLS operation	61
4.2.2	Recovery mechanism in LinkWork Mobile MPLS	65
4.3	Analytical evaluation	72
4.3.1	Signaling cost	73
4.3.2	Handover latency and Packet loss during a session	74
4.3.3	Buffer size	77
4.3.4	Tunnelling overhead	77
4.4	Numerical Results	78
4.5	Concluding remarks	82
5	DM3: A distributed approach	85
5.1	Introduction and contributions	87
5.2	Distributed Mobility Management MPLS	88
5.2.1	DM3 Operation	89
5.2.2	DM3 Recovery mechanism	91
5.2.3	Mobility functions in DM3	92
5.2.4	MDA selection process	93
5.3	Analytical results	94
5.3.1	Signaling cost	94
5.3.2	Packet delivery cost	96
5.3.3	Tunnelling Cost	98
5.3.4	Handover latency and Packet loss during a session	99
5.3.5	Buffer size	100
5.3.6	Mobility Anchors load	100
5.4	Performance evaluation	101
5.4.1	Analytical results	101
5.4.2	Simulation environment and results	106
5.4.3	Experimental results	112
5.5	Concluding remarks	116
6	Hybrid DMM	117
6.1	Introduction	119
6.2	Background and motivation for hybrid solutions	120
6.2.1	CMM and DMM limitations	120
6.2.2	A hybrid mobility management solution	120

6.3	Description of the hybrid mobility management scheme	121
6.3.1	Initial mobility anchoring	121
6.3.2	Registration and data delivery mechanisms	122
6.4	Decision criteria	125
6.4.1	Node-assignment algorithm	125
6.4.2	Link-assignment algorithm	127
6.5	Analytical evaluation	131
6.5.1	Signaling cost	131
6.5.2	Data packet delivery cost	132
6.5.3	Tunnelling Cost	132
6.6	Simulated results	133
6.6.1	Impact of the amount of MNs	134
6.6.2	Impact of Session Duration and Topology Scenarios	139
6.6.3	Impact of the MN's Speed and Topology Scenarios	142
6.7	Concluding remarks	145
7	Conclusions and Future Work	151
7.1	Conclusions	153
7.2	Future Work	155
	Appendix A: Experimental evaluation of CMM protocols	157
A.1	Introduction and contributions	159
A.2	Open source implementations of CMM protocols	159
A.3	Performance evaluation	160
A.3.1	Experimental setup and design	160
A.3.2	Quantitative tests	162
A.3.3	Qualitative tests	169
A.4	Final remarks	170
	Appendix B: IP Mobility Management Simulator	171
B.1	Introduction	173
B.2	Overview of the simulator	173
B.3	Simulation environment and scenario	177
B.4	Metrics for Mobility Management Evaluation	179
B.5	Final Remarks	180
	References	181

LIST OF FIGURES

2.1	Data packet routing in IPv6	14
2.2	IPv6 routing does not allow host mobility	15
2.3	IPv6's packet headers	15
2.4	Overview of Mobile IPv6	19
2.5	Message flow in MIPv6	20
2.6	Overview of Proxy Mobile IPv6	23
2.7	Message flow in PMIPv6	24
2.8	Host-based Distributed Mobility Management	27
2.9	Network-based Distributed Mobility Management	28
2.10	Basic MPLS-TE components	31
2.11	Mobile MPLS operation	32
2.12	Fast Handover Micro Mobile MPLS architecture	33
2.13	Message flow for the MPLS-PMIPv6	34
3.1	A general signaling cost message exchange for a mobility management protocol	49
3.2	A general delivery cost diagram for a mobility management protocol . . .	50
3.3	Generic timing diagram for a mobility management handover	53
4.1	LinkWork Mobile MPLS architecture	62
4.2	Registration procedure in LinkWork Mobile MPLS	63
4.3	LinkWork Mobile MPLS operation	64
4.4	Example data path followed by a packet during the recovery mechanism .	65
4.5	Diagram of the $LSP_{i,n}$ path for $i = 1$ and $n = 5$	66
4.6	Local Recovery for different distances from PELEER to LN. (a) $d = 1$ hop, (b) $d = 2$ hops, (c) $d = 3$ hops, (d) $d = n - i$, i.e., no recovery mechanism is implemented and end-to-end retransmissions are required.	69
4.7	Signaling cost	80
4.8	Signaling cost in LW Mobile MPLS with different lengths between the ELER and the LN	80
4.9	Total packet loss during a session	81
4.10	Impact of network diameter in handover latency.	82
4.11	Buffer size vs. bandwidth in MPLS access network	83
5.1	Overview of the DM3 approach	88
5.2	Handover from PELEER to NELER in DM3	90
5.3	Recovery mechanism operation	91
5.4	Distributed Mobility functions in DM3	93

5.5	Relative distances in hops in the network	95
5.6	Signaling cost of registration updates	102
5.7	Packet delivery cost vs. transmission rate	103
5.8	Tunnelling cost vs. transmission rate	104
5.9	Mobility anchor load vs. number of sessions by MN	104
5.10	Handover latency	105
5.11	Packet loss during a session	106
5.12	Topology used in the simulation	108
5.13	Signaling cost with a human walk mobility model	109
5.14	Signaling cost with a Random Waypoint mobility model	109
5.15	Packet delivery cost with a human walk mobility model	110
5.16	Packet delivery cost with a Random Waypoint mobility model	111
5.17	Tunnelling cost with a human walk mobility model	111
5.18	Tunnelling cost with a Random Waypoint mobility model	112
5.19	Testbed scenario	113
5.20	Testbed executions at different throughputs	113
5.21	Handover Latency	114
5.22	Impact of the location of the Mobility Anchor on packet loss (%) under different conditions. (a) at different throughputs (1, 30 and 50 Mbps) and (b) at different packet sizes (100, 500 and 1470 Bytes)	115
6.1	Overview of hybrid approach	122
6.2	Data path during hybrid solution operation. (a) Initial State. MN1 is attached to AR1. Session 1 is created. (b) MN1 moves to AR2 and session 2 is created. (c) MN1 moves to AR3. Session 1 finishes and session 3 arrives. (d) MN1 moves to AR4. Session 2 finishes and session 4 arrives.	124
6.3	Level of information	126
6.4	Example of mobility anchoring selection by Node 2 using Algorithm 1 . . .	126
6.5	Physical AR and BS structure and Overlay AR graph	130
6.6	Topology used in the simulation	133
6.7	Signaling Cost. (a) Human walk mobility model; (b) Random Waypoint mobility model	135
6.8	Routing cost. (a) Human walk mobility model; (b) Random Waypoint mobility model	137
6.9	Tunnelling cost. (a) Human walk mobility model; (b) Random Waypoint mobility model	138
6.10	The four different topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.	140

6.11	Impact of the signaling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.	141
6.12	Impact of the packet delivery cost in the four topologies used in the ex- periments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.	143
6.13	Impact of the tunnelling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.	144
6.14	Impact of the signaling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.	146
6.15	Impact of the packet delivery cost in the four topologies used in the ex- periments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.	147
6.16	Impact of the tunnelling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.	148
A.1	Architecture of the mobility agents in MIPv6 and PMIPv6 approaches . .	161
A.2	Mobile IPv6 testbed topology	162
A.3	Proxy Mobile IPv6 testbed topology	163
A.4	Handover latency in MIPv6 experiments	164
A.5	Throughput and PSNR in real time experiments	166
A.6	Throughput and PSNR in TCP streaming experiments	167
A.7	Summarize of the obtained PSNR results in the experiments conducted .	168
A.8	Mean Opinion Score of (a) real time and (b) TCP streaming transmissions with a confidence of 95%	169
B.1	Overview of the simulator	173
B.2	Events simulation	174
B.3	Hexagonal grid cellular model	175
B.4	RWP mobility	178
B.5	Human mobility GPS trace data format (sample)	179

LIST OF TABLES

3.1	Summary of the main analytical frameworks in the literature	42
3.2	Parameter definition and description	54
4.1	Protocols Features	60
4.2	Tunnelling Overhead	78
4.3	Parameter settings	79
5.1	Parameter values	95
5.2	Simulation settings	108
5.3	Packet loss percentage during experiments	115
6.1	Simulation settings. Impact of the Session Duration.	139
A.1	Values of the handover latency intervals with a confidence interval of 95%	165

1

INTRODUCTION

This chapter gives an overview of the thesis and outlines the motivation and objectives.

Contents

1.1	Problem statement	3
1.2	Objectives and contributions	4
1.2.1	Objectives	4
1.2.2	Contributions	5
1.3	Thesis Outline	6
1.4	Publications	7

1.1 Problem statement

Over the last few years, Internet data communications have experienced a paradigm shift from the traditional fixed cable access to the wireless and mobile world. From the beginning of the century, wireless technologies have evolved up to approximately a 1000-fold increase in data rate. The 1.2 Mbps maximum rate of the legacy 802.11 standard is now far from the recent high speed wireless networks promising up to gigabit data rates [1, 2]. This evolution, together with the enormous proliferation of powerful mobile devices is showing a high demand of mobile data traffic that grows year by year. In fact, recent reports outline that this traffic will grow nearly tenfold between 2014 and 2019, with approximately half of the traffic offloaded to the fixed network by means of WiFi devices and femtocells each month by 2016 [3]. Moreover, as mobile data traffic increases, the growth in signaling load is expected to increase almost 50% faster than the growth in data traffic over the next few years [4].

As a result, during these years, operators, industry and the research community have been evolving their mobile solutions to tackle such challenge, especially at the standardization organizations. Several new technologies and practices are showing up as hot topics. Fixed-mobile network convergence [5], mobile data offloading [6], decentralized network architecture [7], Software Defined Networks (SDN) [8–10], and many others [11] are ongoing attempts to cope with the challenge of the new era of mobile data networks.

Along this time, mobility support in the Internet has been an active research topic and numerous protocols have been proposed. The main purpose of these mobility management protocols is to provide continuous service to mobile users, even if they change its point of attachment to the network. To support continuous service, mobility protocols should maintain connections during handover and thus, provide seamlessness while the user moves through different wireless networks. From a standard perspective, the most relevant organizations in the field of mobile communications, such as the Internet Engineering Task Force (IETF) [12] and 3GPP [13], have designed different mobility management solutions to be adopted in the mobile network architectures.

Current packet-based mobile architectures, such as the 3GPP Evolved Packet System (EPS) and WiMAX [14], are evolving to an all-IP network for both voice and data communications. Therefore, IP mobility management protocols will inevitably play a key role to address continuity and session persistence across the user movement among different networks. At the same time mobility control at the IP layer has been considered as a network management tool for provisioning load balancing and/or data offloading in heterogeneous wireless networks [15]. Furthermore, the aforementioned mobile architectures are deployed in a hierarchical and centralized manner so, the current IP mobility management

protocols which handle user mobility in these networks also rely on the use of a centralized architecture.

However, a new architectural paradigm is being explored by both research and standards communities due to the necessity for future mobile networks to reduce the load in the core network and to address some well known issues of current deployments. The basic idea of this shift is to introduce flatter system architectures, in which Mobility Anchors (MAs) are placed closer to the mobile users, distributing the mobility functions among several entities in the access networks [16]. The IETF is leading this new distributed approach with the chartering in March 2012 of the Distributed Mobility Management (DMM) Working Group [17].

Although distributed mobility management solutions are suitable candidates for mobility management in future 5G networks, its behavior is not adequate for certain mobility scenarios. In this case, a hybrid solution of centralized and distributed protocols should be preferred.

In this context, this thesis address some of the problems mentioned above by developing new IP-based centralized, distributed and hybrid mobility management solutions.

1.2 Objectives and contributions ---

1.2.1 Objectives ---

The previous section described generally the need for mobility management in IPv6 networks as a mechanism to allow seamless mobility among heterogeneous networks while keeping the service level requirements of the current connections. Consequently, the main objective of this thesis is to develop IPv6 mobility solutions to users, covering the main paradigms in which IP mobility management protocols can be broadly classified: centralized and distributed. Moreover, the proposals need to be analyzed in order to study their performance through analytical evaluation or simulation. Thus, the main objective of this thesis could be summarized in this sentence:

Analyze, design and evaluate IPv6 mobility management protocols following both centralized and distributed paradigms to be suitable for future mobile networks

More specifically, the objectives are described next:

- To review the existing IP mobility management schemes developed by the IETF, as well as other relevant solutions in the literature. Moreover, an important objective is to perform an analytical study of these protocols and to compare their performance in order to understand their strengths and weakness. Apart from the development of an analytical framework, both simulations and experimental evaluations are advisable in order to complement the performance comparison.
- To study the possibilities of coupling Quality of Service (QoS) and mobility management techniques, particularly interesting is MPLS (Multi Protocol Label Switching), because it natively supports tunnelling.
- To design a Centralized Mobility Management (CMM) solution that overcomes the limitations of traditional protocols in terms of classic performance metrics (signaling overhead, packet delivery cost, handover latency and packet loss).
- To design a mobility management mechanism that address the drawbacks of centralized solutions by distributing the mobility functions in the through the access network.

1.2.2 Contributions

The following summarizes the major contributions of this thesis:

- A complete review of the centralized IP mobility management protocols, both host-based and network-based are presented. Moreover, a discussion of the limitations of each of these categories is given. In addition, a comprehensive review of the main distributed mobility management solutions is also included. Some scenarios in which DMM protocols are not adequate are outlined. In those situations, CMM solutions are preferred. We argue that hybrid centralized-distributed solutions provide additional flexibility to the mobile network operators, and can be suitable for future mobile networks.
- In order to assess the performance of the mobility management protocols, an analytical framework has been developed. With this framework it is possible to study the effects of various network parameters on the performance of these protocols. Moreover, a complete review of the most relevant analytical studies available in the literature has been done to summarize the main features of each of these works.
- A mobility management simulator has been developed in Matlab. Several modules have been coded to allow flexibility and an easy software upgrade. The simulator

facilitates the evaluation of the protocols under different mobility conditions (stochastic or realistic mobility models), different traffic models (random or exponentially distributed session arrivals) and different topologies of the access network.

- A mobility management solution called **LinkWork Mobile MPLS** has been proposed for centralized architectures. This solution couples a centralized host-based protocol with MPLS techniques to minimize packet loss and avoid packet disorder. Additionally, based on the analytical framework developed, numerical results are obtained in terms of signaling cost, packet loss, handover latency and buffer size.
- An additional mechanism, which inherits some functionality from LinkWork Mobile MPLS is developed. This solution is called **DM3** (Distributed Mobility Management MPLS) and is based on the DMM paradigm. DM3 distribute the mobility functions through some nodes closer to the users in order to address the new requirements of future mobile networks. Both, analytical and simulated results, are obtained to show the performance of the proposed solution.
- A hybrid centralized-distributed mobility management architecture is proposed (**Hybrid DMM**). In this proposal we develop different decision criteria algorithms that operators can use to handle mobile data traffic in a centralized or a distributed way. Numerical results obtained from analysis and simulation are also shown.
- An experimental evaluation of the main centralized protocols developed by the IETF has been conducted. A real testbed based on the open source implementations of the protocols has been made in order to analyze the handover latency. Moreover, numerical results are also obtained from multimedia transmissions.

1.3 Thesis Outline

This thesis is structured as follows. The related work is introduced in Chapter 2, which provides an overview of IP mobility management, focusing on the Layer 3 mobility protocols standardized at the IETF. The chapter presents these protocols from an evolutionary point of view, starting from centralized solutions, following with distributed approaches and finishing with a discussion about the benefits of hybrid schemes. In addition, we also give an overview of the QoS solutions for IP mobile networks, centering the attention on the couple of MPLS and mobility protocols.

Throughout all of these reviews, we investigate in Chapter 3 the strong and weak points of each scheme. In order to achieve this, we develop a model that analyzes each protocol with a common framework. This analytical framework will be used in the reminder to

analyze the performance of our proposed mobility management solutions, allowing the comparison among all of them.

The following three chapters contain the major contributions of the thesis.

In Chapter 4 we develop LinkWork Mobile MPLS, the first mobility management proposal of this thesis to address the QoS requirements in a centralized mobility architecture.

Chapter 5 describes the second mechanism proposed in this thesis. In this case, we inherit the operation of LinkWork Mobile MPLS to develop a newly distributed solution, called DM3 that overcomes the limitation of centralized approaches. In addition, the simulator developed in Matlab that has been used to obtain the simulated results of this thesis is presented in this chapter. Finally, analytical and simulated results are presented.

Afterwards, once centralized and distributed mobility management solutions have been proposed, we develop a new Hybrid DMM approach. Chapter 6, gives an in-depth description of this proposal. The basic idea of Hybrid DMM is to offer operators a decision criteria based on the network information, that allows to manage the mobile users' traffic in a centralized or distributed way. Analysis and simulations are also provided to measure the performance of our Hybrid DMM proposal.

Chapter 7 concludes the thesis by summarizing the results obtained and discussing some proposals for future work.

Finally, at the end of the thesis, two appendices include additional works related with the topic presented in the thesis body. Appendix A presents a complementary study that evaluates the performance of the main centralized protocols developed by the IETF using their open-source implementations. Thus, we conduct a practical evaluation of multimedia delivery under mobility conditions with real Linux-based prototype implementations. Appendix B presents the software developed to obtain the simulated results.

1.4 Publications

Selected peer-reviewed papers published from this thesis are documented in the following:

International Publications

JOURNALS

- Javier Carmona-Murillo, Vasilis Friderikos and José-Luis González-Sánchez. A Hybrid DMM Solution and Trade off Analysis for Future Wireless Networks. 2015. Under review.

- David Cortés-Polo, José-Luis González-Sánchez, Javier Carmona-Murillo, Francisco-Javier Rodríguez-Pérez. Proposal and Analysis of Integrated PTN architecture in the Mobile Backhaul to improve the QoS of HetNets. *EURASIP Journal on Wireless Communications and Networking*. 2015 (1), pp. 1-13. JCR: 0.800. Article URL: <http://dx.doi.org/10.1186/s13638-015-0341-2>
- Francisco-Javier Rodríguez-Pérez, José-Luis González-Sánchez, Javier Carmona-Murillo, David Cortés-Polo. An OAM function to improve the packet loss in MPLS-TP domains for prioritized QoS-aware services. *International Journal of Communication Systems*. 28-6, pp. 1037 - 1052. April 2015. JCR: 1.106. Article URL: <http://dx.doi.org/10.1002/dac.2742>
- Javier Carmona-Murillo, José-Luis González-Sánchez, David Cortés-Polo, Francisco-Javier Rodríguez-Pérez. DM3: Distributed Mobility Management in MPLS-based Access Networks. *International Journal of Network Management*. 24-2, pp.85-100. March/April 2014. JCR: 0.517. Article URL: <http://dx.doi.org/10.1002/nem.1854>
- Francisco-Javier Rodríguez-Pérez, José-Luis González-Sánchez, David Cortés-Polo, Javier Carmona-Murillo. A delay-oriented prioritization policy based on cooperative lossless buffering in PTN domains. *Journal of Network and System Management*. September 2014. JCR: 0.438. Article URL: <http://dx.doi.org/10.1007/s10922-014-9334-4>
- David Cortés-Polo, José-Luis González-Sánchez, Francisco-Javier Rodríguez-Pérez, Javier Carmona-Murillo. Mobility management in packet transport networks for network convergence. *Transactions on Emerging Telecommunications Technologies*. 26-5, pp. 749-759. May 2015. JCR: 1.354. Article URL: <http://dx.doi.org/10.1002/ett.2705>

BOOK CHAPTERS

- Javier Carmona-Murillo, José-Luis González-Sánchez; David Cortés-Polo, Francisco-Javier Rodríguez-Pérez. QoS in next generation mobile networks. An analytical study. *Resource management in Mobile Computing Environments*. pp.25-41, Springer International Publishing. February 2014. ISBN:978-3-319-06703-2
- David Cortés-Polo, José-Luis González-Sánchez, Francisco-Javier Rodríguez-Pérez, Javier Carmona-Murillo. Mobile-Fixed Integration for Next-Generation Mobile Network: Classification and Evaluation. IGI-Global. 2015. (Accepted for its publication).

CONFERENCE PAPERS

- Javier Carmona-Murillo, José-Luis González-Sánchez and Francisco-Javier Rodríguez-Pérez. Performance evaluation of an architecture for localized IP mobility management. *International Conference on Network and Service Management (CNSM)*, pp. 25-29. October 2010.
- Javier Carmona-Murillo, José-Luis González-Sánchez, David Cortés-Polo, Alfonso Gazo-Cervero and Francisco-Javier Rodríguez-Pérez. Mobility Management in MPLS-based Access Networks. An Analytical Study. *IX Workshop in G/MPLS networks (WGN9)*, pp. 47-59. July 2009.
- Javier Carmona-Murillo, José-Luis González-Sánchez and Isaac Guerrero-Robledo. Handover performance analysis in Mobile IPv6. A contribution to fast detection movement. *International Conference on Wireless Information Networks and Systems (WINSYS 2008)*. pp. 78-81. July 2008.

2

RELATED WORK

In this chapter, we present the different mobility management protocols and their evolution from centralized to distributed mobility approaches. We introduce the need for DMM, its current status at IETF, as well as other related work. Finally, we highlight the limitations of DMM in some scenarios in which centralized solutions should be preferred. Thus, we also draw attention to the hybrid centralized-distributed approach, which is envisioned as a promising solution for the mobility management in future mobile networks.

Contents

2.1	IPv6 Mobility Protocols	13
2.1.1	IPv6 features to support Mobility	14
2.1.2	Centralized Mobility Management	17
2.1.3	Distributed mobility management	25
2.1.4	Distributed hybrid solutions	28
2.2	Quality of Service for IP Mobile Networks	29
2.2.1	Mobility and MPLS integration	31
2.3	Summary	35

2.1 IPv6 Mobility Protocols

IP mobility mechanisms can be seen as the ability to provide a mobility function to IP devices. This mobility concept refers to the movement inside a network (local mobility or micromobility) or between networks (global mobility or macromobility) [18, 19] and its goal is to allow hosts to move around the Internet, changing its point of attachment to the network while keeping active the user's ongoing sessions [20]. Moreover, the mobility of the user must be transparent to the rest of the users in the Internet.

These mobility features can be achieved in various levels of the network stack. In fact, a variety of mobility support protocols have been developed to handle mobility management at different OSI (Open System Interconnection) layers, from link layer to application layer [21, 22]. In [23], an historical survey of the Internet mobility research and standardization evolution since the early 1990s is shown.

The recent fundamental networking trend has been focused mostly on realizing all-IP mobile networks. All-IP mobile networks, are networks in which IP is employed from a mobile subscriber to the access points (APs) that connect the wireless networks to the Internet [24]. For this reason, the main efforts towards efficient IP mobility have been done in the IP layer. Network layer solutions do not rely on or make any assumptions about the underlying wireless access technologies and signaling messages for mobility purposes are carried by IP traffic [25]. Moreover, the mobility solutions that operate at the IP layer are regarded as being more suitable as they do not violate any basic Internet design principles [26]. As we will describe later in this chapter, IPv6 procedures play an important role in the IP mobility management operation.

In the development of IPv6 mobility management protocols, the IETF (Internet Engineering Task Force) has made relevant efforts. In fact, the IETF has standardized the main IPv6 mobility protocols such as Mobile IPv6 (MIPv6) [27] and Proxy Mobile IPv6 (PMIPv6) [28].

Mobile IPv6 is the most representative host-based IPv6 mobility protocol whereas Proxy Mobile IPv6 is the main network-based IPv6 mobility protocol. Host-based protocols require the participation of the host, or Mobile Node (MN), in all aspects of mobility management. On the contrary, in network-based mobility protocols, the host does not participate in any mobility related signaling.

In this section, both host-based and network-based mobility management protocols, relevant in the scope of this thesis, are introduced. The specific details will be given in later chapters, in which the data and control plane procedures are analyzed.

2.1.1 IPv6 features to support Mobility

In the Internet, a node is identified by an IP address that uniquely identifies its point of attachment to the network, and packets are routed to the node based on this address. Therefore, a node must be located on the network indicated by its IP address in order to receive datagrams as can be seen in Figure 2.1. This prohibits the node from moving and remaining able to receive packets using the base IP protocol (see Figure 2.2).

IPv6 protocol [29] was developed by IETF to replace the current IPv4 [30] and solve several issues of IPv6, such as substantially increasing the address space, the mobility support and security by integrating IP security (IPsec) in IPv6 specifications, and more control on the level of quality of service. In addition, some features of IPv6 defined over IPv4, such as IPv6 headers, IPv6 addressed, Neighbour Discovery (ND) protocol and IPv6 Address Autoconfiguration, together with their own mobility protocol mechanisms, make it possible for a MN to roam seamlessly among IPv6 networks.

2.1.1.1 IPv6 headers

The IPv6 protocol uses two types of headers. First, the core IPv6 header, which is similar to the IPv4 header. Second, additional IP layer information may be carried in extension headers, which provide an efficient IPv6 datagram route with further flexibility. The Next Header field in the core IPv6 header specifies the presence of extension headers each of which contain the Next Header field (see Figure 2.3). When multiple extension headers are present, the receiver must process them sequentially. This requirement exists because the contents and semantics of an extension header determine whether to process the next

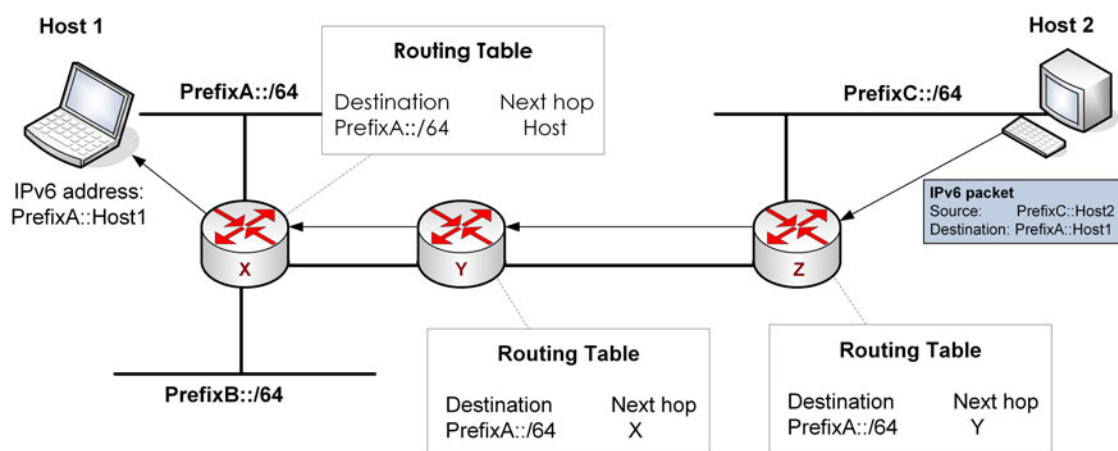


Figure 2.1. Data packet routing in IPv6

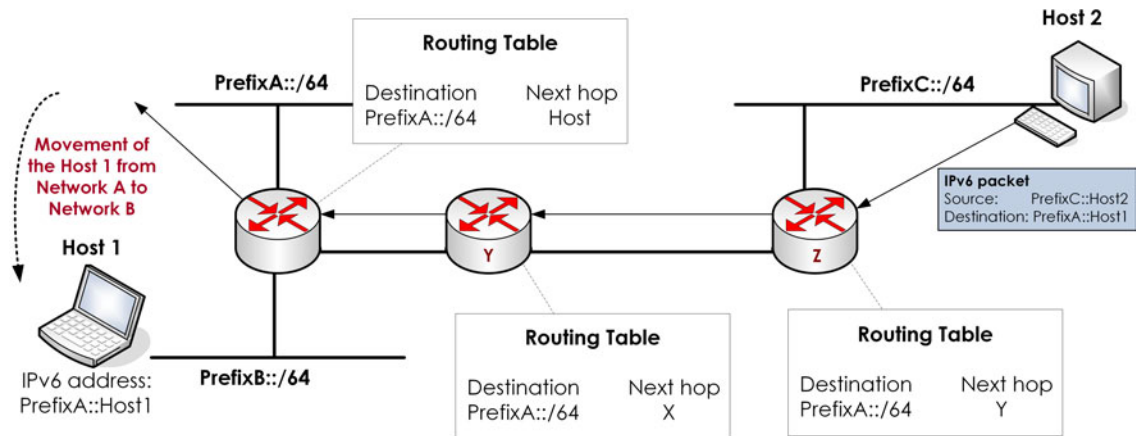


Figure 2.2. IPv6 routing does not allow host mobility

extension header or not. There is also a strict ordering requirement when constructing multiple extension headers [29].

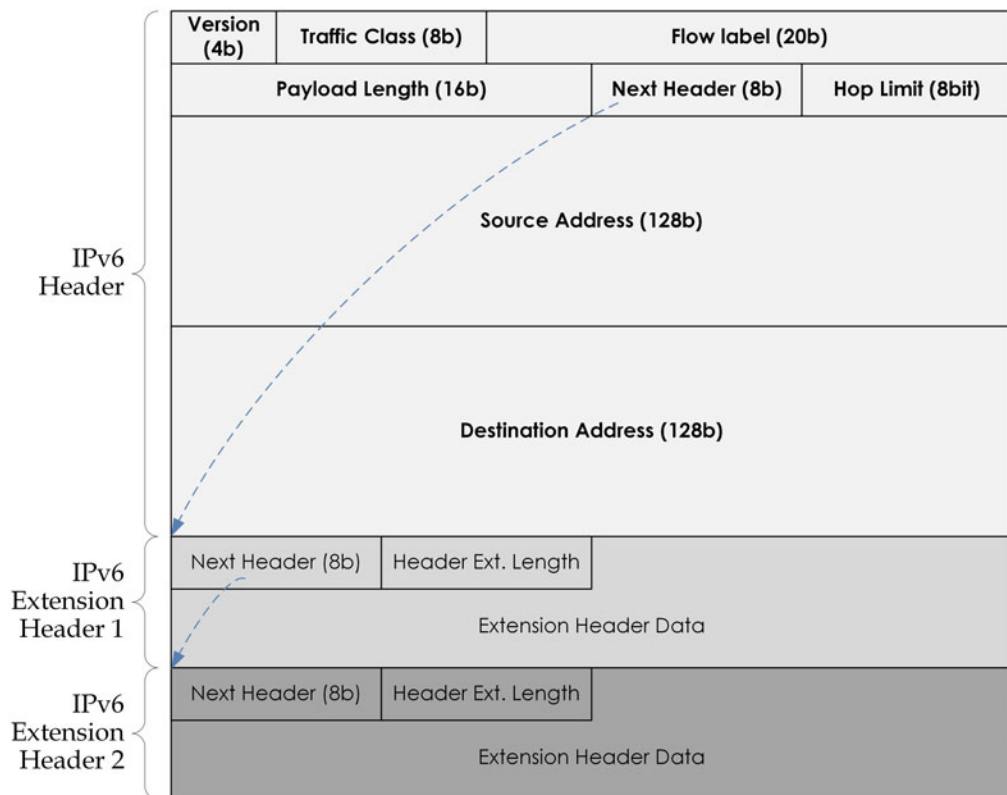


Figure 2.3. IPv6's packet headers

2.1.1.2 IPv6 Addresses

An IPv6 address is assigned to a network interface. If a node has multiple network interfaces, each can have one or more IPv6 addresses of its own, and any address can be used to reach the node subject to address scope and routing availability. There are three types of addresses in IPv6: unicast, anycast, and multicast. Each of these addresses can have different scopes, which limits their applicability and usage. Thus, in IPv6, the addresses can be categorized into three types: link local address, site local address and global address. The link local address can be used to communicate within the node's link and none of the packets with a link local address will be routed outside the link; the Site local address is unique within a site and used to communicate within a defined portion of the site; the global address is globally unique and its packets address can be routed anywhere.

The addresses are represented in text form using the notation $X : X : X : X : X : X : X : X$, where each X is the hexadecimal value corresponding to 16 bits of the overall address (128 bits). Sometimes, the address prefix is also represented using the notation "IPv6-Address/Prefix-Length", where Prefix-Length is the number of leftmost bits in the IPv6-Address that make up the prefix. The remaining bits in the IPv6-Address are assumed to form the host part, the identifier ID.

Next, we describe the Neighbor Discovery Protocol, which is used for hosts to discover routers willing to forward packets for them.

2.1.1.3 Neighbor Discovery Protocol

An important functionality introduced by IPv6, essential in the operation of an IPv6 mobility protocol, is the Neighbor Discovery Protocol [31]. This is a method to determine the link layer addresses for neighbours located on attached access links and also to find neighbouring routers that will forward the data packets on behalf of that node. Moreover, this protocol allows to determine the uniqueness of the configured address on a particular link through the Duplicated Address Detection (DAD) [32]. Neighbor Discovery is also used for a node to obtain or to generate its own address by using the stateful or stateless address auto-configuration method.

Especially relevant for IPv6 mobility are the network discovery mechanisms that are traditionally facilitated by router discovery and movement detection procedures through the utilization of Router Advertisement (RA) and Router Solicitation (RS) messages. The RAs are sent either unsolicited at regular intervals or as replies to RSs. The frequency of unsolicited advertisements, as well as the response time associated with solicited

advertisements are important for IP mobility [33]. Due to the difficulty in determining the optimal RA and/or RS intervals, router discovery hence movement detection contribute long delays to the handover process. Furthermore, the trade-off between the rate of RAs/RSs broadcasting and the bandwidth in the air interface causes inefficiency and unreliability in the RA/RS method of network discovery [34]. Thus, IPv6 movement detection mechanisms can be complemented with quicker and efficient network discovery techniques of lower layers. In particular, the recent media independent information service (MIIS) of the IEEE 802.21 media independent handover (MIH) services 15 provides a technique that reduces the delay due to network discovery in next-generation wireless networks [35–37].

Finally, in this overview of IPv6 protocol we described the methods used by a MN to auto-configure an IPv6 address.

2.1.1.4 IPv6 address auto-configuration

Address auto-configuration in IPv6 is defined in [29]. It defines how the nodes can generate unique link-local and globally routable unicast addresses simply by using a unique identifier local to the host and by using the prefix advertised by the routers in the RA messages. This is often referred to as stateless address autoconfiguration since it avoids the administration overheads (both manual and through a centralized server). The auto-configuration is performed when an interface becomes enabled. Hosts first generate a link-local address by prepending the well-known prefix FE80::0 to the interface identifier. This tentative address is confirmed to be unique in the DAD process. This Duplicate Address Detection uses the Neighbor Solicitation and Neighbor Advertisement messages defined in [31].

As we have reviewed in this section, IPv6 provides crucial features to allow the evolution of the current Internet into a Mobile Internet. Next, specific mobility management approaches are detailed.

2.1.2 Centralized Mobility Management

Traditional IP mobility support protocols developed by the IETF are all based on centralized mobility anchors that facilitate mobility support for all registered Mobile Nodes (MNs) in a highly efficient manner. These approach has been developed in both host-based and network-based solutions. Next, the main Centralized Mobility Management (CMM) protocols are presented.

2.1.2.1 Host-based CMM

Host based mobility approaches involve the mobile node itself in mobility management operations and as such require changes in the network stack of the mobile node. This section will discuss approaches that rely on host-based mobility management protocols.

Mobile IPv6

Until now, Mobile IPv6 is the most representative mobile management scheme developed by the IETF on the way towards next generation mobile networks. However, the IP mobility protocol was first proposed for IPv4 (Mobile IPv4, MIPv4) [38]. Although this protocol could support handover from one network to another and was widely used for handover management [39], Mobile IPv4 was further evolved into Mobile IPv6, due to the weakness it suffers such as extra delays due to triangular routing, ingress filtering issues, tunnelling, high signaling overhead and furthermore, the IPv4 address space is not enough for the many IP nodes that will require ubiquitous wireless Internet connectivity.

Mobile IPv4 is a kind of automatic tunnel establishment protocol. The moving node registers its current location to the proxy node called the Home Agent (HA). All the packets are forwarded once to the home agent and then sent to the final destination. Apparently, if the MN and its communicating node reside nearby and the HA is located far away, the communication path becomes long and redundant. The Mobile IPv4 base protocol does not mention the optimization mechanism for this case [40]. Thus, Mobile IPv6-based mobility management protocols are more suited for next generation wireless networks as they overcome the mentioned problems of MIPv4. In fact, IPv6 specification improves many of the weak aspects of IPv4, e.g., provides an optimal header format, neighbor discovery mechanism, improved security and quality of service, reasonable addressing architecture, and stateless auto-configuration. Thus, MIPv6 benefits from these IPv6 improvements [34].

Mobile IPv6 allows nodes to remain reachable while moving around in IPv6 networks. Without specific support for mobility, packets destined to a mobile node would not be able to reach it while the mobile node is away from its home link. In order to continue communication in spite of its movement, a mobile node could change its IP address each time it moves to a new link, but the mobile node would then not be able to maintain transport and higher-layer connections when it changes location.

MIPv6 supports mobility for the MN by providing it with at least two addresses: a Home Address (HoA) which is a fixed address provided by the Home Agent (HA) and Care-of Address (CoA), which is obtained in the foreign access network and changes when MN moves to a new subnet. Figure 2.4 illustrates an overview of Mobile IPv6 and its basic

terminology. The main mobility functions (location update and packet delivery) in MIPv6 are described next.

Mobile IPv6 location update procedure is as follows. When a mobile node stays in the home domain, it is able to receive packets destined to its Home Address and being forwarded by means of conventional IP routing mechanisms. Periodically, or whenever the user attaches to another Access Router (AR), movement detection is performed in order to identify its new point of attachment and a new CoA is acquired. Once configured with a new CoA, the MN registers with the HA through Binding Update (BU) messages, informing of the user's current location and establishing a tunnel (IP-in-IP or Generic Routing Encapsulation, GRE) between the HA and the MN located in a visited network.

With respect to the packet delivery procedure, when the MN is away from home, the HA has a legal mobility binding and it will act as MN's proxy entity. This means that any packet addressed to the MN will end up at the HA because the HA will respond to all Neighbor Solicitation (NS) request for the MN. Once the HA has intercepted a packet, it

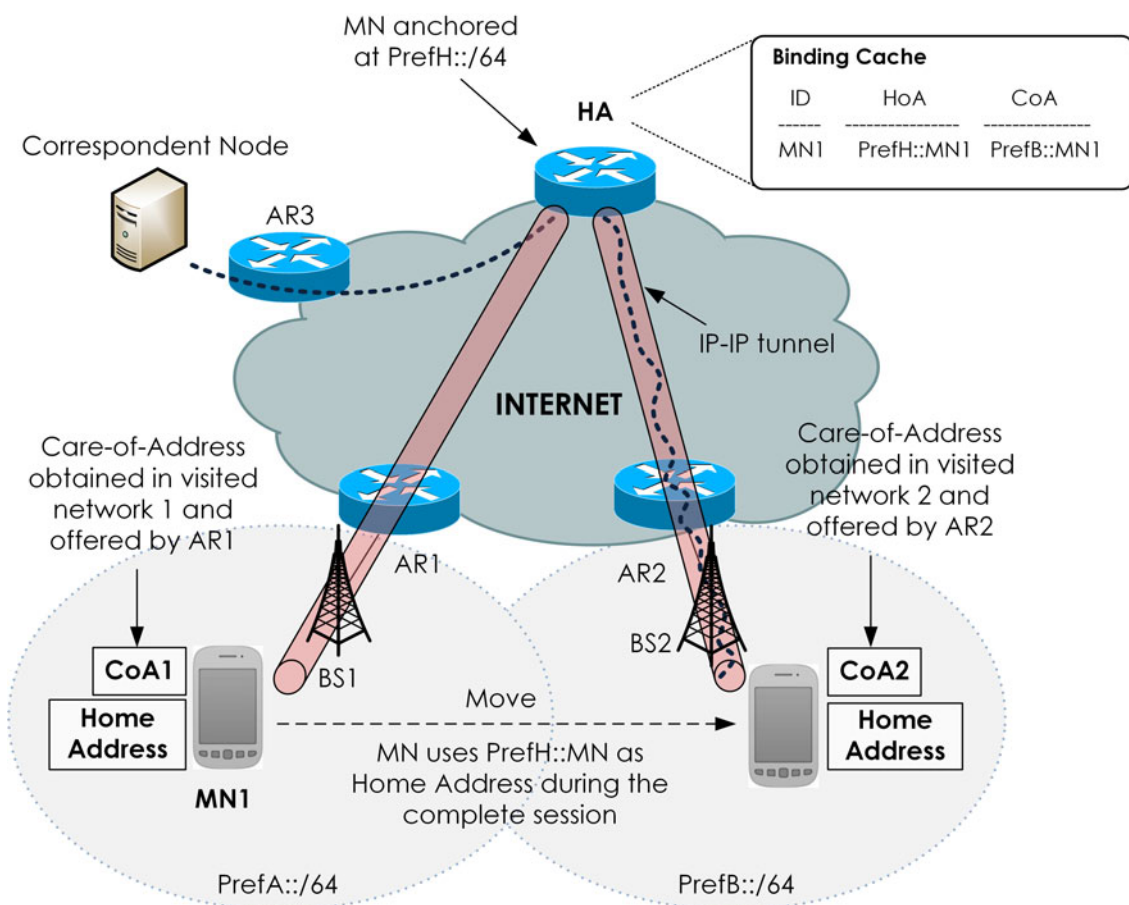


Figure 2.4. Overview of Mobile IPv6

will encapsulate the packet destination address of the MN's CoA. The MN decapsulates the packet upon its arrival to reveal the original packet, as if the Correspondent Node (CN) had sent it directly to the MN (see Figure 2.5). When the MN has not established a connection with its CN, it should send the packets destined to the CN via the HA using the reverse tunnelling procedure [41]. In this operation, the Home Agent is the critical part of the system since it is on the path of both signaling and data for mobile users.

The overall operation of MIPv6 described previously, is shown by its message flows in Figure 2.5.

Finally, it is worth mentioning that the 3GPP Evolved Packet System has adopted Dual Stack Mobile IPv6 (DSMIPv6) [42] for host-based mobility management [43]. The DSMIPv6 is a protocol that operates with IPv6 and IPv4, being the Home Agent the entity that stores the Home Address and the Care of Address. For mobile communication with a corresponding node, the packets are intercepted by HA and passed on to their final

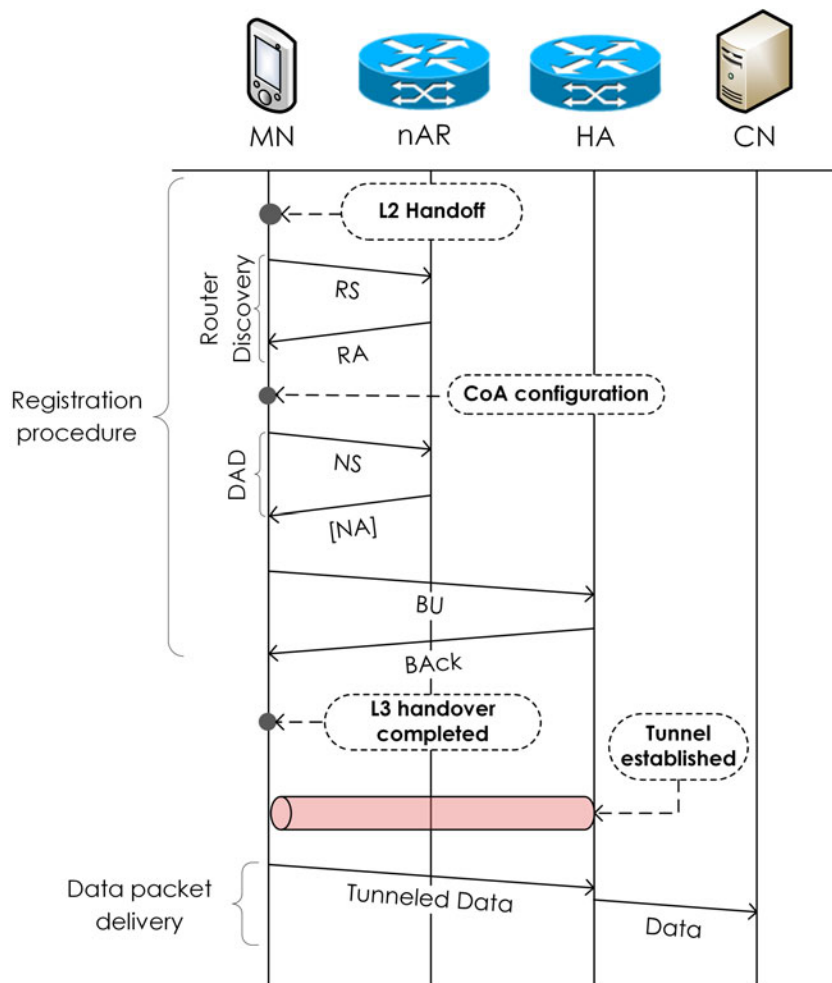


Figure 2.5. Message flow in MIPv6

destination. DSMIPv6 provide an IPv6 transition solution because it can, by definition, offer dual-stack connectivity independently of the address family of the Care-of address obtained within the visited network [44].

In MIPv6 and, in general, in any other mobility management protocol, handover latency results in packet loss that degrades network performance, which is unacceptable and detrimental to real-time traffic causing user perceptible service deterioration [45]. Thus, improving MIPv6 performance is important for wireless networks to provide MNs with seamless mobility, session continuity and guaranteed QoS. Since MIPv6 protocol handles local and global mobility in the same way, mobility management procedure introduces lengthy registration delay and unavoidable packet loss [46].

With similar operation to MIPv6, another approach was proposed to avoid the host's involvement in the mobility process. This approach is called network-based protocols and is explained next.

2.1.2.2 Network-based CMM

As we have detailed in the previous section, host-based mobility management requires client functionality in the IPv6 stack of a mobile node. Exchange of signaling messages between the mobile node and the home agent enables the creation and maintenance of a binding between the mobile node's home address and its care-of address. This implies that the requirement for the modification of mobile nodes may cause them to become increasingly complex. Furthermore, host-based mobility induces high mobility signaling overhead when the mobile node frequently moves between subnets [47]. With these limitations, solutions that provide mobility support within a part of the network by means of functionality residing only on the network infrastructure were developed.

With these design goals, the IETF developed a network-based mobility management protocol which aims to cover [48]:

- Support for unmodified Mobile Nodes: Unlike host-based mobility management protocols, the network-based protocol should not require any software modification for IP mobility support on the mobile nodes.
- Efficient use of wireless resources: The network-based protocol should avoid tunnelling overhead over the wireless link, so it should minimize overhead within the radio access network.
- Reduction in handover-related signaling volume: Considering MIPv6, whenever an MN changes the subnets, various signaling messages are required. Therefore,

in the network-based protocol handover-related signaling should be performed as infrequently as possible.

- Support for IPv4 and IPv6: Although the initial design of the network-based protocol uses an IPv6 host, it is intended to work with IPv4 or a dual-stack host as well. Compared to host-based mobility management approaches such as MIPv6 and its enhancements, a network-based mobility management approach such as Proxy Mobile IPv6 has several advantages.

From a deployment perspective, network-based mobility management does not require any modification of mobile nodes. From a performance perspective, due to the fact that wireless resources are very scarce, the efficient use of wireless resources can result in the enhancement of network scalability. In host-based approaches such as MIPv6, the mobile node is required to participate in mobility related signaling. Thus, a lot of tunneled messages as well as mobility-related signaling messages are exchanged via the wireless links. Considering the explosively increase in the number of mobile subscribers, such a problem would cause serious performance degradation. On the contrary, in a network-based approach the serving network controls mobility management on behalf of the MN, so tunnelling overhead as well as a significant number of mobility-related signaling message exchanges via wireless links can be reduced.

Another advantage is from a network service provider perspective. Network-based mobility management can enhance manageability and flexibility by enabling network service providers to control network traffic and provide differentiated services, among other things. In fact, some cellular systems such as IS-41 and Global System for Mobile Communications (GSM) can be considered network-controlled systems. Moreover, General Packet Radio Service (GPRS) has some resemblance to Proxy Mobile IPv6 in that they are both network-based mobility management protocols and have similar functionalities [24].

Proxy Mobile IPv6

PMIPv6, the main network-based protocol, is based on MIPv6 in the sense that it extends MIPv6 signaling and reuses many concepts such as HA functionality. The new principal functional entities of PMIPv6 are the Mobile Access Gateway (MAG) and Local Mobility Anchor (LMA). The MAG typically runs on the AR. Its main role is to detect the MN's movements and initiate mobility-related signaling with the MN's LMA on behalf of the MN. In addition, the MAG establishes a tunnel with the LMA to enable the MN to use an address from its home network prefix and emulates the MN's home network on the access network for each MN. On the other hand, the LMA is similar to the HA in MIPv6. As in the case of MIPv6, the location update and packet delivery procedures in PMIPv6 are described next.

Starting from a generic architecture of PMIPv6 as is shown in Figure 2.6, in PMIPv6 the mobility support is offered in a portion of the network called Local Mobility Domain (LMD).

When a MN moves into the LMD, it attaches to MAG1, which sends a Proxy Binding Update (PBU) message to LMA to establish a bi-directional tunnel between MAG1 and LMA. This tunnel is used for routing the packets to and from the MN. On receiving the PBU message from MAG1, LMA recognizes that the MN is now under MAG1 so that the LMA can use its binding cache entry of the MN for managing the session and routing information. Then the MN receives a Router Advertisement message from MAG1 which includes the Home Network Prefix (HNP) allocated by LMA. The MN creates its address based on the prefix information. If the MN moves from MAG1 to MAG2, MAG2 also sends a PBU message to LMA and then a bi-directional tunnel between MAG2 and LMA is created for the MN. Because MAG2 also sends the same HNP to the MN, the MN does not observe any IP level mobility, i.e., its IP address remains unchanged. Thus, the MN can move within LMD without participating in any mobility-related signaling. This signaling flow is illustrated in Figure 2.7.

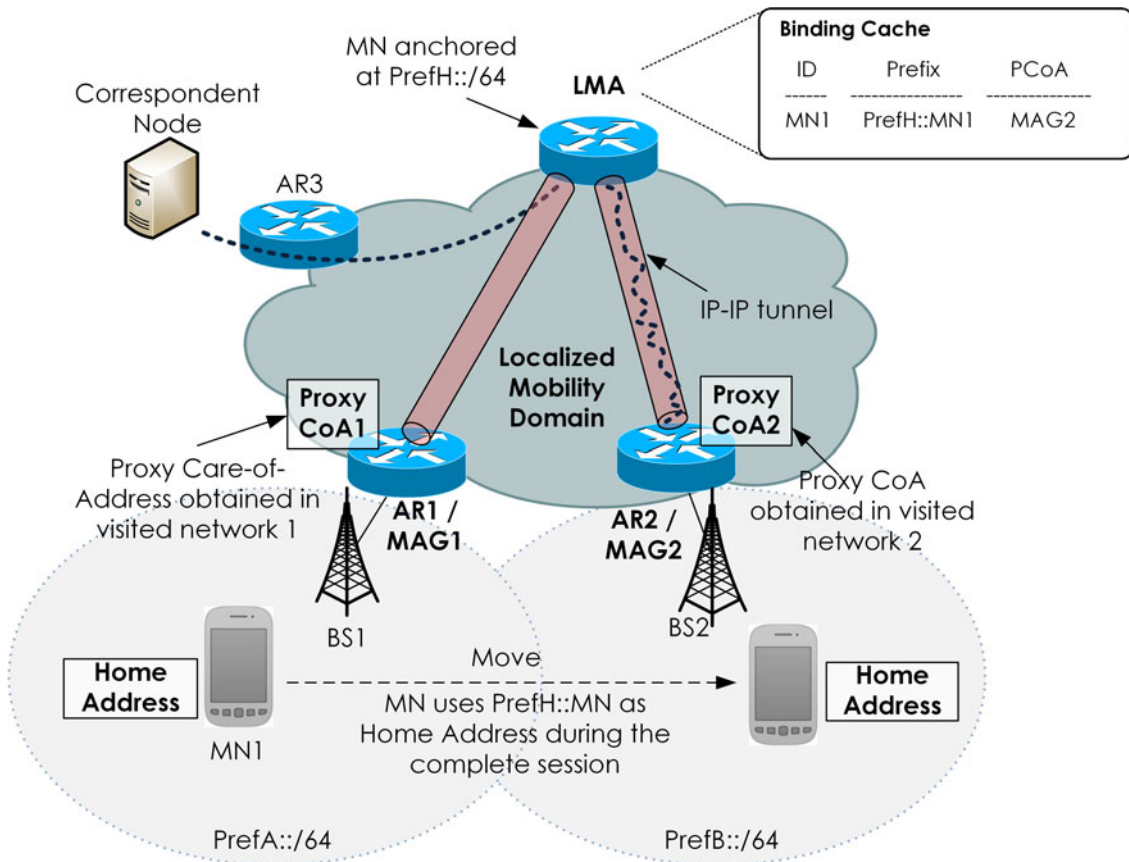


Figure 2.6. Overview of Proxy Mobile IPv6

The packet delivery procedure in PMIPv6 is also based on bi-directional tunnelling. In this case, packets sent from CN are delivered to the LMA according to the IP routing protocol. Based on the binding cache entry of LMA, the packets are forwarded to the serving MAG via the bi-directional tunnel for the MN. Note that the endpoints of the bi-directional tunnel are the address of LMA and the address of MAG, respectively. Finally, the MAG sends the packets to the MN. All the reverse packets are tunneled from MAG to LMA. After removing the tunnel header, LMA routes them to the destination specified in the inner packet header, providing mobility in a transparent way to the IP stack of the mobile node.

Finally, the binding cache is updated whenever the MN roams within a PMIPv6 domain. In that case, the LMA updates the associated Binding Cache Entry (BCE) which contains the MN identifier, the Home Network Prefix and the MN's location, called Proxy Care-of Address (PCoA), which is the MAG's address where the MN is currently attached to.

As well as we mentioned in CMM section, the network-based mobility management protocol adopted for the 3GPP Evolved Packet System has been PMIPv6 [43, 49].

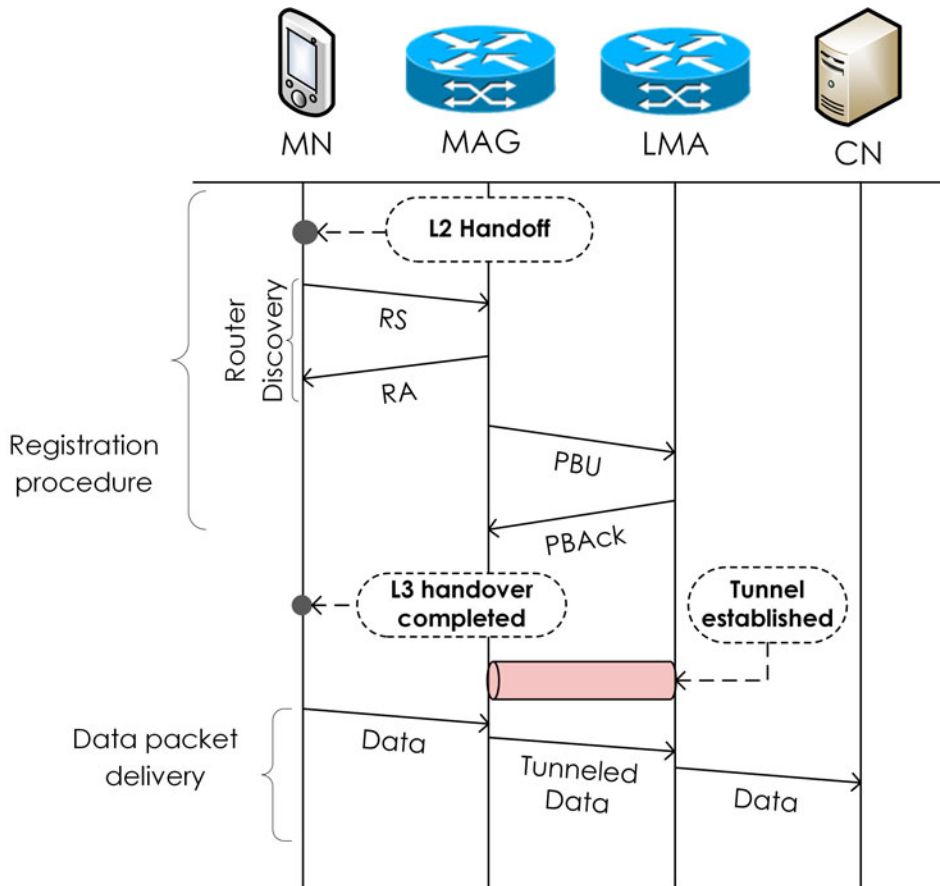


Figure 2.7. Message flow in PMIPv6

2.1.3 Distributed mobility management

The mobility management proposals described in the previous section are based on a centralized mobility agent (HA in Mobile IPv6 or LMA in Proxy Mobile IPv6) that allows a mobile node to remain reachable during its movement. Among other tasks, this anchor point ensures connectivity by forwarding packets destined to, or sent from, the mobile node.

Nowadays, most of the deployed architectures have a small number of centralized anchors managing the traffic of millions of mobile users. This centralized approach brings several limitations such as non-optimal routing [50], scalability problems and reliability [51]:

- **Non-optimal routing:** Since the (home) address used by an MN is anchored at the home link, traffic always traverses the central anchor, leading to paths that are, in general, longer than the direct one between the MN and its communication peer. This is exacerbated by the current trend in which content providers push their data to the edge of the network, as close as possible to the users, as for example by deploying content delivery networks (CDNs). With centralized mobility management approaches, user traffic will always need to go first to the home network and then to the actual content source, sometimes adding unnecessary delay and wasting operator resources.
- **Scalability problems:** Existing mobile networks have to be scaled to support all the traffic traversing the central anchors. This poses several scalability and network design problems, as central mobility anchors need to have enough processing and routing capabilities to be able to deal with all the user traffic simultaneously. Additionally, the entire operator's network needs to be dimensioned to be able to cope with all the user traffic.
- **Reliability:** Centralized solutions share the problem of being more prone to reliability problems, as the central entity is potentially a single point of failure.

Recent IP network usages such as multimedia content access and video streaming contribute to an exponential growth in bandwidth usage [52]. The architectural limitation of centralized topologies requires that data must first be routed to the HA or the LMA (centralized agents) which may be geographically far away from the mobile node, and then tunneled to the mobile node [53]. Therefore, these limitations become clearer when the centralized mobility management needs to support mobile videos, which demand a large volume of data and often require quality of service such as session continuity and low delay.

In order to address the above mentioned limitations of centralized mobility management solutions, a new paradigm has been recently proposed which is gaining momentum: the so-called Distributed Mobility Management. DMM basically develops a new concept for handling mobility, with the main characteristic being that the mobility anchors are placed closer (topologically) to the user, distributing the control and data plane mobility functions among entities located at different places on the core/access network [54].

Similarly to centralized mobility, depending on the role of the mobile node in the handover process, distributed mobility management protocols can be broadly classified in two categories, namely those that require active involvement of the MN and those that do not [55]. Next, we review the main host-based and network-based solutions in DMM.

2.1.3.1 Host-based DMM

With the aim of utilizing existing host-based mobility support protocols, a representative proposal of a DMM solution which is based on Mobile IPv6 is detailed in [56, 57] (Host-based DMM, HB-DMM). HB-DMM extends mobility signaling and reuses many concepts such as the binding cache at the MN, binding cache at the mobility anchor or tunnelling. However the MN's mobility is handled in a different way than in MIPv6. In this case, the authors attempt to improve the performance of mobility support by distributing mobility agents (called Access Mobility Anchor, AMA) at the edge of the access network level and the MN is served by a mobility anchor located in the serving network.

At the beginning, the MN configures its address based on the provided network prefix from the AMA. Then, it registers the configured address to the AMA by sending a BU message. When a MN moves to an adjacent network, the MN configures a new address based on the network prefix obtained from the serving AMA at the new access network, while it keeps the previous address from the origin AMA. When the MN registers by sending a BU message at the new access network, it registers not only the newly configured address, but also informs the previous address to the serving AMA through new signaling messages called Access Binding Update (ABU) and Access Binding Acknowledgement (ABA). Thus, a tunnel is created between the serving AMA and the origin AMA, located in its Home Network and a new address is configured in the MN. As depicted in Figure 2.8, this solution creates multiple tunnels between AMAs and in cases where a high mobility rate exists, the system performance might be critically compromised by the frequent registrations and maintenance of multiple tunnels.

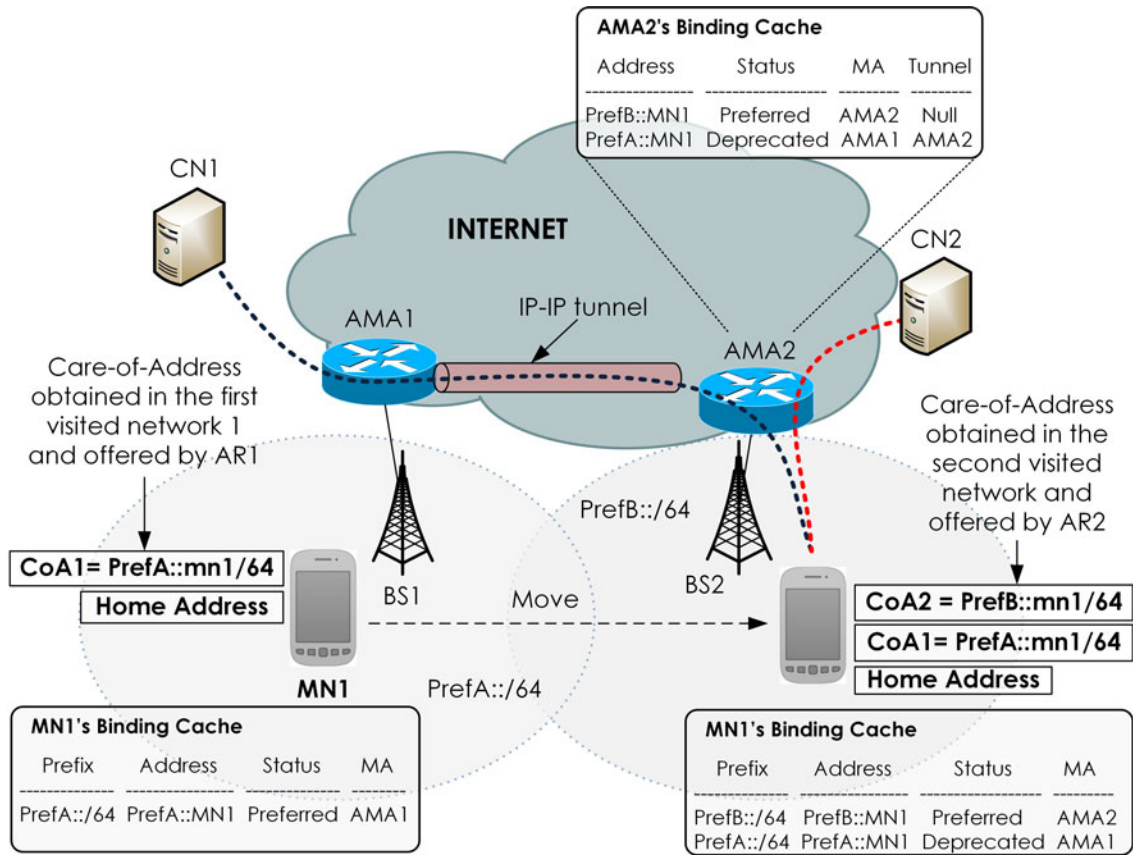


Figure 2.8. Host-based Distributed Mobility Management

2.1.3.2 Network-based DMM

Network-based DMM (NB-DMM) [58] for its part, exempts the MN from participation in any mobility signaling and therefore there is no need for a network software upgrade at the MN for mobility support since distributed mobility anchors perform mobility signaling on behalf of the MN as like PMIPv6. This NB-DMM is one of the early proposals designed in the IETF for network-based DMM at the Distributed Mobility Management Working Group [17].

In NB-DMA, the mobility management functionalities are moved to the Access Routers (ARs) level in order to anchor the traffic closer to the MN. Each AR is required to have both mobility anchoring and location functionalities, and it is referred to as a Mobility capable Access Router (MAR). In NB-DMM, a new session is anchored at the current AR and initiated using the current IPv6 address. When a handover occurs before the end of the session, then the data traffic of this session is tunnelled between the current MAR and the anchoring MAR for this session. In order to achieve a network-based solution without the participation of the MN in the mobility signaling, the architecture is

partially distributed and relies on a centralized database (Mobility Context DB). This DB stores ongoing mobility sessions for the MNs; it stores the home network prefix currently allocated to the MN and their respective anchoring points. Thus, upon a handover, the new MAR retrieves the IP addresses of the anchoring MAR(s) for the MN's ongoing sessions from the database. Then, the new MAR proceeds to location update by sending a Proxy Binding Update to each anchoring MAR. Each anchoring MAR replies by a PBA. The basic operation of NB-DMM is depicted in Figure 2.9.

2.1.4 Distributed hybrid solutions

The evolution from centralized to distributed mobility management approaches has shown clear signs of achieving better utilization of resources in the network, outperforming the traditional protocols and optimizing mobility management performance [57–59]. Despite these facts, there are some scenarios in which the DMM paradigm also incurs high delivery

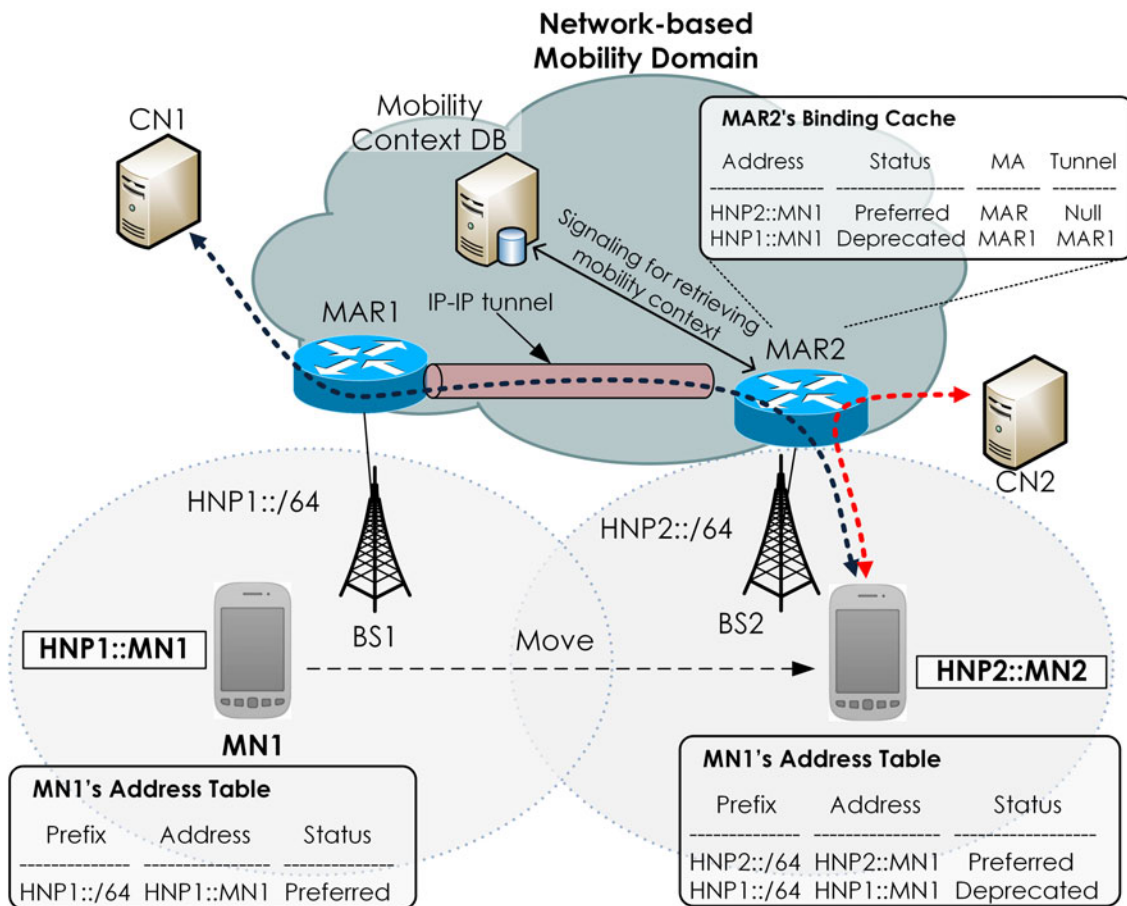


Figure 2.9. Network-based Distributed Mobility Management

costs and possible unacceptable signaling overheads [60].

In the DMM operation, each distributed MAR/AMA that manages an MN keeps a bidirectional tunnel between itself and all MAR/AMA where a session of the MN was originated and is still alive. This situation can be deemed as inefficient in certain circumstances, that generally occur when the movement and the session arrivals are frequent (and its durations are long). In those cases, the DMM approaches set-up through a significant amount of tunnels that can negatively affect performance, and a CMM protocol behavior might be preferred.

To this end, a hybrid CMM-DMM mobility management solution, might provide the benefits of each of them entailing better overall network performance. The mobility management will be managed by centralized or distributed solutions, depending on different parameters, such as, user profile policies, network topology characteristics or the current state of the network in terms of congestion, traffic mix, characteristics of the flow, number of active prefixes, etc).

Thus, from our point of view, hybrid solutions are envisioned as a promising mechanisms to tackle the increasing penetration of mobile devices and the huge amount of data traffic over future mobile networks. Chapter 6 deals with this issue in a detailed way.

2.2 Quality of Service for IP Mobile Networks

One of the major hurdles is in supporting efficient mobility management for the plethora of mobile devices as they move while accessing multimedia rich sessions with stringent QoS requirements [61]. For instance, a disruption of QoS can happen if there is limited capacity in the new cell, slow resource reservation after the handover, etc. Thus, providing a guarantee of resource availability and fast reservation of these resources after the handover is essential to prevent any degradation of ongoing services [62].

The design of future wireless networks has two main goals. First of all, the possibility of maintaining the connectivity while a user moves among heterogeneous networks. Secondly, the ability to provide a similar level of QoS while the node moves between these networks [63]. In order to achieve the first goal, the Internet Engineering Task Force has designed IP mobility management protocols to overcome the problems caused by handover in heterogeneous networks. As we have described in previous sections of this Chapter, both centralized and distributed solutions are being developed for addressing IP mobility management in future wireless mobile networks.

The second problem, related to assuring the provisioning of enough network resources has been largely studied in both wired and wireless environments. There are three general models to provide network resources for quality of service guarantees in the Internet: integrated services (IntServ), differentiated services (DiffServ) and MPLS [64].

IntServ can provide quantitative QoS guarantees to individual flows, DiffServ can provide qualitative QoS guarantees to multiple flows in an aggregate way. For IntServ, a signaling protocol, Resource Reservation Protocol (RSVP), was designed to facilitate reserving resources prior to establishing connections. RSVP is a receiver-initiated protocol which provides QoS guarantees over the Internet. Two main message types exist in RSVP: The Path message and the Resv message. A sender establishes an RSVP session by sending a Path message which contains information regarding the characteristics of the traffic flow to be sent. As this message propagates downstream towards the receiver, it installs Path state in every intermediate router along the way and every router records the IP address of the previous hop router. Once it reaches the receiver, the receiver replies with a Resv message (containing the requested QoS parameters) along the reverse path to the sender. As this message traverses upstream towards the sender, it is intercepted and inspected by every intermediate router. If the required resources are available, an intermediate router sets up a soft-state reservation and forwards the Resv message to the next hop router. Finally, if the required resources on all the links are available, the reservation session with soft state is established, otherwise a ResvErr message will be replied back to the receiver [65].

For its part, MPLS with its traffic engineering (TE) is a QoS technology introduced to enhance the performance of the Internet's datagram model in terms of both management and delivery. MPLS is a scalable routing technique where routing is done by swapping a label on the packet instead of traditional IP destination lookup. In order to distribute the labels, a label distribution protocol is required to maintain the coherence of label bindings across a network. Labels are then used to identify packets through a label switched path (LSP) traversing label switched routers (LSRs). MPLS-TE (Traffic Engineering) attempts to provide a means to manage and enhance network traffic through rigorous analytical studies. RSVP-TE is an RSVP-based label distribution protocol with traffic engineering functionalities. [66]. A typical operation of MPLS-TE is shown in Figure 2.10 and is briefly explained next. In this case, to allow routers in a network to compute Multi Protocol Label Switching Traffic Engineered Label Switched Paths (TE-LSP), the router's addresses must be advertised by OSPF (Open Shortest Path First) with the extensions of Traffic Engineering (OSPF-TE) [67] (1). Once the routes can be calculated with the routing protocol, RSVP-TE [68] is used for resource reservations in the selected path with QoS requirements (2). The results is the instantiation of a label switched tunnel which can be automatically routed away from network failures, congestion, and bottlenecks (3).

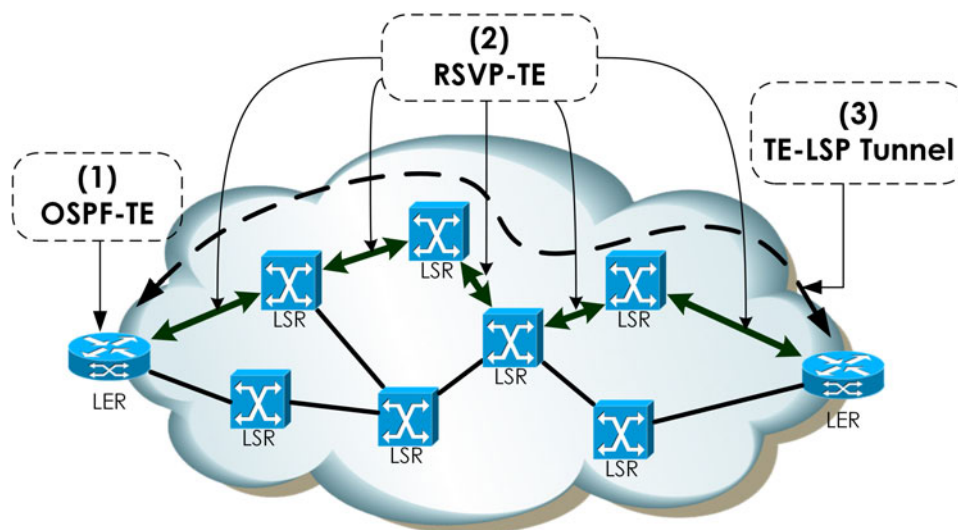


Figure 2.10. Basic MPLS-TE components

2.2.1 Mobility and MPLS integration

The efficient provision of the network resources in the future mobile networks are one of the main goals in the development of the mobility management protocols. These protocols require the use of tunnelling techniques to forward packets during the movement of the mobile nodes. In this perspective, there is an increasing trend towards the use of MPLS in IP-based wireless access networks to benefit from its QoS, traffic engineering and reliability capabilities [69–72].

In order to tackle the QoS provision, Mobile MPLS and Fast Handover Micro Mobile MPLS are interesting solutions proposed in the literature. These schemes are briefly explained next.

Mobile MPLS

Mobile MPLS [73] was one of the first proposals to integrate the Mobile IP and MPLS protocols. It aims to improve the scalability of the Mobile IP data forwarding process by removing the need for IP-in-IP tunnelling from Home Agent (HA) to Foreign Agent (FA) using Label Switched Paths (LSPs). Basically, this solution uses MPLS as a tunnelling technology that outperforms the usual tunnelling technique suggested for Mobile IP.

They first describe an architecture where the FA and HA are edge LSRs and belong to the same MPLS domain. In this architecture, when the MN detects that it is in a foreign network, and after it registers with the local FA, the FA configures its entries, forwards the registration to the HA using regular IP routing, and awaits the HA's LDP (Label

Distribution Protocol) request. After the HA adjusts its MN's binding, it issues the LDP request to the FA, and the LSP is eventually set up. When the MN visits a new FA, the same procedure is repeated. Figure 2.11 illustrates the procedure of Mobile MPLS registration and packet delivery.

Fast Handover Micro Mobile MPLS

Fast Handover Micro Mobile MPLS [63] overcomes some limitations of Mobile MPLS. In this scheme the fast handover mechanism anticipates the LSP procedure setup with an adjacent neighbor subnet that an MN is likely to visit. The main idea behind FH-Micro Mobile MPLS is to set up an LSP before the MN moves into a new subnet to reduce service disruption. In this context, the authors consider active and passive LSPs. The active LSP is the one from the LERG (the root of the MPLS domain) to the current serving LER in the visited network. This LSP is used to transfer data. Passive LSPs are those from the LERG to the neighboring LER of the current foreign agent. These LSPs

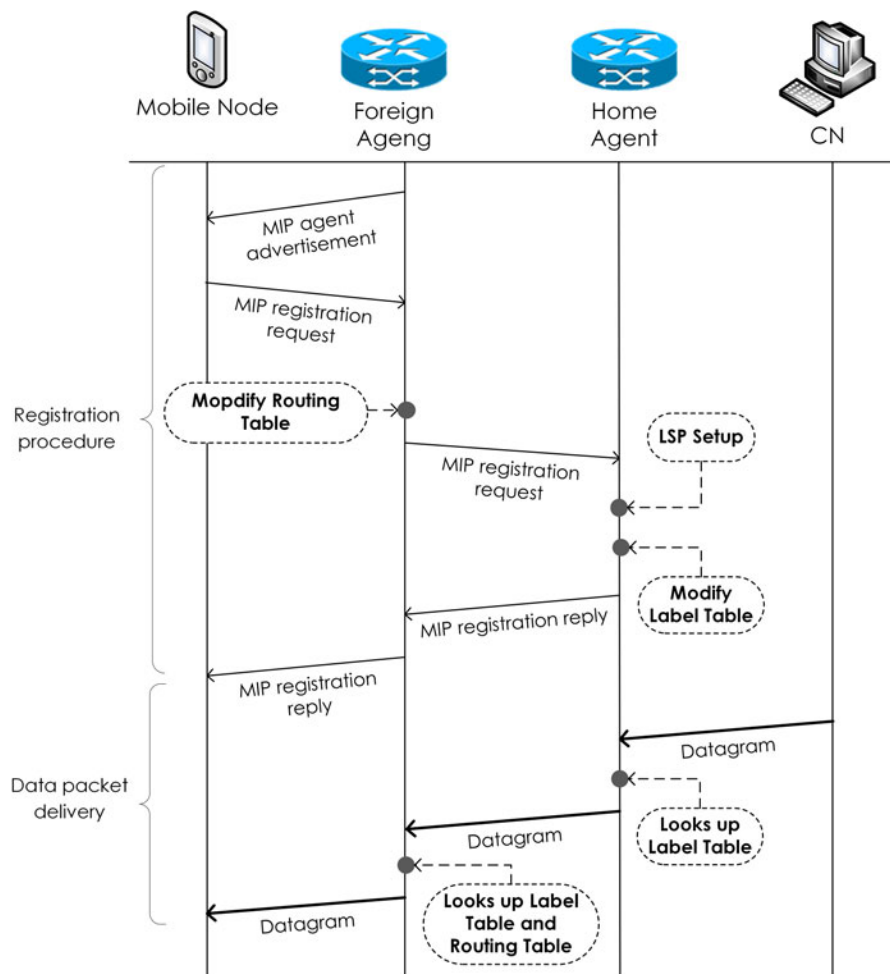


Figure 2.11. Mobile MPLS operation

will not be used except when the MN moves to its own network. In this moment, the MN establishes its new active LSP and passive LSPs with neighboring subnets. A typical architecture of FH-Micro Mobile MPLS is shown in Figure 2.12.

The basic operation of the mechanism is as follows. Once a MN enters an overlapped area of two subnets, it receives an L2 beacon from the possible new BS (Base Station). Immediately, the MN notifies the current FA for possible handover by sending a signaling message. Each LER/FA has a Neighbor Mapping Table (NMT) that binds between the IP and MAC (Media Access Control) address of all neighboring BSs. Hence, when the current FA receives the handover message, it looks into its NMT table to get the new FA's IP address and then informs the LERG for the possible L3 handoff. A passive LSP with the desired QoS requirements will be established between the LERG and the new subnet using the RSVP-TE protocol. At the same time, the current FA informs the MN of the new Regional Care-of Address. Finally, the MN starts the Mobile IP registration process with the LERG.

MPLS-PMIPv6

Both Mobile MPLS and FH-Micro Mobile MPLS are two host-based solutions. A network-based approach that introduces MPLS is MPLS-PMIPv6 [71]. MPLS-PMIPv6 is the first scheme which proposes MPLS as an alternative tunnel technology between a MAG and a LMA. The reasoning behind the idea is that since a tunnel is needed, employing a

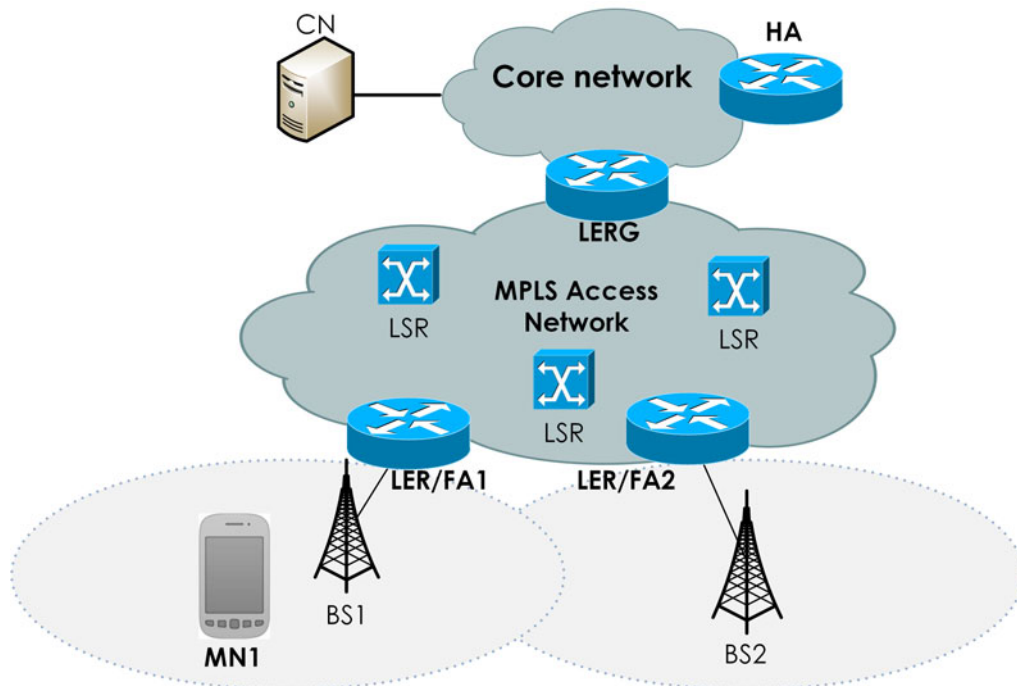


Figure 2.12. Fast Handover Micro Mobile MPLS architecture

technology that natively supports tunnelling seems a natural choice. This work defines some extensions to allow the MAG and the LMA to distribute MPLS labels using Proxy Binding Update and Proxy Binding Acknowledgement messages. Two kinds of labels are employed: a classical tunnel label and a Virtual Pipe (VP) label. The latter is introduced as a means to differentiate traffic with the same MAG-LMA endpoints according to the operators of the various MNs served by the same MAG. The operation of the protocol is described next and the message flow is shown in Figure 2.13.

Once a MN enters a PMIPv6 domain and attaches to an access link, the MAG, after authorise the MN, sends a PBU message to the LMA with the VP label. Next, the LMA records the label as a downstream VP label, which is used for any IP traffic destined to the MN. Based on the MN profile and IP address, the LMA assigns a label for identifying upstream traffic of the MN. Once an IP packet destined to the MN arrives, the LMA locates a Binding Cache Entry based on MN IP address, fetches the downstream VP label, and puts it in front of IP packet. It then identifies the tunnel label based on MN Proxy CoA, it encapsulates the packet with the two labels, and sends it out according to MPLS procedures. Once an IP packets originating from the MN arrives, the LMA pops the tunnel label, stripes the VP label and forwards the packets to the corresponding operator.

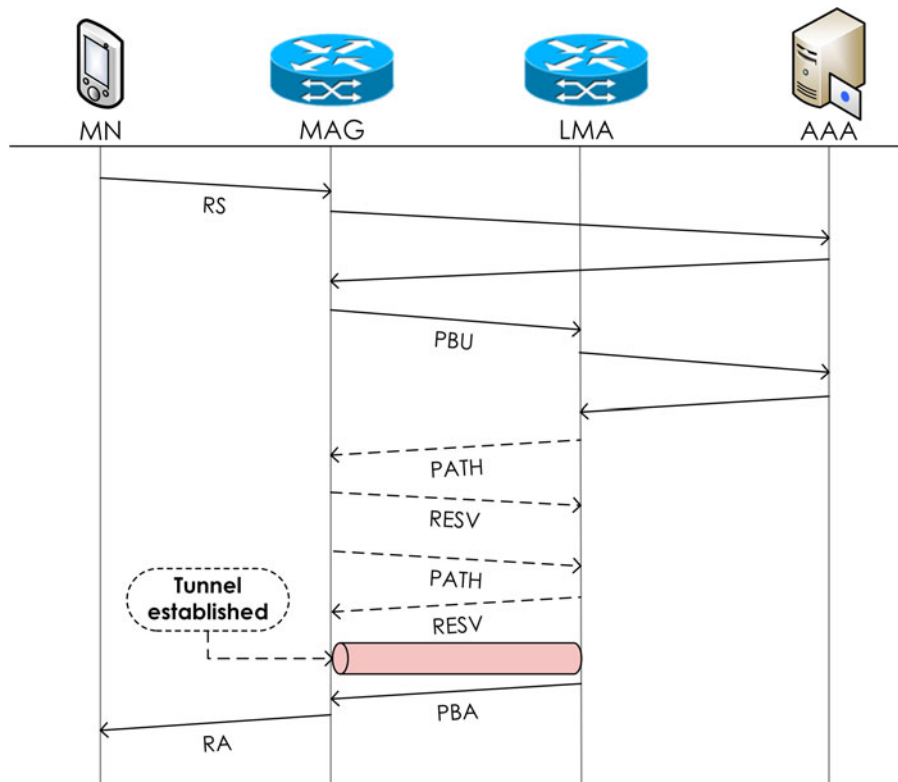


Figure 2.13. Message flow for the MPLS-PMIPv6

2.3 Summary

In this chapter we have presented the concept of mobility management, describing the limitations of the original IP protocols to allow users to roam seamlessly through the Internet. Due to this shortcoming, IPv6 needs mobility protocols in order to manage and handle the IP mobility efficiently. Thus, together with their own IPv6 addressing and some IPv6 procedures such as Neighbor Discovery, these protocols provide a solution to the user's mobility at IPv6 layer.

Then, we have introduced and discussed the main IP mobility protocols developed by the IETF, namely Mobile IPv6 and Proxy Mobile IPv6. These protocols are host-based and network-based respectively and both have a common feature: they implement a centralized approach where end user data traffic is encapsulated between a centralized mobility anchor and the mobile node. This means that all packets associated to a mobile node are first routed to the centralized anchor, so it becomes a single point of failure and a bottleneck affecting network performance by slowing down the end-to-end packet transmission speed.

An alternative mobility management approach that could solve the centralized mobility limitations is to distribute the mobility anchors through the network, locating them close to the users. With this approach, it is expected that the future mobile networks cope with an increasing volume of data, saturating their access links, and triggering the need for additional access technologies to be made available to the users.

Two of the main DMM solutions in the literature have been detailed. HB-DMM and NB-DMM solves the limitations of the centralized solutions but for some scenarios, even CMM are preferred. This is because the DMM deployment also faces several issues such as complex address and tunnel management, high signaling cost and high handover latency as the number of addresses and the number of tunnels associated with the mobile node increase.

Due to the relevance that hybrid CMM-DMM solutions have in this thesis, the benefits that could be achieved in an environment in which the mobility management of some traffic could be kept centralized, while other could be distributed, are discussed.

Finally we focus on the provision of Quality of Service in IPv6 mobile networks. In this case, the integration of mobility and MPLS is also discussed.

In the following chapter we develop the necessary analytical models and analysis metrics. Then, we detail the proposals made in this thesis, each covering an aspect of IPv6 mobility management, seeking to address some of the above mentioned problems:

- Centralized solution: In Chapter 4 we propose the LinkWork Mobile MPLS scheme that deals with the QoS in centralized mobility architectures.
- Distributed solution: Chapter 5 describes our distributed DM3 solution. In this architecture, several nodes are distributed in a MPLS-based access network and the nodes are served by a close-by mobility anchor.
- Hybrid CMM-DMM solution: In Chapter 6, we propose a Hybrid CMM-DMM solution that provides additional flexibility to the mobile network operators, which can decide when and how to combine the CMM or the DMM approach.

3

ANALYSIS AND MODELING

This chapter describes the analytical framework derived in order to model different mobility management schemes. These models are used later for several performance analysis studies.

Contents

3.1	Introduction and contributions	39
3.2	Background in the analysis of IPv6 mobility management	40
3.3	Analytical Model	44
3.3.1	Network model	44
3.3.2	Mobility models	44
3.3.3	Traffic model	45
3.4	Mobility cost	46
3.4.1	Signaling Cost	47
3.4.2	Data packet delivery cost	48
3.4.3	Tunnelling Cost	50
3.4.4	Handover latency	51
3.4.5	Packet loss	52
3.4.6	Buffer size	53
3.5	Concluding remarks	55

3.1 Introduction and contributions

A lot of research in mobility management in IPv6 networks has been carried out in the last years. The design of new and improved protocols, architectures or quality of service techniques are some examples or new research proposals that, day by day, are being developed in the field of computer networks and, particularly, in mobile networking. A crucial step during the design and engineering of communication systems is the estimation of their performance, and the understanding of the behavior of the systems. Typically, this can be realized by applying three different methodologies, such as:

- Experiments with real systems and prototypes
- Mathematical analysis
- Simulation

The development of new research proposals in real systems normally is a complex task that takes a long time, apart from infeasibility due to financial and technical constraints. Although real testbeds deliver detailed and accurate results, sometimes it is difficult to perform many executions varying different parameters due to the long time that the preparation of each requires.

Analytical methods are often adequate to show the borderline behavior of system characteristics of offer upper and lower bounds for specific research questions. Mathematical models can be developed more quickly and result in a good estimation of the performance.

With respect to the latter methodology, simulation is particularly used for systems which are highly dynamic and whose properties are difficult to capture in a mathematical way. The simulated environment offers a controlled environment in which a system can be investigated in more detail. Thus, it can be a powerful and versatile option to analyze the behavior and performance of the communication system.

Although in this thesis we analyze and evaluate the performance of the different IPv6 mobility management protocols through the three mechanisms mentioned above, we mainly focus our efforts in an analytical framework that allows to compare the different mobility management solutions through mathematical models. Thus, we can investigate the strong and weak points of each scheme. In this chapter, we develop such a model capable of comparing the different mobility management solutions through the analysis of relevant metrics that define the behavior of each scheme. The work performed in this chapter is an essential contribution of this thesis.

3.2 Background in the analysis of IPv6 mobility management

With regard to the mobility management in future wireless networks, the literature about performance evaluation of IP-based mobility management schemes is mainly based on simulation and testbed approaches. Only a few works which assess IPv6-based mobility management protocols through analytical models are available. The most representative works that deal with this topic, and on which our analysis is based, are described next.

In [74], a comparative performance analysis studied for MIPv6, FMIPv6 (Fast Handovers for MIPv6) [75], HMIPv6 (Hierarchical Mobile IPv6) [76], and a combination of FMIPv6 and HMIPv6 has been carried out, that identify each mobility management protocol's characteristics and performance indicators. Contrary to previous works, in [74] the authors derive signaling overhead cost, packet delivery cost, binding refresh cost and total signaling cost generated by an MN during its subnet residence time for each protocol. Moreover, the required buffer space, handover latency and packet loss expressions are derived. Numerical results are obtained and the effect of mobility and traffic parameters on these criteria are analyzed. Two of these metrics, signaling traffic and handover latency are the basis of [63], and [77]. In this case, new mechanisms are designed to track efficiently the mobility of the nodes while ensuring the MN's QoS requirements. This is done, in part, through the introduction of MPLS capabilities in the access network. In these two works, the authors develop a new analytical model using Markov chains to derive the protocol performance metrics.

The rapid increase in mobile data traffic and, especially the forecasts that envision an exponential growth over the next few years, has made that analysis of mobility cost a relevant issue in last years. The work carried out in [78] is focused in the analysis and comparison of the different IPv6 mobility management protocols in terms of cost analysis. Thus, the strengths and weaknesses of each mechanism are identified. This analysis also includes a network-based mobility management solution (PMIPv6), as well as MIPv6 and FMIPv6. This work was extended in [57], where authors include FPMIPv6 (Fast Proxy Mobile IPv6) and analyze the performance of the protocols in terms of handover latency, handover blocking probability and packet loss.

In recent years, the analysis of existing approaches in mobility management have followed the mathematical models of the aforementioned works, with the inclusion of distributed mobility management solutions. In [79], the authors perform a quantitative analysis of CMM and DMM in order to compare both paradigms in terms of registration delay, signaling overhead and traffic intensity. A similar analysis is done in [60], where the analytical model looks at signaling and packet delivery cost, and introduces the handover latency. Following the same framework, a similar analysis is done in [58] for both

cost metrics (considering signaling cost, processing cost, data packet delivery cost, and tunnelling cost) and handover metrics (handover latency, handover packet loss and handover failure probability).

Other papers introduce some novelties such as [80], that describes a novel analytical model for comparison of various mobility management protocols in terms of handover latency, as well as packet density, and packet arrival rate during the handover time by applying transport engineering principles. Similarly, the authors in [81] compare distributed versus centralized mobility management through analytical modeling. This analysis is focused on the number of the mobility-related contexts and IP tunnels that need to be managed in the cellular network. The novelty of this work is that the metrics used in this study are not typically considered in earlier evaluations.

Table 3.1 shows a summary of the most relevant works which perform an analytical study of the mobility management protocols. With the phenomenal growth of the mobility research, it is imperative to understand methods to analyze such networks, visualize and extract useful information. This table tries to give a comprehensive understanding about this topic during last years. It can be see how the trends in the analysis of IPv6 mobility management protocols has changed from CMM to DMM whereas the metrics have been quite similar during this period.

In this table, each row corresponds to an specific work in the literature. First column of the table refers to the article in which that analysis has been published. Column *Protocols* refers to the protocols reported in that publication. Next column (*CMM/DMM*) is about the mobility management paradigm investigated, namely centralized mobility management, distributed mobility management or both. The column *Metrics* correspond with the specific parameters analyzed in the analytical framework. With respect to the *Numerical results* column, here it is detailed how researchers try to offer the resulting data from the analytical model. As much *Metrics* as *Numerical results* offer very useful information, because they give an overview about the parameters, methods and results obtained from the investigation of each analytical framework. Lastly, the column *Obtained results* show how the numerical results have been obtained from the mathematical model. They can be derived directly from the analytical model, by experiments with real testbed or prototypes, or through simulations. In this latter case, the simulation tool used is also mentioned.

Table 3.1. Summary of the main analytical frameworks in the literature

	Protocols	CMM/DMM	Metrics	Numerical results	Obtained results
Makaya et al., 2008 [74]	MIPv6, FMIPv6, HMIPv6, F-HMIPv6	CMM	Signaling cost (total, binding update and binding refresh), Packet delivery cost, Buffer size, Handoff latency, Packet loss	Signaling cost vs. session-to-mobility ratio Binding refresh cost vs. binding lifetime period Packet delivery cost vs. packet arrival rate Packet delivery cost vs. prediction probability Buffer space vs. Packet arrival rate Handoff latency vs. Wireless link delay	Analytical
Langar et al., 2008 [63]	FH, FC, MFC-MM MPLS, FMIPv6, MIP-RR, Mobile MPLS, H-MPLS	CMM	Link usage cost, Registration update cost, Hand-off latency, Packet loss, Buffer size	Link Usage cost vs. hop distance Registration Update Cost vs. hop distance Registration update Cost vs. Call-to-Mobility ratio Total lost packets vs. FA Resident time	Analytical and simulated (ns-2)
Langar et al., 2009 [77]	Adaptive MRA, FMIPv6, MIP-RR, Mobile MPLS	CMM	Link usage cost, Registration update cost, Hand-off latency, Packet loss, Buffer size	Link Usage cost vs. hop distance Link usage cost vs. Domain's radius Registration Update Cost vs. Maximum delay Registration update Cost vs. Domain's radius Total signaling cost vs. Maximum delay Total lost packets vs. FA Resident time	Analytical and simulated (ns-2)
Lee et al., 2010 [78]	MIPv6, FMIPv6, HMIPv6, PMIPv6	CMM	Signaling cost, Packet delivery cost, Tunnelling cost, Total cost	Signaling cost vs. MN's velocity Packet delivery cost vs. Session arrival rate Signaling cost vs. Radius Packet delivery cost vs. Indirect path routing ratio Tunnelling cost vs. session arrival rate Tunnelling cost vs. Indirect path routing ratio Tunnelling cost vs. session to mobility ratio Total cost vs. session to mobility ratio	Analytical
Lee et al., 2013 [57]	MIPv6, FMIPv6, FPMIPv6, HMIPv6, PMIPv6	CMM	Handover latency, Handover blocking probability, Packet loss	Handover latency vs. frame error rate Handover blocking probability vs. velocity Handover blocking probability vs. frame error rate Handover blocking probability vs. radius Packet loss vs. frame error rate	Analytical and simulated (ns-2)

Continued on next page

Table 3.1 – continued from previous page

	Protocols	CMM/DMM	Metrics	Numerical results	Obtained results
Lee et al., 2013 [79]	MIPv6, PMIPv6 HB-DMM NB-DMM	CMM and DMM	Registration delay, Sig- naling overhead, Traffic intensity	Registration delay vs. Number of hops MN-HA Signaling overhead vs. Resident time Traffic intensity vs. ongoing communication session	Analytical
Ali-Ahmad et al., 2014 [58]	PMIPv6 DMM proposal (DMA)	CMM and DMM	Signaling cost, Packet de- livery cost, Tunnelling cost, Processing cost, Handover latency, Han- dover failure probability, Packet loss	Signaling cost vs. number of CNs at a time Processing cost vs. number of CNs at a time Packet delivery cost vs. number of CNs at a time Tunnelling cost vs. number of CNs at a time Signaling cost vs. cell's radius Packet delivery cost vs. cell's radius Signaling cost vs. Network scale Packet delivery cost vs. Network scale Handover latency vs. Prob. Failure of wireless link Handover latency vs. Network scale Handover failure probability vs. MN's speed Packet loss vs. Number of CNs at a time	Analytical
Giust et al., 2014 [60]	PMIPv6 DMM proposal	CMM and DMM	Signaling cost, Packet de- livery cost, Handover la- tency, Packet loss	DMM vs. PMIPv6 signaling cost DMM vs. PMIPv6 packet delivery cost DMM vs. PMIPv6 packet loss CDF of DMM handover latencies vs. time	Analytical and ex- perimental testbed
Vasu et al., 2014 [80]	MIPv6 FMIPv6 HMIPv6 PMIPv6 FMIPv6 HB-DMM NB-DMM	CMM and DMM	Signaling cost, packet loss	Average hop latency vs. Wireless link delay Handover delay vs. packet density arrival rate Handover delay vs. packet arrival rate Signaling cost vs. number of link changes Signaling cost vs. session to mobility ratio Signaling cost vs. probability of link failure Packet loss vs. packet arrival rate Packet loss vs. packet density Handover latency vs. packet density	Analytical and sim- ulation
Munir et al., 2014 [81]	PMIPv6 DMM (DMA)	CMM and DMM	Number of contexts, number of tunnels	Required number of anchor contexts Required number of visitor contexts Required number of tunnels in the network	Analytical and sim- ulations

3.3 Analytical Model

3.3.1 Network model

A communication network can be defined as a directed graph $G = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} denotes the set of nodes (vertices) and \mathcal{E} denotes the set of links (edges) interconnecting the nodes. Let $m \in \mathcal{V}$ be the set of routers that serve as the Mobility Anchors for the mobile nodes, $\mathcal{K} \subseteq \mathcal{V}$ be the set of access routers in the network and \mathcal{N} the set of mobile nodes moving around in the network. Each node n_i ($1 \leq i \leq |\mathcal{N}|$) is equipped with network interfaces that enables them to be reachable through wireless technologies. We further assume a number of base stations belonging in the set \mathcal{B} that provide full coverage in the scope geographical area under consideration.

In this scenario, the access routers (AR) are the first hop router, it is the MN's point of attachment at the radio and the IP levels. The CNs, without loss of generality, are assumed to be stationary for simplicity. We denote by h_{x-y} the average hop distance, i.e. average number of hops, between two network entities x and y . The average hop distance is assumed to be symmetric, i.e. $h_{y-x} = h_{x-y}$.

Note that for centralized mobility schemes, the number of mobility anchors is $|m| = 1$.

3.3.2 Mobility models

Based on our network model, an MN undergoes an IP-handover when crossing from one AR to another. The handover probability between adjacent ARs is given by the \mathcal{H} matrix, as follows,

$$\mathcal{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1k-1} & h_{1k} \\ h_{21} & \ddots & & & h_{2k} \\ \vdots & & \ddots & & \vdots \\ h_{k-11} & & & \ddots & h_{k-1k} \\ h_{k1} & h_{k2} & \cdots & h_{kk-1} & h_{kk} \end{pmatrix}$$

Each element of this matrix h_{ij} is the probability of handover occurring between ARs i and j . The handover probability matrix can be obtained for a given network from network traces and statistics and are normally known to a mobile operator.

As with respect to the mobility model, we assume the Random Waypoint (RWP), since it is the most widely used mobility model in the literature, due to its simplicity, realistic and ease of implementation [82]. RWP model is a synthetic model that describes the movement pattern of independent MNs on a finite continuous plane. In RWP, a mobile node moves from one waypoint to the next waypoint by randomly choosing its destination coordinates, its speed of movement, and the amount of time that it will pause when it reaches the destination. On reaching the destination, the node pauses for some time distributed (Θ) according to some random variable and the process repeats itself. Once the pause time expires, the node chooses a new destination, speed, and pause time. In this case, an one-dimensional line segment $[0, \phi]$ is considered to calculate the expected distance between one waypoint to the next waypoint $E(L)$. According to [83], $E(L)$ is as follows:

$$E(L) = \frac{1}{3}\phi$$

If the velocity of an MN v is constant and $v > 0$ during its entire movement process, then the expected transition time $E(T)$ is:

$$E(T) = \frac{1}{v}E(L)$$

Let $E(C)$ denote the number of subnet crossings during the transition. By using the previous equations, we can estimate the average residence time $E(R)$ of the MN in a subnet as follows:

$$E(R) = \frac{E(T) + \Theta}{E(C)}$$

Apart from the RWP mobility model, in order to drive the evaluation in a more realistic scenario, our simulations are also performed with real-world mobility tracks logs obtained from users carrying GPS receivers. This issue will be described in detailed when the simulation scenario is presented.

3.3.3 Traffic model

As with user mobility, traffic models are crucial for efficient system design and performance evaluation. In wireless networks, although the incoming calls or sessions follow the Poisson process (i.e., inter-arrival time are exponentially distributed), the inter-session arrival

times may not be exponentially distributed [84]. Other distribution models like hyper-Erlang, Gamma and Pareto have been proposed to model various time variables in wireless networks. However, performance evaluations reported in [84] show that exponential model can be appropriate for cost analysis. In fact, exponential model provides an acceptable trade-off between complexity and accuracy [74]

Taking these facts into account, we consider a scenario where a MN might be actively engaged simultaneously with several CNs in the Internet, i.e., having several active sessions. Without loss of generality we assume that the sessions from a MN are generated follow a Poisson process with mean rate λ_s (i.e. the inter-arrival time between sessions is exponentially distributed with this rate). We assume also that the duration of a typical session is exponentially distributed with mean rate μ_s . By modeling the scenario as a system under the probability distribution of $M/M/1$ queue, the average number of active sessions at a time is equal to λ_s/μ_s [85], and the average number of users at a time, η_{cn} , can be expressed as follows:

$$\eta_{cn} = \frac{\lambda_s}{\mu_s - \lambda_s}$$

3.4 Mobility cost

One of the most common criteria for evaluating the efficiency of the different IPv6 mobility management solutions is to use some performance metrics as it was described in Section 3.2. Different cost metrics related with the mobility operation of a protocol can be measured by the mathematical models in order to identify the benefits or weakness of each solution. Moreover, these analysis becomes more important due to the rapid growth of mobile Internet traffic, which is expected to continue increasing with an exponential behavior in the next years [52]. This increase in demand is even more significant in the control traffic, which is expected to grow even more than three times faster than mobile data traffic in next years through 2018 [86].

In this environment of traffic explosion, it is more crucial to manage communications resources efficiently. For that reason, the analysis of both control and data planes are critical in order to evaluate the efficiency of each mobility scheme that makes possible the development of solutions that optimize the network resources and their consumption.

The mobility costs considered in this thesis are introduced next. Thus, in this chapter these costs are described in a general way and, the mobility costs for each solution are derived in following chapters.

3.4.1 Signaling Cost

As we have described in previous chapter, one of the main functionalities for any IP mobility management protocol is the process of maintaining the MN's mobility session up to date while a MN moves among subnets. Such tasks require control messages that needs to be sent among the mobility agents in the network. Therefore, an important performance metric is the cost associated with it. In the following discussion, we refer to the aggregate signaling cost of registration update for a mobility management protocol $MMProt$ as $C_s(MMProt)$.

In general, a mobility management protocol requires that an MN sends a location update to its mobility anchor whenever it moves from one subnet to another one. This location registration is required even though the MN does not communicate with others while moving. This signaling cost associated with location updates may become very significant as the number of MNs increases. Moreover, this cost depends on the size of the signaling messages and the number of hops in every level 3 handover process during the time interval that the MN communication remains active. Therefore, a general expression of the signaling cost is represented in Eq. 3.1:

$$C_s(MMProt) = C_{registration} + N_h \cdot C_{handover} + \gamma_{MMProt} \cdot C_{refresh} \quad (3.1)$$

where $N_h = E(R)/t_s$ is the average number of level 3 handover in a session, γ_{MMProt} is the mean rate of the signaling refresh time of the protocol $MMProt$, required to periodically refresh the binding.

$C_{registration}$ is the cost associated to the initial registration to the network.

$C_{handover}$ is the cost associated to all messages related with a handover. In this term is considered the binding update cost after a handover as well as the cost for terminating a prefix that is no longer active.

$C_{refresh}$ is the cost associated to the binding refresh, necessary to maintain the bindings active.

Both costs $C_{handoff}$ and $C_{refresh}$ can be expressed in a general form as follows,

$$C_{process} = \sum_{p=1}^n (s_p \cdot h_{i-j})$$

where n is the number of required messages to accomplish the process, s_p is the size (in Bytes) of the signaling packet p and h_{i-j} is the number of hops traversed by the packet p from i to j .

This generalization of the signaling cost for a common IPv6 mobility management protocol is shown in Figure 3.1.

In the literature, $C_{refresh}$ and the cost associated to the initial registration and de-registration of the MN when it leaves the domain is usually omitted.

3.4.2 Data packet delivery cost

Regarding the data plane, one of the metrics that has a major impact on the overall performance of the network is the data packet delivery cost. During the movement of a MN in a IPv6 mobility scenario, apart from the signaling necessary to manage the mobility process, data packets are sent to the MN by a correspondent node. In centralized solutions, they are addressed to the centralized anchor, causing a bottleneck and a unique point of failure to all the visited nodes in the domain. Distributed mechanisms should mitigate the problems of mobile operators when coping with the foreseen increase in users' traffic.

In packet based networks, the transmission cost between two entities is proportional to the number of hops between these entities. Thus, the routing path that follows a packet between the CN and the MN will be one of the key factors that affects this cost.

With the analysis performed in this section, we develop a framework to evaluate the network load in terms of total data packet delivery cost for a session. This metric is defined as C_{PDC} and its value is controlled by the size of the data messages multiplied by the number of hops needed to forward packets from the CN to the MN or vice versa h_{CN-MN} . In this cost, the tunnelling overhead, if exists, is also included. In general, this cost can be represented as follows,

$$C_{PDC}(MMProt) = \sum_{p=1}^n ((s_d + s_t) \cdot h_{i-j}(p)) \cdot \lambda_d \quad (3.2)$$

where p represents each *pathSegment* in which the complete path between MN and CN can be divided into (n is the number of segments). For its part, s_d and s_t are the average size of a data packet and average overhead added by the tunnelling mechanism respectively. Finally, the λ_d is the transmission rate for a downlink packet.

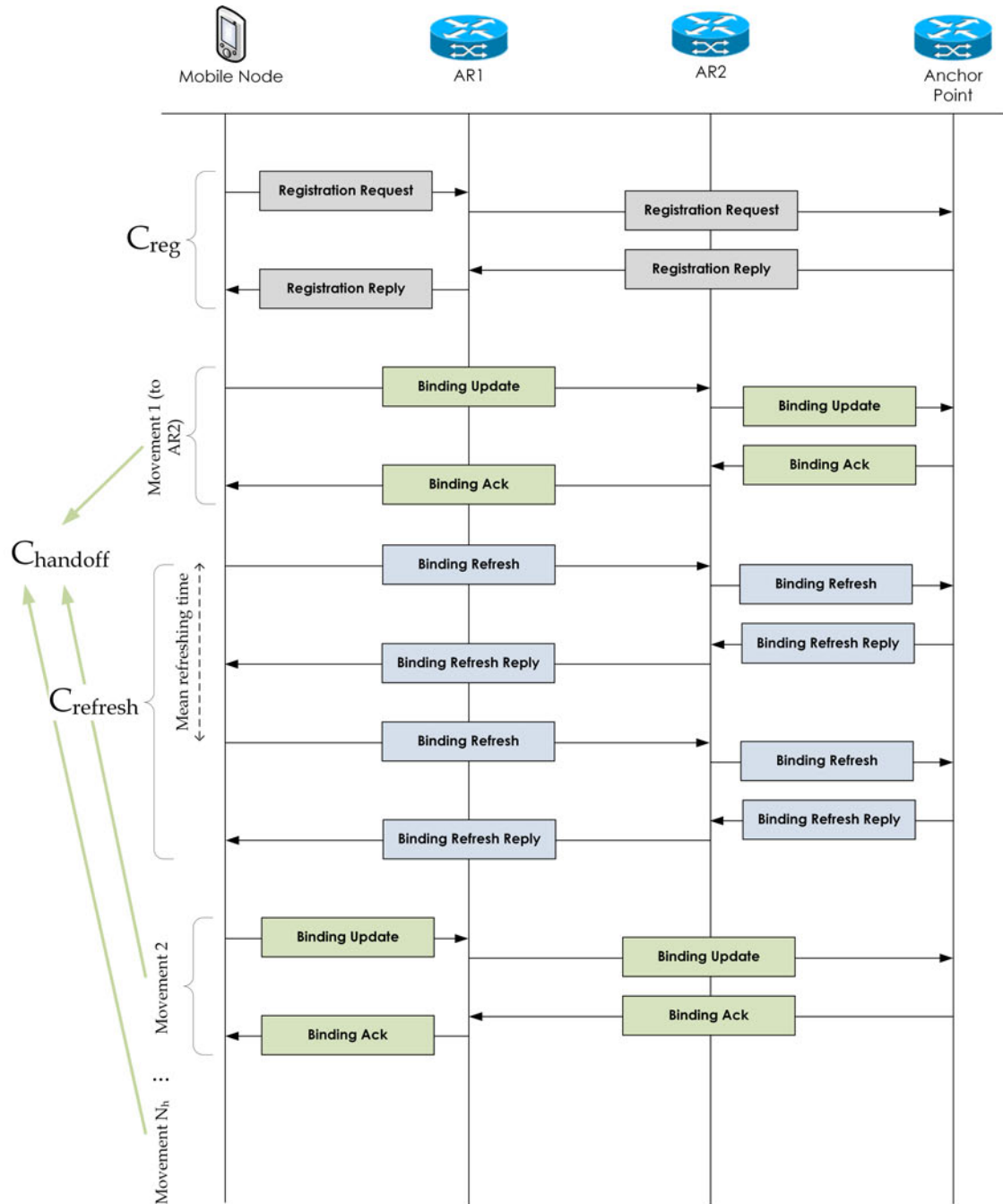


Figure 3.1. A general signaling cost message exchange for a mobility management protocol

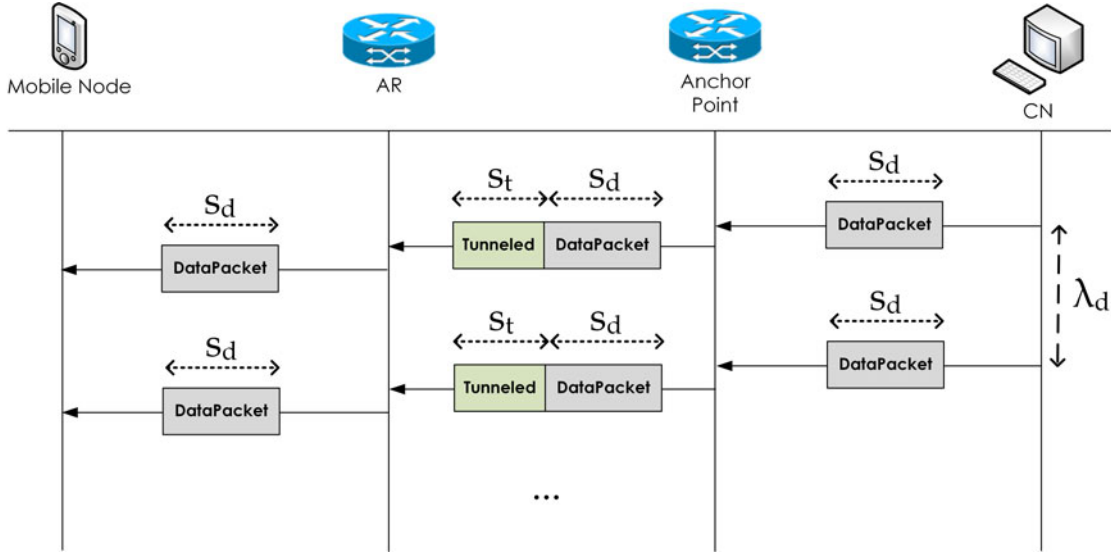


Figure 3.2. A general delivery cost diagram for a mobility management protocol

3.4.3 Tunnelling Cost

To achieve a seamless mobility support, mobility management protocols use a tunnel to forward/re-direct packets. Depending on the operation of each proposal, a certain quantity of those packets will be tunneled and, therefore, the delivery cost will be penalised with the tunnelling overhead.

The tunnelling cost C_t metric represents in essence the cost of adding a tunnelling overhead to the overall data packet delivery cost. So, the tunnelling cost can be derived from packet delivery cost by setting the payload size of the packet to zero, $s_d = 0$

$$C_t(MMProt) = \sum_{p=1}^n (s_t \cdot h_{i-j}(p)) \cdot \lambda_d \quad (3.3)$$

As it was described, packet delivery cost and tunnelling cost are two metrics involved during the communication of data traffic. Figure 3.2 shows a message exchange diagram of a possible data traffic exchange in a generic mobility management protocol.

In this case, the CN sends data packets to the MN at a mean rate λ_d and the complete path between the CN and the MN can be divided into three segments ($p = 3$). In the first one, the data packets are addressed to the mobility anchor point of the MN. Next, the anchor point intercepts these data and encapsulates them inside packets that are addressed to the next agent, in this particular example, the serving AR of the MN. Note that in this segment, the packets are tunneled through the network from the anchor point

to the AR and an extra overhead is considered (s_t). Finally, in the last segment the data packets are de-encapsulated and sent to the MN.

In the example described previously, the tunnelling cost (C_t) only is affected by the second segment, in which the traffic is encapsulated, whereas the packet delivery cost (C_{PDC}) takes into account the cost of the three segments.

3.4.4 Handover latency

Other critical metric that has a huge impact in the performance of the system is the handover latency, T_h , that we can define as the time interval in which an MN does not have global IP connectivity as a result of a handover. This handover process is caused by the nature of the mobility when an MN changes its point of attachment to the network and a disruption time exists.

In order to compute the handover latency, some parameters need to be defined:

- t_{L2} is the L2 handover latency.
- t_{sec} is the time needed to perform the security and authentications tasks required in the system.
- t_e is the time to establish the new route to the MN and to receive the first forwarded data packet through the new path. In this phase, the mobility bindings messages are exchanged.

Thus, T_h can be divided into these three phases as it is shown in 3.3, and can be expressed as,

$$T_h = t_{L2} + t_{sec} + t_e \quad (3.4)$$

If we consider in our analysis the layer 3 operations, both t_{L2} and t_{sec} are out of the scope of this thesis. The value of t_{L2} is heavily dependent on the wireless technology deployed in the system, whereas t_{sec} refers to the security operations that are not mandatory related with the operations of the mobility management protocol. Thus, we can assume that those terms would be identical for any mobility management protocol. Hence, they are omitted in the subsequent analysis.

Moreover, we define $t(s, h_{i-j})$ as the time that takes a message of size s to be forwarded from i to j through the wired and wireless links. $t(s, h_{i-j})$ can be expressed as follows:

$$t(s, h_{i-j}) = c + h_{i-j} \times \left(\frac{s}{B_w} + L_w \right) + (h_{i-j} + 1) \times P_t$$

where

$$c = \begin{cases} \frac{s}{B_{wl}} + L_{wl} & \text{if } i = \text{MN} \\ 0 & \text{otherwise} \end{cases}$$

Other parameters that appear in the previous expression and need to be defined are the following. B_w is the bandwidth of the wired link, B_{wl} the bandwidth of the wireless link, L_w defines the Latency of the wired link (propagation delay) and L_{wl} is the latency of the wireless link (propagation delay). Finally, P_t is the routing or label table lookup and processing delay.

Thus, the establishing time t_e can be defined as the sum of two terms:

- the time used by a MN to send a router solicitation message in order to receive the router advertisement (RA) rapidly. RS and RA are part of the neighbor discovery protocol [31]. This time can be approximated as half RTT (Round Trip Time) between the MN and the Access Router in the new visited network.
- $t_{binding}$ is the time required for the location update message, i.e., binding update request and reply messages.

$$t_e = \frac{1}{2}RTT_{MN-AR} + \sum_{procedure(p)} (N_p \cdot t(s, h_{i-j}))$$

where for each procedure p a certain number of messages N_p are needed to perform the establishment. Moreover, each of these messages of size s traverse h_{i-j} between the nodes i and j and it takes a certain time $t(s, h_{i-j})$ to its completion.

Figure 3.3 shows the timing diagram for a generic mobility management handover. In this figure, the different operations in which a handover can be divided are shown.

3.4.5 Packet loss

Related with the handover, another relevant metric is the packet loss during handover latency or service disruption latency. In fact, the number of packet loss during a handover is directly proportional to the handover latency as analyzed in the previous section, and also proportional to the packet arrival rate.

During the handover process, data packets sent from the CN, will be lost if there is not any buffering mechanism. If this mechanism exists, the number of packets that will be lost during the handover are going to be minimized. In this latter case, the packet loss

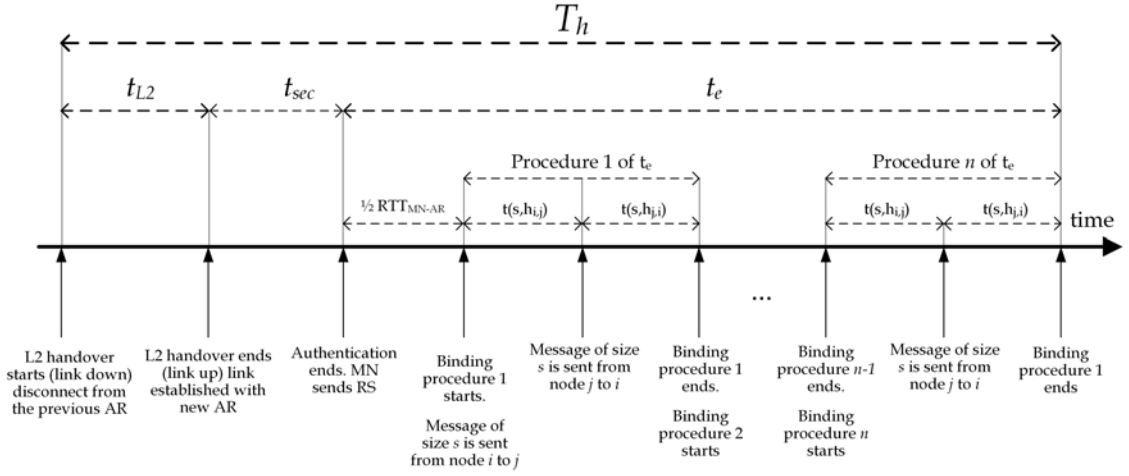


Figure 3.3. Generic timing diagram for a mobility management handover

will be proportional to the time needed to initiate the buffering mechanism. Thus, the metric defined in this section is the total Packet Loss ($P_{loss}(MMProt)$). For a certain mobility management protocol ($MMProt$), it is expressed as follows,

$$P_{loss}(MMProt) = T_h(MMProt) \cdot \lambda_d \cdot N_h \quad (3.5)$$

on the other hand, if a buffering mechanism exists, the expression that describes the behavior of P_{loss} is:

$$P_{loss} = t(s, h_{(i-j)}) \cdot \lambda_d \cdot N_h \quad (3.6)$$

3.4.6 Buffer size

Besides restoring the IP links after a handover, the mobility mechanisms ensure that the ongoing sessions are not disrupted and lost. In general, packet losses may occur during the handover as we have described in previous section. Without any buffering mechanism, data packets sent from the CN to the MN will be lost while the MN performs its handover so, a mechanism that avoids or minimizes those lost is required to store in-flight packets.

For that reason, some protocols include a buffering mechanism in order to minimize the packet loss during the disruption time due to a movement. The buffer space required for a mobility management protocol during a handover is proportional to handover latency [74]. Hence, the buffer size (B_s) can be defined as the time needed to start the buffering mechanism multiplied by the packet arrival rate and is expressed as follows

$$B_s = \frac{1}{2}RTT_{MN-AR} \cdot t(s, h_{(i-j)}) \cdot \lambda_d \quad (3.7)$$

With the aim to facilitate the readability of following chapters and to summarize the symbols introduced in this chapter, Table 3.2 shows the parameters and their description in the scope of this thesis.

Table 3.2. Parameter definition and description

Parameter	Description
$G = (\mathcal{V}, \mathcal{E})$	Direct graph that defines the network
\mathcal{V}	Node set (vertices)
\mathcal{E}	Link set (edges)
$m \in \mathcal{V}$	Set of routers that serve as MA
$\mathcal{K} \subseteq \mathcal{V}$	Set of AR in the network
\mathcal{N}	Set of MNs
\mathcal{B}	Set of BSs
h_{x-y}	Number of hops between x and y
h_{ij}	Handover probability from i to j
λ_s	Inter-arrival time between sessions
λ_d	Transmission rate for a downlink packet
μ_s	Session duration time
η_{cn}	Average number of users at a time
$C_s(MMProt)$	Signaling cost
$C_{PDC}(MMProt)$	Packet delivery Cost
$C_t(MMProt)$	Tunnelling Cost
$T_h(MMProt)$	Handover time
$P_{loss}(MMProt)$	Packet loss
N_h	Average number of level 3 handover in a session ($N_h = t_s/t_r$)
s_p	Size (in Bytes) of the packet p
s_u	Size of the BU message sent from the MN to the HA or vice versa
s_{pu}	Size of the PBU message sent from the LMA to the MAG or vice versa
s_d	Average size of a data packet
s_t	Average size of the IPv6 tunnel header size
s_m	Average size of the MPLS tunnel header size
s_l	Average size of a message for LSP establishment
$E(R)$	Average residence time
$t(s, h_{i-j})$	Time that takes a message of size s to be forwarded from i to j
t_r	Average stay time at a visited network
P_t	Routing or label table lookup and processing delay
B_w	Bandwidth of the wired link
B_{wl}	Bandwidth of the wireless link
L_w	Latency of the wired link (propagation delay)
L_{wl}	Latency of the wireless link (propagation delay)

3.5 Concluding remarks

In this chapter, we have introduced the analytical model used in this thesis. We have proposed a generic comprehensive model to be applied for any IPv6-based mobility management protocol in order to provide a framework over which the overall performance of the protocols can be obtained. We have derived the expressions of several metrics such as signaling cost, data packet delivery cost, tunnelling cost, handover latency, packet loss and buffer size. From this analysis, the numerical results presented in next chapters are obtained.

As we described at the beginning of this chapter, apart from the analytical evaluation, other options exist to evaluate the performance of the mobility management protocols. With the aim of providing a complete performance evaluation of mobility management protocols, both simulation and experimental evaluations are conducted. Thus, the next chapter presents an experimental evaluation of MIPv6 and PMIPv6 protocols. We have performed experiments using a Linux-based prototype, which is used to analyze the handover latency through a realistic scenario.

Then, the next three chapters present our three solutions that can be adequate for next generation wireless networks. First of all, a centralized mobility management protocol based on MIPv6 that can be suitable for 4G or beyond 3G wireless networks. Furthermore, DMM is envisioned as a promising candidate for mobility management in future 5G networks. Based on that distributed paradigm, we propose DM3, a distributed mobility management mechanism and Hybrid DMM, that can be valid solutions to cope with the requirements of future wireless networks.

4

LINKWORK MOBILITY MANAGEMENT MPLS: A CENTRALIZED PROPOSAL

The tunnelling method provided by MPLS can be profitably used in next generation wireless networks. Moreover, service disruption during handoffs cause excessive packet loss that needs to be minimized to support quality of service (QoS) requirements of emerging applications. In this context, we propose a new architecture called LinkWork Mobile MPLS that minimizes packet loss and avoids packet disorder. Additionally, we analyze the signaling cost, packet loss, handover latency and buffer size of our proposal, comparing them with other alternatives. Moreover, we discuss the tunnelling overhead. Through the conducted numerical results we justify the benefits of our proposed architecture.

Contents

4.1	Introduction and contributions	59
4.2	Proposed architecture	61
4.2.1	LinkWork Mobile MPLS operation	61
4.2.2	Recovery mechanism in LinkWork Mobile MPLS	65
4.3	Analytical evaluation	72
4.3.1	Signaling cost	73
4.3.2	Handover latency and Packet loss during a session	74
4.3.3	Buffer size	77
4.3.4	Tunnelling overhead	77
4.4	Numerical Results	78
4.5	Concluding remarks	82

4.1 Introduction and contributions

As already discussed, future mobile networks are expected to be more dynamic and flexible, as well as to provide higher bandwidth at lower costs. In order to achieve these goals, it is necessary to overcome many challenges. One of the most important is how to provide QoS guarantees in such highly dynamic mobile environments. Facing this issue requires discussing mobility management and QoS. Both mobility management proposals and QoS techniques have been introduced in Chapter 2. To this extent, in order to address this challenge, solutions for QoS should be coupled with those for managing mobility in IPv6 networks.

One of the most promising techniques to provide QoS in future mobile IP networks is MPLS. The integration of Mobile IPv6 and Multiprotocol Label Switching has worked successful in many cases [73], [87], [88] due to the ability of MPLS to efficiently engineer traffic tunnels, including constraint-based routing, survivability, and recovery, thus avoiding congestion and enabling an efficient use of the available bandwidth.

These features make MPLS a potential technique to solve MIP's operational and architectural shortcomings such as: high handover latency [89], packet loss, high global signaling load and scalability issues [90–92]. MPLS could also be viewed as an efficient lightweight tunnelling technology that overcomes the tunnelling techniques proposed in Mobile IP standard. Using MPLS tunnels, called label switched paths (LSPs) in MPLS jargon, an overlay network is efficiently created and managed. In MPLS, tunnel redirection, which is a crucial ingredient of any mobility scheme, happens quickly, at the change of a label in a single node in the network. Furthermore, by using this technology, we can directly take advantage of all the capabilities mentioned previously of MPLS [93].

In this context, we propose LinkWork Mobile MPLS (LW-MMPLS) to solve some problems detected in previous works. From our point of view, the setup of a complete LSP after each movement increases the signaling overhead, reducing the overall performance of the network. In our proposal, we efficiently handle mobility reducing the signaling overhead in an MPLS domain. This solution is based on the forwarding chain concept (set of forwarding paths) in a wireless hierarchical network, in which can be costly to manage the mobility process [47]. We also propose some Linkage Nodes (LN) that belong to the MPLS domain, which are responsible of the LSP redirection, improving tunnel rerouting in a Mobile IP-based domain. This way, the LSP is going to be composed by a set of forwarding paths that will allow the signaling to be localized and will adapt its route to track host mobility. In this architecture we also take advantage of MPLS-TE recovery techniques that can be used to optimize the problem of LSP re-establishment after handover and to minimize packet loss during service disruption when a mobile node changes its point

of attachment to the network. The next section gives details about the operation of LW Mobile MPLS and its packet recovery mechanism.

In addition, in order to assess the efficiency of our proposal, a set of protocols are compared. To achieve this, we develop analytical models to evaluate different parameters such as the signaling cost, packet loss and link usage. The objective of this chapter is to provide a comprehensive comparison of the main centralized IPv6 mobility management proposals. Moreover, we discuss a performance evaluation analysis that highlights and addresses the main strong and weak points of each of these well-known solutions through the results obtained.

The centralized protocols included in the analysis are some representative IPv6 mobility management protocols (Mobile IPv6, Proxy MIPv6) as well as other solutions that include MPLS to work in conjunction with IPv6 mobility management mechanisms, such as Mobile MPLS, FH-Micro Mobile MPLS and PMIPv6 MPLS. In Table 4.1, a summary comparison is provided for the main characteristics of these mobility management protocols. Based on different design goals, existing mobility IP protocols are classified into two types, host-based (or MIPv6-based) and network-based (or PMIPv6-based).

Thus, the first row called *Protocol* identifies the mobility management protocol. The *Required infrastructure* refers to the additional entities required in the network in order to support the mobility management related signaling. The *Mobility scope* indicates whether

Table 4.1. Protocols Features

Protocol	MIPv6-based				PMIPv6-based	
	MIPv6	Mobile MPLS	FH-MM MPLS	LW-MMPLS	PMIPv6	PMIP MPLS
Required infrastructure	HA	HA-FA	LERG-FA	ILER PELER-NELER	LMA MAG	LMA MAG
Mobility scope	Global	Global	Local	Local	Local	Local
Tunnelling protocols	IP-IP GRE	MPLS	MPLS	MPLS	IP-IP GRE	MPLS
MPLS integration	No	Yes	Yes	Yes	No	Yes

a given proposal has been designed for local or global mobility management. Next row, *Tunnelling protocols*, refers to the tunnel technology used to forward the data packets. This characteristic is relevant due to the throughput degradation that tunnelling overhead can cause. Finally, the *MPLS integration* refers to whether the mobility protocol is integrated with MPLS.

In the next sections, the proposed LinkWork Mobile MPLS architecture is introduced. The analytical model used to evaluate the performance of our proposal and to derive the numerical results is also presented.

4.2 Proposed architecture

As we mentioned earlier, one of the main research topic in MPLS networks is focused on LSP routing through the network [94]. However, due to the increasing trend towards the introduction of MPLS in wireless environments, it is necessary to develop new LSP re-routing techniques that take in account the continuous movement of nodes across the network.

Based on both Mobile IP protocol and MPLS-TE recovery techniques, we propose a mobility management architecture that offers QoS support in MPLS access networks. Thus, the mobility problem can be seen as a routing problem where the path is broken due to the movement of the mobile node. In this situation, MPLS can be profitably used to complement mobile management protocols, as it enhances the tunnelling paradigm with fast forward and traffic engineering capabilities. Figure 4.1 shows a typical architecture for LinkWork Mobile MPLS.

4.2.1 LinkWork Mobile MPLS operation

The main idea behind this architecture is to handle efficiently mobility by using a set of forwarding paths that can be considered as a new alternative to implement the pointer forwarding technique in an MPLS environment. Moreover, we try to anticipate the L3 handover using L2 functionalities and to setup an LSP before the MN really moves into a new subnet to reduce service disruption.

We assume that an MPLS access network exists between the Label Edge Router/Home Agent (LER/HA) and the Label Edge Router/Access Routers (LER/AR). In this MIP domain, RSVP-TE must be implemented in order to signal paths.

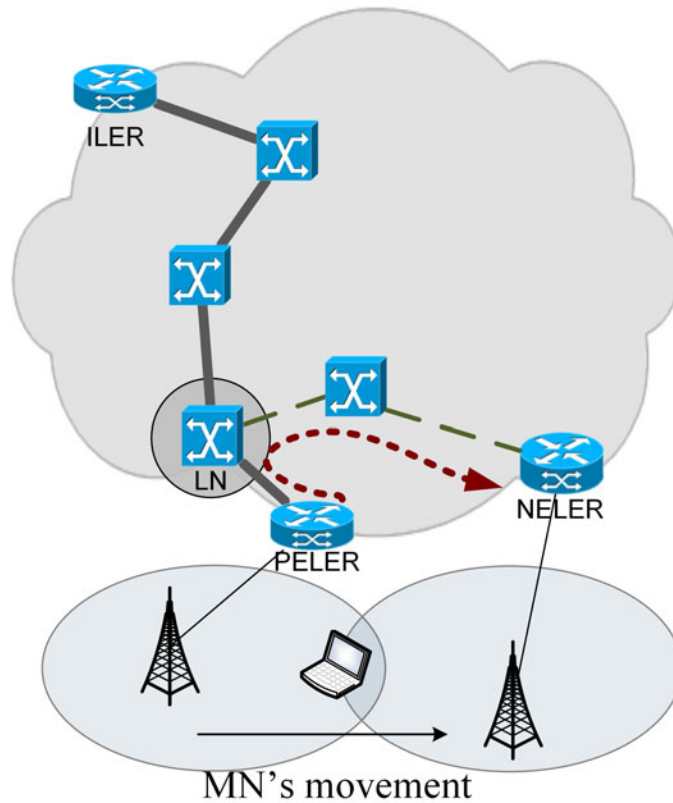


Figure 4.1. LinkWork Mobile MPLS architecture

Initially, when a MN moves for the first time into a LinkWork Mobile MPLS domain, it performs the discovery of new access router through Router Solicitation/Advertisement messages exchange. After acquiring an IP address, the MN performs the initial Registration Request to the home agent. The corresponding PELEER (Previous Egress LER) sends the message to the Linkage Node, that records the MN home address and then relays the registration message to the Ingress LER of the domain. When the ILER (Ingress LER) gets the registration message, it is informed with the IP address of the current LN that serves the MN as its mobility anchor. Then, the ILER establishes an LSP between it and the PELEER, through the LN. Figure 4.2 illustrates the registration procedure for the MN in LW Mobile MPLS.

Once the MN is registered in the domain, it moves while it is communicating with the CN. During this time, the MN can change its point to attachment to the network, performing a handover. This process is explained next and its basic operation is shown in Figure 4.3

Once a mobile node moves to an adjacent network, it disconnects from its previous LER/AR called PELEER and it attaches to a new LER/AR called NELER (New Egress LER), establishing a new LSP towards this router. As mentioned earlier, this architecture

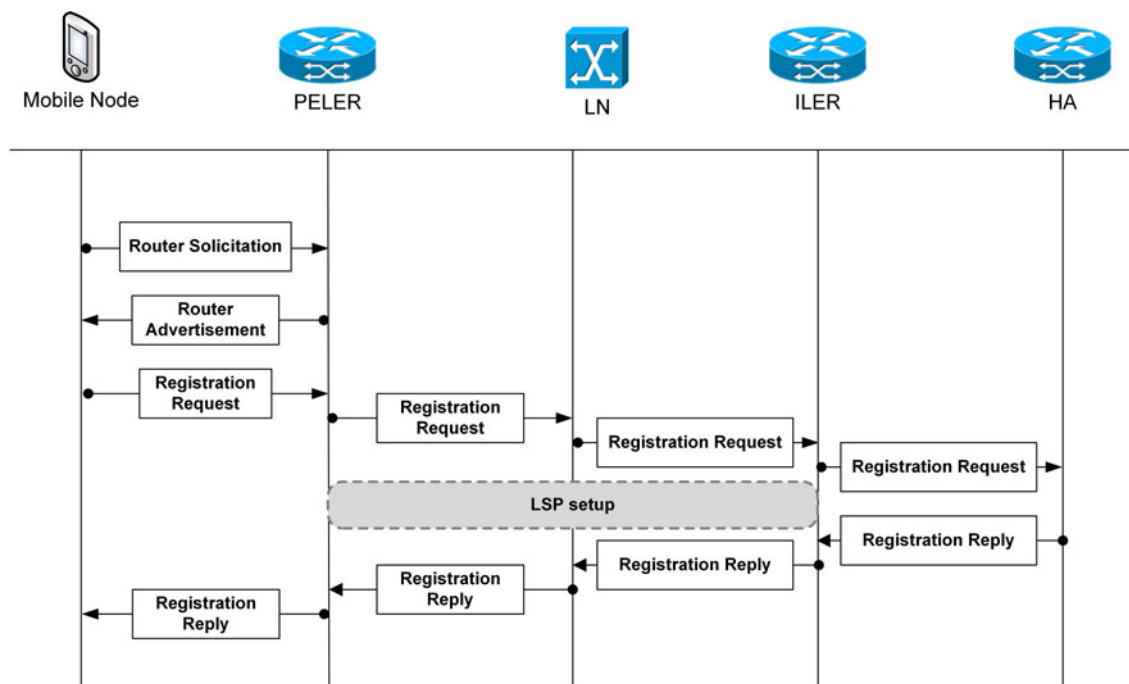


Figure 4.2. Registration procedure in LinkWork Mobile MPLS

provides a mobility environment where LSPs can be rerouted in LSRs nodes. In order to achieve this purpose, the new path could have a common section with the old LSP and other part that forwards packets towards the NELER. The LSR node from where we setup the new section of the LSP that forwards packets to the NELER is called Linkage Node. These LNs are responsible for LSP rerouting and can be considered a point of recovery when the path is disrupted due to a handover. With the introduction of these nodes, it is possible to anticipate the handover and setup the new linkworked LSP as well as start LSP recovery techniques and packet buffering to avoid both data loss and packets disorder. Specifically, when the MN moves to an adjacent network, it proceeds as follows.

At the beginning, the MN is communicating through a LSP tunnel from LER/HA to the LER/AR that serves the MN (PELER). When the MN enters an overlapped area of an adjacent subnet, it receives a L2 signal from the possible new base station (BS) (step 1). Next, the MN notifies the PELER the possibility of a handover by sending a HI (Handover Initiate) message which contains the new Base Station identifier. This information is going to be used to obtain the NELER IP address, thanks to a data structure that keeps a match between this identifier and each adjacent LER IP address (step 2). These 2 steps of the LW Mobile MPLS architecture are similar to those proposed in the FH-Micro Mobile MPLS scheme. Once the PELER knows the subnet which the MN is going to move, it sends a message upstream to the selected LN, notifying about a possible L3 handover, and starting the setup of a new section of the LSP from LN to NELER (step 3) with

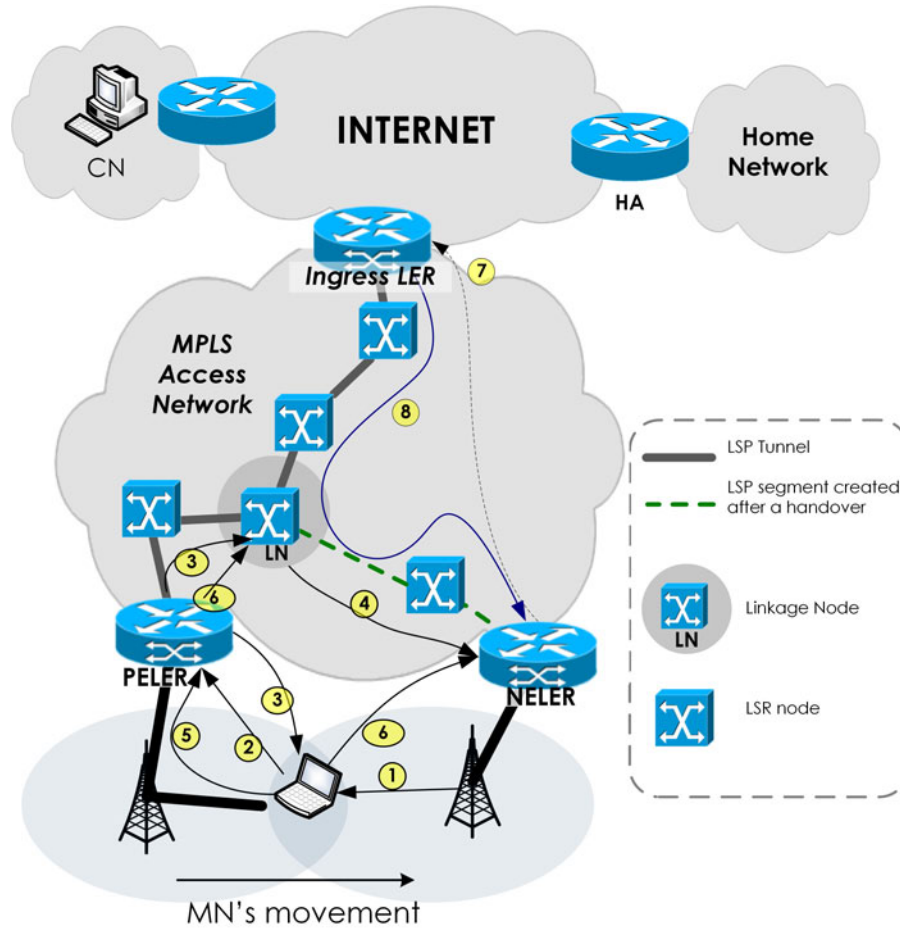


Figure 4.3. LinkWork Mobile MPLS operation

the required QoS, using RSVP-TE. In this step, the PELER also informs to MN about the NELER IP address through a Neighbor Discovery message. At this point, a new section of the LSP tunnel could be set up so, data traffic could be forwarded towards the new location of the MN (step 4). When the signal strength received from the current base station falls below a certain threshold level, the MN notifies the handover to the PELER. In this step, our mechanism responsible for minimizing packet loss is started by the PELER (step 5).

Once the L2 handover is performed, the L3 handover is initiated by the MN with the NELER through MIPv6 registration process (step 6). The new LSP section from LN to the new egress router will be used when the LN is aware of the movement. This happens when the PELER starts to return data packets to the LN, which will be forwarded to the NELER through the new LSP section together with buffered packets according to the recovery mechanism. The operation of this mechanism is going to be detailed in following sections. Finally, the NELER sends the Mobile IPv6 Binding Update message to ILER

(step 7). The ILER will reply to the MN which is located in the new subnet.

At least one LN must exist in the path between the ILER and the corresponding ELER. If there are no internal LN in the LSP, the ILER will consider the LN and it will make the corresponding tasks.

4.2.2 Recovery mechanism in LinkWork Mobile MPLS

MPLS signaling protocols provide control mechanisms to set up, tear down, maintenance and recovery of LSPs. RSVP-TE signaling extensions and have been developed to work out in wired networks, where paths once established, hardly change. In mobile networks, the frequency of a path disruption due to a handover is very high. Packet loss and packet disorder are two important factors which badly affect handover performance of MIP because lost packets during handover are recovered by TCP retransmissions. In this section we present a recovery scheme based on Hundessa mechanism [95] and RSVP-TE signaling to solve the high rate of packet loss and packet delay during a handover as well as the problem of packet disorder after recovery.

Observe that under the architecture explained previously, when the MN informs the PELER node about an L2 handover, this edge router does not send any more packets to the MN. Instead, it sends the packets back to the LN. When the first packet arrives back to the LN, it tags the next packet received and sends it and buffers all incoming packets from the PELER. Once the LN receives the tagged packet from the reverse path, it untags it, forwards it through the new section of the LSP through the NELER. Once all incoming packets from PELER have been sent, buffered packets are forwarded to the NELER. This is how the packet disorder and packet loss minimization is achieved. The described operation is also depicted in Figure 4.4.

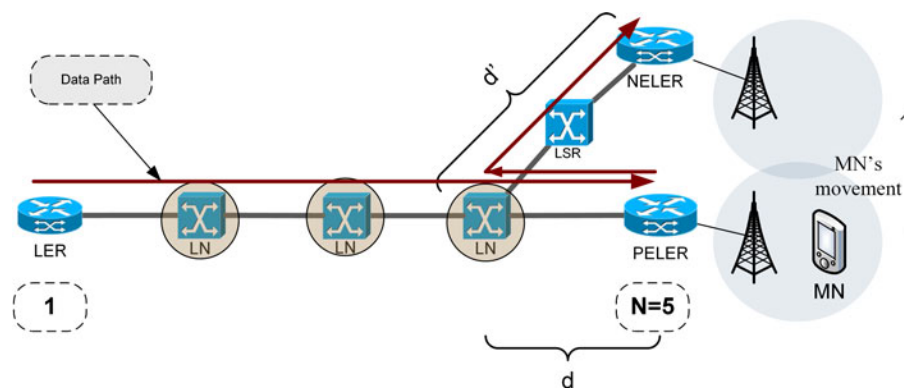


Figure 4.4. Example data path followed by a packet during the recovery mechanism

If no recovery mechanism were implemented, all lost packets during a handover need to be retransmitted end-to-end from the sender. In a general case, the end-to-end retransmission time-out of TCP is triggered. Localized retransmissions close to the MN can offer a better performance in terms of packet delay and packet delivery cost [96].

Apart from the aforementioned benefits, the proposed recovery mechanism offer additional advantages compared to other works. The main one is that the packets are sent in-order towards the new location of the MN so, the tasks to do in the MN to reorder the information is significantly reduced.

Let consider a MPLS domain represented as a graph $G = (X, U)$ with a set of X nodes and U links. Let be $\phi(G)$ a flow in the domain with a path $LSP_{i,n}$, whose source is x_i (an ingress LER node) and the destination x_n (a PELEER node). We assume that a number of LN (x_{LN}), $0 < LN < n - 1$, exist in the path $LSP_{i,n}$ with capacities to perform the recovery and reroute the data flow $\phi(G)$ to the new egress node (NELER).

Thus, in case of a handover, all in-flight packets can be recovered as soon as the PELEER node is informed of the movement of the node to the new network, with the aim to avoid end-to-end retransmissions from the source node.

4.2.2.1 Delay analysis

In this section the LinkWork Mobility MPLS recovery mechanism is analyzed in order to test the improvement with respect to current techniques that do not implement any solution to minimize the packet loss so, end-to-end retransmission is needed. The packet delay difference between end-to-end re-transmission and LW Mobility MPLS recovery is calculated. First, the topology is modeled as a graph. Next, current transport protocols based on end-to-end retransmissions are modeled basing on link cost (re-transmitted packet delay) and restrictions. After that, both models are compared in order to check if there is any improvement. Finally, the enhancement is measured in other to get an equation that permits an extrapolation to different topologies or other traffic samples. The delay analysis is based on our work [97].

Let δ_{ij} be the delay of the link $(x_i, x_j) \in U$ and let $\delta(x_i, x_n)$ be the delay of a path $LSP_{i,n}$ between an ingress node x_i and an access router x_n (see Figure 4.5). The minimum delay

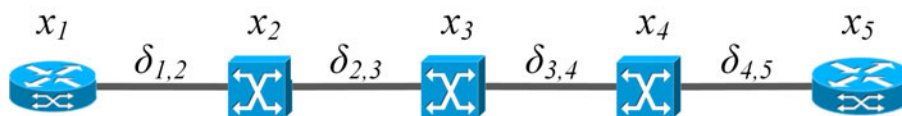


Figure 4.5. Diagram of the $LSP_{i,n}$ path for $i = 1$ and $n = 5$

for any packet that traverses the domain can be represented with the Dijkstra algorithm:

$$\min \delta(x_i, x_j) = \sum_{i=1}^n \sum_{j=1}^n \delta_{ij} \cdot x_{ij}, \quad (4.1)$$

subject to

$$\sum_{l=2}^n x_{1l} = 1 \quad (4.2)$$

$$\sum_{i=1}^n x_{il} - \sum_{j=1}^n x_{lj} = 0, l = 2, 3, \dots, n-1 \quad (4.3)$$

$$\sum_{l=1}^{n-1} x_{ln} = 1 \quad (4.4)$$

where,

$$\delta_i, i = 0, \forall i \in X$$

$$x_{i,j} = 1, \forall (x_i, x_j) \in LSP_{i,n}$$

and

$$x_{i,j} = 0, \forall (x_i, x_j) \notin LSP_{i,n}$$

from these definitions, several parameters can be defined:

- $\Delta_{e-e}(x_i, x_j)$ is the total delay of recovering a certain packet in x_j , whose source is x_i , by using end to end retransmissions.
- $\Delta_{LW}(x_i, x_j)$ is the total delay of recovering a certain packet in x_j , whose source is x_i , by using the LinkWork Mobility MPLS recovery mechanism. In this case, x_j is the PELER node and x_i is a LN. The distance between the PELER and the active LN is called the diameter (d).

- $\delta_{e-e}(x_i, x_j)$ is the time used to recover end-to-end a certain packet whose source was x_i and it was lost in x_j .
- $\delta_{LW}(x_i, x_j)$ is the time used to recover a packet with the LinkWork Mobility recovery mechanism, whose source was x_i and it was lost in x_j . In this case, x_j is the PELER node and the x_i is a LN. The distance between the PELER and the active LN is called the diameter (d).

Particularly, if no recovery mechanism were considered, the retransmission of dropped packets would be performed end-to-end by the upper layers. In this case, when a packet is dropped in the PELER node (x_n), the loss is detected at the source node when the sink does not send the acknowledgement towards the source. This way, the function *Loss Detection Time* (LDD_{e-e}) for each discarded packet that belongs to the flow $\phi(x_i, x_n)$ is defined as:

$$LDT_{e-e}(x_i, x_n) = \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.5)$$

In the best case, if the upper layers perform the end-to-end retransmission of lost data using the Fast-Retransmit mechanism, then it would need to wait for two more disordered packets and the delay of the retransmitted packets would be:

$$\delta_{e-e}(x_i, x_n) = 2 \cdot \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.6)$$

Therefore, the total delay $\Delta_{e-e}(x_i, x_n)$ to retransmit a packet towards the x_n is derived from Eq. 4.5 plus Eq. 4.6:

$$\Delta_{e-e}(x_i, x_n) = LDT_{e-e}(x_i, x_n) + \delta_{e-e}(x_i, x_n) \quad (4.7)$$

$$\Delta_{e-e}(x_i, x_n) = \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + 2 \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.8)$$

$$\Delta_{e-e}(x_i, x_n) = 3 \cdot \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.9)$$

However, in LW Mobility MPLS architecture, if the packet is lost in x_n (PELER) and the LW recovery mechanism is active, at least one LN node (x_{LN}) that belongs to the $LSP_{i,n}$ ($i < LN < n$) (see Figure 4.6). In this case, for each packet lost in the flow $\phi(x_i, x_n)$, the Loss Detection Time LDD_{LW} function is defined as:

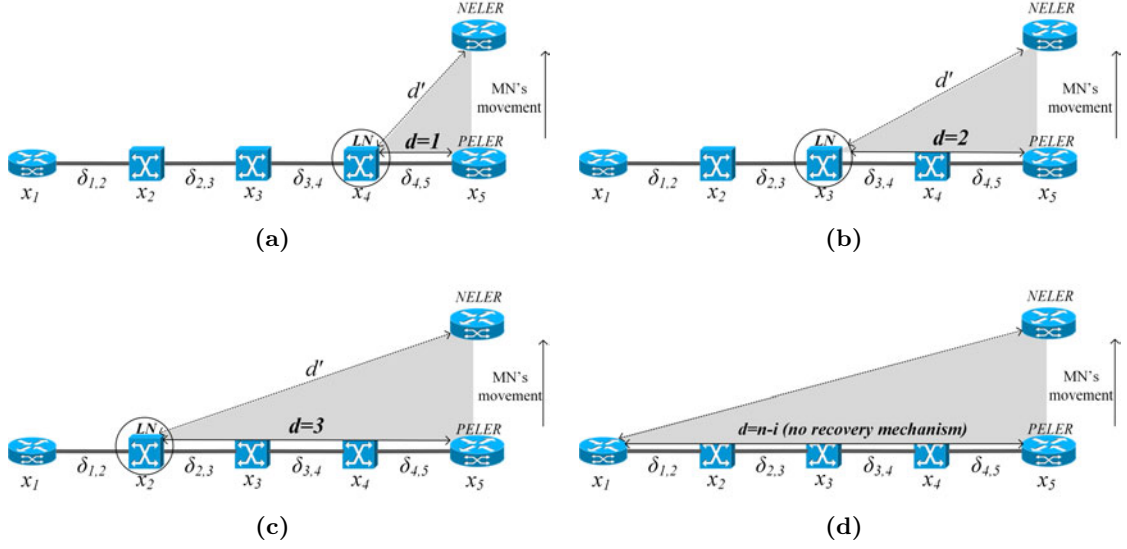


Figure 4.6. Local Recovery for different distances from PELEER to LN. (a) $d = 1$ hop, (b) $d = 2$ hops, (c) $d = 3$ hops, (d) $d = n - i$, i.e., no recovery mechanism is implemented and end-to-end retransmissions are required.

$$LDT_{LW}(x_i, x_n) = \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.10)$$

Moreover, considering that d is the diameter, or distance in number of hops, between the PELEER node and the LN, the time used in the LW recovery mechanism is:

$$\delta_{LW}(x_i, x_n) = \sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.11)$$

subject to:

$$0 < d < n - i$$

$$0 < d + d' < 2(n - i)$$

The first restriction implies that if diameter d in Eq. 4.11 is $n - i$, then $l = n - d = n - (n - i) = n - n + i = i$ we would find that:

$$\sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} = \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=i}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.12)$$

That is, it would be an end-to-end retransmission if we assume that $d = d'$.

Furthermore, if in (4.11) diameter d is bigger than $n - i$, then it would be trying a retransmission from a previous node to x_i , but this one is the source of the $LSP_{i,n}$ so that would be unfeasible.

With respect to the second restriction, that refers to d' , it is considered that the sum of the distance between the LN and the PELEP (d) plus the distance between the LN and the NELER (d') is smaller than $2(n - i)$. Taking into account that $(n - i)$ is the number of hops of the $LSP_{i,n}$, if the sum $d + d' = 2(n - i)$, then we obtain the same delay as in $\delta_{e-e}(x_i, x_n)$ and no improvements are achieved.

As we can see in the two restrictions in which Eq: 4.11 is based on, the diameter d depends only on the distance $(n - i)$, whereas d' depends on both the distance $(n - i)$ and d . This is because the distance d mandatory needs to be defined in a LinkWork Mobile MPLS, even if the MN remains attached to the same AR the whole time. However, d' depends on d because it is defined when a handover is performed and, therefore, d has already been established.

Thus, the total delay $\Delta_{LW}(x_i, x_n)$ needed to retransmit a packet with the LW recovery mechanism is derived from Eq. 4.10 and Eq. 4.11.

$$\Delta_{LW}(x_i, x_n) = LTD_{LW}(x_i, x_n) + \delta_{LW}(x_i, x_n) \quad (4.13)$$

$$\Delta_{LW}(x_i, x_n) = \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} \quad (4.14)$$

At this point, we can test if $\Delta_{LW}(x_i, x_n) < \Delta_{e-e}(x_i, x_n)$:

$$\sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} < 3 \cdot \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1}$$

$$\sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} < \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + 2 \cdot \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1}$$

$$\sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} < 2 \cdot \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1}$$

so, according to Eq. 4.11 and Eq. 4.9, we only need to verify in Eq. 4.14 that $\delta_{LW}(x_i, x_n) < \delta_{e-e}(x_i, x_n)$. The only condition that differentiates the members of Eq. 4.14 is the set of values that the variable l can take. It only needs to be demonstrated that l takes a lesser number of values in $\delta_{LW}(x_i, x_n)$ than in $\delta_{e-e}(x_i, x_n)$.

$$n - 1 - (n - d) + (n - d + d' - 1) - (n - d) < 2(n - 1 - i)$$

where

$n - 1 - (n - d) + (n - d + d' - 1) - (n - d)$ is the rank values of l in $\delta_{LW}(x_i, x_n)$
 $2(n - 1 - i)$ is the rank values of l in $\delta_{e-e}(x_i, x_n)$.

Solving this inequation, we obtain that $d' < 2(n - i) - d$

$$n - 1 - (n - d) + (n - d + d' - 1) - (n - d) < 2(n - 1 - i)$$

$$n - 1 - n + d + n - d + d' - 1 - n + d < 2n - 2 - 2i$$

$$-2 + d' + d < 2n - 2 - 2i$$

$$d' < 2(n - i) - d$$

Thus, we find that the problem remains in the feasibility zone, since optimizing Eq. 4.14 we find that $d' < 2(n - i) - d$, which is one of the restrictions of Eq. 4.11.

This way, it has been demonstrated that $\Delta_{LW}(x_i, x_n) < \Delta_{e-e}(x_i, x_n)$. Therefore, the LW recovery mechanism achieves a benefit in terms of delay for each lost packet that needs to be retransmitted ($\Delta_{e-e}(x_i, x_n) - \Delta_{LW}(x_i, x_n) > 0$). That improvement can be quantified:

$$\begin{aligned} & \Delta_{e-e}(x_i, x_n) - \Delta_{LW}(x_i, x_n) = \\ & = 3 \cdot \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} - \left(\sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} \right) = \\ & = 2 \cdot \sum_{l=i}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} - \left(\sum_{l=n-d}^{n-1} \delta_{l,l+1} \cdot x_{l,l+1} + \sum_{l=n-d}^{n-d+d'-1} \delta_{l,l+1} \cdot x_{l,l+1} \right) \end{aligned} \quad (4.15)$$

From a number of hops point of view, the benefit obtained is:

$$2(n-1-i) - (n-1-(n-d)) - (n-d+d'-1-(n-d)) = 2n-2i-d'-d = 2(n-i) - d - d'$$

4.3 Analytical evaluation

In order to evaluate the performance of LinkWork Mobile MPLS architecture, we make use of the analytical framework developed in Chapter 3 to compare various existing well-known centralized mobility management protocols. In this section, we derive the analytical model of various evaluation criteria such as the cost functions of registration updates total packet loss during a session, buffer size metrics and tunnelling overhead. In order to evaluate the performance of these mobility protocols when a MPLS access network is introduced, some MPLS-based proposals are compared with non MPLS-based ones.

As can be seen in Table 4.1, we compare six mobility protocols. However, not all of them are integrated with MPLS. Despite this, we consider them in our analysis due to their importance since Mobile IPv6 and Proxy MIPv6 have been developed by IETF as the main solutions to provide mobility in the Internet, and 3GPP has focused on them to achieve the mobility in LTE (Long Term Evolution) evolved packet core [98]. Specifically, these protocols are the following. From a host-based schemes point of view, the considered mechanisms are Mobile IPv6 [27] and other proposals based on Mobile IP which integrate mobility and MPLS such as Mobile MPLS [73], FH-Micro Mobile MPLS [87] and LinkWork Mobile MPLS. On the other hand, the network-based schemes investigated in this analysis are PMIPv6 [28] and MPLS-PMIPv6 [71]. MPLS-PMIPv6 is the first scheme which proposes MPLS as an alternative tunnel technology between a MAG and a LMA. Two kinds of labels are employed: a classical tunnel label and a Virtual Pipe label. This last is introduced as a means to differentiate traffic with the same MAG-LMA end-points according to the operators of the various MNs served by the same MAG.

In order to simplify the analytical study, we suppose that every subnet is equidistant from the ILER, with a distance of δ (in terms of number of hops). In the same way, we do not consider the cost of the process that periodically updates the link (Binding Update) between the MN and the HA, in order to update the cache. We analyze the mobility behavior of the MN, keeping in mind a topology where a terminal could move to every neighbour network with the same probability. The parameters to be used in the analysis are described in Table 3.2 in Chapter 3.

4.3.1 Signaling cost

The total signaling cost of registration update for a session can be defined as $C_s(MMPRot)$. As it was described in Chapter 3, this value depends on the traffic load when signaling messages are sent, i.e., this cost depends on the size of signaling messages and the number of hops in every level 3 handover process during the time interval that communication of MN remains active. Therefore, the cost is defined by the message size multiplied by the number of needed hops.

Each movement between neighboring subnets implies sending several signaling messages. This metric was introduced in section 3.4.1 of Chapter 3. With respect to this parameter, in this section we are interested in the evaluation of the cost for the binding update after a handover. In Mobile IPv6, an MN sends Binding Update messages for the HA whenever it changes its point of attachment. We consider a that the communication between the CN and the MN is indirectly through the HA, not considering the direct communication through normal network routing procedures (called Route Optimization). This latter case is not taken into account due to the necessity of the CN to update its IPv6 layer stack in order to perform a direct routing with the MN. It should be emphasized that the BU messages are always exchanged between the MN and the HA irrespective of how far the MN is from its Home Network. Thus, it can be really long especially when the MN is in a visited network that is further away from the home network. Accordingly, the signaling cost of registration update in MIPv6 $C_s(MIPv6)$ is expressed as:

$$C_s(MIPv6) = 2 \cdot s_u \cdot h_{MN-HA} \cdot N_h \quad (4.16)$$

In the Mobile Mobile MPLS case, the registration update with the HA is needed, in the same way as it was performed in MIPv6. In this approach, it is necessary to add the cost of the LSP establishment procedure so, $C_s(MobileMPLS)$ is defined as follows

$$C_s(MobileMPLS) = 2 \cdot s_u \cdot h_{MN-HA} \cdot N_h + 2 \cdot s_l \cdot h_{FA-HA} \cdot N_h \quad (4.17)$$

In FH-Micro Mobile MPLS (FH-), both local registration updates with the LERG and LSPs procedure set-up with the neighboring subnets are performed. In this case we have

$$\begin{aligned} C_s(FH-) &= 2 \cdot s_u \cdot h_{MN-LERG} \cdot N_h + 2 \cdot s_u \cdot h_{FA-FA} \cdot N_h + \\ &+ 2 \cdot s_l \cdot h_{FA-LERG} \cdot N_h \end{aligned} \quad (4.18)$$

In our proposal, LW Mobile MPLS, the record update is done with the ILER, which is the ingress router to the MPLS domain, and the LSP tunnel is established from the LN to the NELER. Thus, the $C_s(LW)$ is expressed as follows:

$$C_s(LW - Mobile MPLS) = 2 \cdot s_u \cdot h_{MN-ILER} \cdot N_h + 2 \cdot s_l \cdot h_{LN-NELER} \cdot N_h \quad (4.19)$$

Contrary to host-based mobility management protocols, in network-based mobility management the MN is not involved in any mobility-related signaling. In fact, a network element or proxy agent performs all the mobility-related signaling on behalf of the MN. The network-based mechanisms considered in this analysis are PMIPv6 and PMIP-MPLS. In general, network-based solutions achieve a better signaling cost due to its localized mobility management.

In contrast with the previously analyzed host-based mobility management protocols, PMIPv6 manages the MN's movement in a localized manner, considering that the signaling is confined between the LMA and MAG agents. Moreover, the MAG node is the responsible for handling all mobility-related signaling on behalf of the MN. It tracks the movement of the MN, and initiates the required mobility signaling. At each MN movement, there is a PBU/PBA exchange between the MAG and the LMA. Thus, $C_s(PMIPv6)$ can be expressed as

$$C_s(PMIPv6) = 2 \cdot s_{pu} \cdot h_{MAG-LMA} \cdot N_h \quad (4.20)$$

The MPLS-PMIPv6 solution in most cases would need the same packets as PMIPv6, plus the RSVP-TE signaling, that adds extra packages to setup the MPLS tunnel, hence

$$C_s(PMIP - MPLS) = 2 \cdot s_{pu} \cdot h_{pMAG-LMA} \cdot N_h + 2 \cdot s_l \cdot h_{nMAG-LMA} \cdot N_h \quad (4.21)$$

4.3.2 Handover latency and Packet loss during a session

Following the analytical framework introduced in Chapter 3, packet loss during a session (P_{loss}) can be defined as the sum of lost packets per MN during all handovers. Let λ_d be the packet arrival rate in unit of packet per time. MIPv6 does not incorporate any buffering mechanism, thus data packets sent from the CN to the MN will be lost while the MN performs its handover. In MIPv6, this cost can be expressed as follows:

$$P_{loss}(MIPv6) = T_h(MIPv6) \cdot \lambda_d \cdot N_h$$

where the $T_h(MIPv6)$ represents the handover delay due to the mobility management mechanisms of *MIPv6*. $T_h(MIPv6)$ can be written as

$$T_h(MIPv6) = \frac{1}{2}RTT_{MN-AR} + 2 \cdot t(s_u, h_{MN-HA}) \quad (4.22)$$

Thus, packet loss can be defined as:

$$P_{loss}(MIPv6) = \left[\frac{1}{2}RTT_{MN-AR} + 2 \cdot t(s_u, h_{MN-HA}) \right] \cdot \lambda_d \cdot N_h \quad (4.23)$$

In Mobile MPLS, all in-flight packets will be lost during the handover disruption time due to the lack of any buffering mechanism. Considering that the handover latency in Mobile MPLS is

$$T_h(M \text{ MPLS}) = \frac{1}{2}RTT_{MN-AR} + 2 \cdot t(s_u, h_{MN-HA}) + 2 \cdot t(s_l, h_{FA-HA}) \quad (4.24)$$

thus,

$$P_{loss}(M \text{ MPLS}) = \left[\frac{1}{2}RTT_{MN-AR} + 2 \cdot t(s_u, h_{MN-HA}) + 2 \cdot t(s_l, h_{FA-HA}) \right] \cdot \lambda_d \cdot N_h \quad (4.25)$$

With respect to FH-Micro Mobile MPLS, this solution incorporates a recovery mechanism that will minimize the packet loss during handover. Thus, all in-flight packets would be lost till the buffering mechanism is initiated. In this case, $T_h(FH - MicroMobileMPLS)$ is

$$T_h(FH-) = \frac{1}{2}RTT_{MN-FA} + 2 \cdot t(s_u, h_{MN-LENG}) + 2 \cdot t(s_u, h_{FA-FA}) + 2 \cdot t(s_l, h_{FA-LENG}) \quad (4.26)$$

whereas $P_{loss}(FH - Micro \text{ Mobile MPLS})$ depends on the time needed to initiate the buffering mechanism

$$P_{loss}(FH-) = t(s_u, h_{MN-FA}) \cdot \lambda_d \cdot N_h \quad (4.27)$$

Our LW Mobile MPLS proposal also has a recovery mechanism that minimizes the packet loss, as it was explained in previous sections. With this technique, in-flight messages between the LN and the PELEER can be recovered and forwarded, in-order, across the new LSP. Therefore, the values of T_h and P_{loss} can be defined as:

$$T_h(LW) = \frac{1}{2}RTT_{MN-NELEER} + 2 \cdot t(s_u, h_{MN-ILER}) + 2 \cdot t(s_l, h_{LN-NELEER}) \quad (4.28)$$

and

$$P_{loss}(LW) = t(s_u, h_{MN-PELEER}) \cdot \lambda_d \cdot N_h \quad (4.29)$$

With respect to the network-based solutions, the behavior of P_{loss} in PMIPv6 is very similar to MIPv6, with the difference that the mobility bindings, necessary to configure the correct routing with the mobility anchor, are sent from the MAG to the LMA instead of from the MN. Hence, the P_{loss} can be expressed as follows

$$P_{loss}(PMIPv6) = T_h(PMIPv6) \cdot \lambda_d \cdot N_h$$

where the $T_h(PMIPv6)$ is

$$T_h(PMIPv6) = \frac{1}{2}RTT_{MN-MAG} + 2 \cdot t(s_{pu}, h_{MAG-LMA}) \quad (4.30)$$

Thus, we have that packet loss in PMIPv6 is:

$$P_{loss}(PMIPv6) = \left[\frac{1}{2}RTT_{MN-MAG} + 2 \cdot t(s_{pu}, h_{MAG-LMA}) \right] \cdot \lambda_d \cdot N_h \quad (4.31)$$

Finally, in PMIP-MPLS, the packet loss is expressed as

$$P_{loss}(PMIP - MPLS) = T_h(PMIP - MPLS) \cdot \lambda_d \cdot N_h \quad (4.32)$$

where $T_h(PMIP - MPLS)$ is as follows

$$T_h(PMIP - MPLS) = \frac{1}{2} \cdot RTT_{MN-nMAG} + 2 \cdot t(s_{pu}, h_{pMAG-LMA}) + 2 \cdot t(s_l, h_{nMAG-LMA}) \quad (4.33)$$

4.3.3 Buffer size

The buffer to store in-flight packets is located at the LN in our LW Mobile MPLS proposal. The packet recovery mechanism is activated when MN notifies the PELER that it is going to change its point of attachment to the network by means of a movement signaling message. With the other proposals, only one implements packet buffering: FH-Micro Mobile MPLS. In this scheme, the buffer is located in the LER/FA nodes. Therefore, the buffer size requirement (B_s) for FH Micro Mobile MPLS, and LW Mobile MPLS is listed as follows:

$$B_{size}(FH - MMMPLS) = \frac{1}{2} \cdot RTT_{MN-AR} + t(s_u, h_{MN-FA} + h_{FA-FA}) \cdot \lambda_d \quad (4.34)$$

$$B_{size}(LW Mobile MPLS) = \frac{1}{2} \cdot RTT_{MN-AR} + t(s_u, h_{MN-LN} + h_{LN-PELER} + h_{PELER-LN}) \cdot \lambda_d \quad (4.35)$$

4.3.4 Tunnelling overhead

As we have seen in previous sections, mobility management protocols establish a tunnel to forward data packets. The IETF advises the use of IP-in-IP or GRE (Generic Routing Encapsulation) as tunnelling methods. In this section we conduct a study of these technologies, and we compare them with MPLS tunnels.

IP-IP (IP in IP) is a protocol by which an IP datagram may be encapsulated (carried as payload) within an IP datagram, by adding a second IP header to each encapsulated datagram. However, IP-in-IP tunnelling increases overhead, because it needs an extra set of IP headers. Typically, a pure IP-over-IP tunnel configured with tunnel mode IP-IP has a 20-byte overhead, so if the normal packet size (Maximum, Transmission Unit, MTU) on a network is 1500 bytes, a packet that is sent through a tunnel can only be 1480 bytes big.

GRE is another tunnelling method that encapsulates any network layer packet. GRE requires the IP-in-IP encapsulation with the extra IP-IP header (20 bytes), but it also adds another 4 bytes of the GRE header to a packet, resulting in 24-byte overhead. After this increase the packet may need to be fragmented because it is larger than the outbound MTU.

On the other hand, an MPLS LSP tunnel has one label (4 bytes) or a stack of labels (for example, when using Link Protection Fast reroute) of overhead. MPLS adds four bytes to

every datagram but, unlike GRE tunnel, MPLS does not change the IP header. Instead, the label stack is imposed on to the packet that takes the tunnel path.

The three approaches can also be compared in terms of the overhead they generate during data packets forwarding operation, i.e. when the MN communicates with the CN while remaining attached to the same foreign network access router. Table 4.2 shows this operational overhead.

From our analysis it emerged that MPLS can be profitably used to complement PMIPv6, as it enhances the tunnelling paradigm with fast forwarding techniques and the possible support of Traffic Engineering. We showed that MPLS adds no extra overhead to MIPv6/PMIPv6; conversely it may even contribute in reducing both handover delay and the operational overhead.

4.4 Numerical Results

In this section, the performance of the LinkWork Mobile MPLS architecture has been evaluated and compared with related proposals such as Mobile IP, Mobile MPLS and FH Micro Mobile MPLS. The parameter settings in our experiments are listed in Table 4.3. Parts of the parameter values are referred to the paper [99]. The settings of the h_{x-y} values are represented by Figure 4.4.

Figure 4.7 presents the comparison of the signaling required in the registration process vs. resident time when parameters have their default settings. LW Mobile MPLS can significantly reduce the registration cost particularly when the MN handover frequently (i.e. the resident time in each subnet is short). As we have stated before, one of the main goals of LW Mobile MPLS is to manage efficient mobility in access networks. The introduction of LN nodes in the MPLS domain allows the reduction of signaling exchange by the creation of a linkworked LSP that allows local registration. In this case, Mobile MPLS scheme is the costliest proposal due to the requirement of establishing a complete LSP tunnel from MN to HA apart from the specific MIP signaling. On the contrary, LW

Table 4.2. Tunnelling Overhead

Tunnelling Mechanism	Overhead
IP-IP	20 Bytes
GRE	24 Bytes
MPLS	4 Bytes

Table 4.3. Parameter settings

Parameter	Value
t_s	1000 sec.
t_r	5 ~ 50 sec.
N_h	t_s/t_r sec.
RTT_{MN-AR}	1 sec.
s_u	56 Bytes
s_{pu}	76 Bytes
s_l	28 Bytes
B_w	100 Mbps
B_{wl}	11 Mbps
L_w	1 msec.
L_{wl}	2 msec.
P_t	10^{-6} sec.
λ_d	64 kbps
h_{x-y}	Figure 4.4

Mobile MPLS uses efficiently the resources in the MPLS access network since it reduces the signaling to an area, not overloading links and nodes near ILER.

As we could see in Figure 4.6, in our LinkWork Mobile MPLS solution, the LN can be located at different distances from the PELER, in the path between the ILER and the own PELER. In order to measure the impact of the location of the LN, Figure 4.8 shows the signaling cost when LN is located at different distances from the PELER. The four lines in the graph correspond to the four scenarios shown in 4.6.

Figure 4.9 shows the amount of lost packets during the whole connection session for different approaches. Both MIP and Mobile MPLS have the largest amount of lost packets due to the lack of buffering mechanisms. In fact, Mobile MPLS scheme has a larger amount of lost packets than MIP since the first proposal requires a higher establishment time to setup an LSP between the HA and the new serving agent in the visited network.

In contrast, both FH Micro Mobile MPLS and our proposal, LW Mobile MPLS, provide the best results thanks to the buffering and recovery mechanism. Notice its similar values of lost packets because both proposals initiate the buffering mechanism at the same time. However, there are significant differences between them that also justify the difference in the performance of registration update cost. First of all, LW Mobile MPLS approach

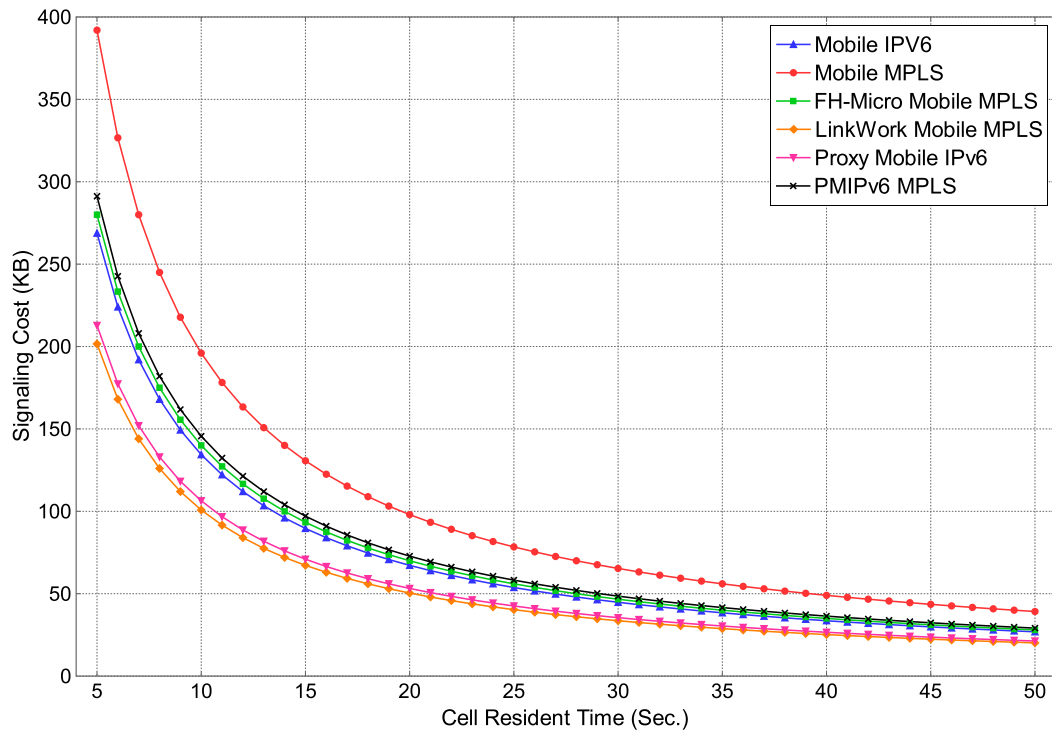


Figure 4.7. Signaling cost

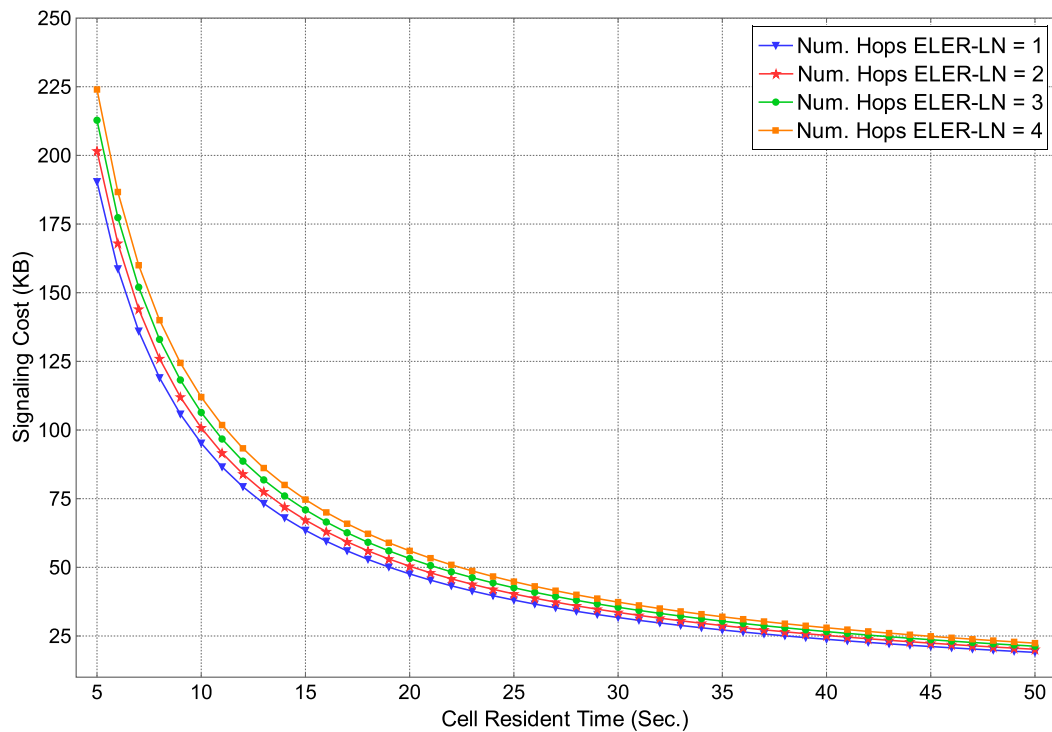


Figure 4.8. Signaling cost in LW Mobile MPLS with different lengths between the ELER and the LN

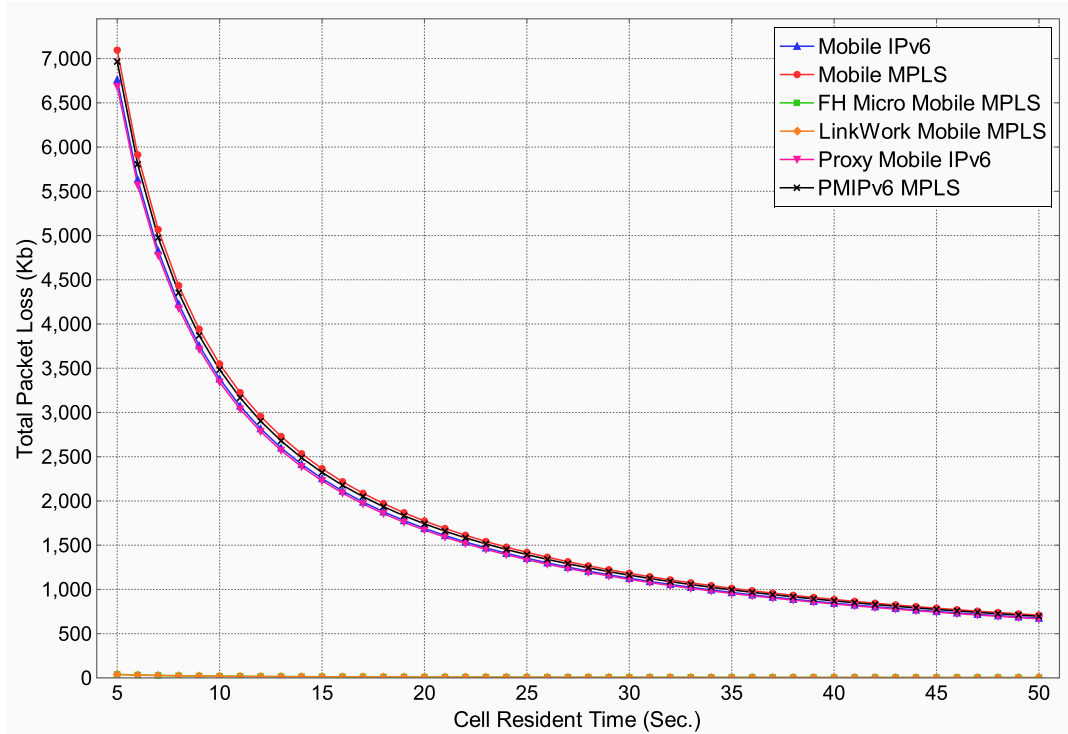


Figure 4.9. Total packet loss during a session

performs the forwarding LSP chain in LN nodes, which are internal routers of the domain whereas FH Micro Mobile MPLS consider several LSPs, one of them is the active LSP (the service is provided by this active one) and the others are passive LSP that are those from the LERG (the root of the domain) to the neighboring LER/FA of the current serving foreign agent. One of this passive LSP is activated when the MN moves to its LER/FA.

In our opinion, the possibility of selecting a few nodes inside the domain as LN can improve the flexibility of the architecture and could be easily adaptable to the needs of the service provider. Secondly, the recovery mechanism proposed in our LW Mobile MPLS architecture is designed to deliver recovered packets in the correct order, this means that our proposal saves the upper transport layer to do this task.

As it was mentioned in Section 4.3.2, packet loss is very dependent the handover latency. In Figure 4.10 we can see that the handover latency increases proportionally with the increment in the access network diameter.

With respect to buffer size requirements, a buffer is needed to store in-flight packets during each handover operation. As stated before, only LinkWork Mobile MPLS and FH Micro Mobile MPLS implement this feature. In this case, the LW proposal needs a buffer greater than the FH proposal one. This difference is based on the fact that our proposal forwards the recovered in-flight packets in the correct order, so the buffer needs to store more

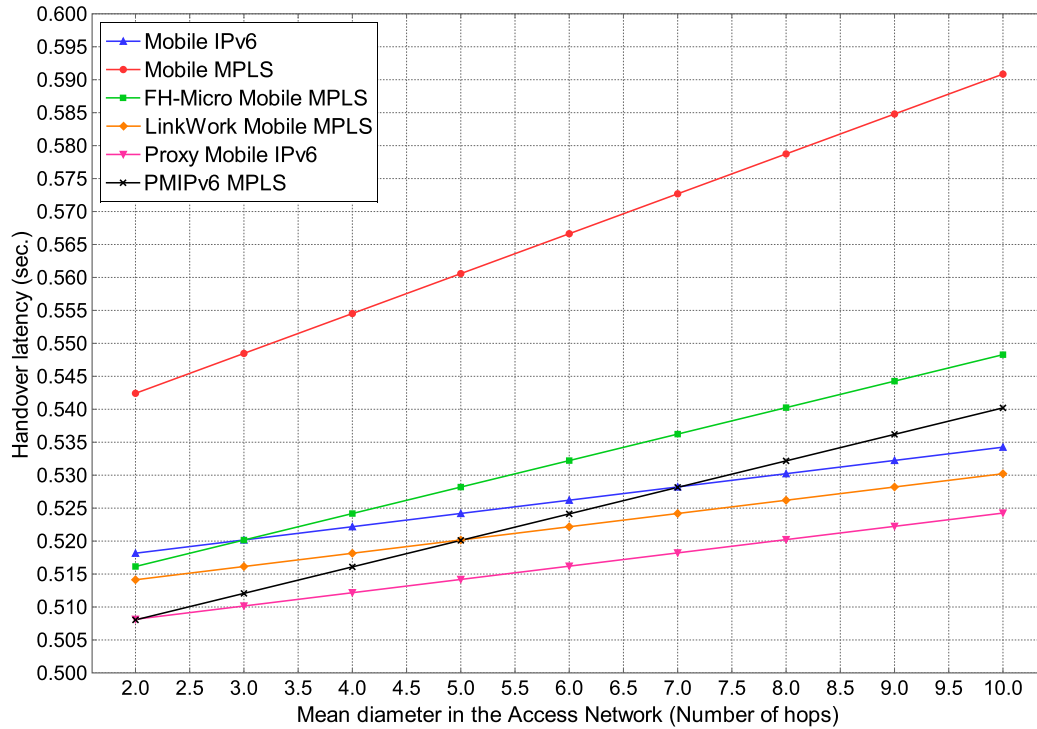


Figure 4.10. Impact of network diameter in handover latency.

packets. Figure 4.11 shows the buffer size vs. the bandwidth of the MPLS access network. In this graph we can observe that from 500 Mbps, the size of the buffer maintains rather stable values around 0.5 ~ 1.50 Kb.

4.5 Concluding remarks

In this chapter we propose a new architecture called LinkWork Mobile MPLS that offers an efficient management of the mobility of the MPLS access networks. The most interesting novelty in this architecture is the use of some special LSR that we call Linkage Nodes. These nodes are responsible for rerouting the LSP tunnel to the LER that serves the mobile node in each handover. Also these nodes retrieve the packets in flight when a service interruption is provoked by a handover. These mechanisms solve some of the problems of the Mobile IPv6.

Through the analytical study carried out we have obtained the behavior of the architecture in relation to the links, the signaling costs, the packets lost during the movements in each session and the ideal sized buffer needed to accomplish the proposed mechanism of retrieval. We compared our work with other previous research such as Mobile IPv6,

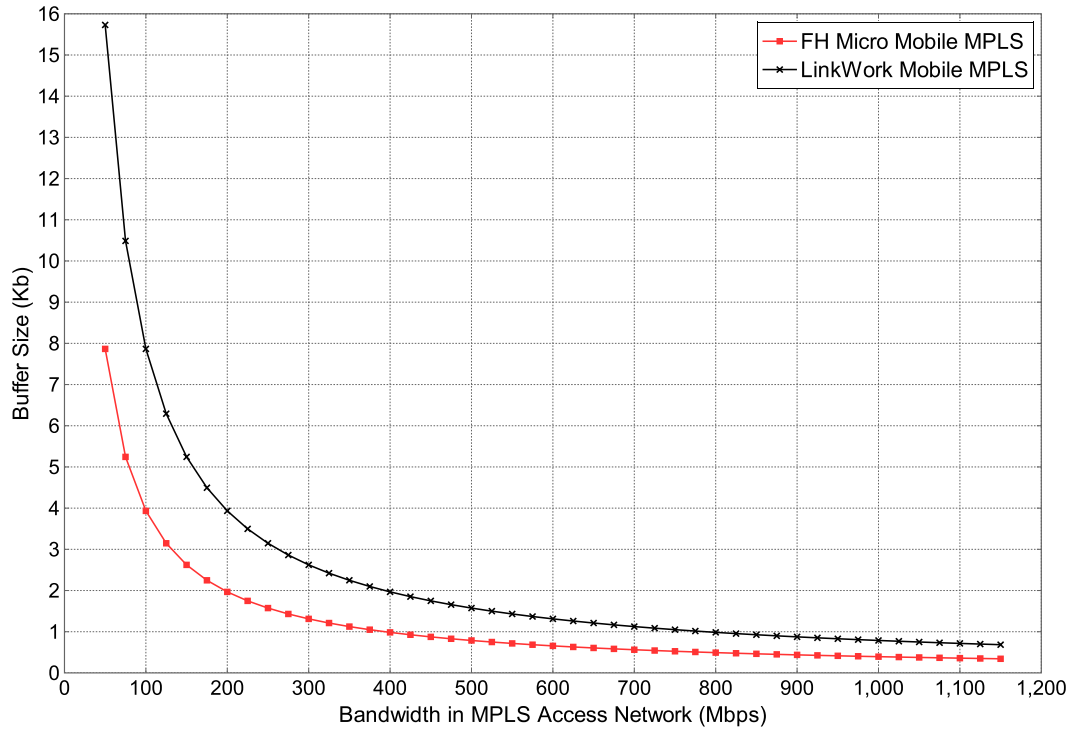


Figure 4.11. Buffer size vs. bandwidth in MPLS access network

Mobile MPLS, FH-Micro Mobile MPLS, Proxy Mobile IPv6 and PMIPv6 MPLS. We highlight the small signaling cost of LW Mobile MPLS and also the great capacity to minimize the packet loss compared to other previous proposal. The analysis proves the need to use a buffer mechanism to store in-flight packets in order to achieve this packet loss improvement.

Finally, from our study it emerged that, in some scenarios, MPLS can be profitably used to complement mobility protocols, as it enhances the tunnelling paradigm with fast forwarding techniques and the possible support of Traffic Engineering. One of the main conclusions of this chapter is that MPLS adds no extra overhead and it may even contribute to reducing both handover delay and the overhead during data forwarding.

Furthermore, some drawbacks still remain in the centralized mobility management protocols. In general, with the increase of mobile Internet traffic and the number of user devices, such centralized models encounter several barriers for scalability, security and performance, such as a single point of failure, longer traffic delays and higher signaling loads. For those cases in which CMM are not adequate due to the aforementioned limitations, a distributed paradigm can be applied. In the next chapter, we propose DM3, a distributed IP mobility approach based on LinkWork Mobile MPLS, where the mobility functions of control and data plane are distributed through the network nodes.

5

DM3: A DISTRIBUTED APPROACH

In order to deal with the limitations that arise from traditional network deployments, in this chapter we introduce DM3, a distributed architecture that inherit its behavior from the LW Mobile MPLS mechanism described in the previous chapter. In DM3, several nodes are distributed in the access network and the MNs are served by a close-by mobility anchor. With this operation, we reduce the routing and signaling cost, and provide a low handover latency with a minimal packet loss rate. Analytical and experimental results are presented to justify the benefits of our proposed architecture.

Contents

5.1	Introduction and contributions	87
5.2	Distributed Mobility Management MPLS	88
5.2.1	DM3 Operation	89
5.2.2	DM3 Recovery mechanism	91
5.2.3	Mobility functions in DM3	92
5.2.4	MDA selection process	93
5.3	Analytical results	94
5.3.1	Signaling cost	94
5.3.2	Packet delivery cost	96
5.3.3	Tunnelling Cost	98
5.3.4	Handover latency and Packet loss during a session	99
5.3.5	Buffer size	100
5.3.6	Mobility Anchors load	100
5.4	Performance evaluation	101
5.4.1	Analytical results	101
5.4.2	Simulation environment and results	106
5.4.3	Experimental results	112
5.5	Concluding remarks	116

5.1 Introduction and contributions

As it was revealed in Chapter 4, future wireless networks are expected to serve a large number of mobile subscribers that is growing exponentially. Moreover, the increasing demand in mobile data traffic highlights new challenges that need to be solved in order to achieve a successful evolution of IPv6 mobility management protocols.

The traditional structure of cellular networks is suffering some problems due to its limitation to deal with the exponential growth of the mobile data traffic in the Internet. Thus, new challenges arise in the evolution of mobility management in the Internet. Firstly, the fact that a small number of centralized anchors manage the traffic of millions of mobile nodes [16]. Secondly, the ability to provide a similar level of QoS while a user moves among heterogeneous networks [47, 63].

Regarding the first challenge, current mobility management schemes developed for IP and cellular networks rely on a centralized mobility anchor entity. This node is responsible for both mobility signaling and user data forwarding. This centralized approach is likely to have several issues or limitations, which require costly network scaling and engineering to resolve. The main problems identified concerning centralized solutions are: non-optimal routing, scalability issues and excessive signaling overhead, that implies longer handover latencies and vulnerabilities due to the existence of a single point of failure [100].

The second goal, related to assuring the provision of enough network resources, has been largely studied in both wired and wireless environments. As we stated in the previous chapter, MPLS natively supports tunnelling and also offers fast forwarding times. In addition, MPLS with its Traffic Engineering is a QoS technology introduced to enhance the performance of the Internet's datagram model in terms of both management and delivery.

In this scenario, and in order to deal with the aforementioned challenges, we continue the work made in the previous chapter to adapt its operation in a distributed way. Thus, starting from the centralized LW Mobility Management MPLS approach, in this chapter we introduce DM3: Distributed Mobility Management MPLS.

Hence, the major contributions of this chapter are fourfold: (1) we present a fully distributed mobility architecture called DM3 that addresses the limitations of centralized mobility architectures and leverage on the Distributed Mobility Management paradigm; (2) The performance of some metrics such as signaling cost, packet loss and buffer size are analyzed through the mathematical framework shown in Chapter 3; (3) an experimental testbed has been built in order to evaluate the impact of the location of the distributed mobility anchor

on the performance of the architecture; (4) Simulation using MATLAB has been performed to measure the behavior of the proposed architecture under different conditions.

The numerical results obtained from analysis, simulation and real experimentation show that DM3 outperforms in most cases the performance of other IPv6 mobility management protocols in terms of routing cost, registration update cost, handover latency and packet loss rate during movements.

5.2 Distributed Mobility Management MPLS

In this section, we introduce DM3, a new DMM architecture that is based on Mobile IPv6. The aim is to achieve an efficient mobility management with QoS support taking advantage of both new distributed mobility management approach and MPLS features. As we can see in Figure 5.1, the architecture of DM3 is inherited from LinkWork Mobile MPLS, with the difference that DM3 relies on the DMM paradigm, whereas LinkWork Mobile MPLS were based on CMM.

We assume that an MPLS domain exists in the access network between the Ingress LER/Egress LER. Both ILER and ELER are the border MPLS routers that define the

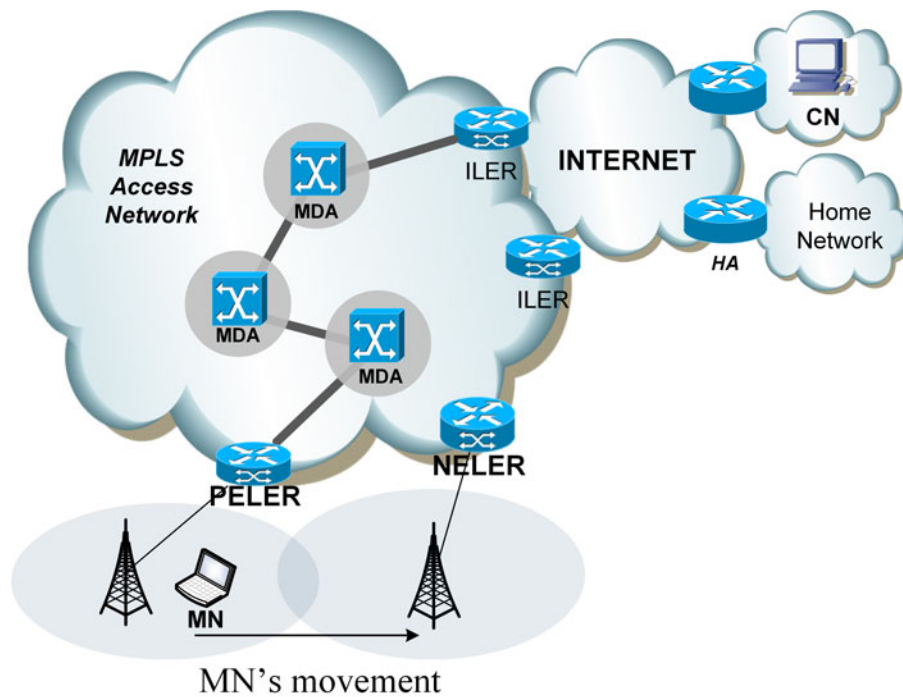


Figure 5.1. Overview of the DM3 approach

limits of the access network. DM3 architecture relies on the distributed mobility agent called MDA (Mobility Distributed Anchor). This node provides mobility management functions [101] and is an intermediate node between the ILER and the serving ELER. The serving ELER is the egress router the MN is currently attached to. Taking into account that the HA not only manages the mobility context, but also manages routing, the main idea behind the DM3 architecture is to benefit from the position of the node that manages the mobility and routing functions in order to reduce the limitations of Mobile IPv6 mentioned in previous sections. As it is stated in [102], the distribution of mobility management based on the decoupling of functionalities can bring several benefits in terms of scalability, security and performance.

Moreover, the proposed architecture is based on the forwarding chain concept (set of forwarding paths). The Mobility Distributed Anchor agent (MDA) is responsible for the LSP redirection when the MN moves to an adjacent network. This way, the LSP is going to be composed of a set of forwarding paths that will adapt to track host mobility and localize signaling in an area close to the location of the MN.

5.2.1 DM3 Operation

The basic operation of the architecture is illustrated in Figure 5.2 in which the MN moves to an adjacent network, changing its point of attachment to the network. Detailed descriptions are as follows:

Initially, when the MN moves to an adjacent network, it proceeds as follows. The MN enters an overlapped area of an adjacent subnet, it receives an L2 signal from the possible new base station (step 1). Next, the MN notifies the PELER of the possibility of a handover by sending a HI (Handover Initiate) message that contains the new Base Station identifier.

This information is going to be used to obtain the NELER IP address, thanks to a data structure that maintains a match between this identifier and each adjacent LER IP address (step 2). It is supposed that the MPLS access network belongs to the same service provider. Once the PELER knows the subnet to which the MN is going to move, it sends a message upstream to the selected MDA, notifying of a possible L3 handover, and starting the setup of a new section of the LSP from MDA to NELER (step 3) with the required QoS. In this step, the PELER also informs the MN about the NELER IP address through an IPv6 Neighbor Discovery message. At this moment, a new section of the LSP tunnel could be set up so data traffic could be forwarded towards the new location of the MN (step 4). When the signal strength received from the current base station falls below a certain threshold level, the MN notifies the handover to the PELER (step 5). Now, our

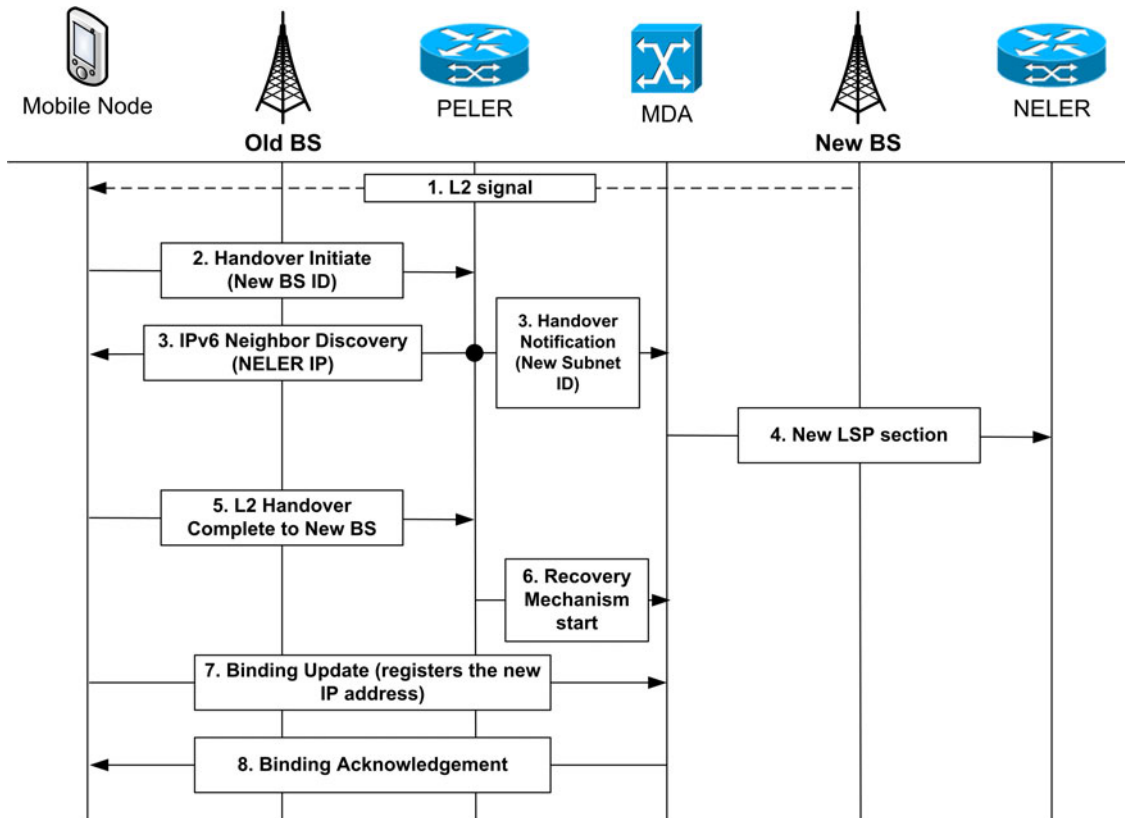


Figure 5.2. Handover from PELER to NELER in DM3

mechanism responsible for minimizing packet loss is started by the PELER (step 6). This threshold can be determined by handover decision algorithms that rely on Received Signal Strength (RSS) and other link-layer parameters. The next section give details about this packet recovery mechanism.

Once the L2 handover is performed, the MN initiates a L3 handover through the Mobile IPv6 registration process (step 7). The new LSP section from MDA to the new egress router will be used when the MDA is aware of the movement. This happens when the PELER starts to return data packets to the MDA, which will be forwarded to the NELER through the new LSP section together with buffered packets according to the recovery mechanism. Finally, the MN sends the Mobile IPv6 Binding Update message to the mobility anchor (MDA) (step 7). The MDA will reply with a Binding Acknowledgement message to the MN that is located in the new subnet (step 8).

The novelty of this handover procedure is the recovery mechanism, as well as the selection of the correct mobility anchor after the movement. Both processes are explained next.

5.2.2 DM3 Recovery mechanism

The recovery mechanism designed for DM3 is similar to the developed for LinkWork Mobile MPLS. As it was detailed in section 4.2.2, the frequency of path disruption due to a handover in mobile networks is very high. With this recovery mechanism, we achieve lower packet loss and avoid packet disorder, two metrics that badly affect the performance of the overall architecture. This mechanism is based on the use of two buffers. In this section we present the DM3 recovery mechanism. Its operation is shown in Figure 5.3 and is detailed as follows.

The MDA nodes have a data structure of 2 buffers (B1 and B2) that allow them to store both in-flight packets (through the path MDA-PELER) and incoming packets. When the MN informs the PELER about an L2 handoff, this edge router does not send any more packets to the MN, instead, it sends the packets back to the MDA as Figure 5.3a illustrates. When the first packet arrives back to the MDA (packet n in the figure 5.3),

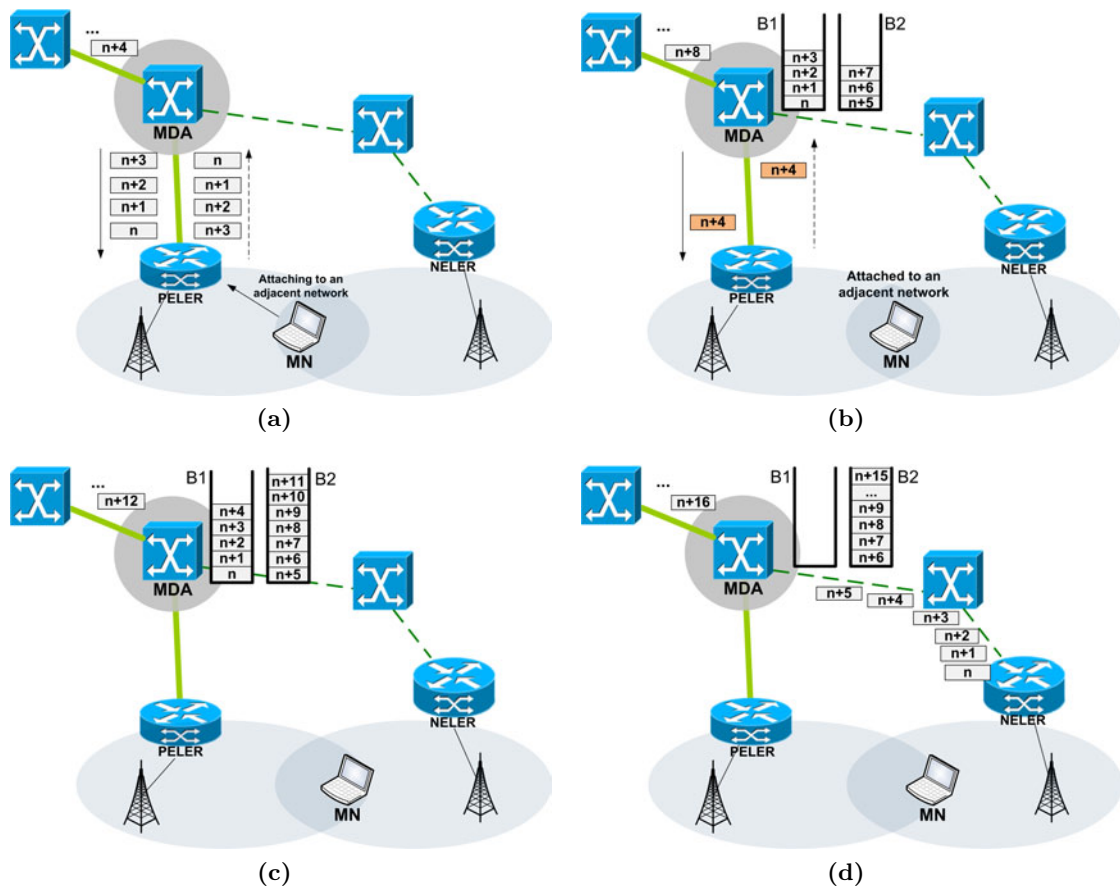


Figure 5.3. Recovery mechanism operation

it tags the next packet received ($n + 4$) and sends it through the PELER again. The incoming packets from the PELER are buffered in B1 while incoming packets from the data path are buffered in B2 (see Figure 5.3b). Once the MDA receives the tagged packet ($n + 4$) from the reverse path, it untags it, and buffers the packet in B1 (see Figure 5.3c).

B2 buffers the rest of the packets until the Binding Update message arrives to the MDA. When the BU is received, all packets can be forwarded through the new section of the LSP (MDA-NELER) as Figure 5.3d shows. In order to avoid packet disorder, packets in B1 are sent before packets in B2. This way, packet loss is minimized and packets are sent in the correct order towards the new location of the MN therefore the work required of the MN to reorder the information is significantly reduced.

The delay analysis of the DM3 recovery mechanism is similar to the detailed in Section 4.2.2.1.

5.2.3 Mobility functions in DM3

In DM3, mobility management is fully distributed. That means that, both the data plane and the control plane are distributed, as it can be seen in Figure 5.4. The data plane is responsible for routing packets to the corresponding peer entity, whereas the control plane maintains the mobility bindings. In the data plane, the mobility routing (MR) function is distributed to multiple locations (MDAs) and each MN is assigned to the closest light load MDA, so the routing can be optimized, avoiding inefficient paths after various movements. In the control plane, the MDAs may signal each other, and the Location Management (LM) function is a distributed database also located at MDAs that maintain the mapping of HoA to CoA.

The LM functions are located in the MDA due to the improvement this provides. To perform mobility routing, the MDA needs the location information and when a MN performs a handover, the LM database needs to be updated. In DM3, when a handover occurs, the MDA is notified so the LM database can be updated at the same time as the mobility anchor is aware of the movement.

The operation of the routing and control functions in DM3 is as follows. When a mobile node attaches to a new LER and initiates an IP communication with a correspondent node, the traffic will be anchored to a MDA. The anchor selection is an important process of the distributed control plane and it is explained in the next section. When performing handover to another network and a different MDA is selected, this movement will be shared by these two MDAs. Once the handover has been made and the previous mobility anchor has been notified of the movement, that MDA can forward packets to the new MN's. Registration

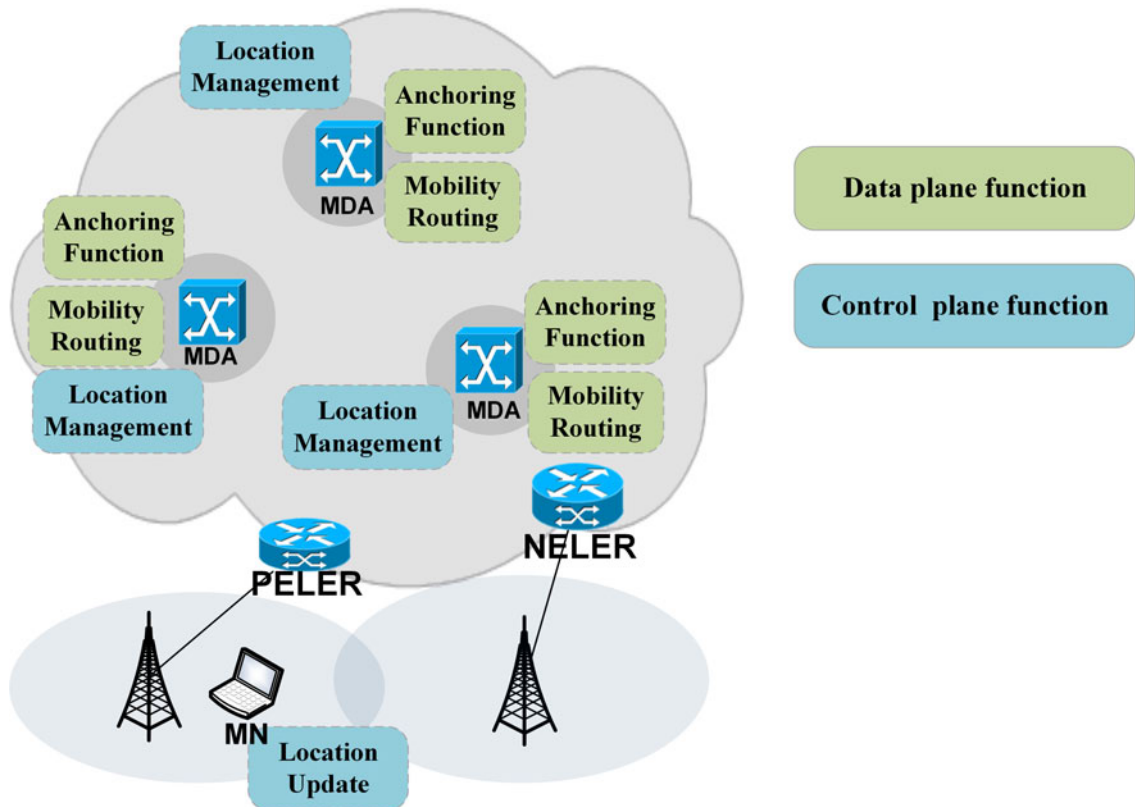


Figure 5.4. Distributed Mobility functions in DM3

update to the control function is initiated by the host. With regards to the location information, when the mobile host moves to another network, the location information needs to be updated and this information may need to be disseminated among MDAs. The user data can be continuously delivered to the MN in the new location by rerouting the tunnel with a new MPLS segment between the old MDA and the NELER (passing by the new MDA). Distributed mobility management architectures require enabling dynamic mobility anchor selection mechanisms.

5.2.4 MDA selection process

One of the most significant issues of the distributed control plane, is that a special mechanism is needed to identify the mobility anchor that maintains the mobility binding of the mobile host. In DM3, every time a MN attaches to the domain or a handover occurs, it is necessary to decide which MDA should be selected for it. In DM3 architecture, mobility anchor selection depends on two main factors. The distance between the NELER and the MDA and the MDA's load. NELER should select the closest light load MDA. For each MDA, let C_i be its capacity in terms of number of flows supported and N the set of

mobile nodes. A traffic flow originating or destined for a mobile node $n \in N$ through an MDA i is given as f_i^n . Hence, we define L_i , as the load of an MDA i as follows:

$$L_i = \frac{\sum_{n \in N} (f_{ni})}{C_i} \quad (5.1)$$

Where f_{ni} represents the total flow of all mobile nodes in MDA i . Let $h_{NELER-MDA_i}$ be the hop count between the LER that MN accesses and the MDA i . We suppose that each LER knows the distance, in number of hops, between it and each MDA, and there are n MDAs in this domain. Thus, the output of the MDA selection can be expressed as follows:

$$MDA_{selection} = \min \{L_i \times h_{NELER-MDA_i}\} \quad (5.2)$$

5.3 Analytical results

In order to evaluate the performance of DM3, we analyze the mobility cost functions of several indicators involved in mobility aspects, such as the cost functions of traffic routing, registration updates, total packet loss during a session, tunnelling overhead and buffer size metrics. Moreover, we also analyze other key parameters to understand the difference between centralized and distributed mechanisms, such as the length of the complete path, the traffic intensity in the mobility anchor and the network load. These indicators have been selected because they are the main limitations of mobility management protocols and they need to be evaluated [24]. Our proposal is also compared with other mobility proposals, such as Mobile IPv6, Proxy Mobile IPv6, FH-Micro Mobile MPLS, Host-based DMM and Network-Based DMM. The analytical study is based on the considerations made in Chapter 3.

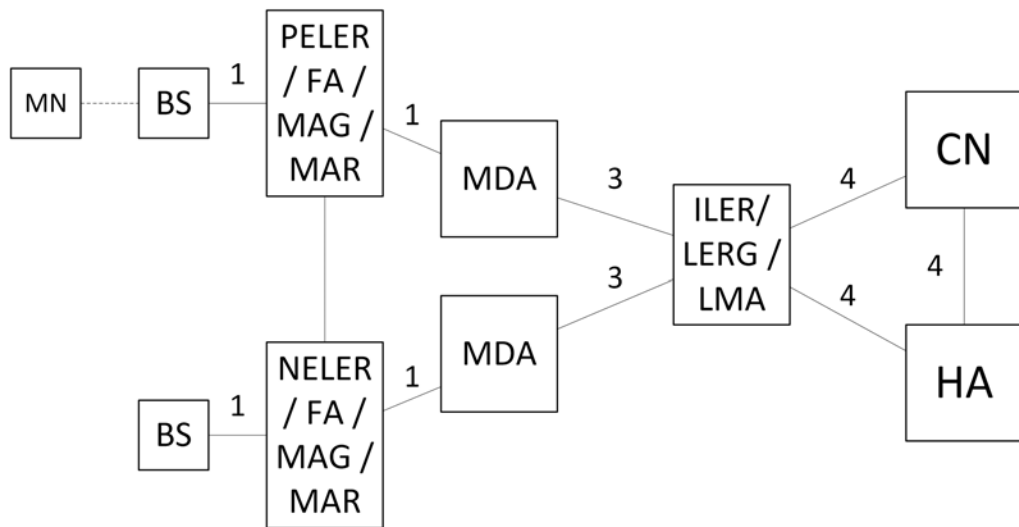
We analyze the mobility behavior of the MN, keeping in mind a topology where a terminal could move to every neighbouring network with the same probability. In the next section, quantitative and experimental results are presented. The parameters and values used in the analysis and in the quantitative results, and are similar to the values shown in Table 4.3. In this case, relative distances in hops (h_{x-y}) are shown in Figure 5.5.

5.3.1 Signaling cost

As it was introduced in Section 3.4.1, one of the key metrics that show the behavior of the data plane is the total signaling cost of registration updates. Defined as C_s , this parameter show the traffic load when signaling messages are sent.

Table 5.1. Parameter values

Parameter	Value
t_s	1000 sec.
t_r	5 ~ 50 sec.
N_h	t_s/t_r sec.
RTT_{MN-AR}	1 sec.
s_u	56 Bytes
s_{pu}	76 Bytes
s_l	28 Bytes
s_t	40 Bytes
s_m	4 Bytes
s_d	1 KByte
B_w	100 Mbps
B_{wl}	11 Mbps
L_w	1 msec.
L_{wl}	2 msec.
P_t	10^{-6} sec.
λ_d	64 kbps
h_{x-y}	Figure 5.5

**Figure 5.5.** Relative distances in hops in the network

The $C_s(MIPv6)$ of centralized approaches was investigated in Chapter 4. Thus, $C_s(MIPv6)$, $C_s(PMIPv6)$, $C_s(Mobile MPLS)$ and $C_s(FH-MicroMobileMPLS)$ were defined in Equations 4.16, 4.20, 4.17 and 4.18 respectively.

The distributed approaches, such as DM3, Host-Based DMM and Network-Based DMM update their movements with the distributed anchor, located closer to the location of the mobile node. In Host-Based DMM, the mobility anchor is called AMA and is located in the access router so AMA is the first IP capable router for the MN. In this approach, when a mobile node moves, the MN registers its movement to the serving AMA and the serving AMA (sAMA) establishes bidirectional tunnels with previous AMA(s) the MN was connected to. Similarly, this occurs with NB-DMM, where upon a handover, the new MAR retrieves the IP addresses of the anchoring MARs for the MN from the DB. The new MAR then registers the MN at all these MARs. Apart from signaling the mobility management protocol, some proposals also add the cost of the LSP procedure set-up. This is the case of Mobile MPLS, FH-Micro Mobile MPLS and DM3. Let n be the number of valid addresses configured at the MN in HB-DMM and NB-DMM (n = number of handover + 1).

This way, we obtain the following values for the signaling cost when the registration update process occurs in the distributed solutions:

$$C_s(DM3) = 2 \cdot s_u \cdot h_{MN-MDA} \cdot N_h + 2 \cdot s_l \cdot h_{MDA-NELER} \cdot N_h \quad (5.3)$$

$$C_s(HB - DMM) = [2 \cdot s_u \cdot h_{MN-sAMA} + \sum_{i=1}^{n-1} 2 \cdot s_u \cdot h_{AMA_i-sAMA}] \cdot N_h \quad (5.4)$$

$$C_s(NB - DMM) = [2 \cdot s_{pu} \cdot h_{AR-DB} + \sum_{i=1}^{n-1} (2 \cdot s_{pu} \cdot h_{MAR_i-sMAR})] \cdot N_h \quad (5.5)$$

5.3.2 Packet delivery cost

The total Packet Delivery Cost for a session (C_{PDC}) has been defined in Section 3.4.2. As it was described, this value represents the transmission cost between the MNs and its CNs in the network, and is influenced by the size of the data messages multiplied by the number of hops needed to forward packets from the CN to the MN or vice versa.

In MIPv6 and PMIPv6, packets are routed from the CN to the MN's anchor, HA or LMA respectively, and forwarded from the anchor to the MN through a tunnel that encapsulates the data packets. Note that the packet delivery mode considered in MIPv6

is the bidirectional IP tunnelling, i.e., without route optimization. Thus, the expressions that represent the cost can be expressed as follows

$$C_{PDC}(MIPv6) = (s_d \cdot h_{CN-HA} + (s_t + s_d) \cdot h_{HA-MN}) \cdot \lambda_d \cdot t_s \quad (5.6)$$

$$C_{PDC}(PMIPv6) = (s_d \cdot h_{CN-LMA} + (s_t + s_d) \cdot h_{LMA-MAG} + s_d \cdot h_{MAG-MN}) \cdot \lambda_d \cdot t_s \quad (5.7)$$

In Mobile MPLS and FH-Micro Mobile MPLS, the data path followed by a packet forwarded from the CN and the MN is similar to the one described in MIPv6. The difference is that the bidirectional tunnel is established by using MPLS in both cases. Thus, the C_{PDC} is

$$C_{PDC}(MobileMPLS) = (s_d \cdot h_{CN-HA} + (s_m + s_d) \cdot h_{HA-MN}) \cdot \lambda_d \cdot t_s \quad (5.8)$$

$$C_{PDC}(FH - MicroMobileMPLS) = (s_d \cdot h_{CN-LENG} + (s_m + s_d) \cdot h_{LENG-MN}) \cdot \lambda_d \cdot t_s \quad (5.9)$$

In HB-DMM and NB-DMM, when a MN moves, the traffic established in the new network will be routed directly to the CN whereas the remaining connections will be tunnelled to its corresponding anchoring MAR and then routed to the CN. The C_{PDC} in these distributed solutions is

$$C_{PDC}(HBDMM) = C_{PDC}(NBDMM) = (P_n \cdot C_{PDC}^d + P_h \cdot C_{PDC}^i) \cdot \lambda_d \cdot t_s \quad (5.10)$$

where P_n and P_h are respectively the probabilities that a traffic is a new or handover traffic. P_{PDC}^d and P_{PDC}^i are the units of cost of delivering one packet in the direct and indirect modes of DMM, respectively. Then these costs are expressed as follows

$$P_{PDC}^d = s_d \cdot h_{CN-sMAR} + s_d \cdot h_{sMAR-MN}$$

$$P_{PDC}^i = \sum_{i=1}^{n-1} [(s_t + s_d) \cdot h_{MAR_i-sMAR} + s_d \cdot h_{CN-MAR_i}] + s_d \cdot h_{sMAR-MN}$$

By its part, DM3 distribute mobility by means of MDA anchors in the access network so, data sent from the CN goes to the MDA through the ILER and finally are delivered to the MN. Hence, the value of Packet delivery cost for DM3 is

$$P_d(DM3) = (s_d \cdot h_{CN-ILER} + (s_m + s_d) \cdot h_{ILER-MDA} + (s_m + s_d) \cdot h_{MDA-MN}) \quad (5.11)$$

5.3.3 Tunnelling Cost

To achieve seamless mobility, all solutions use a tunnel to forward packets. This metric was described in Section 3.4.3 and represents the cost of adding tunnelling overheads to the Data Packet Delivery Cost so, Tunnelling Cost (C_t) can be derived from it by setting data packets size to zero, $s_d = 0$.

In MIPv6 and PMIPv6, the traffic is tunnelled from the centralized anchor to the MN or the MAG agent.

$$C_t(MIPv6) = s_t \cdot h_{HA-MN} \cdot N_h \quad (5.12)$$

$$C_t(PMIPv6) = s_t \cdot h_{LMA-MAG} \cdot N_h \quad (5.13)$$

Mobile MPLS tunnels data from the HA to the MN, whereas FH-Micro Mobile MPLS, do it from the LERG to the LER/FA node. Thus, we obtain the following expressions,

$$C_t PDC(MobileMPLS) = (s_m \cdot h_{HA-MN}) \cdot \lambda_d \cdot t_s \quad (5.14)$$

$$C_t(FH - MicroMobileMPLS) = (s_m \cdot h_{LERG-MN}) \cdot \lambda_d \cdot t_s \quad (5.15)$$

On the other hand, the distributed solutions tunnels traffic between the anchors. In DM3, MPLS tunnelling technology is used instead of IPv6 over IPv6. The reasoning behind the idea is that since a tunnel is needed, employing a technology that natively supports tunnelling, seems a natural choice.

$$C_t(HB - DMM) = \sum_{i=1}^{N_h} s_t \cdot h_{AMA_i - sAMA} \cdot N_h \quad (5.16)$$

$$C_t(NB - DMM) = \sum_{i=1}^{N_h} s_t \cdot h_{MAR_i - sMAR} \cdot N_h \quad (5.17)$$

$$C_t(DM3) = (s_l \cdot h_{ILER-MDA} + s_m \cdot h_{MDA-MN}) \cdot N_h \quad (5.18)$$

5.3.4 Handover latency and Packet loss during a session

Another relevant metric for mobility management protocols is the amount of packet lost during a session (P_{loss}). Defined in Section 3.4.5 as the sum of lost packets per MN during all handovers, this metric depends on the handover latency, already defined. In previous chapter (Section 4.3.2 the Packet Loss and Handover latency expressions for centralized solutions were derived (*MIPv6*, *Mobile MPLS*, *FH - MicroMobileMPLS* and *PMIPv6*). In this section we complete that analysis with the Packet Loss and Handover Latency for the distributed solutions. Hence

$$T_h(HB - DMM) = \frac{1}{2}RTT_{MN-AMA} + 2 \cdot t(s_u, h_{MN-AMA}) + \sum_{i=1}^{n-1} (2 \cdot t(s_u, h_{AMA_i - sAMA})) \quad (5.19)$$

Thus, we have that packet loss in HB-DMM is:

$$P_{loss}(HB - DMM) = T_h(HB - DMM) \cdot \lambda_d \cdot N_h \quad (5.20)$$

Similarly to the HB-DMM, the NB-DMM approach has an analogous operation during handovers, thus

$$T_h(NB - DMM) = \frac{1}{2}RTT_{MN-MAR} + 2 \cdot t(s_u, h_{AR-DB}) + \sum_{i=1}^{n-1} (2 \cdot t(s_{pu}, h_{MAR_i - sMAR})) \quad (5.21)$$

Therefore, packet loss in NB-DMM is:

$$P_{loss}(NB - DMM) = T_h(NB - DMM) \cdot \lambda_d \cdot N_h \quad (5.22)$$

Finally, our DM3 solution introduces a recovery mechanism that minimizes the packet loss. Thus, if a handover occurs, in-flight packets can be recovered and forwarded to the new location of the MN. The values of T_h and P_{loss} can be defined as:

$$T_h(DM3) = \frac{1}{2}RTT_{MN-NELER} + 2 \cdot t(s_u, h_{MN-MDA}) + 2 \cdot t(s_l, h_{MDA-NELER}) \quad (5.23)$$

and

$$P_{loss}(DM3) = t(s_u, h_{MN-PELER}) \cdot \lambda_d \cdot N_h \quad (5.24)$$

5.3.5 Buffer size

As stated in previous section, both FH-Micro Mobile MPLS and DM3 minimize packet loss during handover so they need a buffer to store in-flight packets. In DM3, the buffer is located at the MDA and its size depends on the time needed for the recovery mechanism to store the packets in the correct order. In the FH-Micro Mobile MPLS scheme the buffer is located in the LER/FA node (a border router) and the buffer size depends on the time that takes a message to be forwarded from the MN to the FA plus the time it takes for a message to be forwarded to the FA to the next FA (where the MN is finally attached). In the other mechanisms, all in-flight packets are lost during the handover time due to the lack of any buffering mechanism. Therefore, the buffer size requirement (B_{size}) is listed as follows:

$$B_{size}(FH - MMM) = \left(\frac{1}{2}RTT_{MN-AR} \right) + t(s_u, h_{MN-FA} + h_{FA-FA}) \cdot \lambda_d \quad (5.25)$$

$$B_{size}(DM3) = \left(\frac{1}{2}RTT_{MN-AR} \right) + t(s_u, h_{MN-PELER} + h_{PELER-MDA}) \cdot \lambda_d \quad (5.26)$$

5.3.6 Mobility Anchors load

In order to compare the load of a mobility anchor both in CMM and in DMM, a new metric is analyzed. Mobility Anchor Load (MA_l) is defined as:

$$MA_l = \frac{\text{generated data by a MN}}{\text{number of mobility anchors}} = \frac{N_s \cdot \lambda_s \cdot t_s \cdot (N_h + 1)}{N_m} \quad (5.27)$$

where N_s is the total number of communication sessions of a MN and N_m is the number of mobility anchors associated with the MN. This parameter is $N_m = 1$ in MIPv6 and PMIPv6 and $N_m = N_h + 1$ in HB-DMM and NB-DMM. The value of this parameter in DM3 goes from $N_m = 2..N_h$. Additional discussion about this parameter will be made in the results section.

5.4 Performance evaluation

In this section, numerical results are presented in order to examine the behavior and performance of the different schemes according to the analytical model presented in the previous section. The evaluation measures several metrics regarding both data and control plane and points out the benefits and drawbacks of distributed and centralized models.

These numerical investigations are obtained from analytical results, simulations and from the results of a testbed where we have evaluated the impact of the location of the distributed nodes in the overall performance of the network.

5.4.1 Analytical results

Figure 5.6 shows the comparison of signaling cost of registration update as a function of the cell resident time, that varies from 20 to 140 seconds. As could be expected, the value of C_s achieve the highest values when the cell resident time is low. In conditions of very high mobility (the cell resident time takes low values), only DM3 can significantly maintain an acceptable value for this parameter. In fact, in DM3 this cost remains almost constant regardless of the mobility rate. Apart from DM3, in this zone of frequent handover, the behavior of DMM and CMM solutions can be easily distinguished. Both HB-DMM and NB-DMM (distributed mechanisms are represented in this section with solid lines) reach the highest costs whereas the centralized solutions (represented in this section with dotted lines) take intermediate values. This negative effect suffered by DMM solutions in scenarios of frequent movements among different subnets is due to the fact that the signaling messages in host-based and network-based DMM solutions are exchanged among all the ARs that have been visited by the MN during its movement and remain an active connection. These notifications cause a significant increment in the signaling cost of the distributed protocols. However, the centralized solutions only need to notify to the HA or LMA each time the MN moves to an adjacent network, minimizing the overall cost.

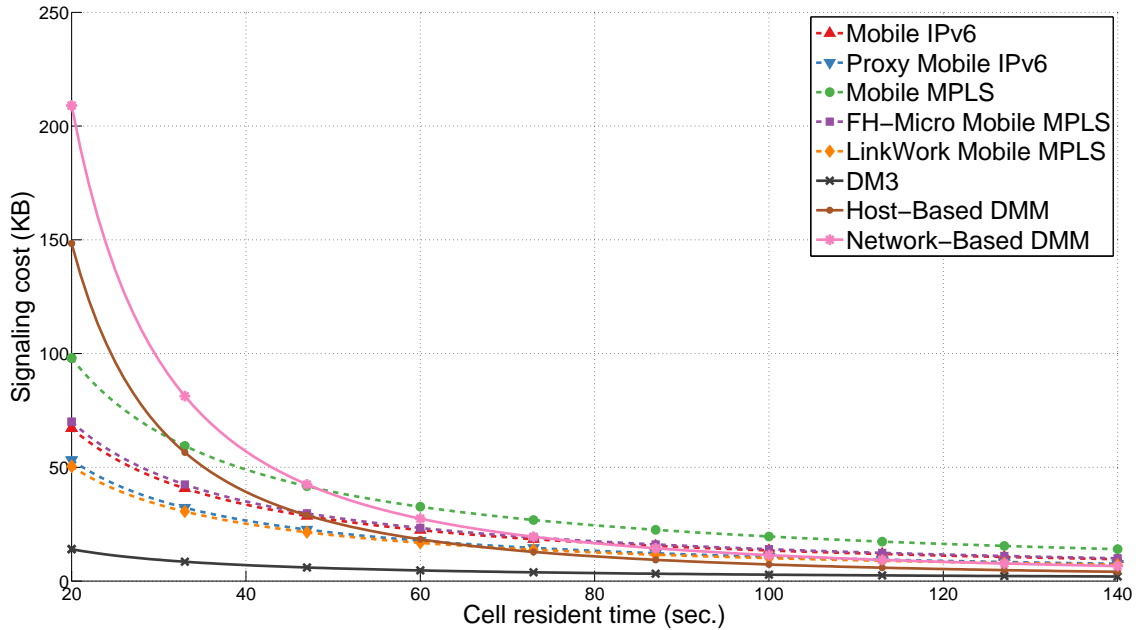


Figure 5.6. Signaling cost of registration updates

However, as cell resident time increases (the mobility of the nodes are lower), the behavior of signaling cost in DMM solutions decreases sharply and the cost remains low, even below the CMM protocols.

Our DM3 solution uses the network resources efficiently since it distributes the HA mobility functions in MDA nodes, not overloading links and nodes near the ILER. This way, DM3 can significantly reduce the registration cost particularly when the MN handoff frequently (i.e. the resident time in each subnet is short). The introduction of MDA nodes in the MPLS domain allows a reduction in the signaling exchange. This figure allows to quantify the trade-off between the different proposals in its control plane.

Data packet delivery cost represents the cost of delivering data packets to an MN per unit time. Figure 5.7 depicts the routing cost of forwarding data traffic during a session as a function of the Transmission rate. The results show that the packet delivery cost increase linearly with the transmission rate. As can be observed, distributed mechanisms outperform CMM solutions due to the operation of DMM protocols, that avoid long routes and forward traffic in an optimized way. In centralized protocols, all packets are routed through a centralized anchors and this often results in longer paths from MN to CN. However, the DM3 scheme optimizes the data path reducing an average of 35% compared to MIPv6 and 21% compared to PMIPv6. Furthermore, the DM3 solution has an average of 11% less packet delivery cost compared to the distributed solutions. This can be attributed to the fact that DM3 has a distributed control plane that allows that each session is anchored to a closer MDA.

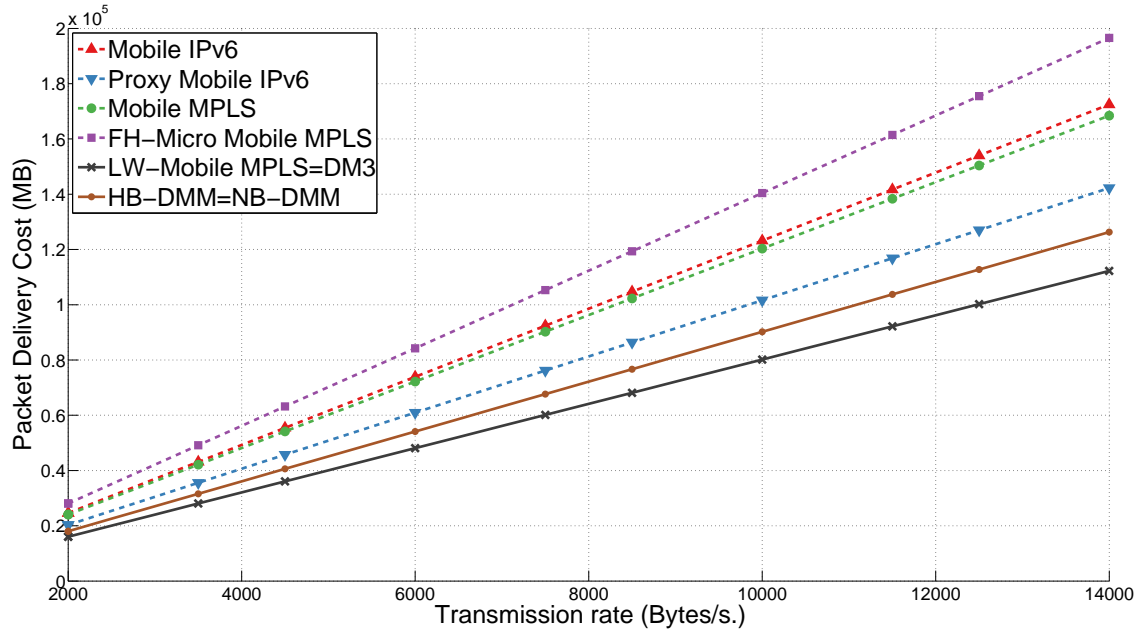


Figure 5.7. Packet delivery cost vs. transmission rate

Another metric closely related to packet delivery cost is the tunnelling cost. Figure 5.8, shows the variation of this cost as a function of the transmission rate. It appears clearly that this cost in DM3 is slower than in the other proposals. This is due to the data path, more optimized than in CMM, and also due to the use of MPLS tunnelling instead of IPv6 tunnelling. MPLS tunnels generate an overhead of 4B, whereas IPv6 tunnels generate an overhead of 40B. For this reason, MIPv6 and PMIPv6 costs rise significantly with respect to the rest of the mechanisms.

Due to the operation of both HB-DMM and NB-DMM, the MN initiates new sessions after the handover using the new IP address. The data traffic of these new sessions is routed in a more optimal way. As a result, the tunnelling cost is low. To obtain the results shown in Figures 5.7 and 5.8, we have considered 60% of the connections to be routed optimally, according to the statistics [3].

Next, the load of the mobility anchors is evaluated. Figure 5.9 shows the Mobility Anchor Load metric as the function of number of sessions established by the MN (n) at different access networks. As n increases, the load of mobility anchors of all protocols increases linearly, but at a different rate. In centralized mechanisms, the mobility anchor is responsible for all traffic forwarded from/to the MN so, its load increases faster than in the distributed schemes. The load of HB-DMM, NB-DMM and DM3 mobility anchors is lower because the traffic is processed in a distributed fashion and the failure impact of a mobility anchor among others is limited locally. Since the DM3 mobility anchors can be distributed at different levels of hierarchy, their behavior in these situations is outlined. This is

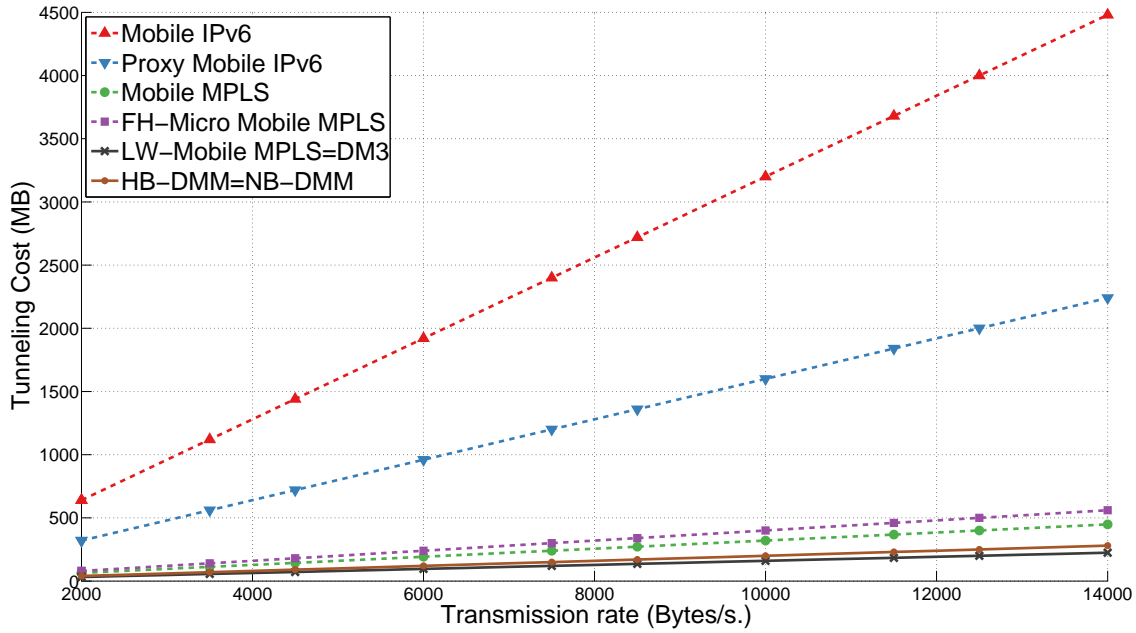


Figure 5.8. Tunnelling cost vs. transmission rate

distinct from HB-DMM and NB-DMM in which all mobility anchors ($Num_{AMA/MAR}$) are located in the access routers. Thus, DM3 can exhibit different number of mobility anchors (Num_{MDA}), depending on the hierarchy level in which they are located. Figure 5.9 also shows the mean traffic load of each mobility anchor in DM3 under these different configurations that vary. It is observed that, although HB-DMM and NB-DMM achieve

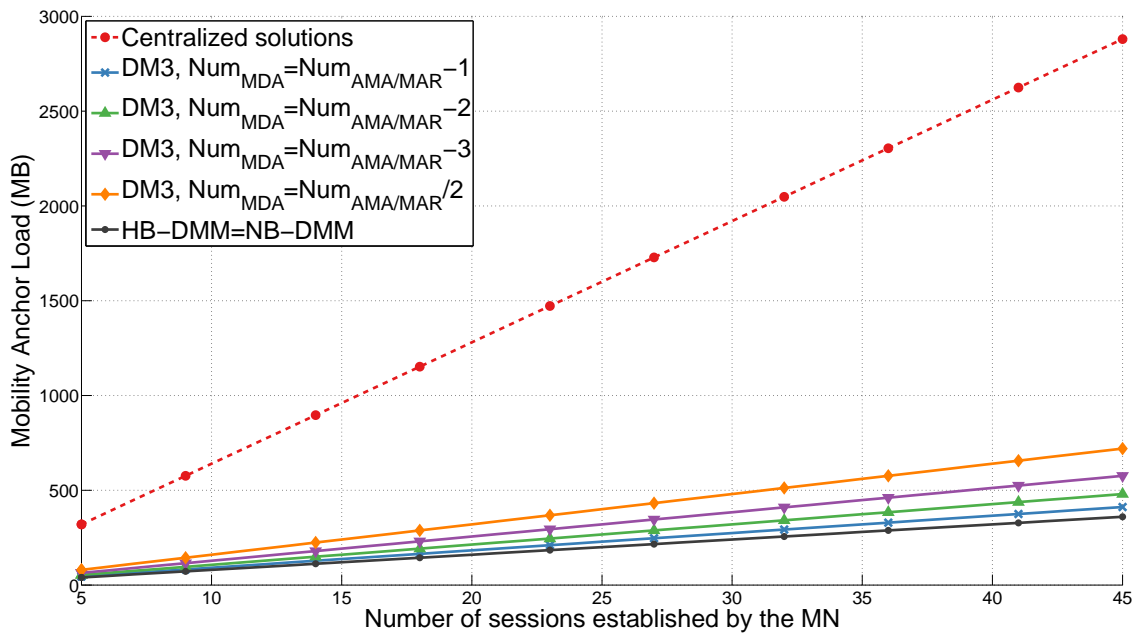


Figure 5.9. Mobility anchor load vs. number of sessions by MN

the lowest load, DM3 also obtains moderate loads, even when introducing only half of the mobility anchors as those in HB-DMM or NB-DMM ($Num_{MDA} = Num_{AMA/MAR}/2$).

Finally, the results of both handover latency and packet loss are shown. As it was stated before in the analytical model section, the handover latency is the sum of three components: the layer 2 handover, T_{L2} ; the time needed to perform the security and authentications tasks, T_{sec} ; and the time to establish the new route to the MN, T_e . Figure 5.10, shows the variation of the handover latency as a function of $n = N_h + 1$. The main difference between the mobility management protocols is that each one establish a procedure based on signaling among the mobility anchors. In CMM solutions, this metric mainly depends on the time needed for establishing a new binding with the HA/LMA agent. Considering that all AR are at the same distance from the centralized anchor, its handover latency is constant. The location management function in DM3 is located at the MDA nodes and a node is always anchored by the same MDA during all its session. That means that the AR needs to signal the MDA in order to complete a handover. That MDA is closer to the MN than the centralized anchor so, the handover latency of DM3 outperforms the centralized solutions. Finally, the value of n has a high impact in HB and NB DMM proposals. This is due to the increasing number of signaling messages that need to be exchanged when the handover rate is elevated and hence they need more time to complete their handover. This figure indicates that DMM protocols are very dependent on the number of movements made by the MN, and more specifically, on the number of connections that still remain established in previous cells.

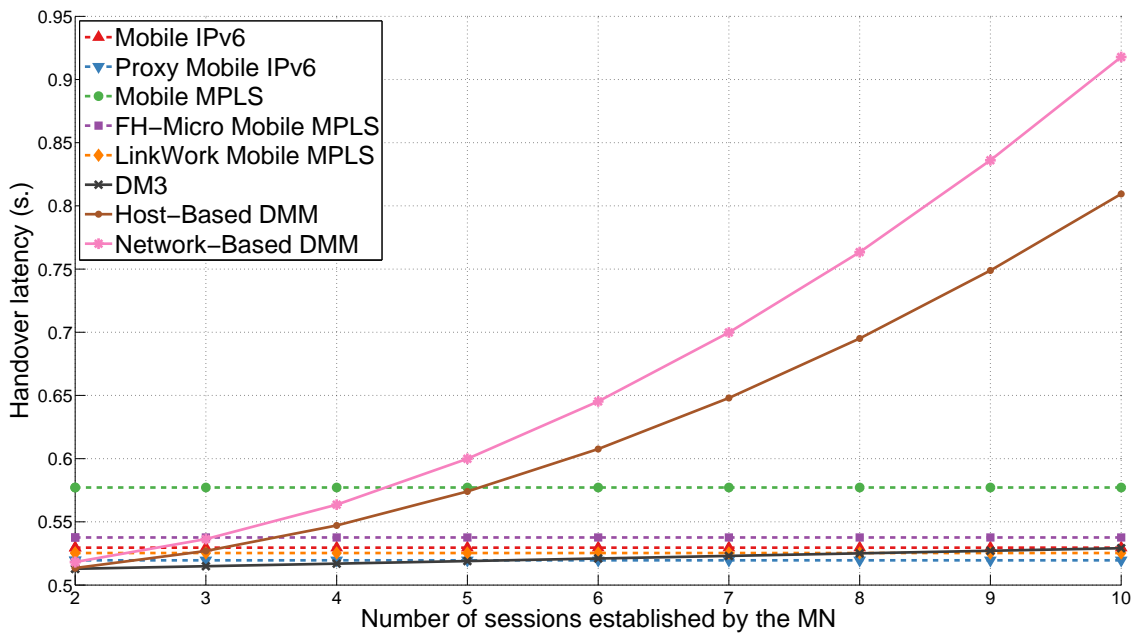


Figure 5.10. Handover latency

Figure 5.11 shows the amount of packet loss during the whole connection session for each approach. These results show the large difference between the proposals which have buffering mechanisms and those which do not. Mobile IPv6, Mobile MPLS and Host-Based DMM have the largest amount of packet loss due to the lack of a buffering mechanism during handover disruption time.

In order to minimize the packet loss, FH-Micro Mobile MPLS, LinkWork Mobile MPLS and DM3 include mechanisms to reduce losses. In all these solutions, the previous serving router of the MN is the one responsible for initiating the buffering mechanism. DM3 also achieves an ordered delivery thanks to the recovery mechanism described in previous sections. However, it is worth noting that handover latency in HB-DMM and NB-DMM are dependent on the number of connections as it was depicted in Figure 5.10. Due to the impact of the handover latency in packet loss, the value of N_h highly affects its behavior. Thus, HB and NB DMM solutions offer acceptable results when the N_h is low, but the packet losses rise dramatically when the number of tunnels established with current AMA/MAR increases.

5.4.2 Simulation environment and results

The evaluation through simulations aims to study IPv6 mobility management approaches in a more realistic environment than the one characterized by the analytical model. The

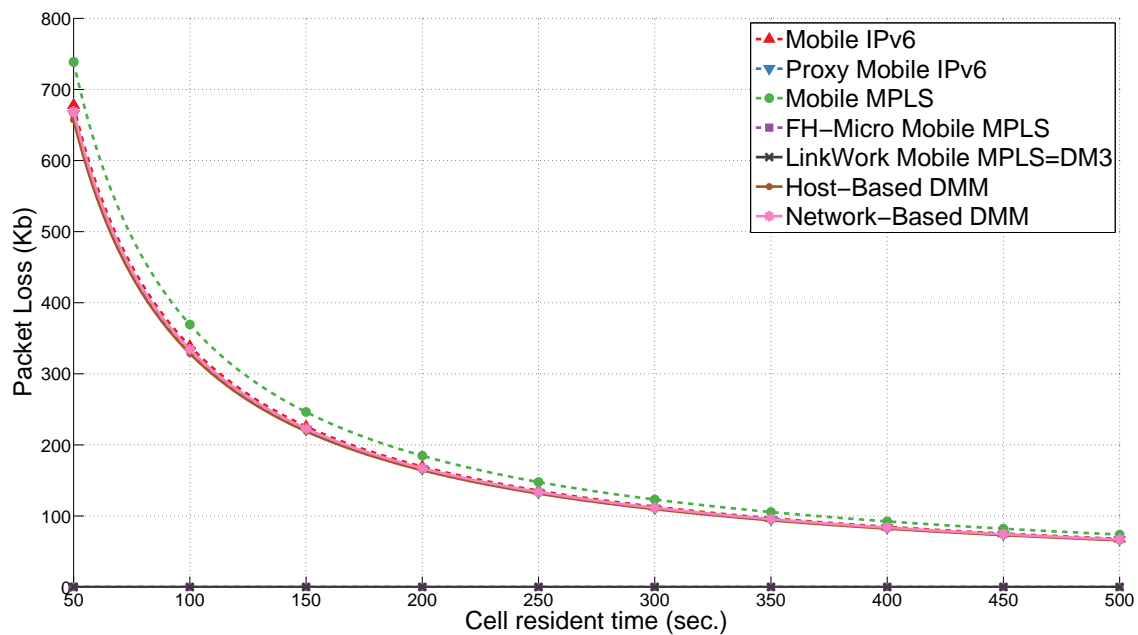


Figure 5.11. Packet loss during a session

platform selected for the evaluation through simulations was MATLAB. The mobility management simulator described in the Appendix, has been extended to model not only the DM3 proposal, but other standardized protocols in order to compare their performance. Thus, the current version of the simulator implements the centralized solutions MIPv6 and PMIPv6 and the distributed host-based (HB-DMM) and network-based (NB-DMM). All these protocols were described in Chapter 2. Next, the traffic model, as well as the mobility patterns and the network topology used in the simulations are explained.

We consider a scenario where an MN may have simultaneous active sessions with several hosts (CNs) in the Internet. We assume that sessions arrivals to a MN follows a Poisson process with mean rate $\lambda_s = 0.01$ (i.e. the inter-arrival time between sessions is exponentially distributed with this rate). We assume also that the duration of a typical session is exponentially distributed with mean session duration $\mu_s = 10$ time units.

Regarding the mobility of the MNs, we consider a Random Way Point (RWP) mobility model. RWP model is simple and straightforward stochastic model. In RWP [103], a mobile node moves on a finite continuous plane from its current position to a new location by randomly choosing its destination coordinates, its speed of movement (from [minSpeed; maxSpeed]), and the amount of time that it will pause when it reaches the destination. On reaching the destination, the node pauses for some time distributed according to a random variable (from [minPause; maxPause]) and the process repeats itself. Once the pause time expires, the node chooses a new destination, speed, and pause time.

Apart from the Random Waypoint Mobility, in order to drive the evaluation in a more realistic scenario, we also run the simulation with real-world mobility track logs obtained from users carrying GPS receivers. The sample settings where traces are obtained are two university campuses (one in Asia and one in the US), one metropolitan area (New York City), one State fair and one theme park (Disney World). The participants walk most of the time and may also occasionally travel by bus, trolley, car, or subway. These settings are selected because they are conducive to collecting GPS readings [104].

Moreover, the simulation is run for different number of MNs, ranging from 1 to 50 and the simulation time is sufficient large (45000 seconds) to avoid 'typical runs' statistical problems. The dimensions of the simulation scenario for the RWP mobility model is a rectangular area of 5×5 km² and the MNs are initially located randomly in the area. With regards to the real mobility tracks, the dimensions of the rectangular simulation area, is set to be the same as in the GPS traces. In all simulation scenarios, we used the same initial positions found in the respective real traces for the same number of users. In the evaluation, the simulations are repeated 25 times to improve the accuracy of the results with a confidence interval of 95%. Table 5.2 gives a summary of the setting values used in the simulations. The scenario defined for the evaluation is illustrated in Figure 5.12. In

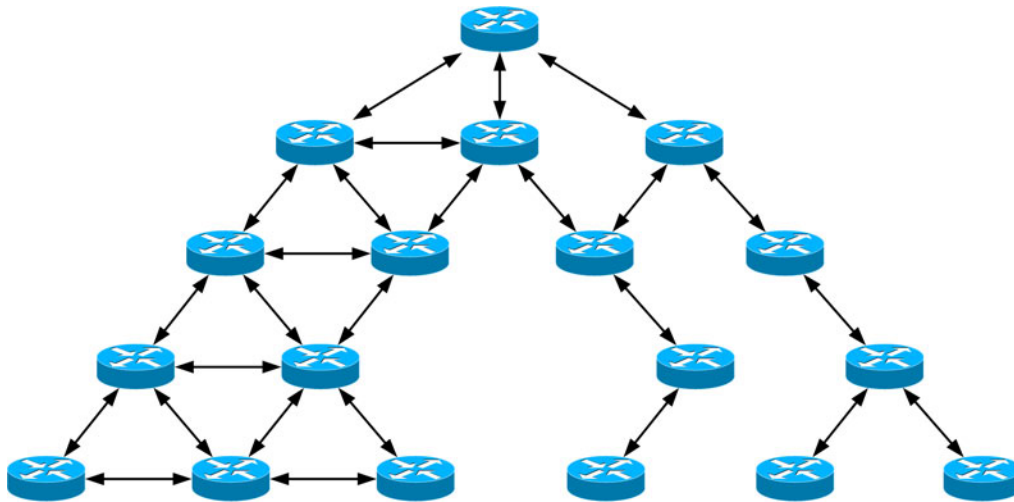


Figure 5.12. Topology used in the simulation

this topology, the leaf nodes act as access routers and the root node act as the ingress node to the domain.

Due to dependence on the topology of DMM protocols, we selected this asymmetric topology due its mixture between a well connected hierarchical network and a sparse network. This will produce more realistic results because the nodes will move around the connected and the sparse areas of the network, avoiding misleading performance by centralized or distributed protocols due to the network topology. In addition such

Table 5.2. Simulation settings

Parameter	Value
Number of MNs	5 to 50
Number of ARs	6
Simulation time	45000 sec.
Session arrival rate (λ_s)	0.01 sec.
Session duration (mean rate μ_s)	10 sec.
MN movement model	RWP and real human mobility
Simulation area	5x5 km ²
RWP speed interval	[1 10] m/s
RWP pause interval	[60 300] sec.
RWP walk interval	[300 1200] sec.
Simulation run repetitions	25

topology will allow us to shed further light on the dependency on network topology on the performance of different mobility management protocols.

Figure 5.13 and Figure 5.14 shows the comparison of accumulated signaling costs of registration update vs. number of MNs during all the simulation executions.

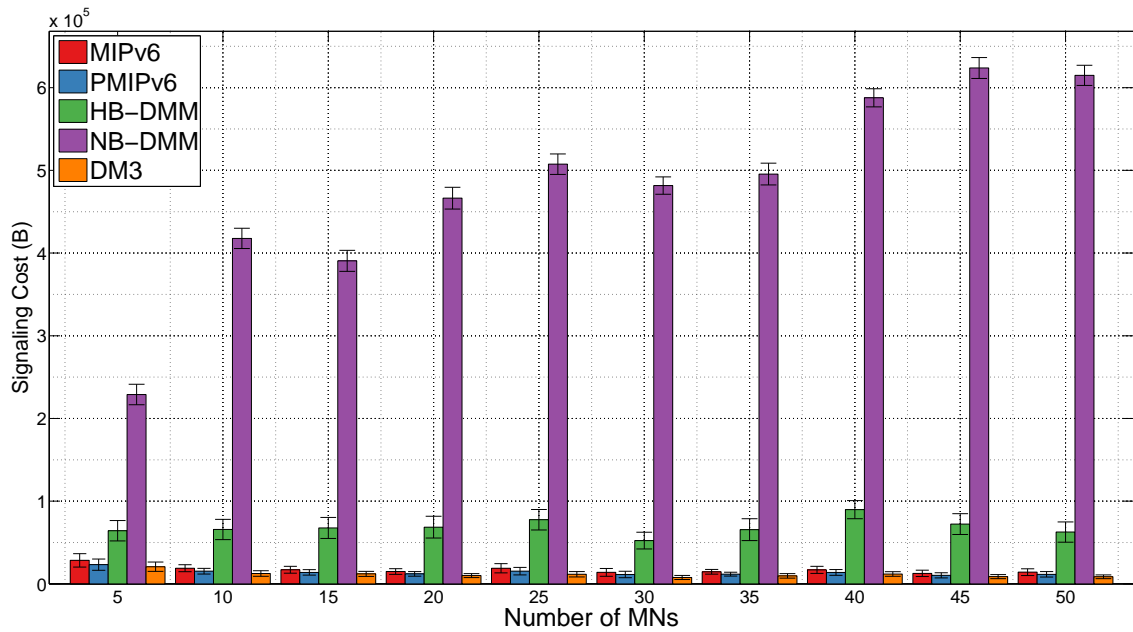


Figure 5.13. Signaling cost with a human walk mobility model

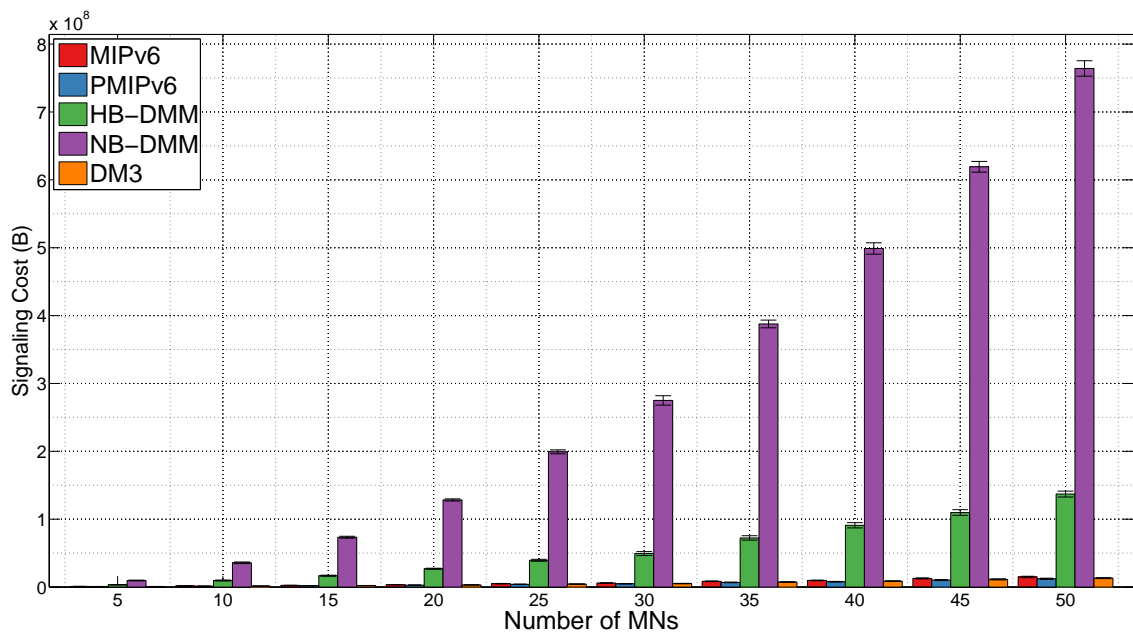


Figure 5.14. Signaling cost with a Random Waypoint mobility model

In this case, HB-DMM and NB-DMM are penalized due to the high number of control messages that they need to send from the serving AR to all remaining mobility anchors in which the MN maintains an active session. Specially, NB-DMM is the costlier protocol because of the necessity of sending an additional control message to the database each time a handover is made. This issue of high signaling cost in distributed solutions at high rates of mobility was also highlighted in Figure 5.6. Centralized solutions obtain the lowest signaling because in these protocols only two messages (Binding Update and Binding Acknowledgement) are needed to update all active sessions. DM3 is an intermediate solution since it distributes the HA mobility functions in MDA nodes, not overloading links and nodes near the Access Routers. This way, DM3 can significantly reduce the registration cost of distributed mechanisms, achieving signaling overhead values similar to centralized protocols.

In Figures 5.15 and 5.16, we present the simulated results of accumulated packet delivery cost (or routing cost) for different executions. In this case, HB-DMM and NB-DMM offer the same results because their data plane is similar, and the data are forwarded through the same path. In this figure can also be observed how centralized solutions perform a non-optimal routing and their cost is higher.

Finally, in Figure 5.17 and Figure 5.18, the tunnelling costs of the mobility protocols are compared. The large difference between CMM and DMM solutions is highlighted. While HB-DMM and NB-DMM introduce an insignificant tunnelling, centralized solutions cause a high overhead in the network due to the tunnelling process. Both HB-DMM and

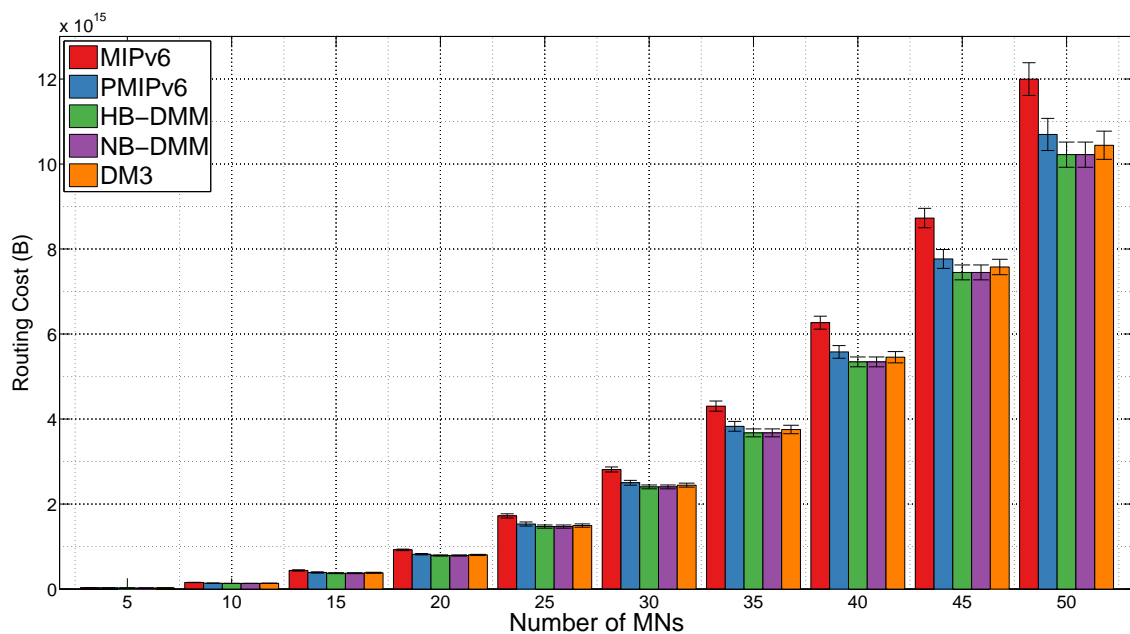


Figure 5.15. Packet delivery cost with a human walk mobility model

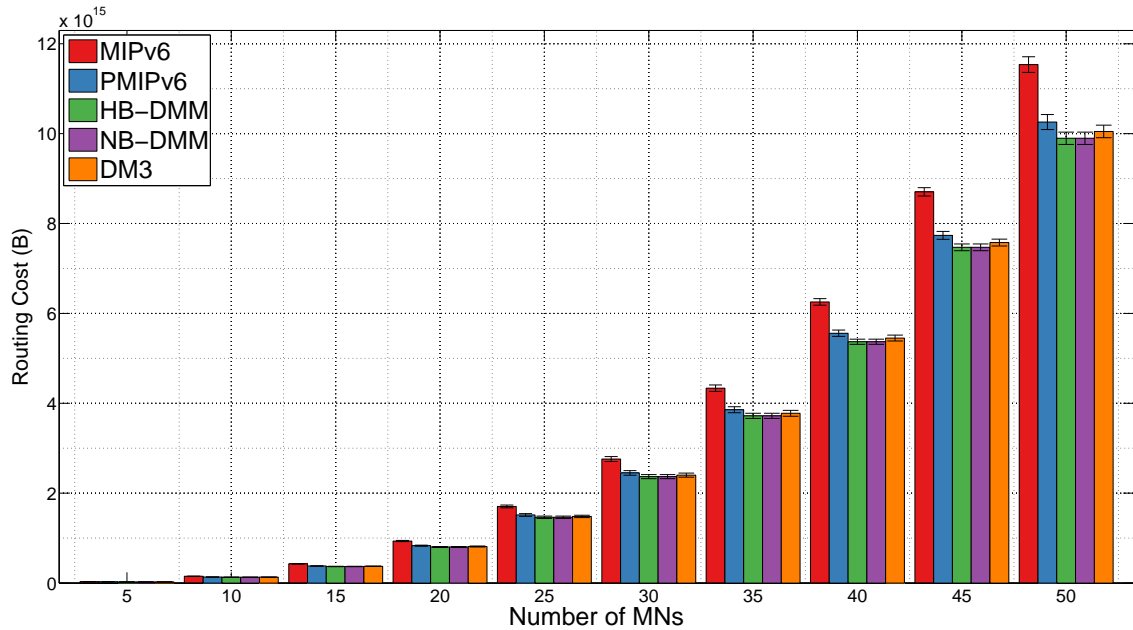


Figure 5.16. Packet delivery cost with a Random Waypoint mobility model

NB-DMM require tunnelling only between the distributed nodes, located at the access routers, whereas the tunnelling in centralized protocols is from the root of the domain. DM3 offers low values, close to the distributed solutions and improves the tunnelling cost of centralized protocols significantly.

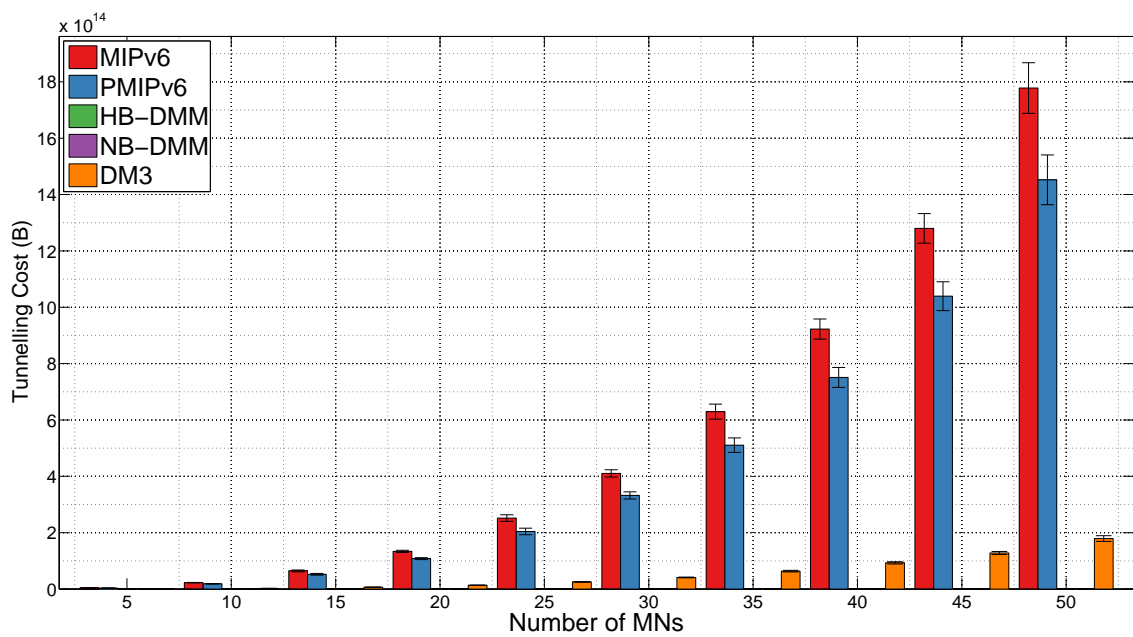


Figure 5.17. Tunnelling cost with a human walk mobility model

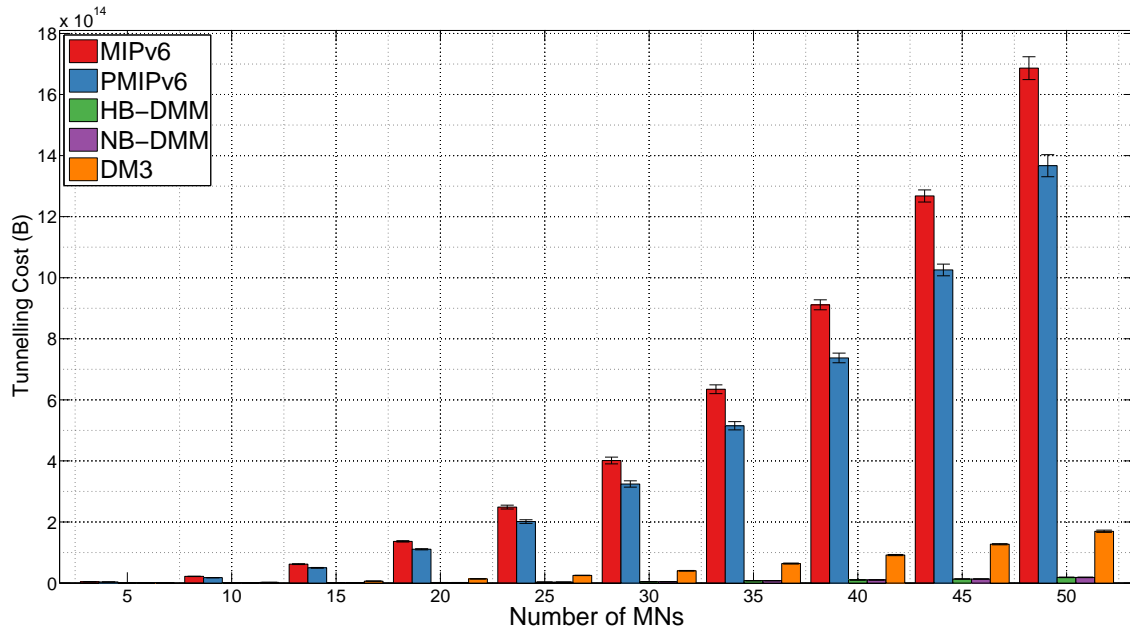


Figure 5.18. Tunneling cost with a Random Waypoint mobility model

5.4.3 Experimental results

In order to complement the analytical and simulated results presented in the previous section, we have built a testbed and developed a set of experiments to analyze the performance of the access network depending on the location of the distributed mobility anchors. This point is one of the most important decisions in the design of a mobility management solution.

In this testbed, the access network is an MPLS domain and the mobility anchors are located at different distances (hops) from the access routers that currently serve the mobile node. The routers of the testbed are Cisco K9/1921 with traffic engineering (TE) capacities. Figure 5.19 shows the layout of the testbed. Both handover latency and packet loss are evaluated in the experiments. Note that L2 information is not given to the routers.

In order to assure the reliability of the results, several experiments were performed using UDP traffic at different rates (1, 30 and 50 Mbps) and with different packet sizes (100, 600 and 1470 Bytes). All the experiments shown in this section were made with a cell resident time = 100 sec.; that is, one handover is performed each 100 seconds. The example in Figure 5.20 shows one of the results from the testbed. In this case, three executions were made with a packet size of 600 Bytes and at different throughput (1 Mbps, 30 Mbps and 50 Mbps). The handover occurs at second 30 approximately and the mobility anchor is located at 1 hop from the old serving node of the mobile node. In the next sections, the experimental results of handover latency and packet loss during the movement of a user

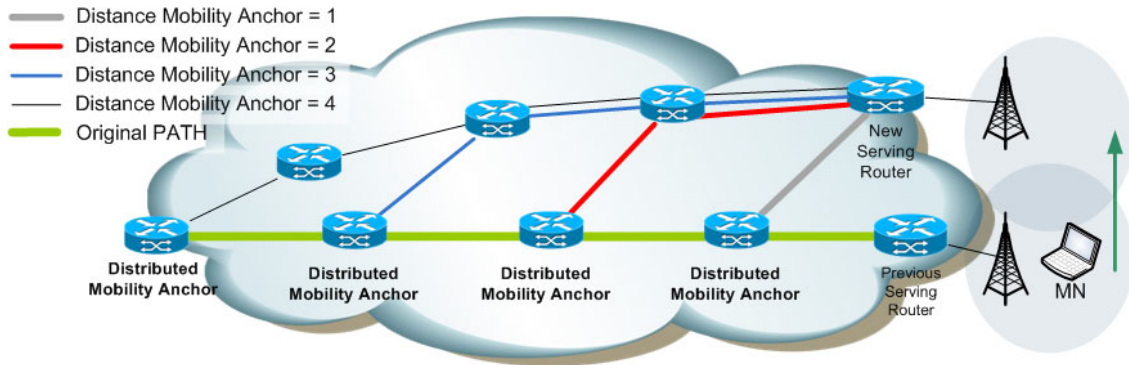


Figure 5.19. Testbed scenario

with different configurations of throughput, packet size and location of the distributed mobility anchor are given.

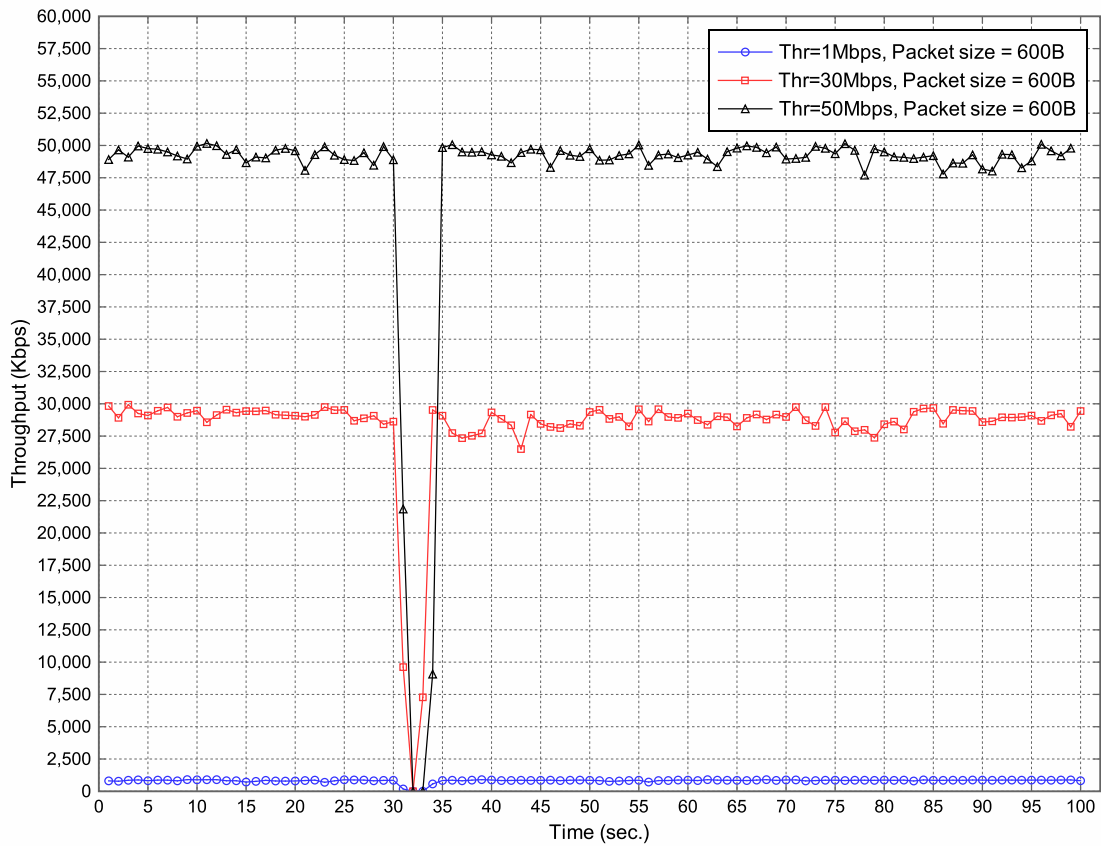


Figure 5.20. Testbed executions at different throughputs

5.4.3.1 Handover latency

Figure 5.21 shows the impact of the distance from the mobile node to the mobility anchor on the handover latency. The results outline the benefits of locating the mobility anchor closer to the user. If the mobility functions are distributed to a node placed at 1 hop from the edge of the access network instead of 4 hops, our experiments outline that the overall handover latency is reduced, on average, by 11.95 %.

With this experiment, we demonstrate that the topological location of the mobility anchor affects the handover performance. In centralized approaches, the distance between the user and the Home Agent is higher so the handover latency would be further increased. Distributed mobility management approaches can reduce this time by locating the mobility anchor closer to the MN.

5.4.3.2 Packet loss

During the handover disruption time, the mobile node cannot receive IP packets. Accordingly, the topological location of the mobility anchor affects not only the handover performance, but also affects the packet loss. Without any buffering mechanism, data packets sent from the CN to the MN will be lost while the MN changes its point of attachment to the network. In this experiment, we analyze the impact of the location of the mobility anchor on the packet loss.

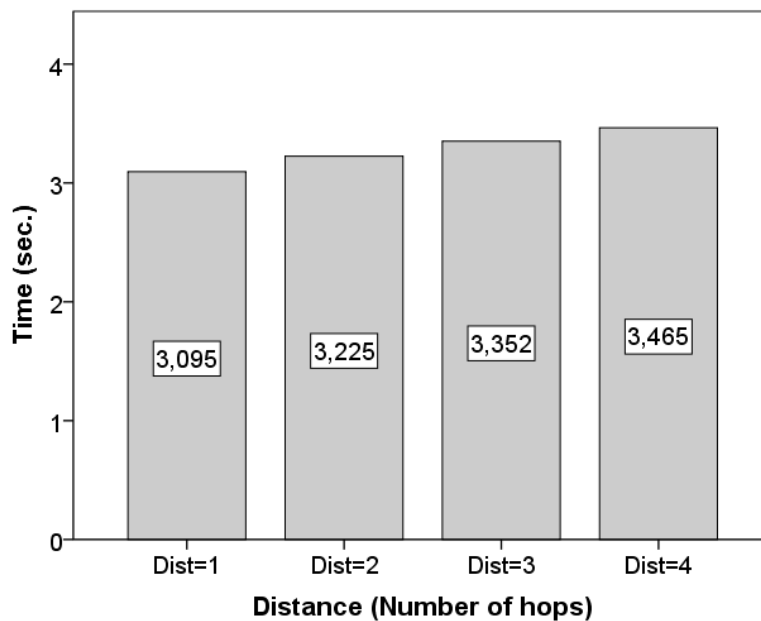


Figure 5.21. Handover Latency

As we mentioned previously, the experiments were made at different bandwidths and using different packet sizes. Moreover, one handover occurs each 100 seconds. The graphs shown in Figure 5.22 illustrates the results of the experiments. In Figure 5.22a, we can observe the packet loss where the mobility anchor is located at different hops based on the throughput and in Figure 5.22b the amount of loss is based on the packet size. A complete summary of the results is shown in Table 5.3. The results shown in the graphs are grouped by the distance to the mobility anchor. Similar to the handover latency experiment, a distributed mobility management solution can reduce the packet loss rate. Moreover, due to the nature of the mobility, where periodically packets are lost, buffering techniques can reduce this limitation. In these buffers, packets can be stored during a short period of time. Our DM3 solution proposes a buffering mechanism in order to minimize the packet loss during handover.

Throughput	Packet Size	Dist=1	Dist=2	Dist=3	Dist=4
1 Mbps	100B	13.1	13	12	8.7
	600B	6.3	3.6	3.4	2.3
	1470B	3	2.7	2.5	2.3
30 Mbps	100B	19	17	14.3	12
	600B	8.7	7.4	5.4	6
	1470B	6.6	5.4	5.3	5
50 Mbps	100B	24	22	20	18
	600B	9.2	6.7	5.7	5.1
	1470B	7	6.2	5.9	5.3

Table 5.3. Packet loss percentage during experiments

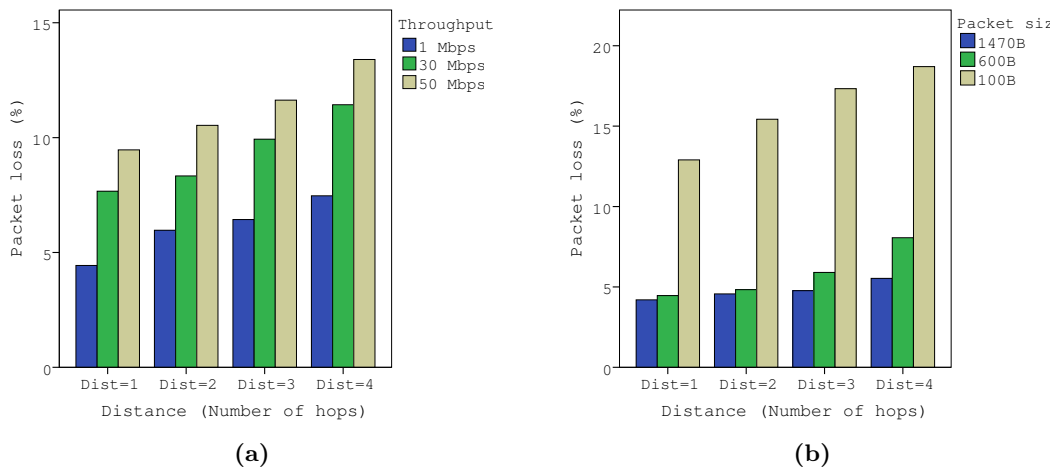


Figure 5.22. Impact of the location of the Mobility Anchor on packet loss (%) under different conditions. (a) at different throughputs (1, 30 and 50 Mbps) and (b) at different packet sizes (100, 500 and 1470 Bytes)

5.5 Concluding remarks

This chapter has provided a detailed mobility management architecture called DM3 (Distributed Mobility Management MPLS) to solve the limitations of existing IP mobility management solutions relying on centralized mobility anchors. DM3 suggest the inclusion of distributed mobility anchors in the access network, close to the user that they are serving. This way, the path between the MN and the Correspondent Node follows an LSP path composed by a set of forwarding chains. We propose a flexible access network where data plane routing is optimized with the location of the mobility anchors in order to build an efficient routing path to track the MN's movement.

To this end, we carried a wide set of numerical evaluations that allow us to shed further light and compare the DMM approaches with existing CMM protocols both in terms of control and data plane metrics.

Special attention has been paid to the location of the mobility anchor, a crucial decision in DMM architectures. We have built a testbed to analyze the impact of the location of the mobility anchor on the performance of handover latency and packet loss. Our experimental results demonstrate the benefits of distributing the mobility anchors close to the user. This is one of the reasons why mobile networks are moving towards flatter designs.

The results reveal that our DM3 mechanism solves the limitation of centralized mobility proposals and offers a small registration update cost and minimizes the packet loss rate. In general terms we can agree that DMM solutions outperform CMM protocols. However, this conclusion comes with a warning, distributed mobility management solutions also leads to several issues such as complex address and tunnel management, high signaling cost and long handover latency as the number of addresses and the number of bi-directional tunnels associated with the MN increase, for example, in case of users moving at a high speed and/or with long lasting sessions, this is, when session duration is significant larger than cell residence time. This effect has been demonstrated with the results shown in this Chapter.

As a result, DMM may not be a suitable scheme in some scenarios. Thus, in order to achieve a balanced behavior, future mobile network architectures might benefit from hybrid CMM-DMM solutions, in which operators would be able to handle the traffic in a distributed or centralized way depending on topological network characteristics. The next chapter deals with this issue. Hybrid solutions are discussed and two mechanisms are proposed to handle mobility from this new perspective.

6

HYBRID DMM

From the previous chapters, we can understand current efforts in the evolution of IP mobility management protocols towards a more distributed operation to tackle shortcomings that stem from fully centralized approaches. However, and as will be detailed hereafter, there are instances where distributed mobility management result in lower performance which might affect real time and several over the top (OTT) applications. In this chapter, we propose a Hybrid DMM solution that overcomes, in terms of mobility costs, both centralized and distributed mobility management protocols. Our results indicate the significant benefits in terms of mobility costs that hybrid CMM-DMM solutions might bring.

Contents

6.1	Introduction	119
6.2	Background and motivation for hybrid solutions	120
6.2.1	CMM and DMM limitations	120
6.2.2	A hybrid mobility management solution	120
6.3	Description of the hybrid mobility management scheme	121
6.3.1	Initial mobility anchoring	121
6.3.2	Registration and data delivery mechanisms	122
6.4	Decision criteria	125
6.4.1	Node-assignment algorithm	125
6.4.2	Link-assignment algorithm	127
6.5	Analytical evaluation	131
6.5.1	Signaling cost	131
6.5.2	Data packet delivery cost	132
6.5.3	Tunnelling Cost	132
6.6	Simulated results	133
6.6.1	Impact of the amount of MNs	134
6.6.2	Impact of Session Duration and Topology Scenarios	139
6.6.3	Impact of the MN's Speed and Topology Scenarios	142
6.7	Concluding remarks	145

6.1 Introduction

In the previous chapters we have examined both centralized and distributed IPv6 mobility management solutions as the mechanism responsible for maintaining the ongoing communications while the user roams among distinct networks (changing points of attachment) and also to provide reachability to the mobile users in such heterogeneous environment in terms of access. Moreover, future mobile networks need to cope with the recent trends in mobile Internet, as well as with the current increasing mobile data traffic demand [3]. This new emerging communication paradigm has brought challenges in flexibility, efficiency, and scalability to the current mobile networks [105]. Thus, operators can benefit from solutions that offer a greater flexibility and a certain freedom to deploy different configurations of the mobile network.

As we have seen in previous chapters, centralized mobility management protocols are specially affected by these new environments, and distributed solutions are currently being developed to efficiently handle the current mobile traffic explosion [54]. However, despite the fact that a number of mobility management approaches are being designed towards a more distributed operation aiming to mitigate problems related to centralized operation, there are instances where DMM incurs higher costs and the performance of the network might be compromised. In fact, in some of these cases, CMM seems to solve the mobility problem more efficiently and therefore should be preferred.

As is stated in [60], future mobile network architectures might potentially exhibit hybrid centralized-distributed behavior in which the mobility management of some traffic will be kept centralized, while mobility support for other applications can be distributed. This paradigm change, from centralized to distributed mobility support is clearly not easily accomplished by operators, because current deployments are mainly centralized. Hence, instead of a drastic change, a more gradual deployment should be preferred by the telecommunications industry.

Therefore in this chapter, a hybrid CMM-DMM mobility management architecture (Hybrid DMM) is proposed, that adapts to the specific topological characteristics of the infrastructure network of mobile operators, in which the data and signaling traffic is forwarded following a centralized or distributed scheme depending on, hereafter detailed, decision criteria. As a network-based approach, it provides the mobility support for all legacy devices. The key benefit of the proposed hybrid solution is that it manages to reduce significantly both signaling and routing cost.

6.2 Background and motivation for hybrid solutions

6.2.1 CMM and DMM limitations

As has been described, CMM requires a single handed mobility anchor, e.g., HA at MIPv6 and LMA at PMIPv6 to allow for session continuity when MNs are moving across different networks. Nowadays, most of the deployed architectures have a small number of centralized anchors managing the traffic of thousands of mobile users. These centralized approaches have certain limitations for handling a large volume of mobile data traffic such as non-optimal routing, scalability problems and reliability/robustness. These limitations have been identified in [51] and summarized below,

- **Non-optimal routing:** In CMM, all traffic is routed through a central mobility anchor, resulting in a longer path and thereby, increasing the end-to-end delay.
- **Scalability problems:** With the increase of mobile nodes and the traffic explosion of mobile data, the centralized anchor needs to be able to deal with all the user traffic simultaneously. Here arises scalability problems due to the processing and routing resources that the mobility anchor needs to manage that huge amount of traffic.
- **Reliability:** Centralized anchors are more vulnerable to single point failures and attacks than a distributed system.

The evolution from CMM to DMM approaches has shown clear signs of achieving better utilization of resources in the network, outperforming the traditional protocols and optimize mobility management performance [57], [58], [59], [50]. Despite these facts, there are some scenarios in which the DMM paradigm also incurs high delivery costs and possible significant signaling overheads [60].

In the DMM operation, each distributed MAR/AMA that manages a MN keeps a bidirectional tunnel between itself and all MAR/AMA where a session of the MN was originated and is still alive. This situation can be deemed as inefficient in certain circumstances. In these cases, a CMM based behavior might be preferred.

6.2.2 A hybrid mobility management solution

To this end, and in order to maintain the advantages of DMM while mitigating its drawback in comparison with a centralized management approach, a hybrid CMM-DMM mobility management solution might provide the benefits of each one of them resulting in better

overall network performance. The mobility management will be managed by centralized or distributed solutions, depending on different parameters, such as, user profile policies, network topology characteristics or the current state of the network in terms of congestion, traffic mix, etc. Hence, under the proposed scheme, depending on the network topology selected flows will be distributed among the different mobility anchors, while mobility anchoring for other flows will be kept centralized.

In the next section the proposed network-based hybrid mobility management scheme is detailed and a set of decision criteria that an operator could use to determine how the traffic will be managed and anchored is discussed.

6.3 Description of the hybrid mobility management scheme

The proposed hybrid solution can be deemed as an amalgamation of previous schemes, where mobility management in an IPv6 network can be handled by a centralized protocol such as Proxy MIPv6 or by a distributed one such as NB-DMM.

One key motivation of our hybrid solution is to take advantage of some aspects from PMIPv6 and others from NB-DMM, minimizing the limitations of centralized and distributed mechanisms, developing a solution that allows the network to decide, depending on some network performance based criteria, the appropriate approach to be applied to a session, i.e., when to manage the traffic in a distributed way or when to keep it centralized. It should be mentioned that although the proposed hybrid mechanism relies on network-based mobility management solutions, it could be applied to host-based environments.

Hence, our Hybrid DMM architecture is based on combining PMIPv6 and NB-DMM, and allows intelligently selecting a suitable mobility scheme in an appropriate manner. In the current state of the proposal, two algorithms perform this selection depending on the available information. The first algorithm uses only the topological information of the core network whereas the second algorithm makes use of the topological information plus the the location of the Base Stations.

6.3.1 Initial mobility anchoring

At an initial state of the network operation, each AR is selected to manage the traffic with one of the two possible protocols (PMIPv6 or NB-DMM) although they can offer support in both mechanisms, if needed. This means that the operators, based on a decision criteria,

select each AR to operate in a centralized or distributed way. In the next sections we discuss in detail about the decision criteria used in this approach.

When a mobile node attaches to a base station, the AR in charge of that BS assigned for the new session provides the MN an IP address ensuring IP reachability and/or IP session continuity, as well as the respective routing/forwarding support from the establishment of the session.

As we can see in Figure 6.1, when MN1 joins the access subnet, the responsible AR (AR1) retrieves the IP address for the MN following the registration procedure of PMIPv6 so, AR1 will act as MAG and the LMA will be the anchor point for the MN. When MN2, located in a different subnet, attaches to a BS, it initiates the NB-DMM registration procedure. In this case, AR5 will act as MAR so, AR5 will be the origin anchor router for the MN2. While MNs remains attached to the same AR, all connections will be managed with the initial mobility management protocol. This initial mobility anchoring is transparent to the user.

6.3.2 Registration and data delivery mechanisms

This section describes the hybrid solution operation when a mobile node moves among several points of attachment associated to different mobility management protocols, and the forwarding data mechanisms.

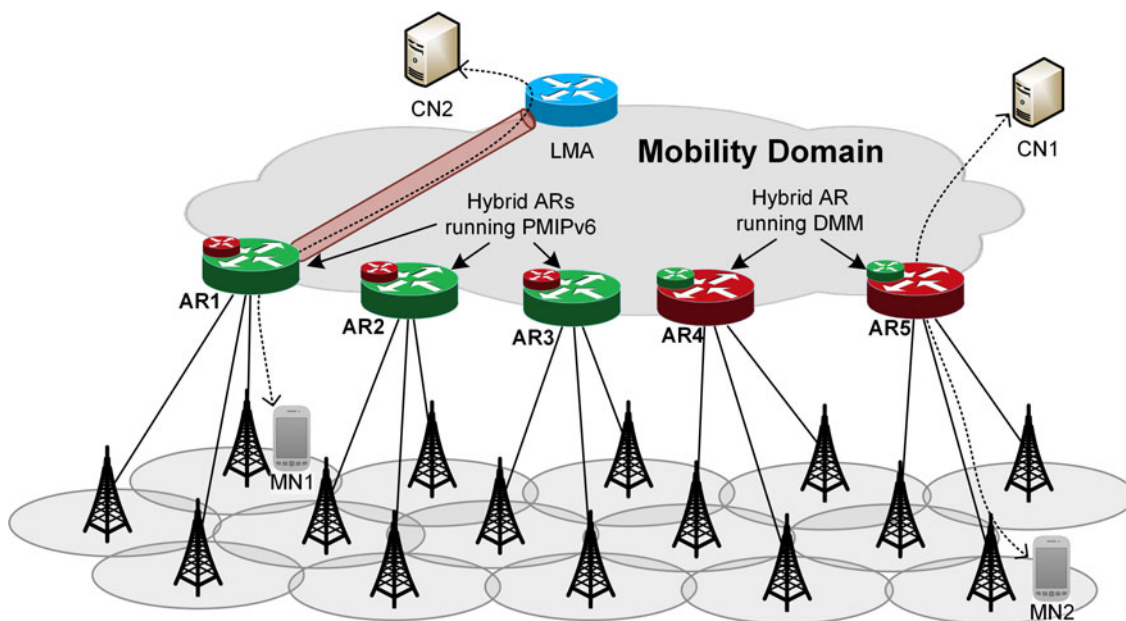


Figure 6.1. Overview of hybrid approach

Once the MNs are initially anchored to a centralized or distributed agent (LMA or MAR respectively), they can deliver data with their correspondent nodes following the PMIPv6 or NB-DMM operation. No changes are needed to the initial attachment or the initial forwarding mechanism.

Figure 6.2 shows the MN's movement from the access network of AR1 to the access network of AR5 and how new sessions are managed. In Figure 6.2(a), MN1 is initially registered to AR1 and a session with CN1 is active. Upon an IP handover from one MAG (AR1) to another (AR2) as is shown in Figure 6.2(b), the registration procedures starts. The binding is updated at the LMA as follows. The previous MAG (AR1) sends a proxy binding update to the LMA deregistering the MN, and the new MAG (AR2) sends a PBU to the LMA registering it. The LMA replies to each by a proxy binding acknowledgement. With regard to the data plane of PMIPv6, when a new session arrives (session 2 in this case) from CN2, the centralized sessions of MN1 are anchored to the LMA and all data traffic of the MN passes through the LMA. A tunnel is established for this purpose between the LMA and the serving MAG (AR2).

When the MN1, that is managed by centralized anchors moves to a DMM AR (AR3) as it is shown in Figure 6.2(c), if any centralized connection remains active, the AR registers them with the PMIPv6 procedure. This is possible because all ARs in the hybrid approach are both PMIPv6 and DMM capable. If during its stay in AR3 the MN establish a new connection with CN3 (session 3), it will be anchored following the DMM procedure and allowing centralized and distributed connections at the same time. For the DMM signaling, the protocol relies on a database (DB) that stores ongoing mobility sessions for the MNs. Specifically, it stores the home network prefixes currently allocated to the MN and their respective anchoring points.

Finally, in Figure 6.2(d), the MN moves to a new DMM AR (AR4). In this case, the new serving AR will act as a MAR and performs the handover management following the regular operation of NB-DMM as follows. The new MAR (AR4) retrieves the IP address of the anchoring MAR (AR3) from the database and sends a PBU to each anchoring MAR, in this specific case it would be AR3. Then, AR3 replies by a PBA and a bidirectional tunnel is created among AR4 and each anchoring MAR, in this case between AR4 and AR3.

If a new session arrives (session4), the new traffic is routed directly between the MN and the correspondent node (CR4). The old traffic (session 3) is routed through the anchoring MAR (AR3) using the previous established tunnel.

This packet delivery process aims to improve the overall packet delivery cost or, at least, in the worst cases, not introduce an additional cost in the forwarding operation than distributed solutions.

Having as a starting point the model presented in Chapter 3, and in order to demonstrate that a hybrid CMM-DMM outperforms previous proposed DMM solutions, we obtain the following,

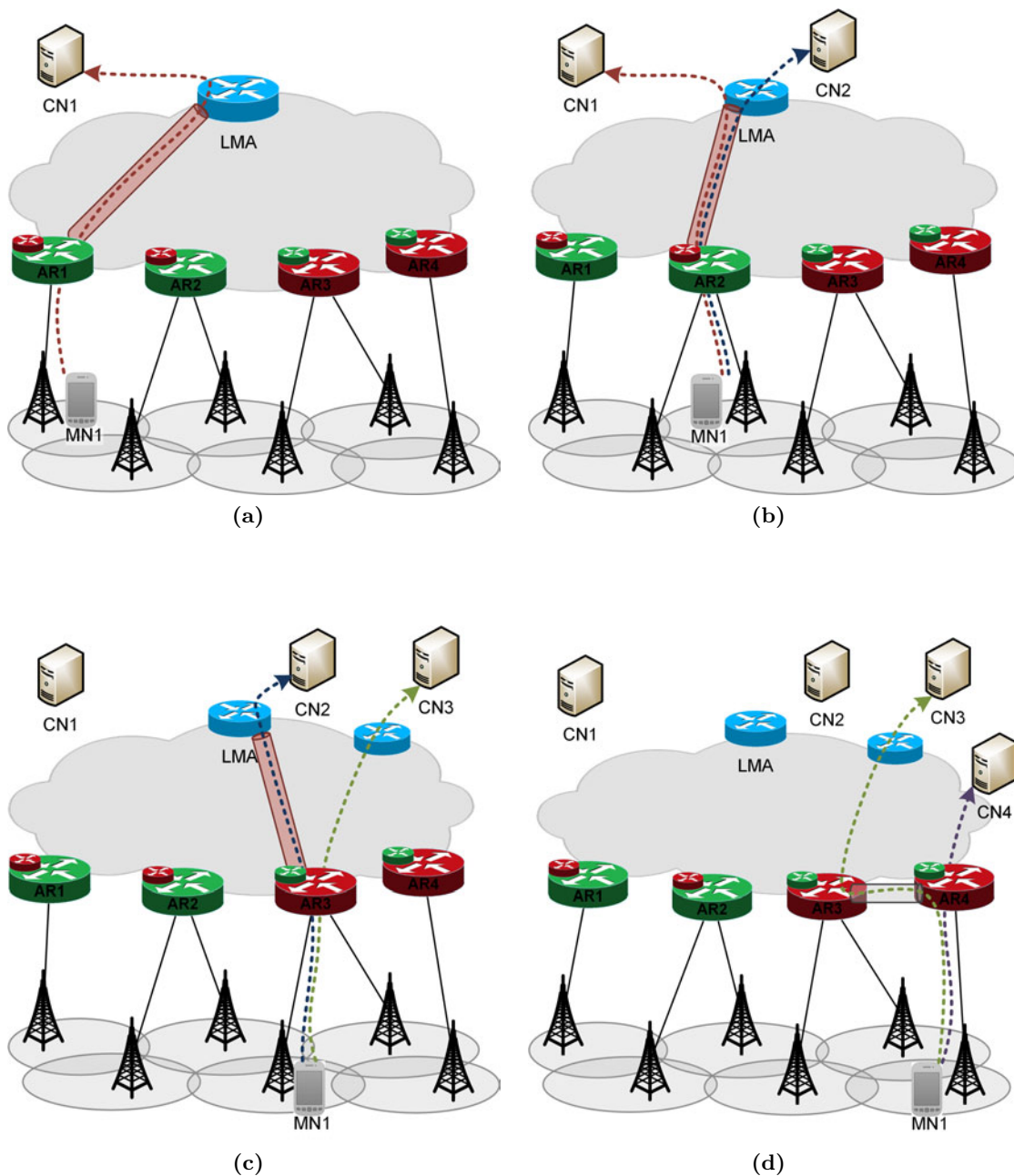


Figure 6.2. Data path during hybrid solution operation. (a) Initial State. MN1 is attached to AR1. Session 1 is created. (b) MN1 moves to AR2 and session 2 is created. (c) MN1 moves to AR3. Session 1 finishes and session 3 arrives. (d) MN1 moves to AR4. Session 2 finishes and session 4 arrives.

Lemma 1. Let $C_i^{\mathcal{H}}$ be the total routing cost of the network using the hybrid DMM mechanism, and let $C_i^{\mathcal{D}}$ be the total routing cost of the network using DMM mechanisms, then $\sum C_i^{\mathcal{H}} \leq \sum C_i^{\mathcal{D}}$.

Proof. Let \mathcal{K} denote the set of access routers in the network and \mathcal{M} be the set of routers that serve as the centralized mobility anchors (LMA) for the mobile nodes. Let $g \in \mathcal{M}, k_i \text{ and } k_{i+1} \in \mathcal{K}, k_i$ and k_{i+1} being two adjacent routers. Since $C_{g,i} \leq C_{g,i+1} + C_{i,i+1}$ because all paths are SP (Shortest Path), then Lemma 1 holds, as required. \square

6.4 Decision criteria

As we have eluded in previous sections, the ARs in the hybrid approach are both PMIPv6 and NB-DMM capable. However, they are initially selected to operate in a centralized or distributed way depending on some criteria that the operator can select. In this section we introduce two decision criteria algorithms, that carry out this protocol selection. These algorithms make decisions based on some network information. As explained previously, DMM provides some clear benefits with respect to CMM, but the performance of DMM is very topology dependent [57] since the flows have to be tunnelled from the old AR to the new AR. This means that, in some cases, depending on the routing path between the two ARs that are involved in the handover phase CMM should be preferred instead of DMM. For this reason, the main criteria that is used in our algorithms, is the topology information.

As it is shown in Figure 6.3, the first algorithm (node-assignment) uses only information about the network topology, whereas the second one (link-assignment) uses both topology and location information of the BSs. In the sequel we detail two algorithms based on these criteria aiming to optimize network performance and provide specific benefits to network operators by allowing various degree of freedom on deciding mobility management procedures.

6.4.1 Node-assignment algorithm

In Algorithm 1, which is called hereafter as node-assignment, the decision is made according to the actual operation of the mobility management protocols. The data plane of distributed protocols extend the data path during the movement of the MN through the ARs at the edge of the network, whereas centralized protocols anchor all sessions of an MN to the same entity, the mobility anchor. Hence, the algorithm in essence declares for each

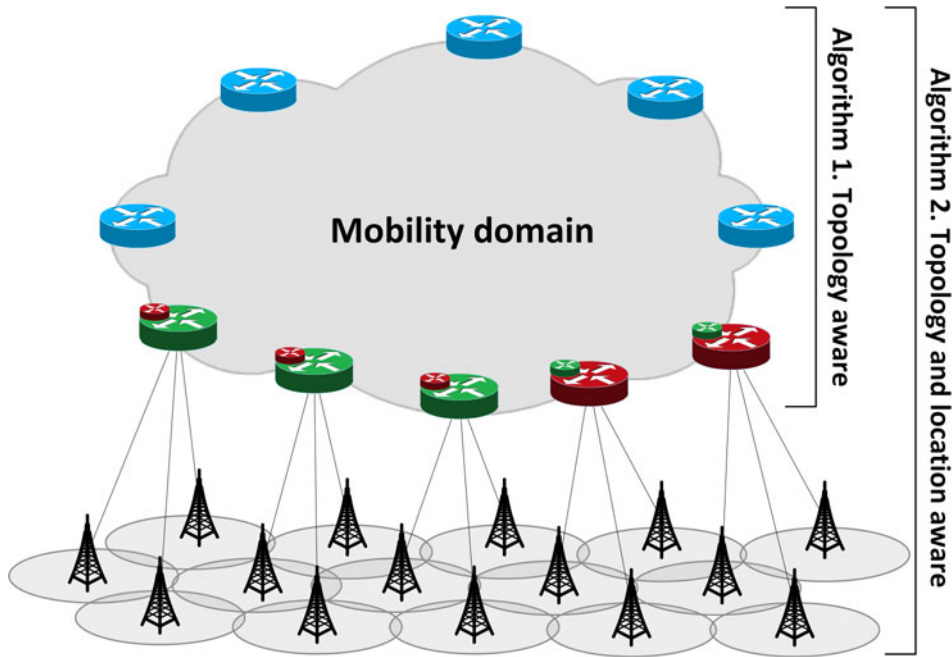


Figure 6.3. Level of information

candidate AR if will be acting as a mobility anchor or not, which will as a result define the degree of mobility function decentralization in the network.

Figure 6.4 shows an example of the decision criteria process that, as a result, select the optimal mobility anchoring that the AR (Node 2) should use according to the topological information.

In this example, Node 2 has three neighbour ARs, so we consider the routing cost to connect with each one of them. Let $c_{i,j}$ be the cost to reach node j from node i , for the Node 2 we consider $c_{2,1}$, $c_{2,3}$ and $c_{2,4}$. With these three costs, we obtain α_2 , which

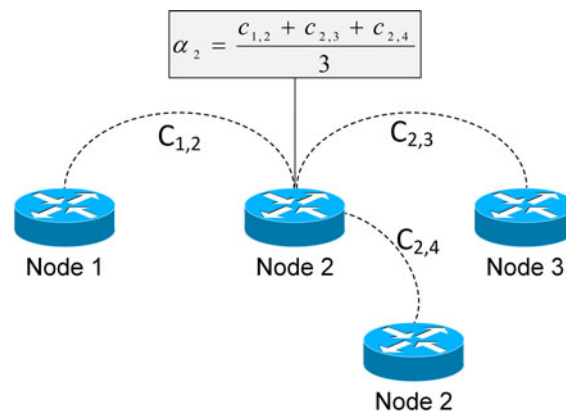


Figure 6.4. Example of mobility anchoring selection by Node 2 using Algorithm 1

Algorithm 1: Node-assignment decision criteria

Input: $AdjMatrix \leftarrow Adjacencymatrix$; $K \leftarrow SetofARs$;**Output:** Array $protocolAR$ with the protocol to use in each AR;

```

1 initialization;
2 foreach  $AR\ i \in K$  do
3   | foreach  $AR\ j \in V_i, j \neq i$  do
4   |   |  $routingCost(i, j) \leftarrow dijkstra(i, j, AdjMatrix)$ ;
5   |   end
6   end
7 foreach  $AR\ i \in K$  do
8   |  $\alpha(i) \leftarrow \frac{\sum_{j=1}^{V_i} routingCost(i, j)}{V_i}$ ;
9   end
10  $\bar{\alpha} \leftarrow mean(routingCost)$ ;
11 foreach  $AR\ i \in K$  do
12   | if  $\alpha(i) > \bar{\alpha} + threshold$  then
13   |   |  $protocolAR(i) \leftarrow PMIPv6$ ;
14   |   else
15   |   |  $protocolAR(i) \leftarrow NB - DMM$ ;
16   |   end
17 end
18 return  $protocolAR$ ;
```

represents the mean routing cost for Node2 to reach any of its AR neighbours. The algorithm checks if this aggregate cost is above or below the average value ($\bar{\alpha}$) of all the α_i . There can be a number of different policies about the threshold cost value that can be used in order to take into account current network conditions and/or provide a weight for the different mobility anchoring options.

If the value α_i is larger than ($\bar{\alpha}$), it means that the network is sparse in that area and it would be preferable to use a centralized protocol. On the other hand, a lower value of α_i means that the AR is well connected to its neighbours and a distributed approach is a more efficient option.

6.4.2 Link-assignment algorithm

The previous algorithm establishes a simple mechanism to select the mobility management protocol to be used at each AR, using topological information from the network. As we have previously detailed, the way of selecting a centralized or distributed protocol for

each AR is based on the routing cost to its neighbours given by a link state algorithm. Intuitively speaking, this means that the base stations located in a well connected area of the network will be assigned to operate with the DMM protocol, whereas base stations in topological areas of the network which are not well connected will be assigned to operate in a centralized way.

However, using only the topological information may cause some sessions to not be managed properly, especially when the value of α_i does not represent correctly routing cost $c_{i,j}$ values, i.e, one of the values $c_{i,j}$ is significantly higher and the rest of the cost $c_{i,k}$ costs are low (in other words there is high deviation from the calculated mean cost that is used in the node-assignment algorithm). In these cases, it might occur that the movement of the MNs between the ARs i and j could be enhanced by differentiating on the use of the appropriate mobility anchoring selection. For this reason, we propose a second algorithm that, instead of deciding the protocol that will be used per node (at each AR), provides a decision on selecting mobility anchoring per link. Using this link-assignment algorithm we can obtain a more efficient decision, avoiding the problems that the first algorithm could entail, at the cost of using a more detailed network topology information.

The main idea behind this second algorithm (see Algorithm 2) is to make the decision closer to the mobile node, considering the area where the MN is moving. Thus, the decision considers the link cost between the current AR where the MN is attached to, and the AR at which the MN is moving to.

In order to define formally an algorithm that takes into account the above mentioned information, we make use of an overlay network, where the nodes are the ARs and the links are the connection between them following the next model. Starting from a network architecture where each AR manages several BSs, we define an Overlay AR Graph (see Figure 6.5) obtained with the following process. Let be $G = (V, E)$ the graph that defines the network as presented in Chapter 3, $K \in V$ the set of access routers in the network, let G' be the overlay graph, $G' = (W, X)$, where W is the set of ARs in this overlay network and X is the set of links, $X_{i,j}$ denote a link $(0, c_{i,j})$ integer variable that is set to $c_{i,j}$ if there is any BS $i \in K$ from which a MN can move to any other BS $j \in K$, otherwise it is equal to 0. In this case, $c_{i,j}$ represents the routing cost from i to j obtained with Dijkstra algorithm over the graph G .

Finally, to identify the mobility anchoring scheme that will be assigned to each link in the overlay graph, we calculate the average value of all routing costs, defined as \bar{c} . If the value of this cost $c_{i,j}$ is greater than \bar{c} , the AR will use PMIPv6 as the mobility management protocol to handle the connections associated to the link between AR i and AR j , and NB-DMM otherwise.

Algorithm 2: Link-assignment decision criteria

Input: $AdjMatrix \leftarrow Adjacencymatrix$;
 $K \leftarrow SetofARsinG$;
Output: Matrix $mLinks$ with the protocol to use at each link;

- 1 initialization;
- 2 **foreach** $AR\ i \in K$ **do**
- 3 $V_i \leftarrow ARNeighbours(i)$;
- 4 **foreach** $AR\ j \in V_i, j \neq i$ **do**
- 5 $routingCost(i, j) \leftarrow dijkskra(i, j, AdjMatrix)$;
- 6 **end**
- 7 **end**
- 8 $G' = (W, X) \leftarrow buildOverlayARGraph(K, routingCost)$;
- 9 $\bar{c} \leftarrow mean(routingCost)$;
- 10 **foreach** $link\ l_{i,j}$ **in** X **do**
- 11 **if** $c_{i,j} > \bar{c} + threshold$ **then**
- 12 $mLinks(i, j) \leftarrow PMIPv6$;
- 13 **else**
- 14 $mLinks(i, j) \leftarrow NB - DMM$;
- 15 **end**
- 16 **end**
- 17 **return** $mLinks$;

This operation is detailed in the Algorithm 2. The code shown is mainly divided in two main loops. The first one is used to determine the set of neighbors of each access router i . For each of these neighbors, we determine its routing cost through the Dijkstra algorithm. Once all costs have been calculated and the matrix $routingCost$ is filled with these values, the overlay graph $G' = (W, X)$ can be built and the value of \bar{c} calculated. Finally, for each link belonging to the overlay network, the matrix $mLinks$ is created and the mobility management protocols selected.

With the overlay AR graph G' , we are able to identify the related location information needed to make better decisions in the proposed hybrid mobility management mechanism. The underlay graph G corresponds to the physical network topology, and the overlay level is conceptual, and contains the connections and associated costs between neighboring ARs.

Moreover, the degree of each node in the overlay network G' indicates the number of connections with adjacent ARs. Each one of these links relates to a mobility anchoring assignment that will be used to determine how the AR that manages that link handle the mobility of the MNs moving around the BSs.

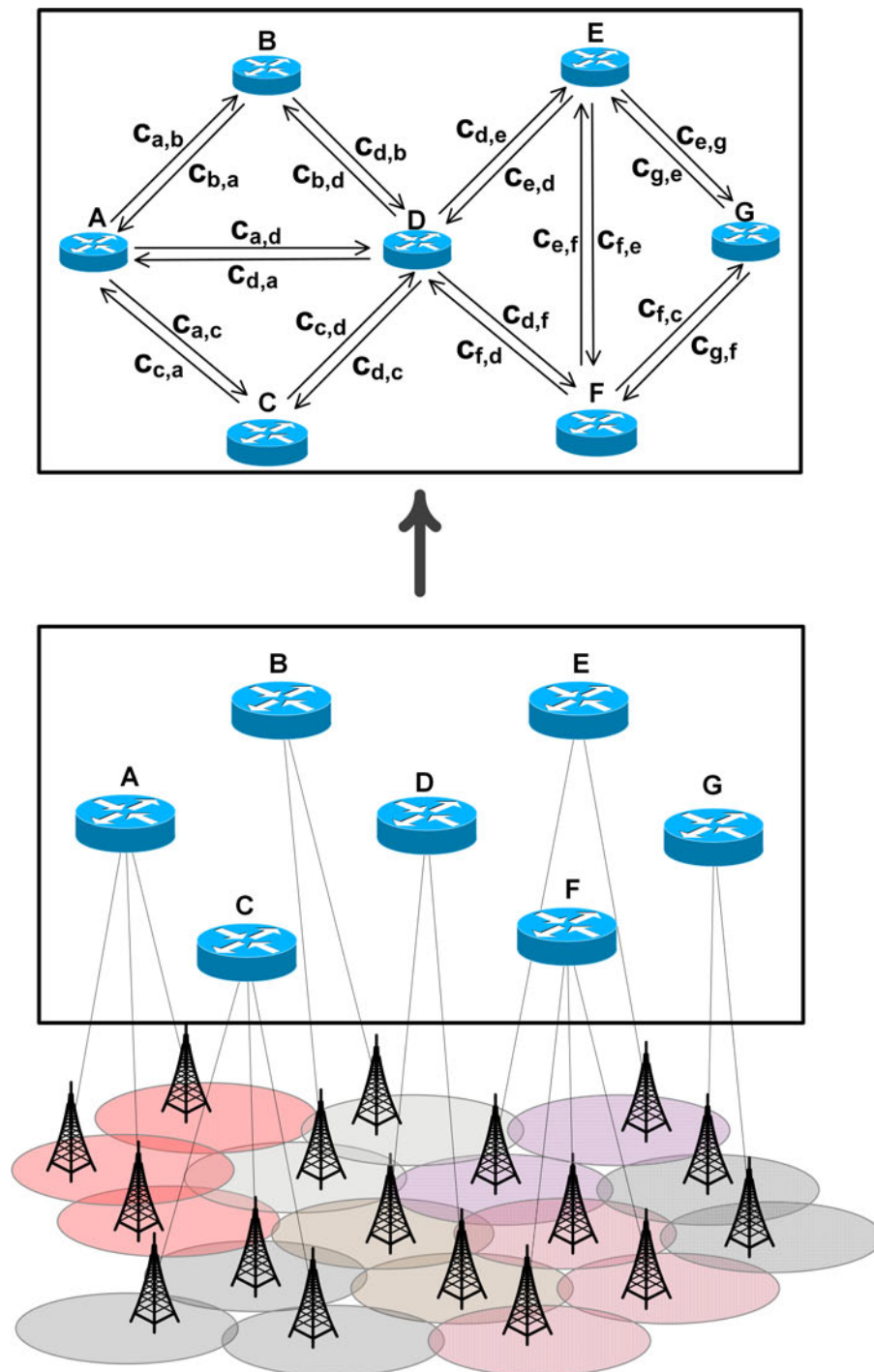


Figure 6.5. Physical AR and BS structure and Overlay AR graph

6.5 Analytical evaluation

When a MN moves, new signalling between the network entities is introduced and a change in the routing path is needed in order to deliver the data packets to the new point of attachment of the MN. Moreover, mobility protocols uses tunnelling mechanisms and the associated overhead by such encapsulation needs to be taken into account for the different schemes. As we introduced in Chapter 3, different parameters are involved in mobility aspects, such as the cost functions of signaling (C_s), packet delivery cost (C_{PDC}) and tunnelling overhead (C_t). We first define the quantitative analysis of the hybrid solution and in the sequel we carry out numerical investigations to provide a more in-depth analysis.

The rest of the protocols that will be used in the comparative analysis are the following. Both host-based and network-based distributed mobility management (HB-DMM, NB-DMM) and the centralized mobility management solutions; namely MIPv6, PMIPv6 and DM3. The analysis of these protocols was conducted in Chapter 5.

6.5.1 Signaling cost

When a MN moves among different networks, its location and routes must be updated. Such operations require dedicated signaling that, in the scope of this document has been called C_s . This value is defined as the size of the control messages multiplied by the number of IP hops those messages cross in the network.

Regarding the proposed hybrid proposal, the average probability that the sessions are managed by centralized or distributed protocols are denoted by \mathcal{P}_C and \mathcal{P}_D respectively. Moreover $C_u(Hybrid)$ consists of the sum of the cost of the sessions that are managed in a distributed manner and those that keep centralized.

$$\mathcal{P}_C + \mathcal{P}_D = 1$$

Hence,

$$C_s(HybridDMM) = \mathcal{P}_C(C_u(PMIPv6)) + \mathcal{P}_D(C_u(NB - DMM)) \quad (6.1)$$

6.5.2 Data packet delivery cost

One of the key challenges that mobile networks have to deal with is the increasing amount of data traffic mainly generated by high quality multimedia and video/audio streaming sessions. The introduction of distributed solutions aims to reduce this cost by avoiding all user data to traverse a centralized anchor. However, DMM might not be adequate in certain situations due to their topology dependent performance. As we described in Section 6.2, distributed mechanisms create multiple tunnels between the anchors and in cases where a high mobility rate exists, the system performance might be critically compromised by the frequent registrations and maintenance of multiple tunnels.

With the analysis developed in this section, together with the numerical results obtained from the simulations, we evaluate and measure the network load in terms of total data packet delivery cost for a session. This metric was already defined in Chapter 3 as C_{PDC} and its value is controlled by the size of the data messages multiplied by the number of hops needed to forward packets from the CN to the MN or vice versa.

With our hybrid DMM proposal, the mobility is managed by PMIPv6 or NB-DMM depending on the decision criteria described previously. Hence, the values of packet delivery cost for each solution are as follows:

$$C_{PDC}(HybridDMM) = \mathcal{P}_C(C_{PDC}(PMIPv6)) + \mathcal{P}_D(C_{PDC}(NB - DMM)) \quad (6.2)$$

6.5.3 Tunnelling Cost

To achieve a seamless mobility support, mobility management protocols use a tunnel to forward/re-direct packets. The tunnelling cost metric (C_t) represents in essence the cost of adding a tunnelling overhead to the overall data packet delivery cost. So, the tunnelling cost can be derived from packet delivery cost by setting the payload size of the packet to zero, $s_d = 0$. In that respect, and similarly to the signaling cost and the packet delivery cost, our hybrid DMM proposal adapts its operation to the network topology due to the decisions taken at the beginning of the network operation.

$$C_t(HybridDMM) = \mathcal{P}_C(C_t(PMIPv6)) + \mathcal{P}_D(C_t(NB - DMM)) \quad (6.3)$$

6.6 Simulated results

Following the analysis presented above, this section aims to provide insights about the impact of several mobility costs on the overall network performance. We compare the introduced hybrid CMM-DMM solution in its two versions, node-assignment algorithm (Hybrid Solution 1) and link-assignment algorithm (Hybrid Solution 2), with MIPv6 and PMIPv6 as centralized approaches, as well as HB-DMM and NB-DMM as distributed protocols in terms of registration update cost, packet delivery cost and tunnelling overhead.

The evaluation through simulations aims to study the distributed approaches in a more realistic environment. The platform selected for the evaluation through simulations was MATLAB. The scenario defined for the evaluation is illustrated in Figure 6.6. Due to the dependence on the topology of DMM protocols, we select this asymmetric topology due its mixture between a well connected hierarchical network and a sparse network. This will produce more realistic results because the nodes will move around the connected and the sparse areas of the network, avoiding misleading performance of centralized or distributed protocols due to the network topology. In addition such topology will allow us to shed further light on the dependency of network topology on the performance of different mobility management protocols.

The traffic and mobility parameters values used in the simulations, as well as the numerical results of mobility costs are presented next.

We consider a scenario where an MN may traversal several simultaneous active sessions with several CNs in the Internet. We assume that sessions arrivals to an MN follows a homogeneous Poisson process with mean rate $\lambda_s = 0.01$ (i.e. the inter-arrival time between

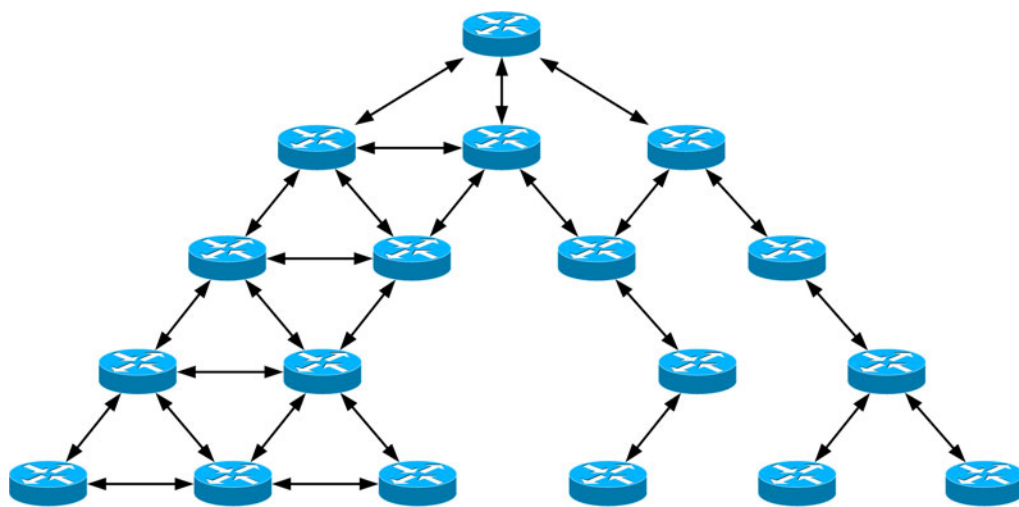


Figure 6.6. Topology used in the simulation

sessions is exponentially distributed with this rate). We assume also that the duration of a typical session is exponentially distributed with mean session duration $\mu_s = 10$ time units.

We validate our model using a Random Waypoint mobility model with the following parameters. Speed: uniformly distributed between 1 and 10 m/s; Pause interval: uniformly distributed between 1 and 5 min.; Walk interval: uniformly distributed between 5 and 20 min. In order to drive the evaluation in a more realistic scenario, we also run the simulations with real-world mobility track logs obtained from users carrying GPS receivers. The sample settings where traces are obtained are two university campuses (one in Asia and one in the US), one metropolitan area (New York City), one State fair and one theme park (Disney World). The participants walk most of the time in these locations and may also occasionally travel by bus, trolley, car, or subway. These settings are selected because they are conducive to collecting GPS readings [104].

Moreover, the simulation time is sufficient large (45000 units) to avoid "typical runs" statistical problems. The dimensions of the simulation scenario for the RWP mobility model is a rectangular area of 5x5 km² and the MNs are initially located randomly in that area. With regards to the real mobility tracks, the dimensions of the rectangular simulation area, is set to be the same as in the GPS traces. In all simulation scenarios, we used the same initial positions found in the respective real traces for the same number of users. In the evaluation, the simulations are repeated 25 times to improve the accuracy of the results with a confidence interval of 95%.

6.6.1 Impact of the amount of MNs

We present and discuss in this section the numerical results showing the impact of the amount of MNs on the mobility costs. The interest in this metric is due to the fact that higher values reflect more probability of encountering scalability issues, which are a major concern in current mobility protocols. In this case, the simulation is run for different number of MNs, ranging from 1 to 50.

Figure 6.7 shows the accumulated signaling cost of registration delay update vs. the number of MNs during all the simulation execution. Both mobility models are shown in order to be compared, human mobility is shown in Figure 6.7(a), whereas the Random Waypoint results are shown in Figure 6.7(b).

In this case, in MIPv6 and PMIPv6 protocols the control messages are exchanged between two entities, the serving AR (MAG in PMIPv6) and the centralized mobility anchor (HA in MIPv6 and LMA in PMIPv6). In these centralized solutions, all sessions are anchored

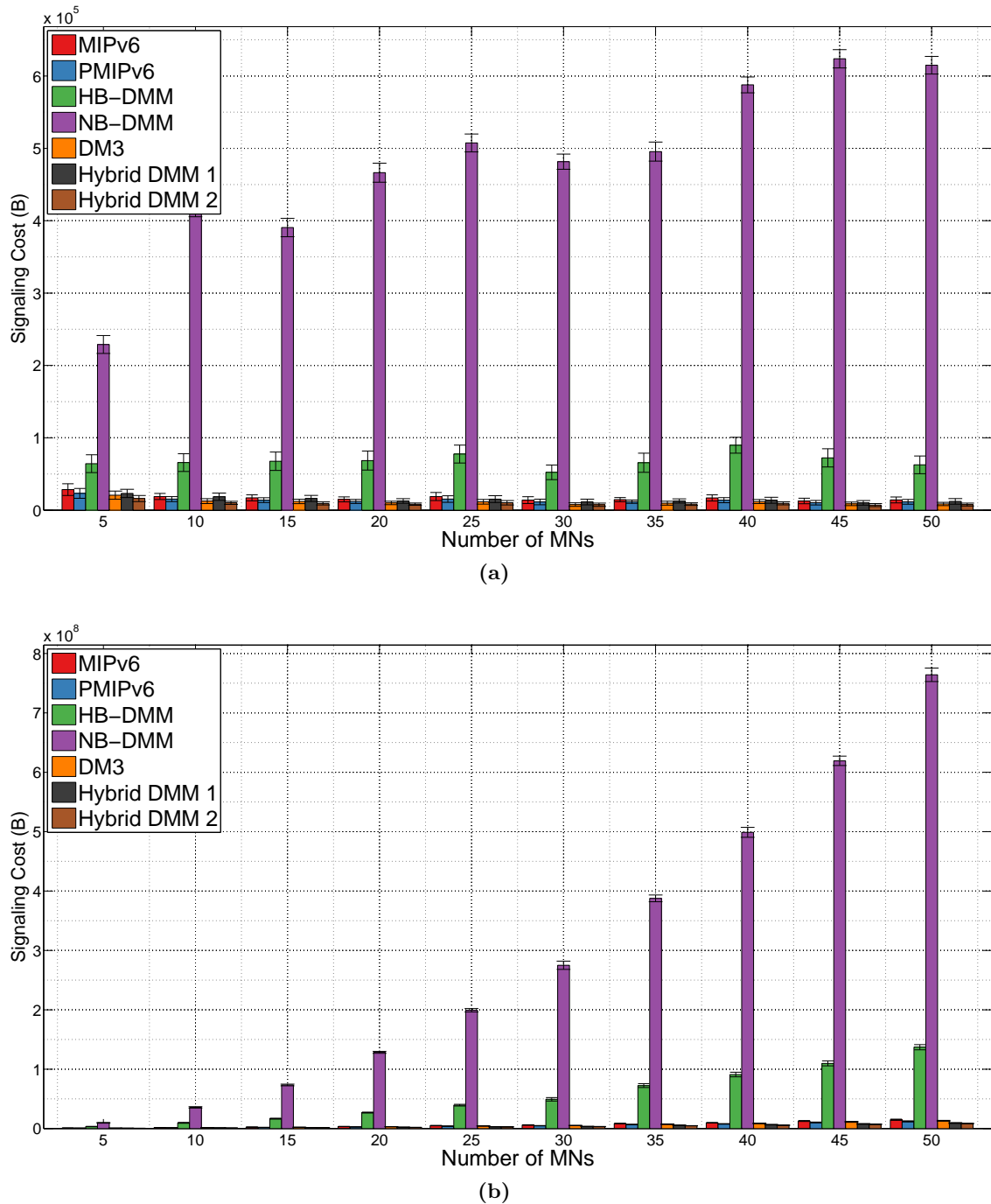


Figure 6.7. Signaling Cost. (a) Human walk mobility model; (b) Random Waypoint mobility model

to the same agent, therefore all of them are updated in the same signaling message, thereby introducing a low overhead. On the other hand, in HB-DMM and NB-DMM, the control messages are exchanged between the distributed nodes for each connection that remains active for MNs during their movement; thus, there can be scalability concerns since the

signaling overhead increases in proportion to the number of the MN's IP prefixes/addresses anchored at ARs other than the serving AR. The signaling in DM3 is done with the distributed mobility anchor (MDA), located near the edge of the access router. This location gives more flexibility and avoids complex tunnelling management. With respect to the proposed hybrid mobility proposals, both algorithms manage the signaling overhead in an efficient way, offering an equivalent performance in that matter with centralized approaches. As it is shown in Figure 6.7, the signaling cost in distributed approaches is significantly higher than the other protocols, especially NB-DMM. This high cost in NB-DMM is produced because it additionally requires the exchange of a control message with the database each time a handover is produced, adding an extra signaling overhead.

With respect to the mobility models, the RWP mobility shows a clear trend as the number of MNs increments. This effect is not visible when using the realistic human based mobility traces because when some users are added to the scenario, they do not add any additional movement so, the signaling cost of registration update is not increased, whereas in RWP mobility model, all users have a much more similar behavior from a mobility point of view.

With respect to the accumulated packet delivery cost, the results obtained are shown in Figure 6.8. It can be observed that both mobility models present a similar behavior because the packet delivery cost is not highly dependent on the mobility model. In the same figure it can also be observed how centralized solutions perform a non-optimal routing and therefore the overall cost is higher. DMM protocols on the other hand, outperform the cost of centralized protocols although when a handover occurs, packets are routed through a suboptimal path. As we can see, the value of both DMM protocols is the same. This occurs because the data plane in both HB-DMM and NB-DMM operates in the same manner.

Based on the above observation we can conclude from Figure 6.8 that the benefits obtained from the deployment of hybrid solutions in mobile networks can produce substantial improvements, up to 60% reduction in routing cost with respect to centralized approaches and up to a 45% improvement with respect to DMM protocols. Taking into account the increasing traffic expected for future mobile networks, this saving in resources due to the inclusion of the proposed hybrid approaches can facilitate the deployment of next generation mobile network architectures.

Finally, in Figure 6.9 the tunnelling cost of the mobility protocols is compared. The significant difference between CMM and DMM solutions is highlighted. While HB-DMM and NB-DMM introduce an insignificant tunnelling, centralized solutions cause a high overhead in the network due to the tunnelling process. With the proposed hybrid approaches, we try to keep all mobility sessions with a good performance, in terms of

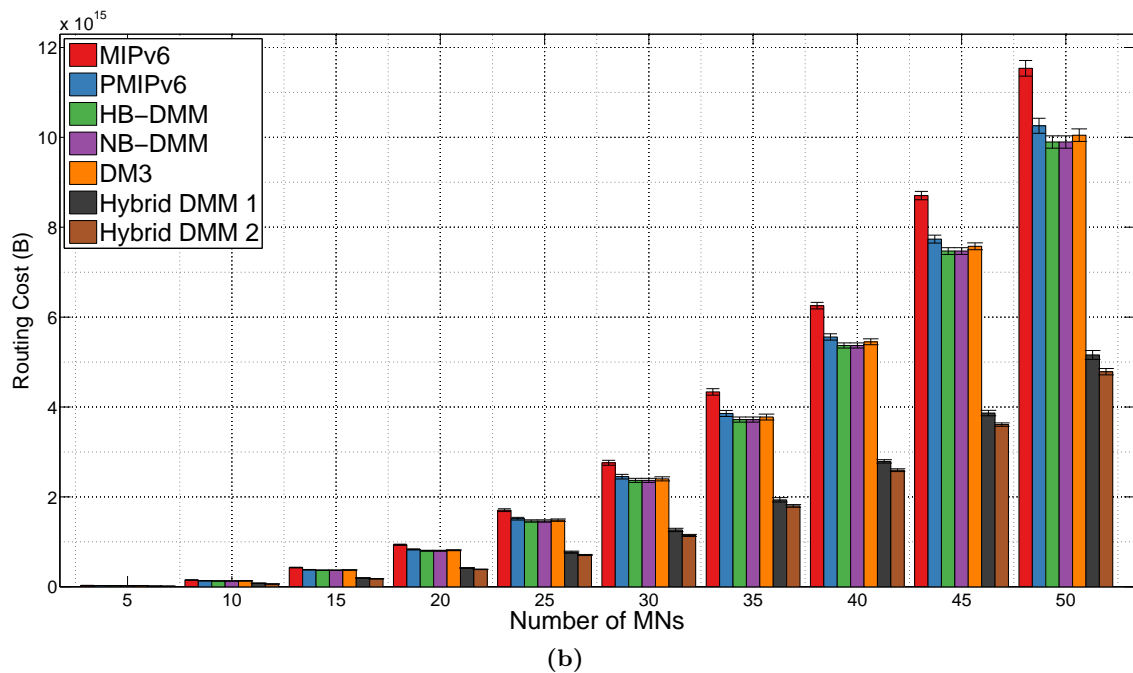
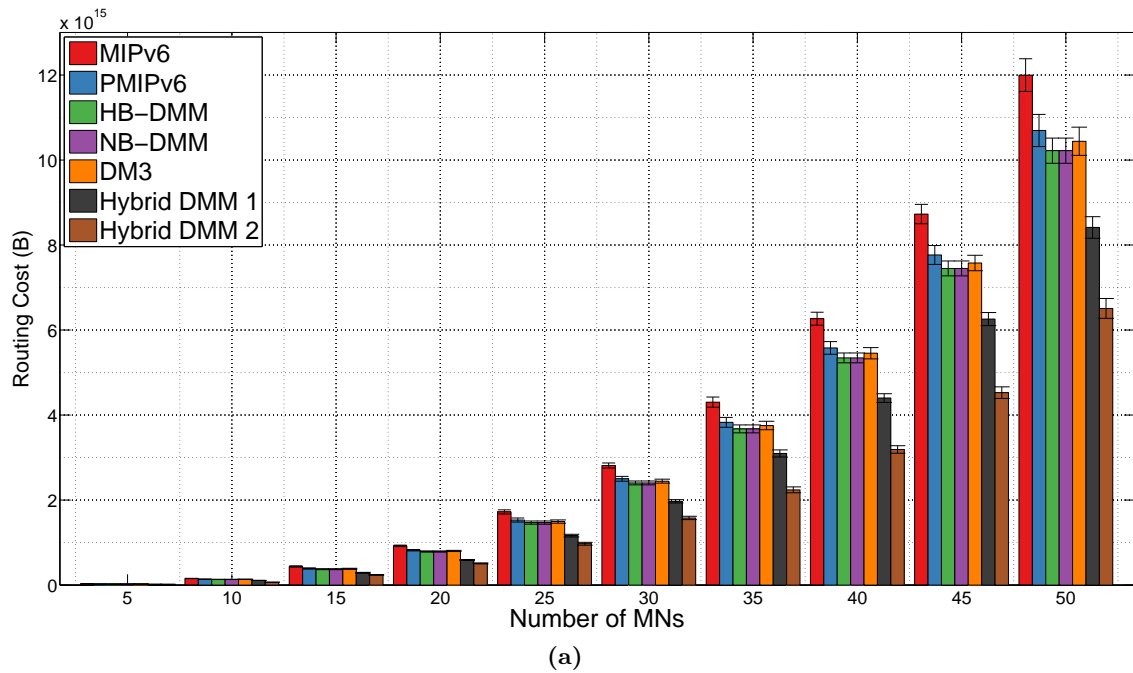


Figure 6.8. Routing cost. (a) Human walk mobility model; (b) Random Waypoint mobility model

signaling and packet delivery cost so, an acceptable tunnelling cost is generated. Thus, Figure 6.9 illustrates that the tunnelling overhead introduced by both Hybrid DMM solutions is minimal.

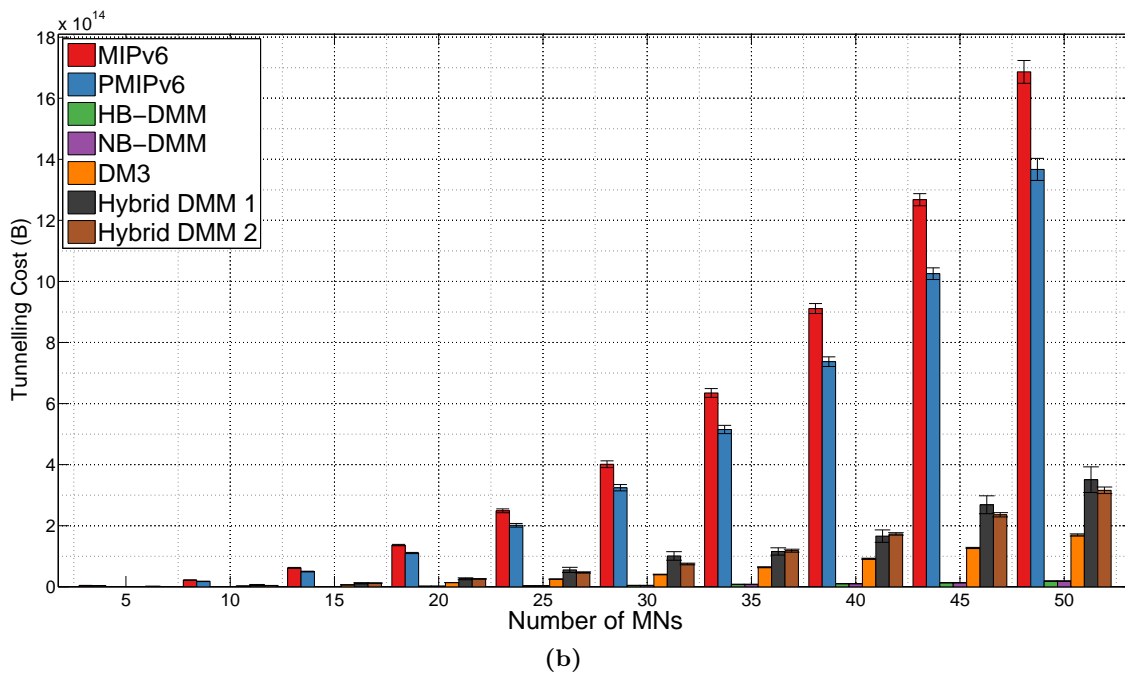
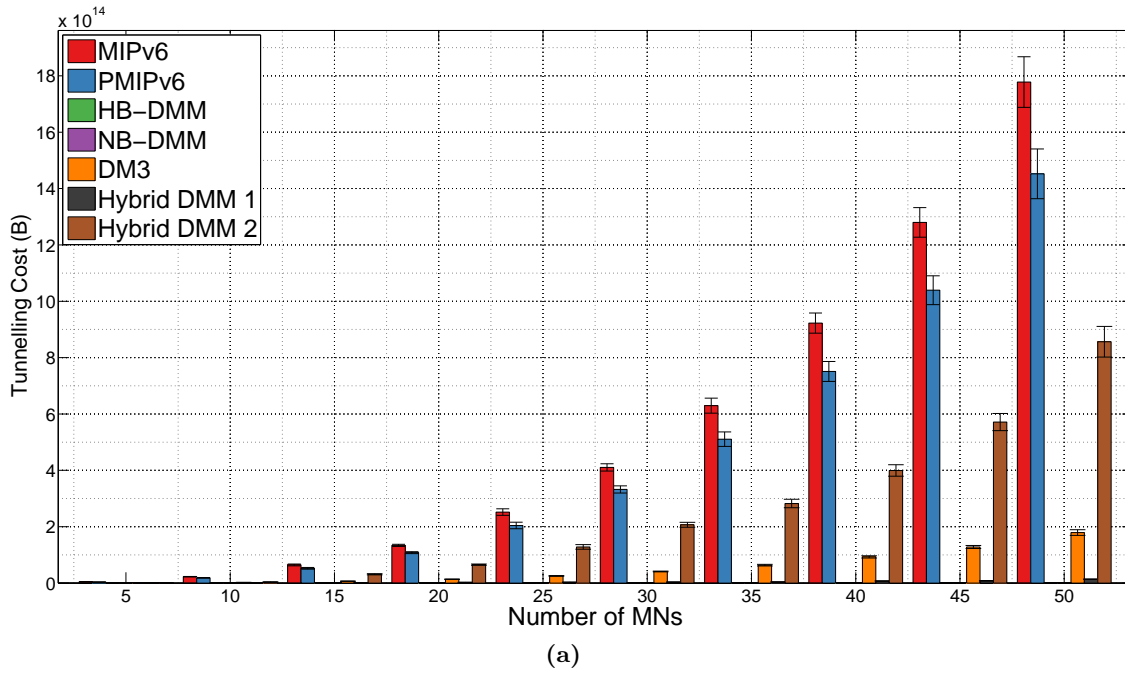


Figure 6.9. Tunneling cost. (a) Human walk mobility model; (b) Random Waypoint mobility model

As we have mentioned previously, DMM deployments faces several issues such as complex address and tunnel management, high signaling cost and high signaling cost as the number of MN increase, for example, in the case of users moving at high speed and/or long-lasting sessions. Consequently, DMM may not be a suitable scheme in certain situations. For this

reason, in the next sections we investigate the impact of different key parameters for IP mobility management protocols and the possible benefits that hybrid solutions can offer.

6.6.2 Impact of Session Duration and Topology Scenarios

Now we examine the impact of the session duration μ_s . We consider that the duration of a typical session is exponentially distributed with mean rate μ_s . The rest of the parameters are shown in Table 6.1). We vary μ_s from 5 to 1750 seconds. These values are been considered to cover the typical session duration values for mobile users [106] as it shown in [107], in which can be seen the average mobile application session length sorted by app category. In this case, the longest sessions corresponding to music apps, lasted 8.9 minutes whereas social networking sessions only took an average of 2.5 minutes. Moreover, in this experimentation we also investigate the topology dependence. Thus, the experiments have been conducted over four different topologies, as can be seen in Figure 6.10.

Figure 6.11 shows the variation of the signaling cost as a function of the session duration (μ_s). As μ_s increases the overall signaling cost increases in the four topologies. As it appears in the figure, the cost in HB-DMM and NB-DMM increases at a higher rate than the other solutions. This is due to the fact that for long-lasting sessions, the distributed solutions need to manage complex address and tunnel management that require extra signaling in each movement. However, both MIPv6 and PMIPv6 achieve similar cost values,

Table 6.1. Simulation settings. Impact of the Session Duration.

Parameter	Value
Number of MNs	15
Number of ARs	6
Simulation time	45000 sec.
Session arrival rate (λ_s)	0.01 sec.
Session duration (mean rate μ_s)	5 to 1750 sec.
MN movement model	RWP mobility model
Simulation area	5x5 km ²
RWP speed interval	[1 10] m/s
RWP pause interval	[60 300] sec.
RWP walk interval	[300 1200] sec.
Simulation run repetitions	25

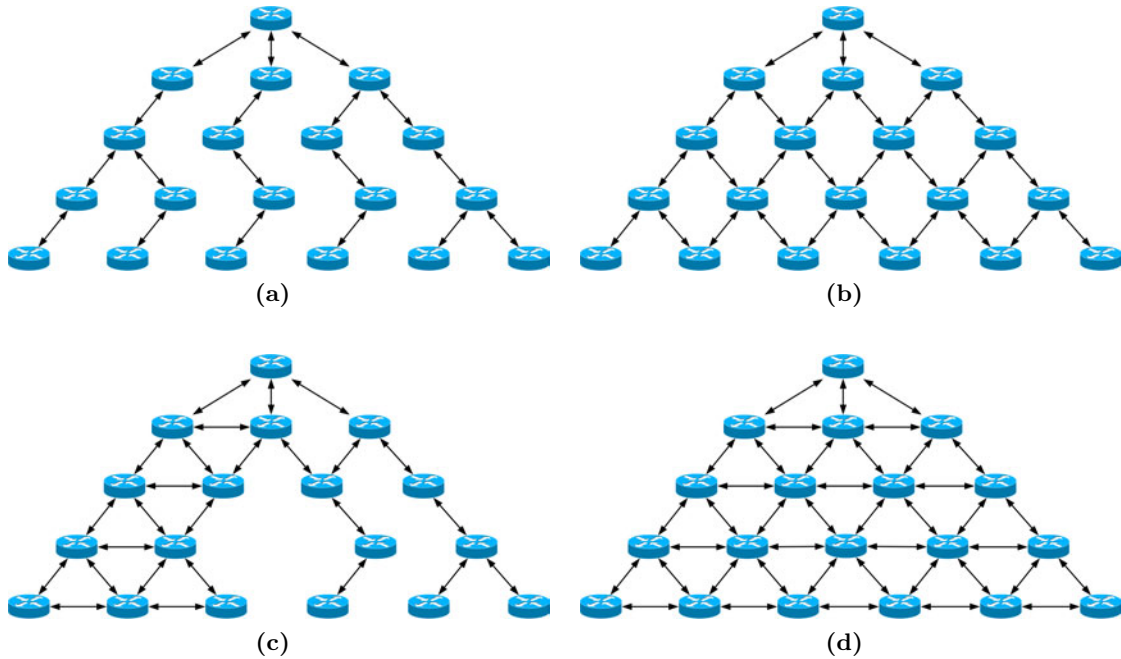


Figure 6.10. The four different topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.

because they only need to send a unique update message (BU or PBU respectively) to the central anchor per movement. Similarly, DM3 introduces low signaling overhead because although several anchors are distributed through the access network, only one message with the serving MDA node is necessary to notify the movement. Finally, The hybrid solutions are very dependent on the topology. Whereas in tree-based and fully connected topologies the sessions are mostly distributed, achieving similar values to NB-DMM, in low connected scenarios the cost reduction with respect to NB-DMM is very significant.

Figure 6.12 shows the variation of the session duration with respect to the routing cost, also called packet delivery cost. In this case, routing cost in all solutions provides similar results. This parameter is the most relevant for the Hybrid DMM solutions, because the decision criteria algorithms presented in this chapter were based on it. Packet delivery cost is a key metric for future mobile networks due to the exponential growth that current and future networks are experiencing in terms of mobile data traffic. Thus, the results shown in this plot demonstrate the benefits that our two Hybrid DMM solutions can offer. These benefits are especially significant in low connected topologies or in topologies that mix areas well connected with sparse connected areas (as the Trade-Off topology represents). It is worth noting that the second Hybrid DMM solution provides lower values of routing cost. This is because it optimally selects the paradigm that a new connection must follow. But this improvement comes with a penalty that can be seen in the previous Figure 6.11,

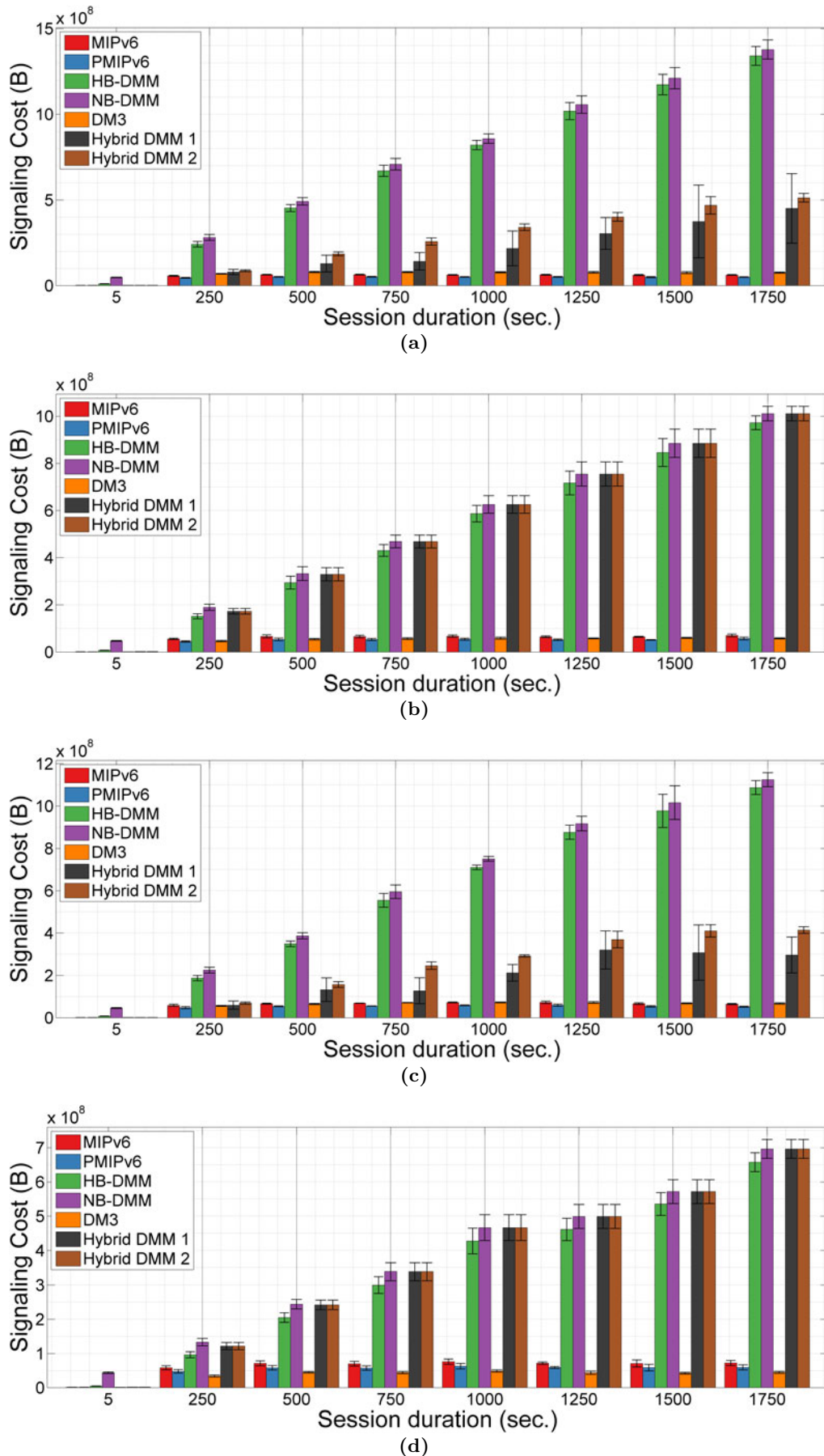


Figure 6.11. Impact of the signaling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.

in which this solution needs an extra signaling overhead with respect to the first Hybrid DMM algorithm. This also demonstrates that our algorithms work correctly, because they are designed to optimize the routing cost as much as possible.

Finally, Figure 6.13 shows the variation of the tunnelling cost with respect to (μ_s). In fact, the tunnelling cost is included in the packet delivery cost but we separate it here in order to compare the effects of tunnelling overhead introduced by IP mobility management protocols to perform the seamless mobility process. We note that the tunnelling cost in centralized solutions is higher than the distributed ones. Only when the topology is low connected, the distributed solutions do not introduce any benefits. This is due to the fact that the tunnel in DMM is set-up between the edge nodes, but if those edge nodes are not connected directly, which is the case of this topology, the tunnel is as costly as in CMM protocols. As it happened in the previous routing cost experiment, Hybrid DMM solutions achieve significant improvements with respect to both CMM and DMM approaches. With respect to DM3, which uses MPLS, the introduction of a technology that natively tunnels the data traffic, makes it possible to achieve the lowest values.

6.6.3 Impact of the MN's Speed and Topology Scenarios

Now we examine the impact of the MN's speed in the mobility costs. This parameter affects the Cell Residence Time in the sense that the higher the velocity of a mobile user, the lower the Cell Residence Time is experienced. Moreover, the simulation parameters used in this chapter are similar to those introduced in the previous section and shown in Table 6.1). In this case, we vary the MN's speed from 1 to 40 m/s (in Km/h, these range goes from 3.6 to 144). These values are considered to cover a wide range from a typical human walk speed to users in moving vehicles. Thus, the experiments have been conducted over four different topologies, as can be seen in Figure 6.10.

Figure 6.14 shows the variation of the signaling cost as a function of the MN's speed. As the velocity increases, the signaling cost becomes higher. In this case, signaling overhead increase in centralized solutions is significantly lower than in distributed solutions. This effect is due to the fact that for sessions with a high speed, the MN moves frequently, requiring extra signaling for each movement. However, both MIPv6 and PMIPv6 are less affected by this speed variation because they only send a unique update message (Binding Update or Proxy Binding Update respectively) to the central anchor independently of the remaining active connections. Similarly, DM3 introduces low signaling overhead because although several anchors are distributed through the access network, only one message with the serving MDA node is necessary to notify the movement. Hence, DM3 helps prevent the significant increase in terms of signaling cost compared to the centralized-based

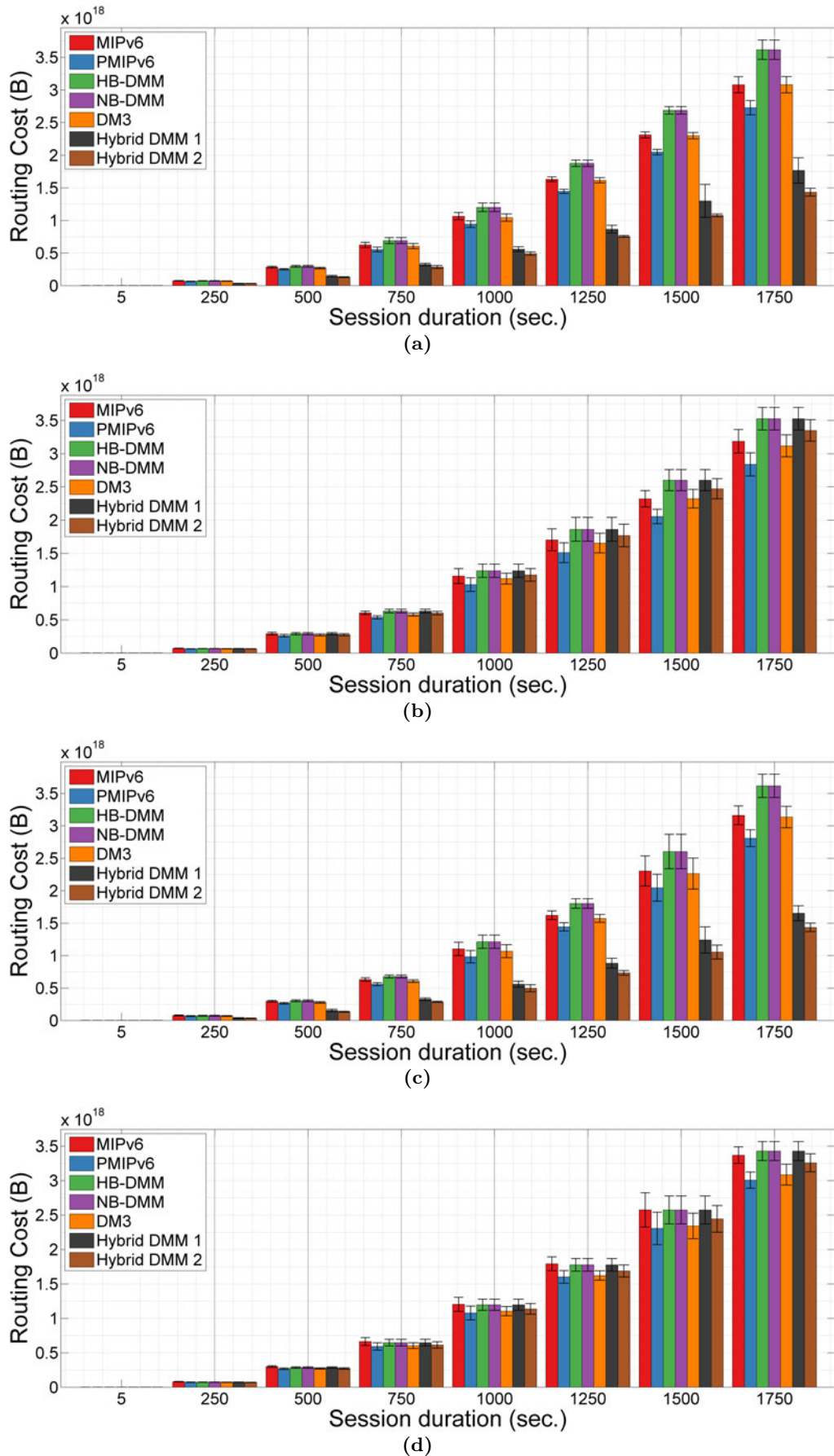


Figure 6.12. Impact of the packet delivery cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.

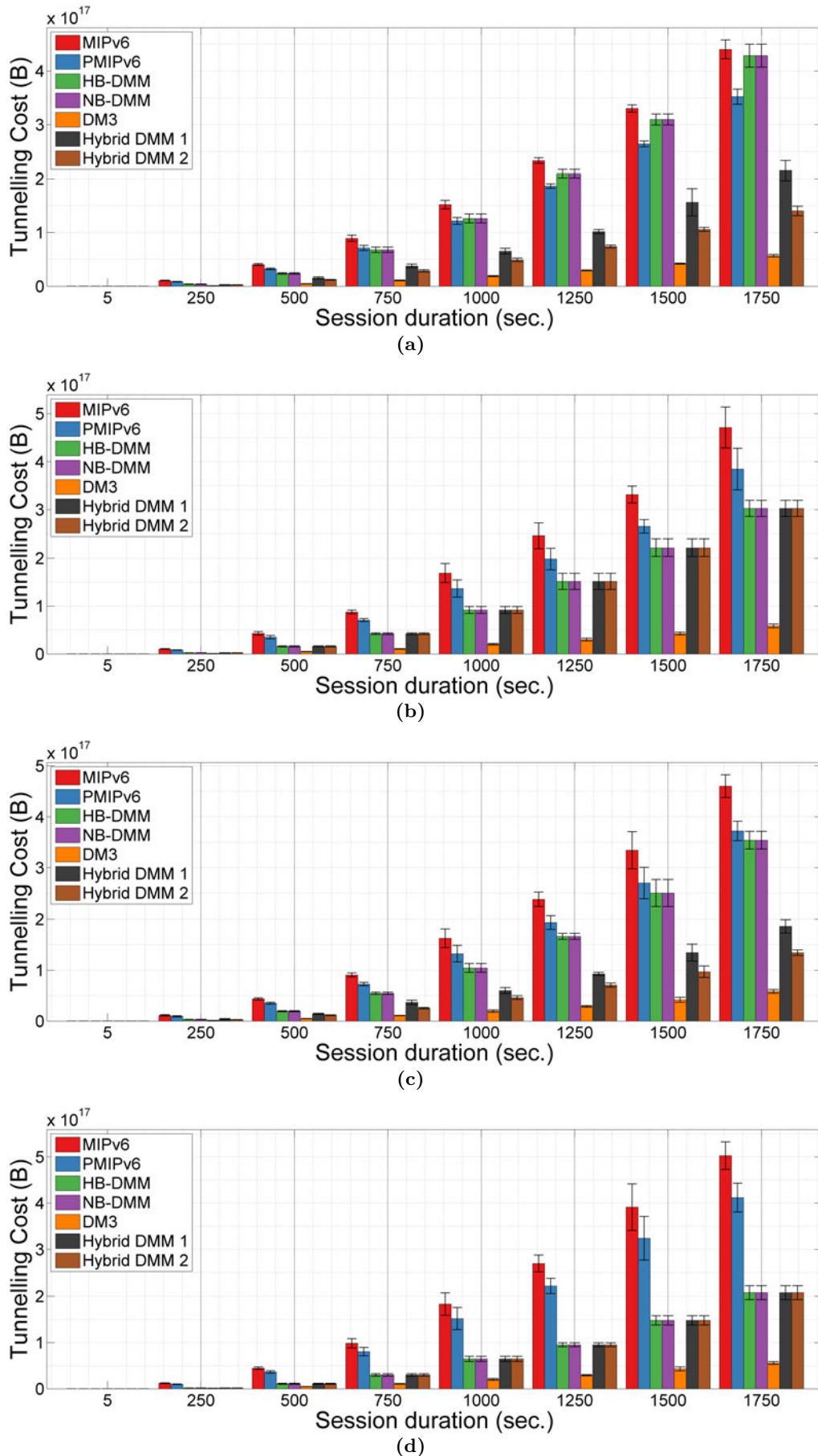


Figure 6.13. Impact of the tunnelling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.

schemes and the other distributed-based solutions as the velocity increases. The case of our Hybrid DMM solutions is different and depends on the topology. Whereas in tree-based and full connected topologies all sessions are distributed, in low connected scenarios the cost reduction with respect to the DMM solutions is very significant.

Figure 6.15 shows the variation of the speed of the MN with respect to the packet delivery cost (or Routing Cost). The results obtained demonstrate that the routing cost is not affected by the MN's speed, achieving a similar value since the packet is routed through a path of a similar length and the end-to-end distance is supposed to be constant. The Hybrid DMM solutions are the unique mechanisms that obtain a significant reduction in this cost. As we have described previously in this chapter, the decision criteria algorithms consider the routing cost the key parameter to perform the decisions. Thus, improvement can be achieved especially in low connected topologies and in topologies that combine well connected areas with sparsely connected ones (as the Trade-Off topology represents). The benefits obtained in terms of routing cost in these topologies are around 50 %.

With respect to the tunnelling cost, Figure 6.16 shows the variation of this metric with respect to the MN's speed. Data packets are encapsulated and de-capsulated, while they travel through the tunnels between some of the nodes that perform the mobility process. We note that the tunnelling cost is rather stable for centralized solutions independently of the MN's speed, whereas in distributed solutions this cost grows significantly at low velocities and remains stable from 10 m/s. Additionally, the value of the tunnelling cost in distributed solutions is very dependent on the topology scenarios. The fully connected topology achieve the lowest values. This is because the edge nodes are connected directly and the tunnels among the serving AMA/MAR and the other AMAs/MARs from which there is still an active connection are short in number of hops. On the contrary, under low connected topologies, both HB-DMM and NB-DMM achieve high values. With respect to DM3, the introduction of a technology such as MPLS that natively tunnel the data traffic, make possible to achieve the lowest values.

6.7 Concluding remarks

Mobility management protocols are evolving towards a distributed operation in order to deal with the continuous increasing of mobile Internet traffic. However, as we have detailed in some scenarios, the performance of DMM falls dramatically due a number of different factors, such as the high session arrivals rate, long and frequent movements (i.e, short residence time) and long-lasting sessions. In these situations the operation of DMM

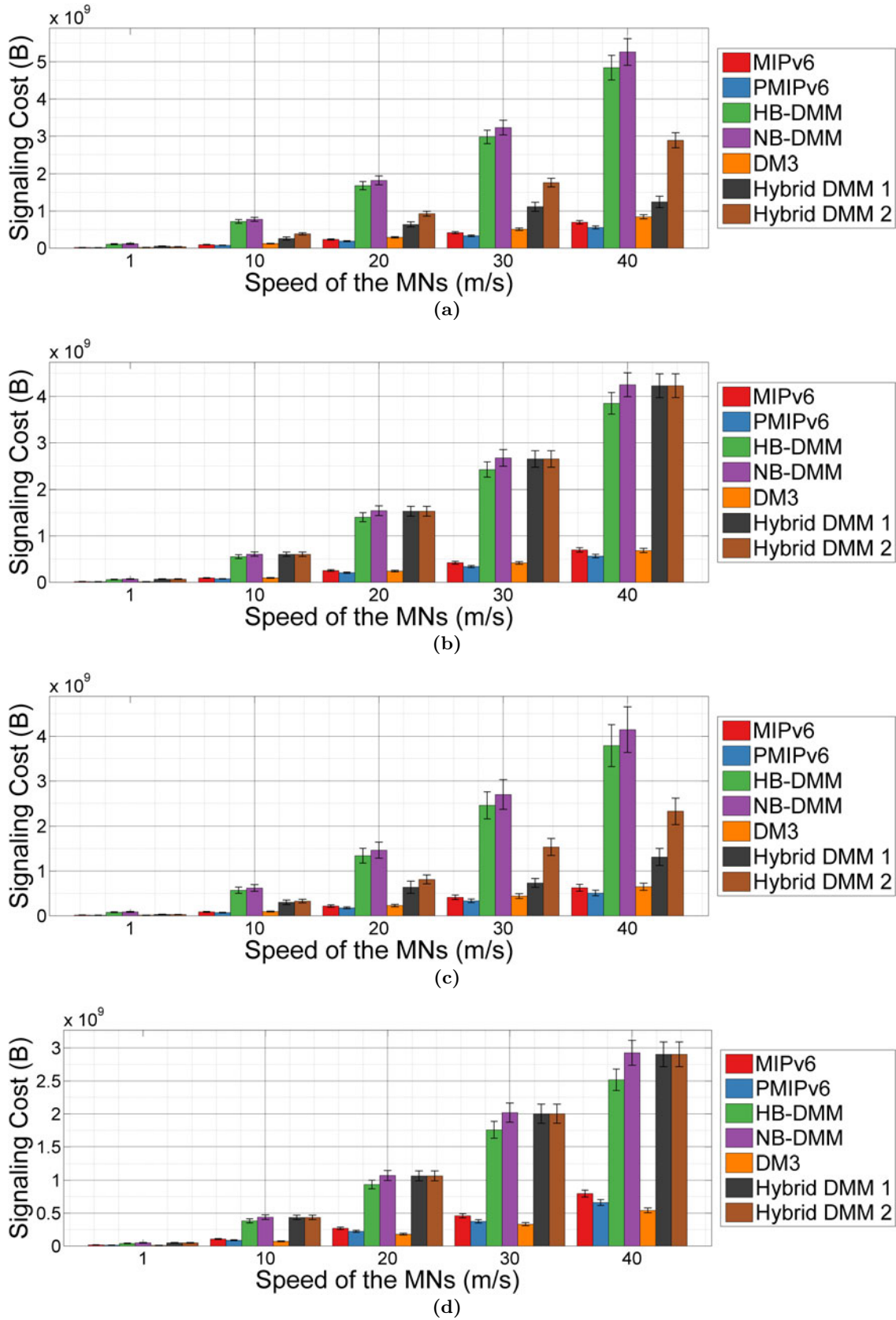


Figure 6.14. Impact of the signaling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.

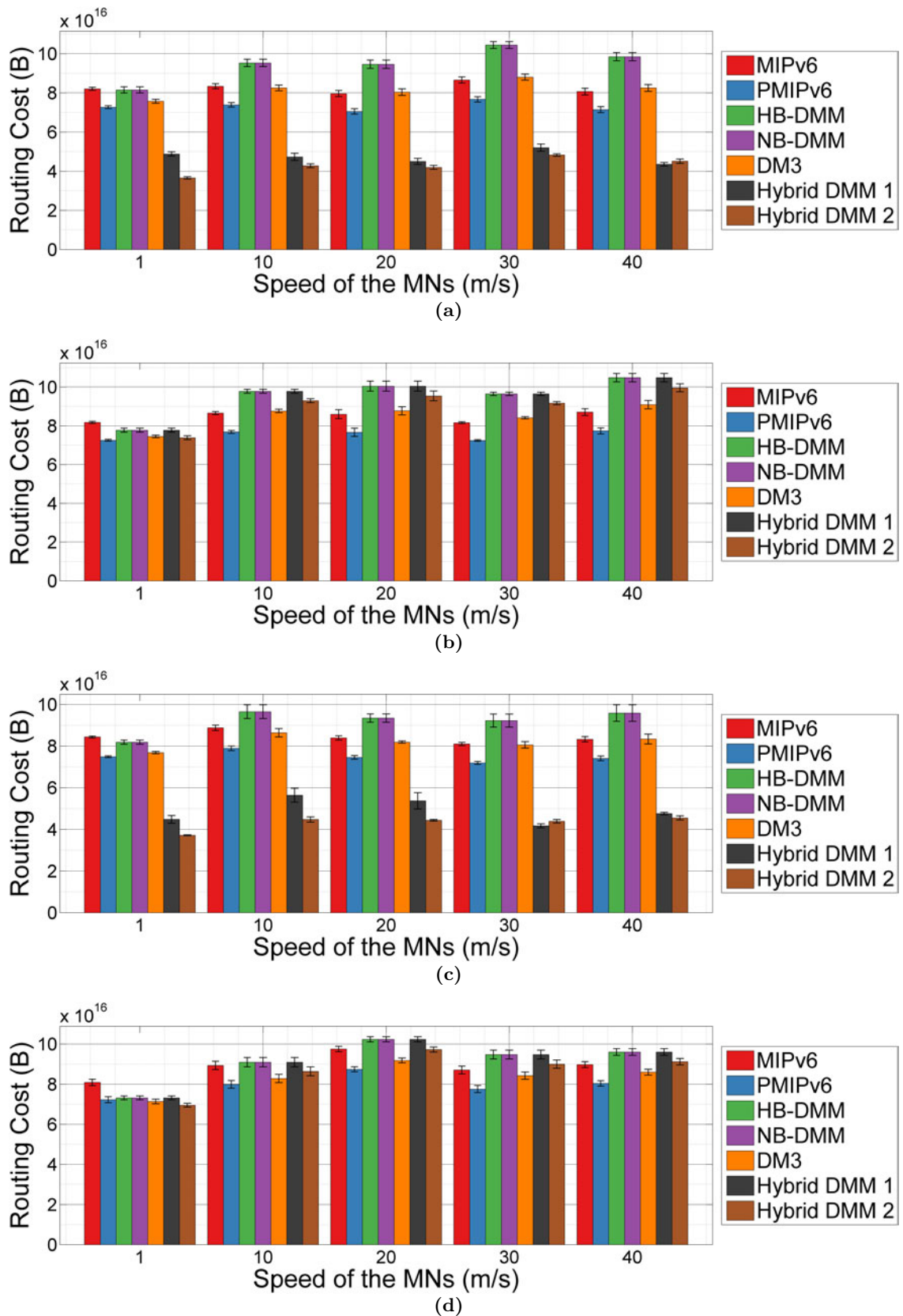


Figure 6.15. Impact of the packet delivery cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.

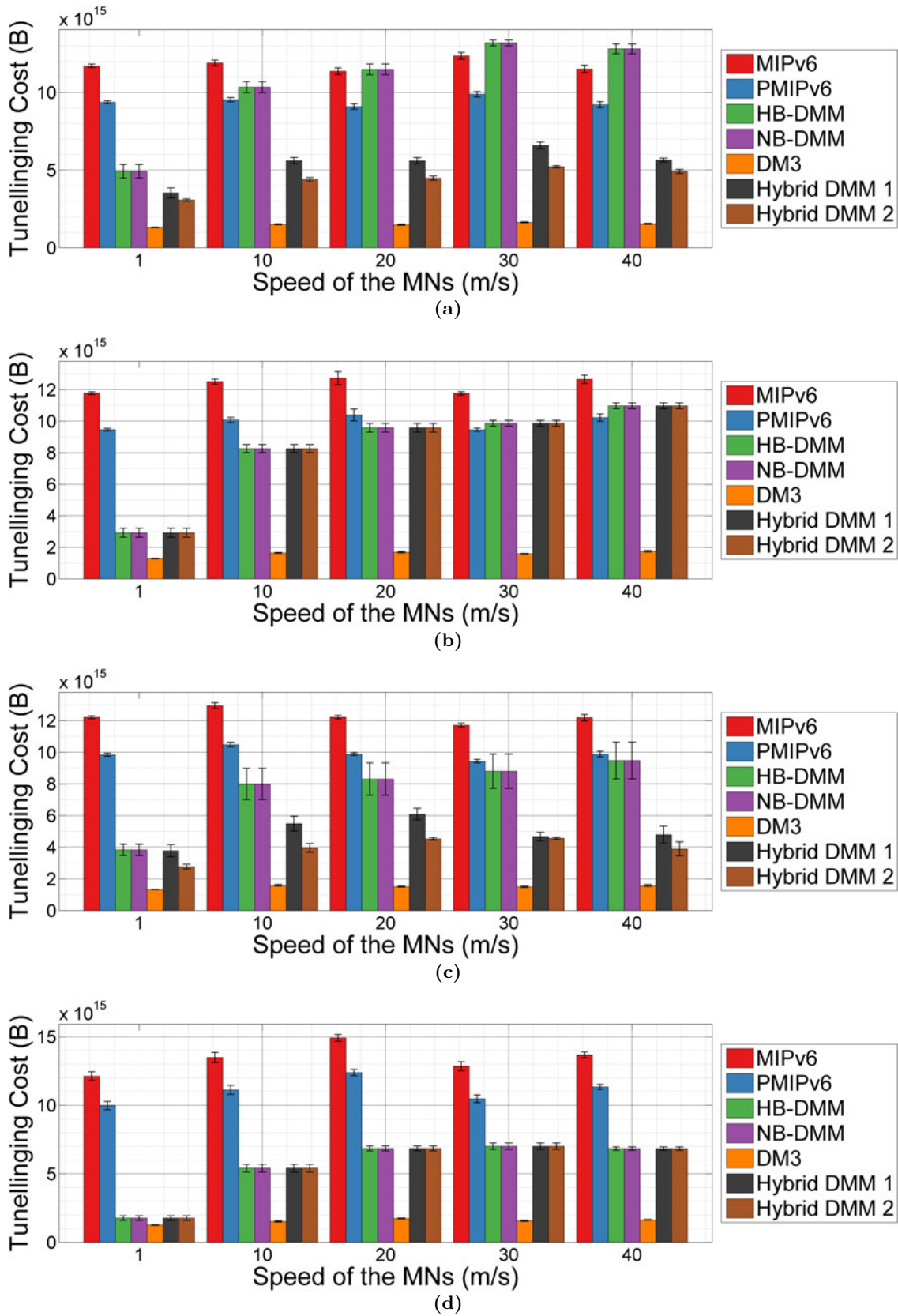


Figure 6.16. Impact of the tunnelling cost in the four topologies used in the experiments. (a) Low connected topology. (b) Tree topology. (c) Trade-off topology. (d) Full connected topology.

might lead to a lower performance and hence the use of a centralized mobility management solution would be a more preferred option.

In this chapter and based on the above mentioned observations a hybrid centralized-DMM solution is proposed in which the mobility can be managed by a centralized protocol such as PMIPv6 or by a distributed mechanism (NB-DMM). The decision of selecting one protocol or the other is made by some decision criteria which can be adapted based on network conditions. Two decision criteria algorithms are proposed depending on the level of available information. The first one is denoted as node-assignment algorithm and it takes the network topology information to take decisions about the protocol that an AR should use. The second algorithm is called link-assignment and in addition to information about network topology, it uses BS location information in order to decide the protocol to use according to the path the MN is moving into. Additionally, we have conducted an extensive set of analytical and numerical investigations of CMM, DMM and hybrid solutions. After defining the analytical models, the expressions of mobility costs have been derived.

The numerical results showed that the proposed Hybrid DMM can retain the advantages of DMM while limiting its drawbacks in terms of mobility costs from both the control and the data plane.

7

CONCLUSIONS AND FUTURE WORK

This chapter concludes the thesis and provides future avenues of research in the scope of this thesis.

Contents

7.1	Conclusions	153
7.2	Future Work	155

7.1 Conclusions

Over the last few years, IP-based mobility in the Internet has been one of the most active research fields in communications. The management protocols responsible for maintaining ongoing communications while the user roams among distinct networks have been evolving together with the rapid development of new wireless technologies and, in parallel, with the necessities of users, that require reachability in such heterogeneous environments in terms of access. This trend is predicted to continue in the future, due to the increase in mobile data traffic demand and the increasing number of smart mobile devices.

The existing IP mobility support protocols developed by the IETF have been based on centralized anchors that manage the traffic and signaling of the MNs. However, in order to cope with the recent trends in mobile Internet and the current increasing mobile data traffic demand, vendors and the research community are investigating possible research guidelines to carry out effective solutions for future mobile networks.

One of the current research areas focuses on the redesign of these IP mobility protocols, aimed at addressing the limitations of traditional schemes. In this context, this thesis covers the topic of Mobility Management, including both Centralized Mobility Management and Distributed Mobility Management. The reason behind covering both approaches is due to the fact that current mobile architectures are heavily centralized but the network operators are looking towards a new paradigm that relies on distributing the entities in charge of managing users' mobility. Thus, DMM is envisioned as a promising mobility management scheme that provides an efficient use of network resources through better traffic distribution among several mobility anchors.

Thus, in this thesis, an in-depth review of the existing centralized and distributed IP mobility management schemes has been conducted. The review highlights some of the limitations of the existing schemes. The drawbacks of centralized protocols have been thoroughly discussed in several works regarding current mobile Internet issues. However, the review has revealed some limitations of the existing DMM schemes. This review motivated the research carried out in this thesis, that can be summarized in three main novel IP mobility management designs:

1. One mobility solution based on a host-based centralized approach called LinkWork Mobile MPLS (LW-MMPLS).
2. One mobility solution, called DM3, based on LW-MMPLS that distributes mobility functions through different nodes in the access network.

3. One hybrid DMM scheme that couples a centralized network-based solution (PMIPv6) and a distributed network-based one (NB-DMM).

Besides the design of these novel solutions, we have dedicated a significant part of the thesis to the comparative performance analysis of the existing mobility management schemes as well as the proposed ones, focusing on the key mobility costs, such as the signaling overhead, the packet delivery cost, the tunnelling overhead or the handover latency. This performance comparison has been made using the following methodology. After reviewing in detail the existing mobility protocols, we have developed an analytical framework in which we have modelled the operation of these protocols defining the network model and the mobility model. We have also defined the analysis criteria and developed their expressions in a general form that can be applied to any mobility protocol. Thus, based on this framework it is possible to compare existing solutions with new designs. Therefore, we have also carried out a comparative performance analysis to compare each of our proposals with different representative mobility management schemes.

We have discussed the benefits and limitations of our LW-MMPLS and DM3 compared to the existing protocols. Through analytical and simulated evaluation, we show that DM3 reduces significantly the signaling overhead as well as the routing cost, and provides a low handover latency with a minimal packet loss rate. Additionally, we have conducted experimental evaluations in order to demonstrate that the handover latency depends on the distance of the network entities involved in the signaling.

Our investigations showed that in some scenarios, DMM solutions face several issues that cause higher mobility costs than centralized protocols, especially those in which the MN is moving at a high speed and/or with lost-lasting sessions. To cope with this issue, the Hybrid DMM solution has been proposed. This Hybrid DMM is based on a decision criteria algorithm that an operator could use to determine how the traffic will be managed and anchored. Accordingly, two algorithms have been developed. The algorithms make decisions based on some network information. The first one (node-assignment) uses only information about the network topology, whereas the second algorithm (link-assignment) uses both topology and BSs location information. In this case, both analytical and simulated evaluations are conducted in order to measure the behavior of our Hybrid DMM scheme compared to the existing mobility protocols. The results obtained show that the use of the proposed hybrid solutions outperform significantly previous mobility management schemes in terms of network resource consumption (control plane) as well as current mobility management performance (data plane). Based on the evaluations made, we can conclude that it might be beneficial for future mobile network architectures to utilize hybrid CMM-DMM solutions, in which operators would be able to handle the traffic in a distributed or centralized way, depending on topological network characteristics.

7.2 Future Work

Following the development of different schemes during the research for this thesis, there are some interesting issues that need further investigation, in order to improve the design, applicability and performance of the proposed schemes. Moreover, other issues that have been considered out of the scope of this thesis might be worth investigating in future work.

- *Optimize the location and amount of MDAs in DM3.* Current design of our DM3 proposal needs further investigations in the dynamic location of the distributed nodes, as well as the optimal number of these anchors. Based on given network traces and statistics, normally known to a mobile operator, the next step in the evolution of this solution is to determine the best network configuration in order to achieve a better performance.
- *Mobility management for non-mobility aware applications.* As we have mentioned in the Conclusions section, DMM approaches exhibit some limitations due to the complex address and tunnel management, which provokes high signaling cost and high handover latency. Moreover, recent reports mention that signaling traffic in mobile networks is growing faster than data traffic. Considering that a mobile user runs two kinds of applications, aware and non-aware mobility, new solutions should cope with this issue avoiding some of the signaling overhead generated by those non-mobility aware applications.
- *Hybrid DMM decision criteria.* This aspect is related to the decision criteria algorithms that select the mobility protocol that will be applied for each MN or each session. With the aim of offering more flexibility for operators, these decisions are taken based on a set of metrics (e.g., access network topology, characteristics of the flow, number of active prefixes, etc.). In this thesis we have proposed two algorithms that reduce the overall routing cost, however other mobility costs can be also considered as the metric to optimize.
- *Mobility management prototypes.* The performance evaluation and analysis of the proposed schemes have been carried out using analytical modelling, simulation and, in some specific aspects, experimental testbeds. Although the efforts needed to develop functional prototypes are very high, it will be interesting to extend the real-life implementations to validate the proposed schemes. Thus, one of the main future works derived from this thesis is to implement and to produce a completely functional prototype of the DM3 proposal, as well as the Hybrid DMM solution.

- *SDN mobility for 5G networks.* Software Defined Networking (SDN) is a new paradigm of networking whose key idea is to separate the control and data planes. Recently there is a growing trend to apply SDN to the mobile Internet. SDN makes the network programmable, and simplifies the deployment of mobile solutions by facilitating the distribution of mobility-related functionalities. As future work we consider studying on integrating SDN and mobility in order to optimize the mobility support in 5G mobile networks.

APPENDIX A: EXPERIMENTAL EVALUATION OF CMM PROTOCOLS

This appendix presents a complementary work carried out in the scope of this thesis. It concerns an experimental study of Mobile IPv6 and Proxy MIPv6, based on their open source implementations. We analyze the handover latency characteristics in both protocols and provide quantitative and qualitative performance measures of multimedia communications under different network conditions.

Contents

A.1	Introduction and contributions	159
A.2	Open source implementations of CMM protocols	159
A.3	Performance evaluation	160
A.3.1	Experimental setup and design	160
A.3.2	Quantitative tests	162
A.3.3	Qualitative tests	169
A.4	Final remarks	170

A.1 Introduction and contributions

In previous chapter we introduced the analytical framework for performance evaluation of IPv6 mobility management protocols that is used in this thesis. As it was stated, there are different methods for evaluating the performance of a networking system: analytical, simulation and testbed experiments. In order to offer a complete performance and to understand the real behavior of the main centralized mobility management protocols such as Mobile IPv6 and Proxy Mobile IPv6, in this chapter we report on the experimental evaluation of these protocols through the evaluation of handover latency in a real scenario.

Due to the proliferation of multimedia mobile devices and the variety of mobile applications that generate an enormous amount of data traffic over mobile networks, mobile video is becoming the key driver of the mobile traffic growth. Moreover, this growth over the last few years is expected to intensify in the near future, specifically all those related to video over mobile networks [3].

In addition, the Linux based mobile platforms market is increasing rapidly and various open source solutions that offer seamless IP mobility in a device equipped with a mobility ready Linux kernel have been developed.

Given this scenario, we have conducted an experimental study of the handover latency characteristics of MIPv6 and PMIPv6, based on the open source implementations of these protocols (UMIP [108] and Open Air Interface PMIPv6 [109] respectively) in the Linux TCP/IPv6 stack. This experimentation also includes a quantitative and qualitative study of multimedia communications during the movements of mobile users that indicate the behavior of the MIPv6 and PMIPv6 implementations and the user experience for real-time and streaming multimedia applications. The results have been obtained in a real scenario with real implementations of the protocols in Linux based devices and open source tools have also been used in order to generate, analyze and measure the mobile data traffic. Moreover, Netem (NETwork EMulator) has been used in order to emulate the characteristics of the end to end path in the access network. This way, a wide variety of conditions such as delay, jitter and packet loss have been tested. Other previous implementations of MIPv6 and PMIPv6 are introduced in [110].

A.2 Open source implementations of CMM protocols

Linux implementation of mobility is divided into both kernel space and user space. This implementation strategy has been made in order to allow easily the extension of other IP

mobility protocols with a minimal kernel side and a user space where different protocols can be implemented. This way, the changes in the kernel can be kept to a minimum, which means more robust implementations. The kernel side support consists of a single module (`mip6.ko`) that has been newly integrated into the kernel sources since version 3.8.2 and most mobility protocol functionality such as MIPv6 or PMIPv6 are implemented in the user space.

In this work, MIPv6 support is provided by UMIP under a Linux mobility ready kernel. UMIP is an open source implementation of the MIPv6 protocol for the Linux operating system. To support PMIPv6, the Open Air Interface (OAI) PMIPv6 was developed extending UMIP functionality in order to support all necessary PMIPv6 messages and events. In both cases, the user space implements a daemon that takes care of the MIPv6 or PMIPv6 logic, such as tunnelling, binding signaling, security associations and IPv6 extensions, if needed. Some of these features are configured in each agent by means of configuration files that the daemon reads.

A.3 Performance evaluation

A.3.1 Experimental setup and design

This section describes the setup and design of the MIPv6 and PMIPv6 scenarios used in the experimental study presented in this article. Figure A.1 illustrates the modules of the mobility agents involved in both approaches from an architectural point of view. The bottom layer represents the hardware and the wireless infrastructure; the Operative System (OS) layer includes the kernel space and the module to provide the mobility mentioned previously. In these experimental configurations, Ubuntu 12.04 OS is running in the mobility agents. The Linux 3.8.2 kernel with the mobility support has been compiled and installed in the OS to provide the basis to execute the mobility management daemons.

As can be observed in Figure A.1, the two open source approaches are implemented in the Linux user space. MIPv6 approach implements only UMIP as well as the MIPv6 daemon to provide the mobility management in the MN and the HA. The HA also requires the `radvd` daemon to send the Router Advertisement in the Home Network of the MN. The HA is integrated as a router in the access network, so it must exchange the necessary IPv6 routing information with other routers in the network. This function is implemented with Quagga routing software suite. This suite provides the implementation of Routing Information Protocol (RIP) for IPv6 using `zebra` and `ripngd` daemons.

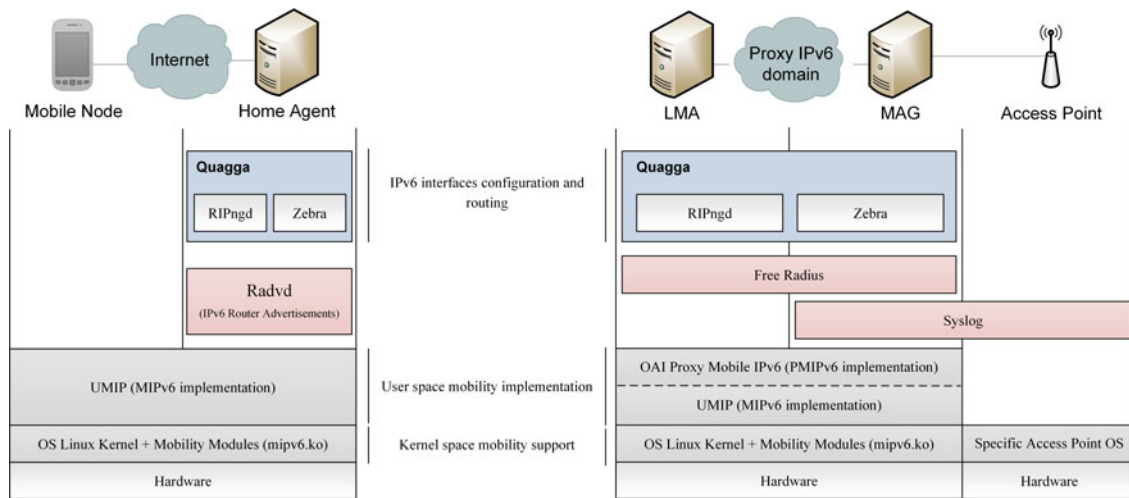


Figure A.1. Architecture of the mobility agents in MIPv6 and PMIPv6 approaches

For its part, PMIPv6 approach requires the OAI Proxy Mobile IPv6 implementation, in which the PMIPv6 functionality is included. Similarly to UMIP, OAI PMIPv6 provides a daemon that controls the protocol operation in the network entities (LMA and MAG). This daemon is implemented over UMIP and takes advantage of the similarities of both protocols. The LMA and MAG agents must also implement Free Radius server and client respectively to manage the IPv6 addresses and also implement the security mechanism to register the users in the access network.

Both LMA and MAG entities are located at the edges of the Proxy IPv6 domain and must implement the Quagga¹ routing software suite in order to exchange the routing information with other access network routers. Finally, The MAG and the access point must implement a Syslog server and client respectively to track the movement of the MN in the wireless network and to signal the movements to the LMA using the PMIPv6 Binding Update/ACK messages. Figure A.2 depicts the Mobile IPv6 scenario used for the tests made in the experimental study. The MIPv6 scenario consists of three Cisco 1921/K9 routers which support IPv6 and RIPng [111] that interconnect the HA, and the access points which provide the wireless access to the MN. These APs are Cisco Aironet 1130 AG series that support IEEE 802.11b/g specifications.

The PMIPv6 scenario (see Figure A.3) is similar to the previous MIPv6 one with the difference that in PMIPv6 two new agents (LMA and MAGs) must be configured with the OAI PMIPv6 implementation. LMA functionality is similar to HA whereas MAGs are introduced instead of the access routers that serve the MN. In PMIPv6 the MN does not need any additional configuration to its default IPv6 stack. One important point is the requirement to emulate the conditions of the access network in order to evaluate how the

¹Quagga Routing Suite - <http://www.nongnu.org/quagga/>

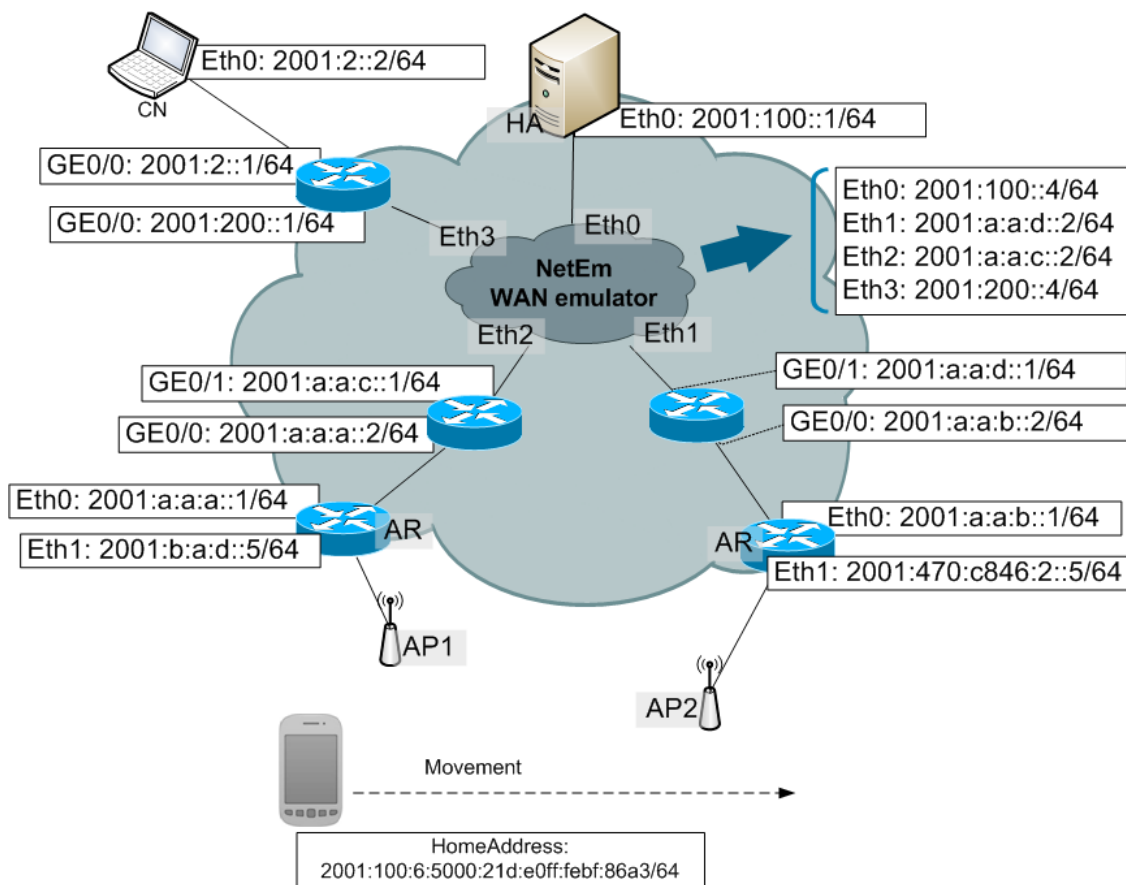


Figure A.2. Mobile IPv6 testbed topology

access network characteristics affect mobile communication. For this purpose, the Netem network emulator is used. Netem allows a single Linux PC, set up as a router, to emulate a wide variety of network conditions (e.g. latency, jitter, packet loss,...). Both scenarios are similar so that the experiments obtain comparable results. The MN's handover is performed between two WLAN cells. These cells have enough overlapping surface so there is no possibility of the MN being unable to communicate with either of them. Notice that cell overlap is a requirement for seamless handover. Each WLAN cell belongs to a different IPv6 subnet.

A.3.2 Quantitative tests

This section provides a description of the test performed, as well as the open source tools used and the measurements done. The goal of these quantitative experimentations is to evaluate the performance of both MIPv6 and PMIPv6 Linux based open source implementations in terms of handover delay and the behavior of multimedia traffic under

different circumstances such as packet delay or packet loss using the topologies observed in Figure A.2 and Figure A.3.

The handover latency (disruption time) is defined as the time elapsed between the last data frame being transmitted/received through the old interface and the first data frame being transmitted/received through the new interface. This time is a common parameter evaluated in a mobility scenario because the handover is one of the most critical processes. There are several comparative, analytical or simulation based studies of the behavior of handover latency in different mobility management protocols, however we have evaluated and compared the performance of both open source implementations using real infrastructure. The data given in this work allows to measure the gap between analytical or simulated data and real implementations. We have evaluated this parameter by modifying the routing advertisement interval from 0.5 to 4 seconds.

The measurement of the handover latency consisted in sending packets at a high Constant Bit Rate (CBR), so that we could measure the time between the last packet received by

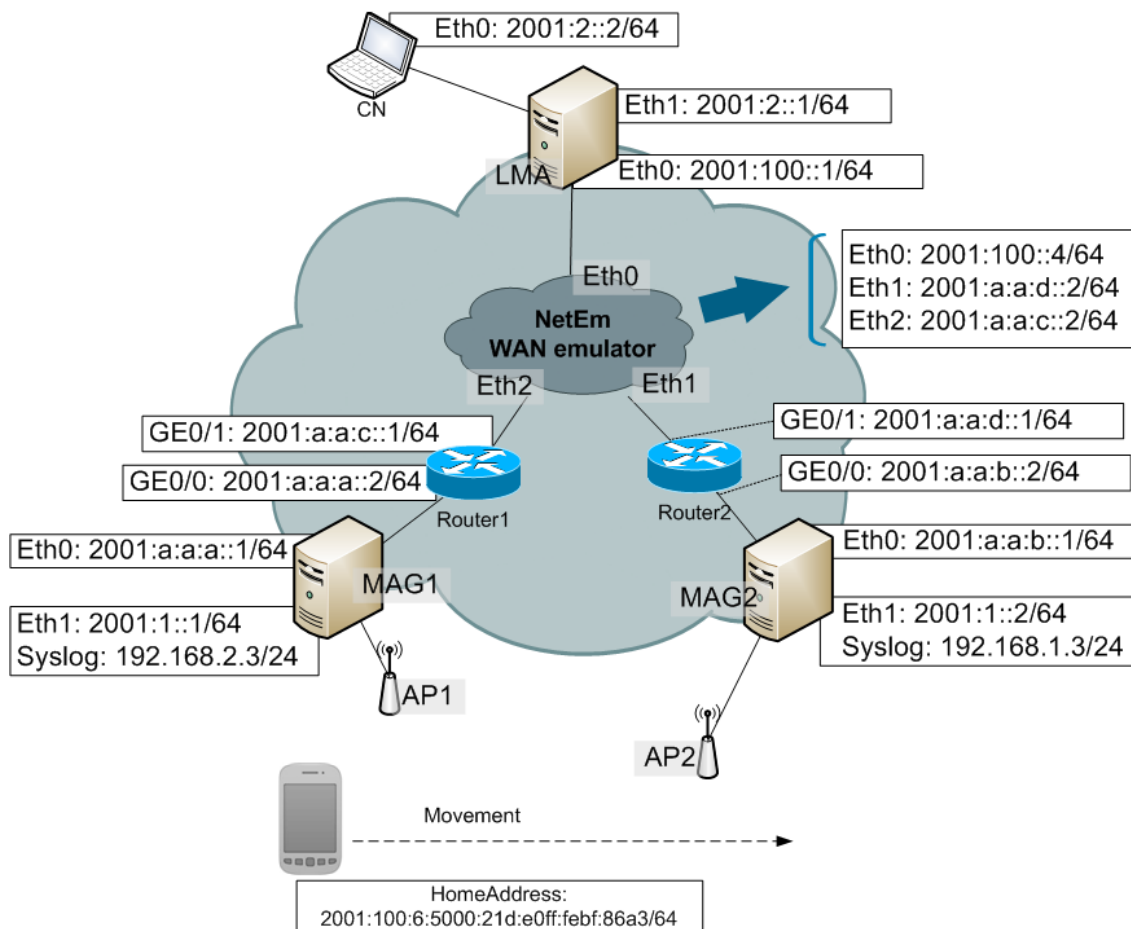


Figure A.3. Proxy Mobile IPv6 testbed topology

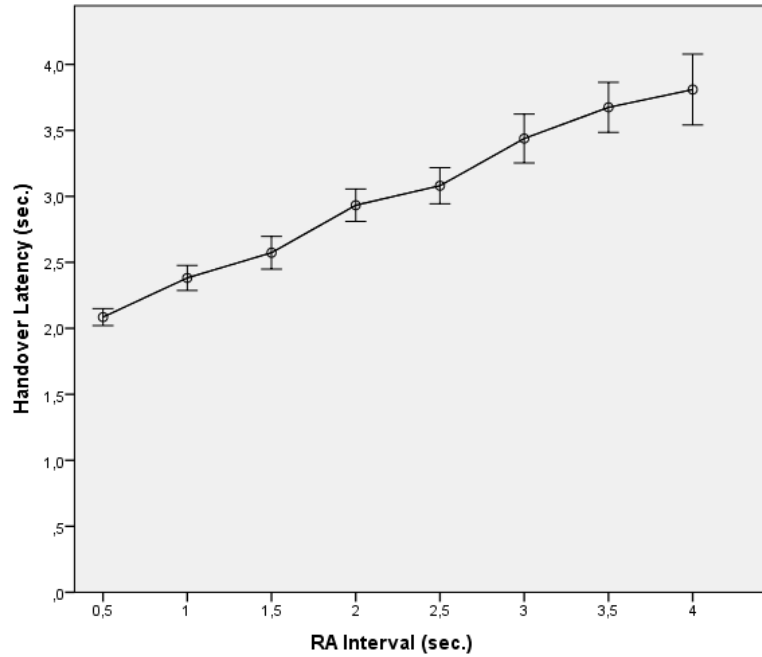


Figure A.4. Handover latency in MIPv6 experiments

the MN before the handover and the first packet received after the movement. The open source tool used to generate CBR traffic was Ostinato ⁱⁱ, whereas Wireshark ⁱⁱⁱ was used to analyze the time stamped CBR traffic received by the MN during its movement. The handover delay for each of the RA intervals in MIPv6 is shown in Figure A.4, while Table A.1 represents the numerical values of the handover delay parameter for the MIPv6 and PMIPv6 scenarios with a confidence interval of 95% for the different router advertisement values. In order to assure the confidence of the results, for each RA value 100 repetitions of the test have been made.

From the results shown in Figure A.4 we can quantify the dependency of the MIPv6 latency with the different Router Advertisement intervals, from 0.5 to 4 seconds. The minimum handover latency has been obtained with a RA interval of 0.5 and is 2.02 sec. The confidence interval for the mean at 95% limits is short as we can observe in Table I. As the RA interval increase, the confidence interval also increases due to the variability of the arrival to the MN of the unsolicited RA from the access router. As can be seen, PMIPv6 latency has not been included in Figure A.4. It is necessary to note that PMIPv6 is a network based approach and the home network is responsible for detecting that a new MN has been attached. In PMIPv6, the movement detection mechanism is not dependent on the RA messages. Table I demonstrates that the open source implementation of the

ⁱⁱOstinato is an open-source, cross-platform network packet/traffic generator and analyser - <http://ostinato.org>

ⁱⁱⁱWireshark is a network protocol analyser for Unix and Windows - <https://www.wireshark.org>

Table A.1. Values of the handover latency intervals with a confidence interval of 95%

RA interval (s)	Hand. interval in MIPv6(s)		Hand. interval in PMIPv6(s)	
	Min.	Max.	Min.	Max.
RA interval = 0.5	2.020	2.149	2.788	2.788
RA interval = 1	2.287	2.477	2.789	2.790
RA interval = 1.5	2.449	2.697	2.779	2.780
RA interval = 2	2.810	3.055	2.783	2.784
RA interval = 2.5	2.944	3.218	2.781	2.784
RA interval = 3	3.252	3.624	2.794	2.797
RA interval = 3.5	3.485	3.864	2.807	2.811
RA interval = 4	3.540	4.078	2.790	2.790

protocol follows this behavior. The result obtained of the mean handover delay value for the 800 repetitions is, 2.788 seconds. The 95% confidence interval for this mean ranges from 2.788 to 2.790.

As we previously mentioned, we have also evaluated the effect caused by IP mobility in multimedia communications. In this case, two types of multimedia transmissions have been considered and compared under different conditions. On the one hand, UDP real time communications based on RTP (Real Time Protocol) and, on the other hand, TCP multimedia streaming. Some network situations can affect the UDP multimedia stream such as network congestion, packet loss during handover and RTP packets arriving out of the playout time. In the TCP streaming case, the multimedia server sends the stream flow content to the receiver that is buffered in the client side. This mechanism avoids packet loss and minimizes the delay and jitter effects in the stream flow.

For each type of multimedia traffic, several parameters have been modified in the MIPv6 access network or in the PMIPv6 domain in order to evaluate its effect on the overall performance of the communication. These parameters are delay and packet loss, which were introduced by Netem emulator.

The experiments consist of the same movement of a terminal as in our testbed (causing a handover) during a real time or a streaming multimedia transmission under the different aforementioned conditions. The multimedia transmission is the Open Source Film Big Buck Bunny (duration, 120 s., bitrate approximately: 1626 kbps, resolution: 640x360). A handover is performed in each test at second 53 approximately. In order to measure the transmission and to evaluate the received video quality by the MN, the Peak Signal to Noise

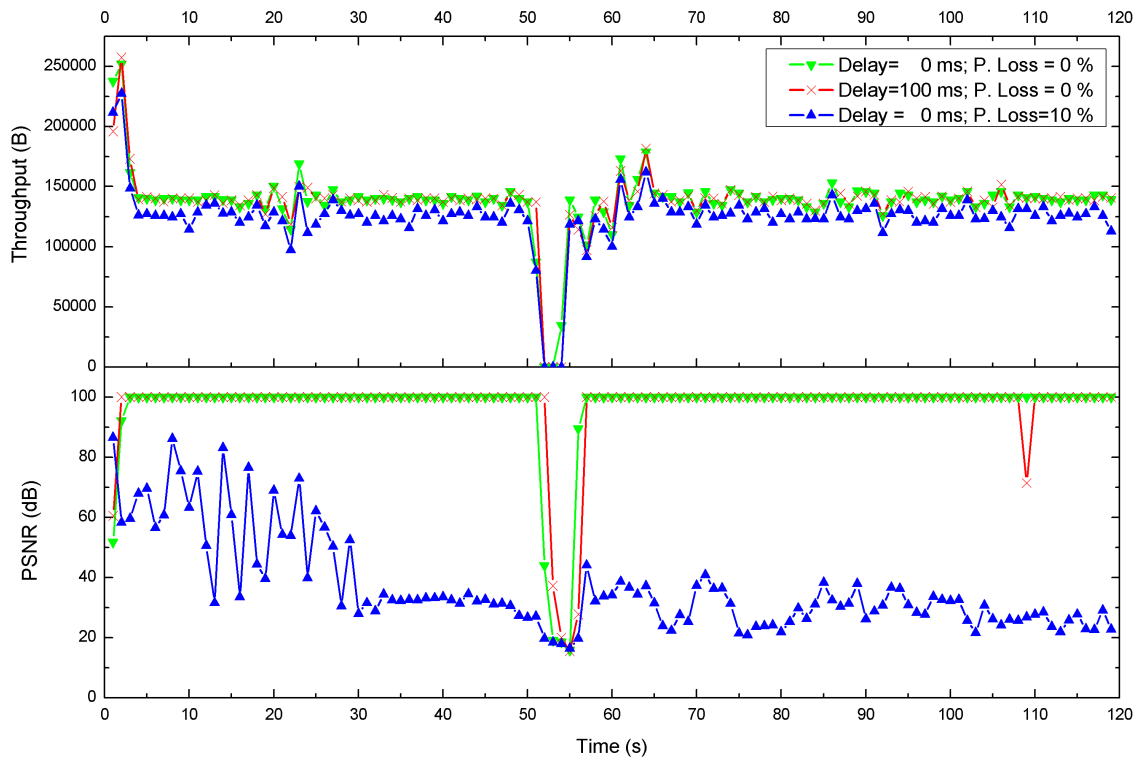


Figure A.5. Throughput and PSNR in real time experiments

Ratio (PSNR) indicator has been used. PSNR is generally considered to be a reference benchmark for developing objective perceptual video quality assessment models [112]. In this case, the reference model required by PSNR is the video transmitted in the testbed without the influence of the mobility management protocol and without any access network parameter modification. This reference is compared with the video received by the MN during the test using one of the evaluated mobility management protocols and changing the access network parameters. The MSU Video Quality Measurement Tool^{iv} was used for the PSNR evaluation.

Figure A.5, Figure A.6 and Figure A.7 shows a summary of the multimedia experimentation results. Figure A.5, and Figure A.6 show the behavior of both real time and streaming multimedia communications respectively in MIPv6. In this case, just three representative tests are present in the graph. The first one is a communication without any access network parameter modification, the access network in the second one introduces a delay of 100 ms, and finally, a test with a packet loss of 10 %. The result of the PSNR evaluation of all the experiments is presented in Figure A.7.

^{iv}MSU Video Quality Measurement Tool is a program for objective video quality assessment. It provides functionality for both full-reference (two videos are examined) and single-reference (one video is analyzed) comparisons- http://compression.ru/video/quality_measure/video_measurement_tool_en.html

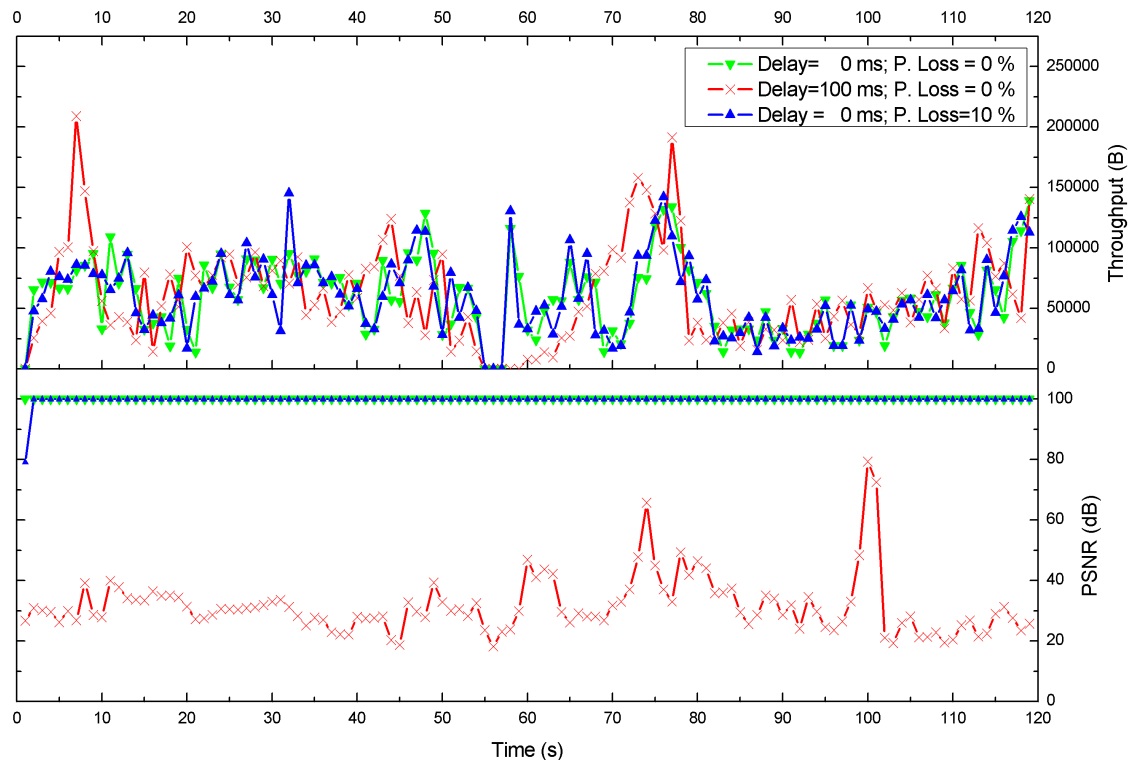


Figure A.6. Throughput and PSNR in TCP streaming experiments

Figure A.5 presents packet loss and delay influence in a real time communication transmitted in the MIPv6 testbed. The results show that the PSNR is extremely sensitive to packet loss in RTP. When this parameter is increased, many RTP packets are lost and the video frames must be predicted using previous frames. This prediction reduces the PSNR value of the experiment. In this Figure, a high packet loss rate can also be observed and the throughput of the communication decreases compared with the delay test.

In Figure A.6 we can observe a different behavior by the PSNR. In this case, three TCP streaming communications are compared in the same conditions as stated before. The streaming application provides a buffer that avoids packet loss and reduces the delay and the jitter of the packets. As can be observed in Figure A.6, when the delay is high the application must wait to recompose the video transmission and the PSNR is penalized due to long wait times. In contrast, when the packet loss is high, the video frames are retransmitted by the transmitter and the PSNR value is higher compared with the real time transmission. In this case there are no wait times and the PSNR is not penalized. This implies that PSNR parameter in streaming transmission decrease when the delay of the access network increases. As could be observed, although the handover is produced, the PSNR is not affected because of the buffer. Only when the delay is increased, the PSNR decreases.

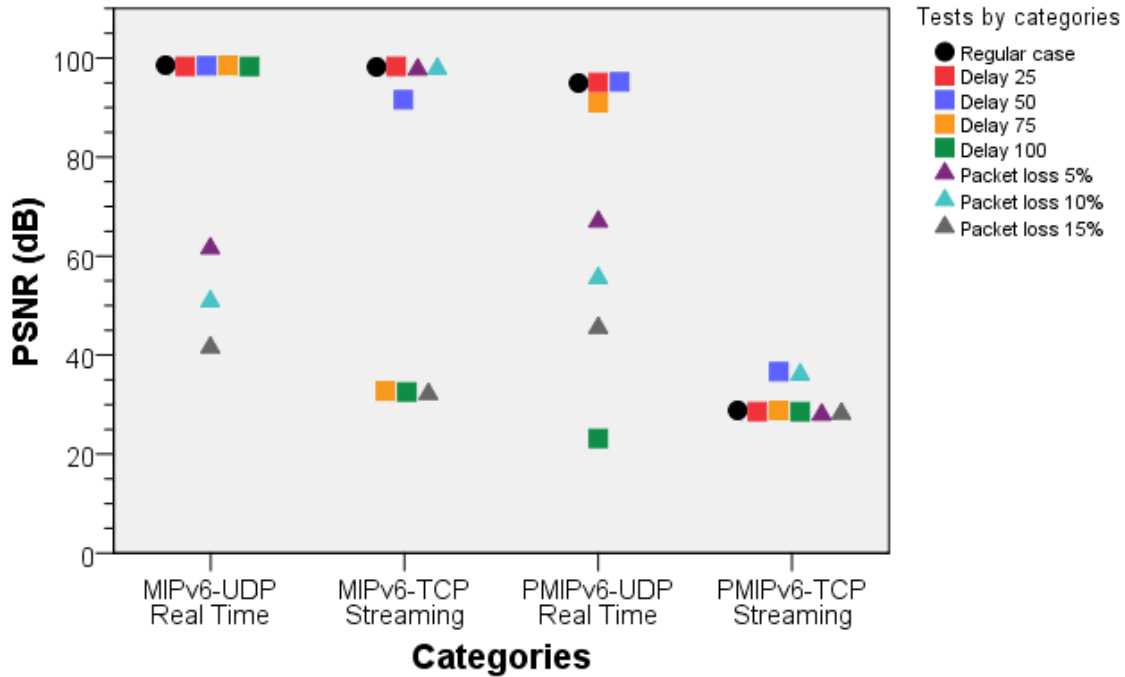


Figure A.7. Summarize of the obtained PSNR results in the experiments conducted

Figure A.7 shows the different experiments made in both MIPv6 and PMIPv6 testbed using UDP real time and TCP streaming communications. In each of these four categories, eight tests have been made and each one represents a PSNR value in the plot (5 repetitions have been made for each PSNR value). Regular case refers to tests made with the configuration of Netem at Delay = 0 and without packet loss. The behavior of MIPv6 corresponds to the one explained previously in Figure A.7 and Figure A.6 whereas in PMIPv6 we can observe some differences. Real time transmission in the PMIPv6 testbed demonstrates the same behavior as in MIPv6 when the delay of the access network increases. When delay exceeds 75 ms, the PSNR decreases because the packets arrive too late to the MN and the video frames are discarded by the player as the time limit to be reproduced has been exceeded. In our experiments, when the handover occurs, the serving MAG of the visited network introduces long delays to forward packets to the MN. These latencies penalize the communication and the PSNR values. The same effect is produced in PMIPv6 streaming, where the PSNR achieves low values because the reproduction pauses until it has enough information to continue the reproduction due to buffer starvation. Despite the PSNR values, that reflect the comparison between the original video and the received one, the videos seems to be reproduced at good quality. In the next section, a user experience valuation is presented in order to compare qualitative versus quantitative results.

A.3.3 Qualitative tests

In the previous section, some quantitative tests and experimental results have been presented. The quantitative tests offer a measurement of various well known parameters which affect communication. In contrast, the qualitative tests deal with user expectation, satisfaction and overall experience. A typical user related measure is the Mean Opinion Score (MOS) [113], which can be determined from subjective ratings by real users. With the purpose of reflecting how users perceive multimedia content in mobility scenarios managed by open source implementations of MIPv6 and PMIPv6, in this section we show the results obtained from user valuations. In total, 30 real users were asked to evaluate the perceived quality according to the five point MOS scale (MOS scores goes from 1 to 5, 5 being the best). Each test was conducted in a similar way as the one presented in the previous section. In this case, Netem is configured in the access network in order to introduce different delays (0, 25, 50 and 100 ms). Figure A.8 shows the user valuation of both real time and streaming transmissions under different delay conditions in the MIPv6 and PMIPv6 experimental testbed.

Figure A.8a shows the results obtained to transmit UDP real time content. As can be observed, the MOS rate decreases when the delay of the access network increases. With the introduction of delay, some packets arrive to the MN out of the reproduction timestamp and they are discarded producing image pixelation and a gap between frames. As can be observed in the figure, the users penalize these effects.

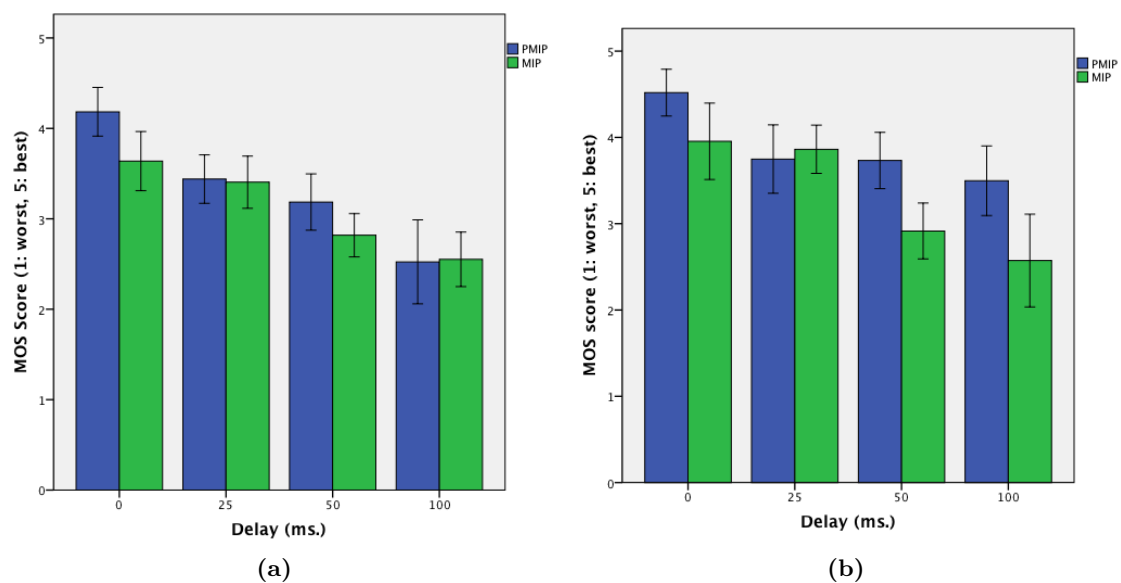


Figure A.8. Mean Opinion Score of (a) real time and (b) TCP streaming transmissions with a confidence of 95%

The TCP streaming transmission increases the overall MOS score in both management protocols because all packets are stored in a buffer avoiding image pixelation, and a gap in the video reproduction produced by packet loss or packet discard. The inconvenience introduced by TCP streaming is the long wait time produced by the buffer starvation. This effect can be observed when the delay of the access network increases. The handover latency is increased and this means that the wait time until the buffer is filled with enough packets to continue the reproduction is also increased. This effect is shown in Figure A.8b.

A.4 Final remarks

Testbed are, by far, the most realistic method for evaluating performance, as they practically use the protocol implementations and the hardware that is the same, or very similar, to the one used for the production networks. However, large testbed are expensive to build and manage and for very large (or highly mobile) networks, practically impossible to implement. Furthermore, the degrees of freedom in a testbed is significantly reduced in comparison with a mathematical analysis or network simulation. Assuming these limitations, we have conducted a study related to the performance of real IPv6 mobility management implementations in the IPv6 Linux stack, evaluating the current state of these protocols by means of an experimental testbed.

We have focused on the most representative IETF solutions: MIPv6 and PMIPv6. Qualitative and quantitative results have been obtained from both real time and streaming multimedia applications under different network conditions. Quantitative results have shown the handover latency produced by both open source implementations. We quantify the dependency of the MIPv6 handover delay in the Routing Advertisement interval, whereas the PMIPv6 implementation is not dependent of RA messages, as expected. Moreover, the PSNR video quality indicator has been used to evaluate the multimedia transmission. It should be noted that streaming traffic in the PMIPv6 testbed give low PSNR values due to the short additional delays produced by the MAG which penalizes the PSNR value because some wait times are introduced due to buffer starvation.

Finally, in order to compare these quantitative results of the multimedia transmissions with a subjective valuation, user perceptions are evaluated. In this case, streaming multimedia provides a better valuation than real time due to penalization of the handover disruption time. A better user experience is achieved when handover latency is reduced.

APPENDIX B: IP MOBILITY MANAGEMENT SIMULATOR

This appendix concerns the design of a mobility management simulator in Matlab, used to obtain the simulation results presented in this thesis. The implementation of this simulator greatly simplifies the generation of mobility results, and the evaluation of different mobility protocols. The modular design enables it to be extended with additional functionality easily, (e.g. new mobility models or new mobility protocols).

Contents

B.1 Introduction	173
B.2 Overview of the simulator	173
B.3 Simulation environment and scenario	177
B.4 Metrics for Mobility Management Evaluation	179
B.5 Final Remarks	180

B.1 Introduction

With the aim to obtain additional results that allow a better understanding of the behavior of the proposed mechanisms, a simulation study has been carried out to compare our proposed solutions (DM3 and Hybrid DMM) with the existing mobility management mechanisms. This section describes briefly some details of the simulator developed, as well as the general simulation scenario used to model the proposed scheme, the simulation configuration parameters and the simulation environment.

B.2 Overview of the simulator

The Mobility Management Simulator is a Matlab-based tool used to simulate and analyze the performance of IPv6 mobility management protocols. The main idea behind the design of the simulator is to develop different modules involved in the mobility process and to ease the further development of additional mobility functionalities.

The simulator relies on a core controller that performs the simulation of the events. These events depend on the selected topology, the traffic model and the mobility model that defines the movement of the mobile nodes in the scenario. This design allows an easy way to incorporate different topologies, generate the traffic with new traffic models or simulate the mobility with any other model developed in the specific input module. Figure B.1 shows the basic components of the simulator.

The main role of the core controller is as follows. It receives the data structures of the input modules and is responsible for generating and scheduling the discrete events provided by

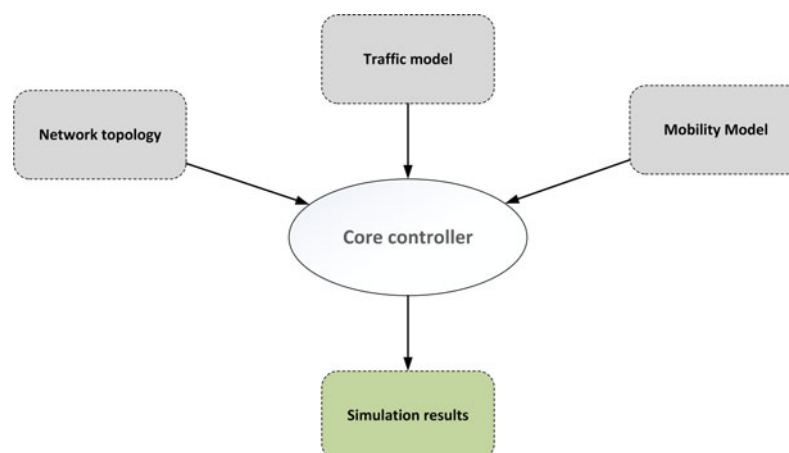


Figure B.1. Overview of the simulator

these modules. The events can be mainly differentiated as traffic sessions or as movements of the nodes. Therefore, the core controller needs to organize an event list that splits into simulation period intervals of different durations, as is shown in Figure B.2

Therefore, the simulation is performed in the following way. At the beginning of a simulation run, the input modules are initialized. The three modules considered in the simulation (network topology, traffic model and mobility model) are described next.

- Network topology: This module is responsible to create the access network topology that serves the mobile nodes. It is represented as a matrix $N \times N$ where N is the number of nodes in the access network. In this matrix, each position (i, j) represents the path between the node i and node j , $path(i, j)$, and takes the following values:

$$path(i, j) = \begin{cases} 1 & \text{if there exists a path between node } i \text{ and node } j \\ 0 & \text{else} \end{cases}$$

- Traffic model: The traffic model module generates the demands arrival. These connections arrive independently for each MN in the scenario following random models. The simulations performed in this thesis follows a model in which both the

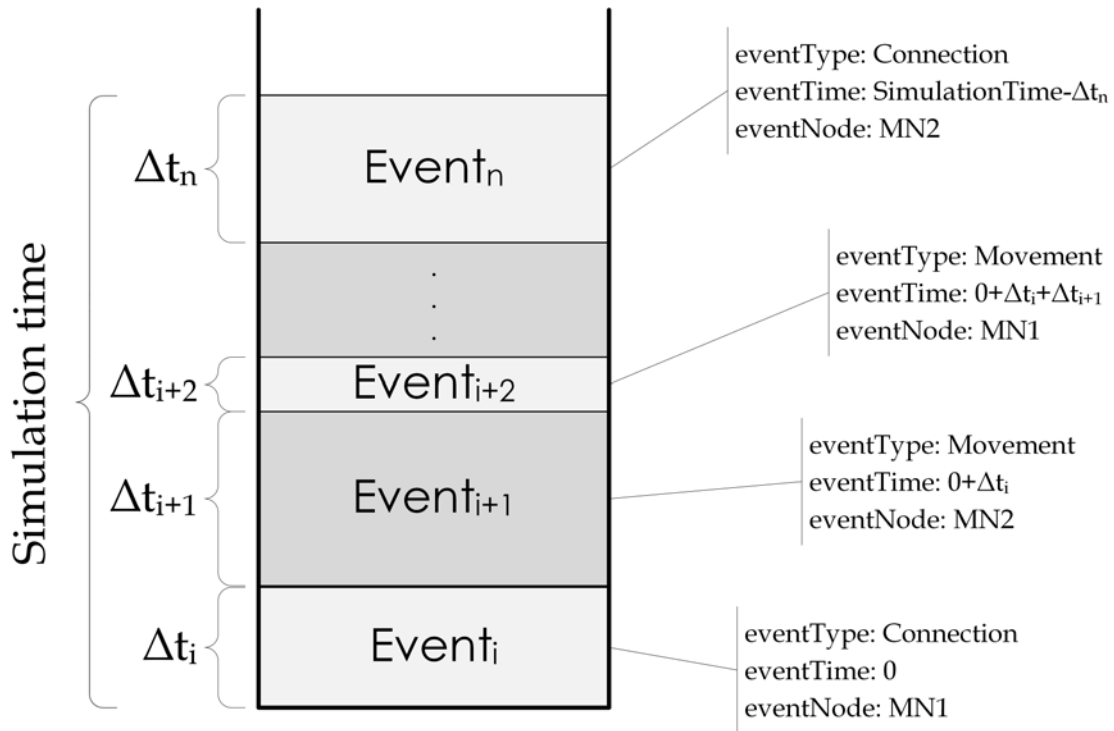


Figure B.2. Events simulation

session arrival and the inter-arrival time follows an exponential distribution with a given rate λ_d and λ_i respectively.

- **Mobility model:** In this module, the mobility model is considered. To generate the mobility of all the nodes in the scenario, each entity moves independently from each other and their movements are predicted independently from the rest of the nodes. Also the number of mobile nodes in the simulation area is not considered in the algorithm that predicts a single entity's movement. In this thesis two mobility models are considered. One stochastic model such as the Random Waypoint Model, and other real-life pattern, achieved by gathering traces from real moving users. Such traces can also be used to verify the mobility approximation of synthetic mobility models against real user behavior.

The mobility model determines the initial position of all participating nodes in the simulation area. In order to represent a cellular network, we consider a hexagonal grid. In future developments of the simulator, other tessellations can be considered in the Mobility model module, such as Poisson-Voronoi [114, 115] or other emerging approaches proposed for future wireless networks. An example of the cellular grid scenario created by the mobility model module is shown in Figure B.3. Here, five MNs (identified by a red asterisk) move in a grid area following a specific mobility model, in this case a Random Waypoint Mobility model.

Once each module returns the data, the core controller performs the simulation following the main loop shown in Algorithm 3.

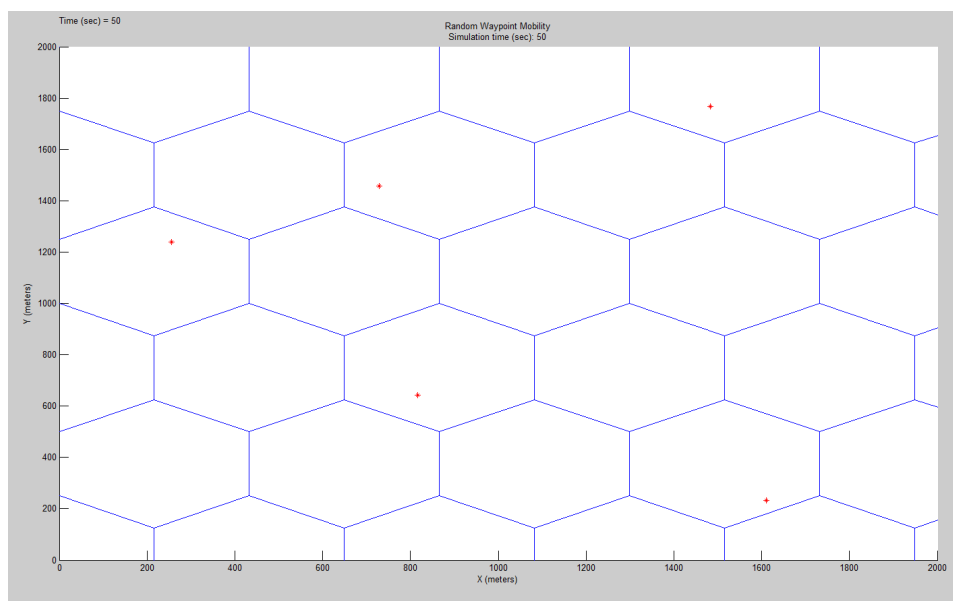


Figure B.3. Hexagonal grid cellular model

Algorithm 3: Pseudocode of the main loop in the core controller

Input : matrixTopology $mTopology$;

trafficModel $mobiSim.node(i).sessionArrivals$, $\forall i \in$ the set of MNs;

mobilityModel $mobiSim.node(i).movementData$, $\forall i \in$ the set of MNs;

Output: simResults

```

1 listEvents=sortEvents(mobiSim);
2 while  $i < size(listEvents)$  do
3   event=processEvent(i);
4   foreach  $j \in activeSessions$  do
5     if  $activeSessions(j)$  has finished then
6       | releaseResources(j);
7       | endSession(j);
8     else
9       | saveCostEvent(j);
10    end
11  end
12  if  $event.type == movement$  then
13    | updateLocation(event.MN);
14  else
15    session=newSession(event.MN);
16    pathCalculation(session,mTopology);
17    if  $pathCalculation(session,mTopology)$  then
18      | addSession(session, activeSessions);
19    else
20      | blockSession(session);
21    end
22  end
23 end
24 return  $simResults$ ;
```

B.3 Simulation environment and scenario

This subsection presents the generic simulation scenario used to evaluate the performance of the IP mobility management protocols, including our DM3 and Hybrid DMM schemes. The mobility management simulator developed has been extended to model not only DM3 and Hybrid DMM, but other standardized protocols in order to compare their performance. Thus, the current version of the simulator implements the centralized solutions MIPv6 and PMIPv6 and the distributed host-based (HB-DMM) and network-based (NB-DMM). All these protocols were described in Chapter 2.

Next, the mobility patterns, as well as the traffic model and the network topology used in the simulations are explained.

To investigate the performance of wireless networks through simulation, it is necessary to consider the movement of mobile nodes within the simulated environment. The mobility of these nodes is a key attribute in the behavior of the mobility protocols and the performance of these protocols needs to be studied in the presence of mobility. Actually, the results of network simulations that include mobility can vary significantly when the mobility patterns of moving nodes are changed. There are different possibilities to incorporate mobility in the simulation environment, a first option, it is to gather real movement data traces, but often, this possibility is difficult due to the necessity of obtaining a sufficient number of traces for simulations. To overcome this problem, synthetic mobility models have been developed that are generating simplified virtual movement data for a number of entities. There are several mobility models with different properties.

The random waypoint mobility model has been widely used in mobile network simulations. This mobility model is simple and straightforward stochastic model. In RWP [103], a mobile node moves on a finite continuous plane from its current position to a new location by randomly choosing its destination coordinates, its speed of movement (from [minSpeed; maxSpeed]), and the amount of time that it will pause when it reaches the destination. On reaching the destination, the node pauses for some time distributed according to a random variable (from [minPause; maxPause]) and the process repeats itself. Once the pause time expires, the node chooses a new destination, speed, and pause time. The movement of a node from the starting position (waypoint) to its next destination (waypoint) is defined as one movement epoch, movement period, or transition time (from [minWalkInterval; maxWalkInterval]). The distance traveled between the movements of a node from the starting waypoint to its next waypoint is defined as transition length. The destination points (waypoints) are uniformly and randomly chosen in the selected system area.

Figure B.4 shows an example of a travelling pattern of a mobile node using the Random

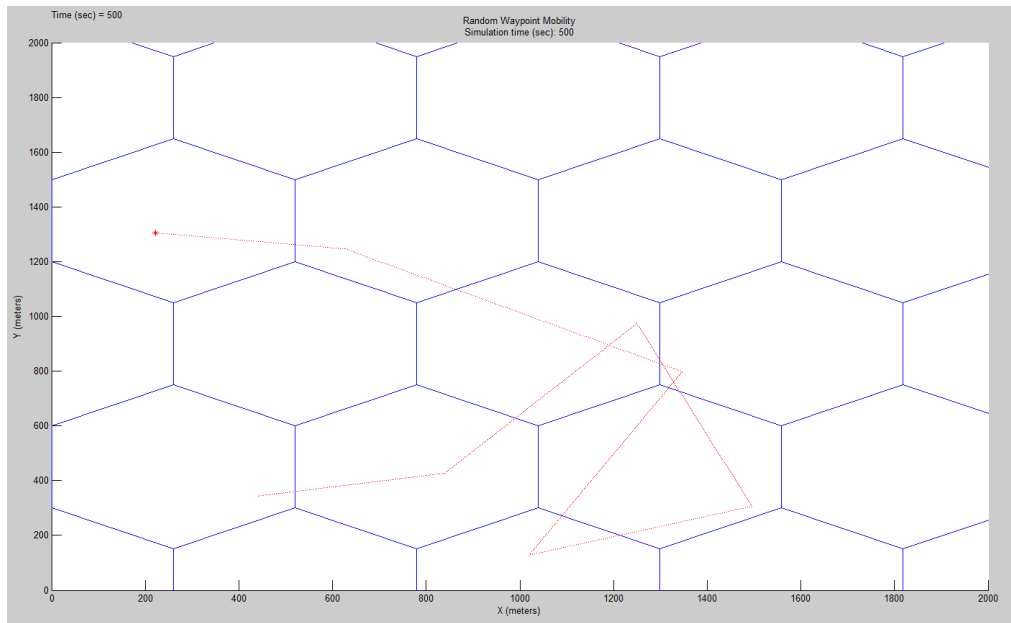


Figure B.4. RWP mobility

Waypoint mobility model implemented in our IP mobility management simulator. In this case, the node starts at a randomly position; the speed of the MN in the figure is uniformly chosen from an interval of between 5 and 10 m/s. The pause time between subsequent trips is uniformly distributed between 0 and 10 seconds to simulate a short stop at a destination point. Moreover, the time walk interval is [50 100] seconds. Finally, the simulation time is 500 seconds. In order to observe the track of the MN, the mobility pattern is highlighted with a red line.

Apart from the Random Waypoint Mobility, in order to drive the evaluation in a more realistic scenario, we also run the simulations with real-world mobility track logs obtained from users carrying GPS receivers. The sample settings where traces are obtained are two university campuses (one in Asia and one in the US), one metropolitan area (New York City), one State fair and one theme park (Disney World). The participants walk most of the time and may also occasionally travel by bus, trolley, car, or subway. These settings are selected because they are conducive to collecting GPS readings [104].

One of the main challenges with the real-world mobility traces is the need to acquire a sufficient number of samples to be used in the simulations. If the number of samples is too small, only few MNs can be simulated. One approach to overcome this challenge is to aggregate traces from several similar scenarios. In our simulations, we aggregate the traces from the scenarios mentioned above. Of course, to aggregate traces, it is important that the parameters in the traces fit together.

The data traces used in our simulator is retrieved from the CRAWDDAD repository [116],

%Time (seconds)	X-coordinate (meters)	Y-coordinate (meters)
0.0000000000000000e+000	5.9016633952322616e+002	1.5767217702011691e+004
3.0000000000000000e+001	5.8687563956937754e+002	1.5766661116464113e+004
6.0000000000000000e+001	5.9050385716008293e+002	1.5770779848858925e+004
9.0000000000000000e+001	5.9176950898610448e+002	1.5770223263310951e+004
1.2000000000000000e+002	5.9176950898610448e+002	1.5769666677879379e+004
1.5000000000000000e+002	5.9362580952548649e+002	1.5771781703115668e+004
1.8000000000000000e+002	5.9421643880972306e+002	1.5771559068672719e+004
2.1000000000000000e+002	5.9193825443366416e+002	1.5771781703115668e+004
2.4000000000000000e+002	5.9396330785355167e+002	1.5767662970338804e+004
2.7000000000000000e+002	5.9489146965892473e+002	1.5766438482472464e+004
3.0000000000000000e+002	6.1100749167381844e+002	1.5769555361013237e+004
3.3000000000000000e+002	5.7362843683861934e+002	1.5774787264895516e+004
3.6000000000000000e+002	6.0037597160414805e+002	1.5776679655229096e+004
3.9000000000000000e+002	5.9219139684256788e+002	1.5775789118873816e+004
4.2000000000000000e+002	5.8358492635418247e+002	1.5776568338607811e+004
4.5000000000000000e+002	5.9311954459971889e+002	1.5775900436126660e+004
4.8000000000000000e+002	6.1117624593657217e+002	1.5763210286066833e+004
5.1000000000000000e+002	6.0991059858953486e+002	1.5761763164289032e+004

Figure B.5. Human mobility GPS trace data format (sample)

although several traces from different measurements are available in other repositories apart from CRAWDAD^v such as UNC/FORTH^{vi}, and MobiLib^{vii}. A complete summary of the available movement traces can be found in [117]. The mobility trace for a specified mobile user is extracted from the files in which the GPS coordinates are given. The GPS data format shown in Figure B.5 and the three data field given are the (*i*) time in seconds; (*ii*) the position of the *x* coordinate in meters and; (*iii*) the position of the *y* coordinate in meters.

B.4 Metrics for Mobility Management Evaluation

As we have described in the previous section, mobility models are used to generate movements of mobile nodes in a cellular network. Position, speed and moving direction of nodes are defined by the mobility model during the whole simulation. However, in the scope of the IP mobility management protocols, there are several key metrics that show the behavior of the protocols and have a major impact on the performance of the network. The most relevant metrics, resulting from the simulation of the mobility management protocols are described next:

- Signaling cost: This metric shows the control data exchanged among the entities to perform the mobility process. The simulator calculates this value for each MN and, therefore for the overall simulation.

^vCRAWDAD <http://crawdad.cs.dartmouth.edu/>

^{vi}UNC/FORTH <http://netserver.ics.forth.gr/datatraces>

^{vii}MobiLib <http://www.cise.ufl.edu/~helmy/MobiLib.htm>

- Packet delivery cost: This is the most relevant metric for the data plane of the mobility protocols. It represents the cost of forwarding a data packet from the source to the destination. The packet delivery cost is mainly affected by the routing path and the tunnelling overhead.
- Tunnelling cost: This metric is similar to the data packet delivery cost but it is focused on representing the cost of adding tunnelling overheads.

B.5 Final Remarks

In this section a mobility management protocol has been briefly described. The fully functional implementation of this simulator in Matlab greatly simplifies the generation of numerical results to study the performance of different IP mobility protocols. As a result of its modular design, the results can be obtained with different mobility models, traffic models and topologies.

The current implementation of our simulator can be enhanced with additional features to further improve the representativeness of IP mobility management. From one side, the development of new mobility models and protocols should increase the functionality of the software. In addition, the development of a friendly user interface could make this tool adequate for use in some networking courses.

REFERENCES

- [1] Michelle X. Gong, Brian Hart, and Shiwen Mao. Advanced Wireless LAN Technologies: IEEE 802.11AC and Beyond. *SIGMOBILE Mob. Comput. Commun. Rev.*, 18(4):48–52, January 2015.
- [2] E. Halepovic, C. Williamson, and M. Ghaderi. Wireless data traffic: a decade of change. *Network, IEEE*, 23(2):20–26, March 2009.
- [3] Cisco Systems Inc. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014–2019 - Cisco, February 2015.
- [4] Nokia Siemens Networks. Signaling is growing 50% faster than data traffic. White Paper, 2012.
- [5] David Cortés-Polo, José-Luis González-Sánchez, Francisco-Javier Rodríguez-Pérez, and Javier Carmona-Murillo. Mobility management in packet transport networks for network convergence. *Transactions on Emerging Telecommunications Technologies*, 26(5):749–759, 2015.
- [6] A. Aijaz, H. Aghvami, and M. Amani. A survey on mobile data offloading: technical and business perspectives. *Wireless Communications, IEEE*, 20(2):104–112, April 2013.
- [7] F. Giust, L. Cominardi, and C. J. Bernardos. Distributed mobility management for future 5G networks: overview and analysis of existing approaches. *Communications Magazine, IEEE*, 53(1):142–149, 2015.
- [8] B.A.A. Nunes, M. Mendonca, Xuan-Nam Nguyen, K. Obraczka, and T. Turletti. A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks. *Communications Surveys Tutorials, IEEE*, 16(3):1617–1634, Third 2014.
- [9] Fei Hu, Qi Hao, and Ke Bao. A Survey on Software-Defined Network and Open-Flow: From Concept to Implementation. *Communications Surveys Tutorials, IEEE*, 16(4):2181–2206, Fourthquarter 2014.
- [10] R. Trivisonno, R. Guerzoni, I. Vaishnavi, and D. Soldani. SDN-based 5G mobile networks: architecture, functions, procedures and backward compatibility. *Transactions on Emerging Telecommunications Technologies*, 26(1):82–92, 2015.
- [11] Woon Hau Chin, Zhong Fan, and R. Haines. Emerging technologies and research challenges for 5G wireless networks. *Wireless Communications, IEEE*, 21(2):106–112, April 2014.
- [12] Internet Engineering Task Force (IETF), <http://www.ietf.org/>.

- [13] Third Generation Partnership Project (3GPP), <http://www.3gpp.org/>.
- [14] WiMAX forum, <http://www.wimaxforum.org/>.
- [15] Kyunghan Lee, Joohyun Lee, Yung Yi, Injong Rhee, and Song Chong. Mobile Data Offloading: How Much Can WiFi Deliver? *Networking, IEEE/ACM Transactions on*, 21(2):536–550, 2013.
- [16] J. C. Zuniga, C. J. Bernardos, A. de la Oliva, T. Melia, R. Costa, and A. Reznik. Distributed mobility management: a standards landscape. *Communications Magazine, IEEE*, 51(3):80–87, 2013.
- [17] IETF Distributed Mobility Management (DMM) Working Group, March 2015.
- [18] Mohammad Reza HeidariNezhad, Zuriati Ahmad Zukarnain, Nur Izura Udzir, and Mohamed Othman. Mobility support across hybrid IP-based wireless environment: review of concepts, solutions, and related issues. *Annals of Telecommunications*, 64(9-10):677–691, 2009.
- [19] D. Saha, A. Mukherjee, I.S. Misra, and M. Chakraborty. Mobility support in IP: a survey of related protocols. *Network, IEEE*, 18(6):34–40, Nov 2004.
- [20] F. Richard Yu, Vincent W. S. Wong, Joo-Han Song, Victor C. M. Leung, and Henry C. B. Chan. Next generation mobility management: an introduction. *Wireless Communications and Mobile Computing*, 11(4):446–458, 2011.
- [21] Ibrahim Al-Surmi, Mohamed Othman, and Borhanuddin Mohd Ali. Mobility management for IP-based next generation mobile networks: Review, challenge and perspective. *Journal of Network and Computer Applications*, 35(1):295–315, 1 2012.
- [22] Mariem Zekri, Badii Jouaber, and Djamel Zeghlache. A review on mobility management and vertical handover solutions over heterogeneous wireless networks. *Computer Communications*, 35(17):2055–2068, 10/1 2012.
- [23] Z. Zhu, R. Wakikawa, and L. Zhang. A Survey of Mobility Support in the Internet. RFC 6301 (Informational), July 2011.
- [24] Kong Ki-Sik, Lee Wonjun, Han Youn-Hee, Shin Myung-Ki, and You HeungRyeol. Mobility management for all-ip mobile networks: mobile ipv6 vs. proxy mobile ipv6 [architectures and protocols for mobility management in all-ip mobile networks]. *Wireless Communications, IEEE*, 15(2):36–45, 2008.
- [25] I.F. Akyildiz, Jiang Xie, and S. Mohanty. A survey of mobility management in next-generation all-IP-based wireless systems. *Wireless Communications, IEEE*, 11(4):16–28, Aug 2004.

- [26] Marjory S. Blumenthal and David D. Clark. Rethinking the Design of the Internet: The End-to-end Arguments vs. The Brave New World. *ACM Trans. Internet Technol.*, 1(1):70–109, August 2001.
- [27] C. Perkins, D. Johnson, and J. Arkko. Mobility Support in IPv6. RFC 6275, July 2011.
- [28] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil. Proxy Mobile IPv6. RFC 5213, August 2008.
- [29] S. Deering and R. Hinden. Internet Protocol, Version 6 (IPv6) Specification. RFC 2460 (Draft Standard), December 1998. Updated by RFCs 5095, 5722, 5871, 6437, 6564.
- [30] J. Postel. Internet Protocol. RFC 791 (INTERNET STANDARD), September 1981. Updated by RFCs 1349, 2474.
- [31] T. Narten, E. Nordmark, W. Simpson, and H. Soliman. Neighbor Discovery for IP version 6 (IPv6). RFC 4861 (Draft Standard), September 2007. Updated by RFC 5942.
- [32] S. Thomson, T. Narten, and T. Jinmei. IPv6 Stateless Address Autoconfiguration. RFC 4862 (Draft Standard), September 2007.
- [33] Rajeev S. Koodli and Charles E. Perkins. *Mobile IPv6 Protocol*, pages 59–75. MOBILE Inter-Networking with IPv6. John Wiley & Sons, Inc., 2007.
- [34] Linoh A. Magagula, H. Anthony Chan, and Olabisi E. Falowo. Handover approaches for seamless mobility management in next generation wireless networks. *Wireless Communications and Mobile Computing*, 12(16):1414–1428, 2012.
- [35] Antonio De La Oliva, Albert Banchs, Ignacio Soto, Telemaco Melia, and Albert Vidal. An overview of IEEE 802.21: media-independent handover services. *Wireless Communications, IEEE*, 15(4):96–103, 2008.
- [36] G. Lampropoulos, A.K. Salkintzis, and N. Passas. Media-independent handover for seamless service provision in heterogeneous networks. *Communications Magazine, IEEE*, 46(1):64–71, January 2008.
- [37] Behrouz Shahgholi Ghahfarokhi and Naser Movahhedinia. A survey on applications of {IEEE} 802.21 Media Independent Handover framework in next generation wireless networks. *Computer Communications*, 36(10-11):1101 – 1119, 2013.
- [38] C. Perkins. IP Mobility Support for IPv4. RFC 3344 (Proposed Standard), August 2002. Obsoleted by RFC 5944, updated by RFC 4721.

- [39] G. Lampropoulos, N. Passas, L. Merakos, and A. Kaloxylos. Handover management architectures in integrated WLAN/cellular networks. *Communications Surveys Tutorials, IEEE*, 7(4):30–44, Fourth 2005.
- [40] Qing Li, Tatuya Jinmei, and Keiichi Shima. *Mobile IPv6: Protocols and Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2009.
- [41] Zhiwei Yan, Huachun Zhou, Hwang-Cheng Wang, Hongke Zhang, and Sidong Zhang. Design and implementation of a hybrid MIPv6/PMIPv6-based mobility management architecture. *Mathematical and Computer Modelling*, 53(3-4):421 – 442, 2011.
- [42] H. Soliman. Mobile IPv6 Support for Dual Stack Hosts and Routers. RFC 5555 (Proposed Standard), June 2009.
- [43] Dong-Hoon Shin, D. Moses, M. Venkatachalam, and S. Bagchi. Distributed mobility management for efficient video delivery over all-ip mobile networks: Competing approaches. *Network, IEEE*, 27(2):28–33, 2013.
- [44] 3GPP. Mobility management based on Dual-Stack Mobile IPv6; Stage 3 TS 24.303, 3rd Generation Partnership Project (3GPP), 2014.
- [45] Hyon-Young Choi, Kwang-Ryoul Kim, Hyo-Beom Lee, Sung-Gi Min, and Youn-Hee Han. Smart Buffering for seamless handover in Proxy Mobile IPv6. *Wireless Communications and Mobile Computing*, 11(4):491–499, 2011.
- [46] LiJun Zhang, WeiKian Chen, Yan Zhang, Alejandro Quintero, and Samuel Pierre. Seamless Mobility Management Schemes for IPv6-based Wireless Networks. *Mobile Networks and Applications*, 19(6):745–757, 2014.
- [47] Konstantinos Samdanis and A. Hamid Aghvami. Qos provision in wireless access networks: a routing perspective considering mobility. *International Journal of Network Management*, 19(5):445–456, 2009.
- [48] J. Kempf. Goals for Network-Based Localized Mobility Management (NETLMM). RFC 4831 (Informational), April 2007.
- [49] 3GPP. Proxy Mobile IPv6 based Mobility and Tunnelling protocols; Stage 3. TS 29.275, 3rd Generation Partnership Project (3GPP), 2015.
- [50] T. Condeixa and S. Sargento. Centralized, Distributed or Replicated IP Mobility? *Communications Letters, IEEE*, 18(2):376–379, February 2014.
- [51] H. Chan, D. Liu, P. Seite, H. Yokota, and J. Korhonen. Requirements for Distributed Mobility Management. RFC 7333 (Informational), August 2014.

- [52] Cisco Systems Inc. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018 - Cisco, June 2014.
- [53] P. Bertin, S. Bonjour, and J. M. Bonnin. Distributed or Centralized Mobility? In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–6, 2009.
- [54] D. Liu, J. C. Zuniga, P. Seite, H. Chan, and C. J. Bernardos. Distributed Mobility Management: Current practices and gap analysis. RFC 7429, January 2015.
- [55] PetroPasha Ernest, H. Anthony Chan, Olabisi Emmanuel Falowo, and Linoh A. Magagula. Distributed Mobility Management with Distributed Routing Management at Access Routers for Network-Based Mobility Support. *Wireless Personal Communications*, pages 1–25, 2015.
- [56] Jong-Hyouk Lee, Jean-Marie Bonnin, and Xavier Lagrange. Host-based distributed mobility management support protocol for ipv6 mobile networks. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2012 IEEE 8th International Conference on*, pages 61–68, 2012.
- [57] Jong-Hyouk Lee, J. Bonnin, Ilsun You, and Tai-Myoung Chung. Comparative Handover Performance Analysis of IPv6 Mobility Management Protocols. *Industrial Electronics, IEEE Transactions on*, 60(3):1077–1088, 2013.
- [58] Hassan Ali-Ahmad, Meryem Ouzzif, Philippe Bertin, and Xavier Lagrange. Performance Analysis on Network-Based Distributed Mobility Management. *Wireless Personal Communications*, 74(4):1245–1263, 02/01 2014.
- [59] Tiago Condeixa and Susana Sargento. Context-aware adaptive IP mobility anchoring. *Computer Networks*, 71(0):84–99, 10/4 2014.
- [60] F. Giust, C. J. Bernardos, and A. de la Oliva. Analytic Evaluation and Experimental Validation of a Network-Based IPv6 Distributed Mobility Management Solution. *Mobile Computing, IEEE Transactions on*, 13(11):2484–2497, 2014.
- [61] A. Dev Pragad, Vasilis Friderikos, Paul Pangalos, and A. Hamid Aghvami. QoS aware dynamic route optimization for Proxy Mobile IPv6 networks. *Wireless Communications and Mobile Computing*, 11(4):508–521, 2011.
- [62] Esam Alnasouri, Andreas Mitschele-Thiel, and Ali Diab. Handling Mobility Management and QoS Aspects in All-IP Networks. *Mobile Networks and Applications*, 16(1):83–95, 2011.

- [63] R. Langar, N. Bouabdallah, and R. Boutaba. A comprehensive analysis of mobility management in mpls-based wireless access networks. *Networking, IEEE/ACM Transactions on*, 16(4):918–931, 2008.
- [64] E. Rosen, A. Viswanathan, and R. Callon. Multiprotocol Label Switching Architecture. RFC 3031, January 2001.
- [65] A. Belhoul, Y.A. Sekercioglu, and Nallasamy Mani. Interaction of RSVP and Mobile IPv6 Protocols: An Evaluation of Performance on Wireless Networks. In *Communications, 2006. ICC '06. IEEE International Conference on*, volume 3, pages 973–978, June 2006.
- [66] A-E M. Taha, H. S. Hassanein, and H. T. Mouftah. Extensions for Internet QoS paradigms to mobile IP: a survey. *Communications Magazine, IEEE*, 43(5):132–139, 2005.
- [67] R. Coltun. The OSPF Opaque LSA Option. RFC 2370 (Proposed Standard), July 1998. Obsoleted by RFC 5250, updated by RFC 3630.
- [68] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow. RSVP-TE: Extensions to RSVP for LSP Tunnels. RFC 3209 (Proposed Standard), December 2001. Updated by RFCs 3936, 4420, 4874, 5151, 5420, 5711, 6780, 6790.
- [69] Francesco Palmieri. An MPLS-based architecture for scalable QoS and traffic engineering in converged multiservice mobile IP networks. *Computer Networks*, 47(2):257–269, 2/4 2005.
- [70] Victor C. M. Leung Kaiduan Xie, Vincent W. S. Wong. Support of micro-mobility in mpls-based wireless access networks. *IEICE Transactions on Communications*, E88-B No.7:2735–2742, 2005.
- [71] R. G. Garroppo, S. Giordano, and L. Tavanti. Network-based micro-mobility in wireless mesh networks: Is mpls convenient? In *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pages 1–5, 2009.
- [72] Sahel Alouneh, Saed Abed, Mazen Kharbutli, and BassamJamil Mohd. MPLS technology in wireless networks. *Wireless Networks*, 20(5):1037–1051, 2014.
- [73] Zhong Ren, Chen-Khong Tham, Chun-Choong Foo, and Chi-Chung Ko. Integration of mobile IP and multi-protocol label switching. In *Communications, 2001. ICC 2001. IEEE International Conference on*, volume 7, pages 2123–2127 vol.7, 2001.
- [74] C. Makaya and Samuel Pierre. An Analytical Framework for Performance Evaluation of IPv6-Based mobility Management Protocols. *Wireless Communications, IEEE Transactions on*, 7(3):972–983, 2008.

- [75] R. Koodli. Mobile IPv6 Fast Handovers. RFC 5568, July 2009.
- [76] H. Soliman, C. Castelluccia, K. ElMalki, and L. Bellier. Hierarchical Mobile IPv6 (HMIPv6) Mobility Management. RFC 5380, October 2008.
- [77] R. Langar, N. Bouabdallah, R. Boutaba, and B. Sericola. Proposal and analysis of adaptive mobility management in ip-based mobile networks. *Wireless Communications, IEEE Transactions on*, 8(7):3608–3619, 2009.
- [78] Jong-Hyouk Lee, T. Ernst, and Tai-Myoung Chung. Cost analysis of IP mobility management protocols for consumer mobile devices. *Consumer Electronics, IEEE Transactions on*, 56(2):1010–1017, 2010.
- [79] Jong-Hyouk Lee, J. M Bonnin, P. Seite, and H. A. Chan. Distributed IP mobility management from the perspective of the IETF: motivations, requirements, approaches, comparison, and challenges. *Wireless Communications, IEEE*, 20(5):159–168, 2013.
- [80] Kantubukta Vasu, Sudipta Mahapatra, and CheruvuSiva Kumar. A Comprehensive Framework for Evaluating IPv6 Based Mobility Management Protocols. *Wireless Personal Communications*, 78(2):943–977, 09/01 2014.
- [81] Kashif Munir, Xavier Lagrange, Philippe Bertin, Karine Guillouard, and Meryem Ouzzif. Performance analysis of mobility management architectures in cellular networks. *Telecommunication Systems*, pages 1–17, 12/31 2014.
- [82] E. Hyytia, P. Lassila, and J. Virtamo. Spatial node distribution of the random waypoint mobility model with applications. *Mobile Computing, IEEE Transactions on*, 5(6):680–694, 2006.
- [83] Christian Bettstetter, Hannes Hartenstein, and Xavier Pérez-Costa. Stochastic Properties of the Random Waypoint Mobility Model. *Wireless Networks*, 10(5):555–567, 09/01 2004.
- [84] Yuguang Fang. Movement-based mobility management and trade off analysis for wireless mobile networks. *Computers, IEEE Transactions on*, 52(6):791–803, 2003.
- [85] Leonard Kleinrock. Queueing systems, volume I: theory. 1975.
- [86] Oracle Corporation. LTE Diameter Signaling Index. An Oracle White Paper, September 2014.
- [87] Rami Langar, Samir Tohme, and Nizar Bouabdallah. Mobility management support and performance analysis for wireless MPLS networks. *International Journal of Network Management*, 16(4):279–294, 2006.

- [88] W. Sa-Ngiamsak, P. Krachodnok, and R. Varakulsiripunth. A Recovery Scheme for QoS Guaranteed Mobile IP Over MPLS Network. In *Wireless Pervasive Computing, 2006 1st International Symposium on*, pages 1–5, 2006.
- [89] R. Koodli. Fast Handovers for Mobile IPv6. RFC 4068 (Experimental), July 2005. Obsoleted by RFC 5268.
- [90] O. Tipmongkolsilp, S. Zaghloul, and A. Jukan. The evolution of cellular backhaul technologies: Current issues and future trends. *Communications Surveys & Tutorials, IEEE*, 13(1):97–113, 2011.
- [91] Z. Ghebretensae, J. Harmatos, and K. Gustafsson. Mobile broadband backhaul network migration from tdm to carrier ethernet. *Communications Magazine, IEEE*, 48(10):102–109, 2010.
- [92] N. Passas, A. K. Salkintzis, K. D. Wong, and V. K. Varma. Architectures and protocols for mobility management in all-ip mobile networks [guest editorial]. *Wireless Communications, IEEE*, 15(2):6–7, 2008.
- [93] F. M. Chiussi, D. A. Khotimsky, and S. Krishnan. Mobility management in third-generation all-IP networks. *Communications Magazine, IEEE*, 40(9):124–135, 2002.
- [94] T. Anjali and C. Scoglio. Traffic routing in MPLS networks based on QoS estimation and forecast. In *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, volume 2, pages 1135–1139 Vol.2, 2004.
- [95] L. Hundessa and J. Domingo-Pascual. Reliable and fast rerouting mechanism for a protected label switched path. In *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, volume 2, pages 1608–1612 vol.2, 2002.
- [96] Francisco-Javier Rodríguez-Pérez, José-Luis González-Sánchez, Javier Carmona-Murillo, and David Cortés-Polo. An OAM function to improve the packet loss in MPLS-TP domains for prioritized QoS-aware services. *International Journal of Communication Systems*, 28(6):1037–1052, 2015.
- [97] Francisco-Javier Rodríguez-Pérez, José-Luis González-Sánchez, David Cortés-Polo, and Javier Carmona-Murillo. A Delay-Oriented Prioritization Policy Based on Cooperative Lossless Buffering in PTN Domains. *Journal of Network and Systems Management*, pages 1–18, 2014.
- [98] I. Ali, A. Casati, K. Chowdhury, K. Nishida, E. Parsons, S. Schmid, and R. Vaidya. Network-based mobility management in the evolved 3GPP core network. *Communications Magazine, IEEE*, 47(2):58–66, 2009.

- [99] Shou-Chih Lo, Guanling Lee, Wen-Tsuen Chen, and Jen-Chi Liu. Architecture for mobility and QoS support in all-IP wireless networks. *Selected Areas in Communications, IEEE Journal on*, 22(4):691–705, 2004.
- [100] H. Chan, D. Liu, P. Seite, H. Yokota, and J. Korhonen. Requirements for Distributed Mobility Management. Internet-Draft, August 2013.
- [101] PetroPasha Ernest, H. A. Chan, Jiang Xie, and OlabisiEmmanuel Falowo. Mobility management with distributed mobility routing functions. *Telecommunication Systems*, pages 1–18, 01/07 2015.
- [102] Tiago Condeixa and Susana Sargento. Studying the integration of distributed and dynamic schemes in the mobility management. *Computer Networks*, 60(0):46–59, 2/26 2014.
- [103] RadhikaRanjan Roy. *Mobility Model Characteristics*, pages 23–32. Springer US, 01/01 2011.
- [104] Injong Rhee, Minsu Shin, Seongik Hong, Kyunghan Lee, and Song Chong. On the Levy-Walk Nature of Human Mobility. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, 2008.
- [105] H. Wang, S. Chen, H. Xu, M. Ai, and Y. Shi. SoftNet: A software defined decentralized mobile network architecture toward 5G. *Network, IEEE*, 29(2):16–22, March 2015.
- [106] Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '11*, pages 47–56, New York, NY, USA, 2011. ACM.
- [107] Statista. Average mobile app session length 2014, by category, <http://www.statista.com/statistics/202485/average-ipad-app-session-length-by-app-categories/>, September 2014.
- [108] UMIP. Mobile IPv6 and NEMO for Linux, 2014.
- [109] Open Air Interface Proxy Mobile IPv6 (OAI PMIPv6), 2014.
- [110] Jianfeng Guan, Huachun Zhou, Zhiwei Yan, Yajuan Qin, and Hongke Zhang. Implementation and analysis of proxy MIPv6. *Wireless Communications and Mobile Computing*, 11(4):477–490, 2011.
- [111] G. Malkin and R. Minnear. RIPng for IPv6. RFC 2080 (Proposed Standard), January 1997.

-
- [112] Huynh-Thu Quan and Ghanbari Mohammed. The accuracy of PSNR in predicting video quality for different video scenes and frame rates, 2012.
 - [113] M. Fiedler, T. Hossfeld, and P. Tran-Gia. A generic quantitative relationship between quality of experience and quality of service. *Network, IEEE*, 24(2):36–41, 2010.
 - [114] François Baccelli, Maurice Klein, Marc Lebourges, and Sergei Zuyev. Stochastic geometry and architecture of communication networks. *Telecommunication Systems*, 7(1-3):209–227, 1997.
 - [115] Xueli An, Fabio Pianese, Indra Widjaja, and Utku Günay Acer. DMME: A Distributed LTE Mobility Management Entity. *Bell Labs Technical Journal*, 17(2):97–120, 2012.
 - [116] D. Kotz and T. Henderson. CRAWDAD: A Community Resource for Archiving Wireless Data at Dartmouth. *Pervasive Computing, IEEE*, 4(4):12–14, Oct 2005.
 - [117] Nils Aschenbruck, Aarti Munjal, and Tracy Camp. Trace-based mobility modeling for multi-hop wireless networks. *Computer Communications*, 34(6):704 – 714, 2011.